

Cite this: DOI: 10.1039/d1ob00698c

Deciphering protein microenvironment by using a cysteine specific switch-ON fluorescent probe†

Jessy Mariam,^a Anila Hoskere Ashoka,^{‡b} Vandana Gaded,^a Firoj Ali,^{§b} Harshada Malvi,^a Amitava Das^{||b,c} and Ruchi Anand^{||*a}

Fluorescent probes provide an unparalleled opportunity to visualize and quantify dynamic events. Here, we employ a medium-size, cysteine specific coumarin based switch-ON fluorescent probe 'L' to track protein unfolding profiles and accessibility of cysteine residues in proteins. It was established that 'L' is highly selective and exhibits no artifact due to interaction with other bystander species. 'L' is able to gauge subtle changes in protein microenvironment and proved to be effective in delineating early unfolding events that are difficult to otherwise discern by classic techniques such as circular dichroism. By solving the X-ray structure of TadA and probing the temperature dependent fluorescence-ON response with native TadA and its cysteine mutants, it was revealed that unfolding occurs in a stage-wise manner and the regions that are functionally important form compact sub-domains and unfold at later stages. Our results assert that probe 'L' serves as an efficient tool to monitor subtle changes in protein structure and can be employed as a generic dye to study processes such as protein unfolding.

Received 9th April 2021,

Accepted 21st May 2021

DOI: 10.1039/d1ob00698c

rsc.li/obc

Introduction

Labeling proteins with fluorescent probes is a rapid and economic approach to selectively track their dynamics and interactions *in vitro* as well as *in vivo*,^{1,2} especially when proteins lack intrinsic fluorescent amino acids or harbor moieties with very low intrinsic quantum yields. The side chains of cysteine,^{3,4} lysine,⁵ and histidine^{6,7} are generally used as the target sites for fluorophore attachment in proteins. Among these, cysteine chemistry is the most exploited for achieving the desired site-specificity in tagging. The reactivity of the free sulfhydryl moiety of a cysteine residue in a protein allows itself to participate in a range of chemical transformations. Additionally, site-directed mutagenesis offers scope for site-specific labels for studying the protein structure and establishes a structure-activity correlation.^{8,9} Thus combining the above two approaches, we can probe structural fluctuations in

real-time and specifically study select processes such as protein structure and folding. Over the years, several cysteine specific probes have been utilized to monitor protein microenvironment, such as fluorescein-5-maleimide,¹⁰ 5-iodoacetamido-fluorescein,¹¹ 1,5-IAEDANS¹² *etc.* However, the major drawback of these probes is that they remain constitutively ON. This property adds to background noise and therefore careful post labeling purification becomes important. An alternative approach involves the use of fluorescent probes that show switch-ON fluorescence response. Such probes selectively turn ON, only upon specific reaction (*e.g.* with the free sulfhydryl groups in cysteine residue of a protein) and the unreacted free molecular probe remains fluorescently silent.^{13–19} This eliminates the need for post labeling purification steps, thereby minimizing protein loss. There are several cysteine specific switch-ON fluorescent probes that are employed for endogeneous cell targeting^{13,18,20–22} but only a handful of such switch-ON fluorescent probes^{14,23–25} (listed in ESI Table S1†) have been established for tracking protein unfolding. Thus, there is an emerging trend to develop and design such turn ON probes that would be highly suitable for studying protein folding and dynamics.

Consequently, in this study, we have used a recently developed medium-sized switch-ON coumarin based probe, named L²⁶ (Fig. 1A). This probe has a coumarin-based signaling unit and a nitro-olefin trigger that favours the Michael-type addition reaction with cysteine.²⁶ Probe L exhibits high selectivity towards cysteine while being unreactive towards structurally similar molecules such as glutathione, homocysteine, *etc.*²⁶ In fact probe L labels well in a physiological buffer

^aDepartment of Chemistry, IIT Bombay, Mumbai-400076, India.

E-mail: ruchi@chem.iitb.ac.in

^bAnalytical Science Discipline, CSIR-Central Salt and Marine Chemicals Research Institute, G.B. Marg, Bhavnagar: 364002, Gujarat, India

^cDepartment of Chemical Sciences, Indian Institute of Science and Education Research, KolkataMohanpur: 742246, India. E-mail: amitava@iiserkol.ac.in

†Electronic supplementary information (ESI) available. See DOI: 10.1039/d1ob00698c

‡Present address: Laboratoire de Bioimagerie et Pathologies, UMR 7021 CNRS, Université de Strasbourg, 74 route du Rhin, 67401, Illkirch, France.

§Present address: CSIR – Central Institute of Mining and Fuel Research, Barwa Road, CIMFR Colony, Dhanbad, Jharkhand 826015, India.

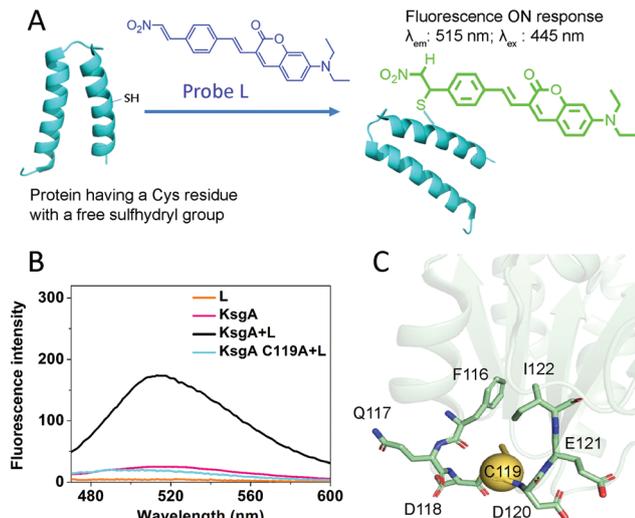


Fig. 1 Fluorescence ON response of L in a single cysteine protein. (A) Schematic representation of the reaction of L with a Cys-residue containing a free sulfhydryl group. (B) Fluorescence emission spectra of L, KsgA, KsgA + L and KsgA C119A + L (equimolar concentration of L and protein *i.e.* 20 μ M) in appropriate aqueous buffer [25 mM HEPES (pH 7.4); 300 mM NaCl]. (C) Zoom view of the crystal structure of KsgA (PDB ID: 6IFS) showing the local environment of the cysteine (golden sphere) C119.

medium and importantly, this reagent is silent to all other amino acids, cations, and anions commonly used in physiological buffers.²⁶ Although L has been demonstrated to be internalized in cells,²⁶ the focus of the current work is to establish its utility in elucidating *in vitro* protein microenvironments. We monitor unfolding profiles of single and multidomain proteins, by determining their cysteine accessibility profiles as a function of temperature, thereby proposing a strategy to augment the existing knowledge gap.

Results and discussion

To demonstrate the selectivity and wide applicability of this probe in a protein environment we have chosen three model proteins [KsgA (Kasugamycin resistance protein),²⁷ TadA (tRNA-specific adenosine deaminase) and PurL (phosphoribosylformylglycinamide synthase)²⁸] containing varying number of cysteine residues. These proteins were chosen as they play critical roles in important RNA/DNA modifications, essential for cell survival and hence tracking their unfolding profiles provide important insights. The three proteins were purified to near homogeneity (Fig. S1†) and subsequently, it was confirmed that they do not show any background fluorescence in the buffer systems and salts chosen for our study (Fig. S2A†). Probe L also displayed negligible background fluorescence in its free form (Fig. 1B, and S2B†).

Site-specific labeling of cysteine residues in proteins by probe L

To check the efficacy of the chemodosimetric probe molecule L toward the Cys-residue and the associated fluorescence ON

response, we first monitored the interaction of L with protein KsgA that contains a single solvent accessible cysteine residue (C119) (Fig. 1C). Fig. 1B shows the fluorescence emission spectra of KsgA and L with the signature of the L-Cys adduct. Efficient tagging is evident from the 170-fold enhancement in fluorescence intensity upon reaction. It appears that since C119 is present in a surface exposed loop region that connects an α -helix and a β -strand (Fig. 1C, and S3†), probe L is efficiently able to tag. Moreover, since C119 is proximal to residues F116 and I122 (Fig. 1C), the hydrophobic environment may additionally favor tagging by allowing L to stack. To further prove that this response is specific to the free sulfhydryl moiety of the respective cysteine residue, a control experiment was performed by mutating the only cysteine, C119 to an alanine. The C119A mutant, of KsgA shows (Fig. 1B) no fluorescence, thus asserting the high selectivity and sensitivity of this probe.

Performance of probe L in a multi-cysteine, multidomain protein environment

To gain deeper insights we next used a multi-cysteine protein PurL (phosphoribosylformylglycinamide synthase) which is a very large 140 kDa, multidomain protein system with 13 cysteine residues (Fig. 2A).²⁸ PurL, is part of the purine biosynthetic pathway and catalyzes a crucial step in the production of DNA base inosine.²⁸ It is an allosterically regulated bi-functional enzyme that harbors a transient ammonia tunnel^{29,30} and has four domains namely, N-terminal, linker, FGAM synthetase, and glutaminase domain with different number of cysteine residues (Fig. 2A). The N-terminal domain has a single cysteine residue (C98) while no cysteines are present in the small linker domain. FGAM synthetase domain has 9 cysteines wherein six of them are buried (C481, C488, C544, C567, C680, C765) and the other three residues (C217, C903, and C951) are close to the surface (Fig. 2A). The C-terminal glutaminase domain has three cysteines with a single cysteine (C1001) being solvent accessible and the other two (C1091 and C1138) are deeply buried. This protein was subject to tagging and on incubation with L, enhancement of fluorescence intensity for a characteristic L-Cys-residue adduct appeared at 25° C (Fig. 2B). We believe that this fluorescence could be attributed to the tagging of cysteines such as C217 in the FGAM synthetase domain and other cysteine residues such as C903, C951, and C1001. Thus, the initial fluorescence response can be accredited to these surface exposed cysteines. As the temperature was raised to 40 °C, approximately four-fold enhancement in fluorescence was observed (Fig. 2B). The rise in fluorescence at 40 °C can be attributed to the local unfolding events. Thus after 40 °C the protein rapidly starts to unfold, exposing the buried cysteine residues, leading to the creation of additional tagging sites by the chemodosimetric reagent L. The steep rise in fluorescence is indicative of the local opening of the tertiary structure/increased dynamics that is captured *via* progressive tagging by probe L of buried cysteines as the protein core becomes loose (Fig. 2C). This information is exclusively gained

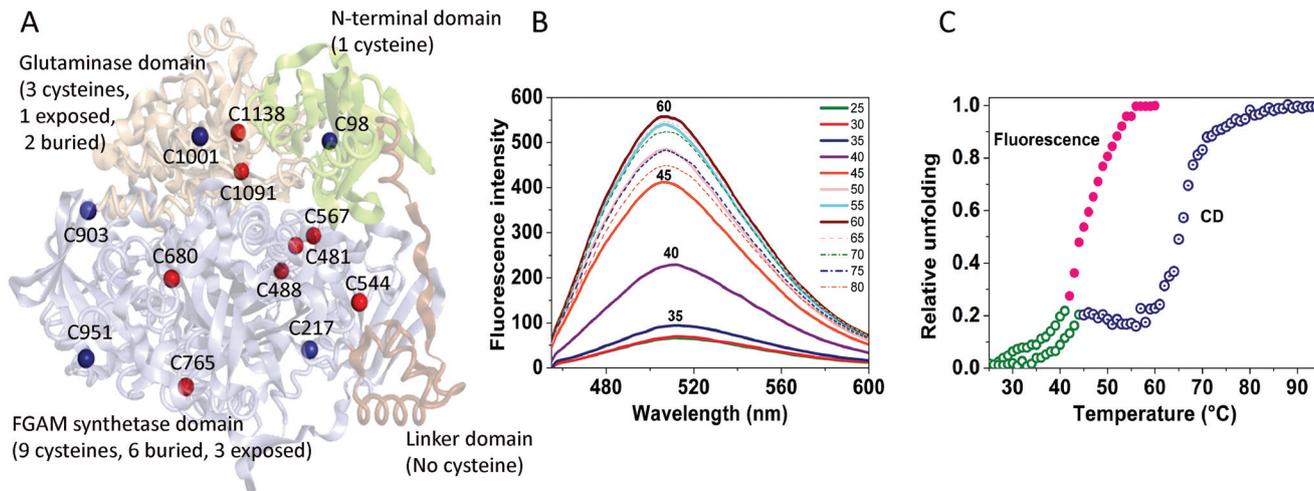


Fig. 2 Monitoring unfolding of multi-cysteine protein PurL with probe L. (A) Crystal structure of PurL (PDB ID: 1T3T) showing the four domains with surface exposed (navy blue spheres) and buried (red spheres) cysteine residues. (B) Fluorescence emission spectra of L and PurL (5 μ M) at varying temperatures (25 $^{\circ}$ C–80 $^{\circ}$ C), λ_{ex} = 445 nm (C) Thermal unfolding profile of PurL monitored using fluorescence (PurL tagged with L) and circular dichroism (untagged PurL). Sharp rise in fluorescence intensity is attributed to the local unfolding of protein structure.

by probe L and is missing from CD measurements that can only apprehend secondary structure changes (Fig. 2C). To confirm that no interference from temperature induced fluorescence of L is observed, we monitored switch-ON fluorescence of free probe L, in the temperature range of 25–80 $^{\circ}$ C and show that without complexation with cysteine the probe is fluorescently silent (Fig. S4 \dagger).

Tracking accessibility of cysteine residues in protein *via* probe L, a structure-guided approach

To develop deeper insights into understanding specific tagging of cysteines and their accessibility profiles we decided to pursue a comprehensive study of the protein TadA, with a relatively manageable number of cysteines. Moreover, it was considered an apt choice as TadA is a central protein in RNA editing that deaminates the wobble position, of adenosine 34 (A34) in the anticodon stem-loop of tRNA to inosine (Fig. S5 \dagger).³¹ Thus, understanding how its protein structure responds to temperature based denaturation could provide clues to the regions that are important for function. This protein consists of five cysteine residues *i.e.* C63, C83, C86, C109, and C137 (Fig. 3A). In order to get insights into the spatial configuration and structural coordinates of these cysteines, we first solved the X-ray crystallographic structure of TadA from *Bacillus subtilis*. The structure of TadA (PDB ID: 7CPH) was determined by molecular replacement to a resolution of 2.3 \AA (details of the data collection statistics are given in Table 1, Fig. 3A). Structural inspection shows that TadA exhibits the typical cytidine deaminase-like (CDA) fold consisting of a central five-stranded mixed β -sheet flanked by helices (Fig. 3A).³² The active site harbors an ordered catalytic zinc ion that forms a tetrahedral coordination engaging two of the cysteine residues, C83 (2.3 \AA), C86 (2.3 \AA) (Fig. S6 \dagger). The other cysteines are scattered over different regions of the protein.

C63 is located in α 3, C109 is present in a loop connecting strands β 4– β 5 and C137 is present on helix α 5. Since these cysteines are positioned at distinct sites, TadA seemed to be an appropriate model to track temperature dependent changes in protein structure. On comparing the fluorescence of L upon incubation with TadA at varying temperatures, it could be seen that the CD and fluorescence graphs are mostly overlapping till 50 $^{\circ}$ C with no dramatic changes (Fig. 3B). After this temperature, similar to the scenario observed for PurL, the tertiary structure starts opening up, which is indicated by a rapid increase in the fluorescence signal. A non-overlap of the unfolding transitions was observed at higher temperatures (Fig. 3B) implying that each of these techniques captured different transition intermediates of the unfolding process.

To further probe which regions are most sensitive to temperature fluctuations, cysteines were progressively replaced by alternative amino acids. Mutations were designed using the local environment of these cysteines as a guide, Fig. 3C–E. Three single cysteine mutants, *viz.* C63S, C109A, C137S, and a double mutant C63A C109A (DM) were created. Mutants with the two zinc bound cysteines (Fig. S6 \dagger) could not be pursued as mutation in this region resulted in no protein production likely due to the instability of the core induced due to improper coordination of the zinc ion. Fig. 4A represents the thermal unfolding profile of native TadA and its mutants as a function of temperature. The average fluorescence intensity values of TadA and its mutants at four representative temperatures are depicted in Fig. 4B while their standard deviations are represented in the form of a box plot in Fig. S7 \dagger Fig. 4B reveals that at 25 $^{\circ}$ C, C109 and C63 contribute to a large percentage of the total fluorescence and are the two most fluorescently active amino acids. To confirm this, the spectra where these cysteines are rendered silent were compared with that of the native. The additive fluorescence spectra of both these

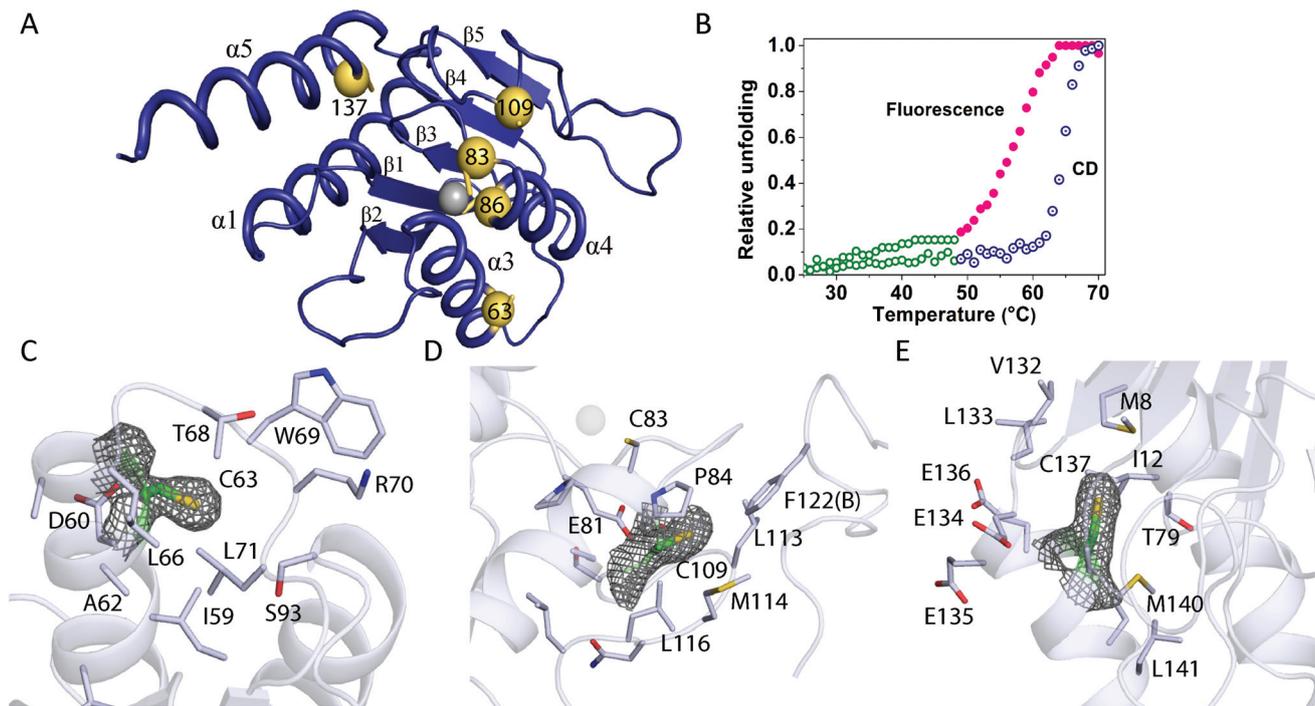


Fig. 3 Structural insights of tRNA-specific adenosine deaminase (TadA). (A) Crystal structure of TadA (PDB ID: 7CPH) showing the five cysteine residues. (B) Thermal unfolding profile of TadA obtained from fluorescence (tagged with L) and circular dichroism spectroscopy (untagged TadA). (C–E) Local environment around cysteine residues in TadA (PDB ID: 7CPH): C63 (C), C109 (D) and C137 (E).

Table 1 Data processing and refinement statistics of native TadA from *Bacillus subtilis*

Protein	Native TadA PDB ID – 7CPH
Space group	$P2_12_12_1$
Resolution (Å)	2.3
Multiplicity	4.4 (4.9)
Completeness ^a (%)	91.5 (95.4)
R_{sym}^a (%)	12.6 (51.4)
I/σ	7.9 (2.8)
Total no. of reflections	61 225
No. of unique reflections	13 799
Refinement	
Resolution range (Å)	38.19–2.30
No. of reflections total	13 741
No. of reflections test set	985
$R_{\text{work}}/R_{\text{free}}$ (%)	17.67/23.96
No. of atoms	
Total	2321
Protein	2269
Ligands	12
Ion	2
Water	38
RMSD	
Bond lengths (Å)	0.007
Bond angles (°)	0.890
Ramachandran plot	
Most favored region (%)	96.05
Additionally allowed region (%)	3.95
Outliers (%)	0.0

^a Values in parentheses represent the data in the highest-resolution shell.

mutants C109A and C63A coincided with the spectra of native TadA (Fig. 5A) corroborating that maximal signal arises from these two cysteines at 25 °C. Briefly, to explain the observations, in C109A mutant the signal could arise from C63 or C137 or both while in C63S the signal could arise from C109 or C137 or both. This implies that C63, C109, and C137 are the only potential tagging sites for L. In order to mask the signal originating from C137, we created C137S mutant. On comparing the fluorescence intensity of C137S it was observed to be close to that of native TadA (Fig. 4B) Thus, we can conclude that the signal from C137S emanates solely from C63 and C109, and C137 is essentially untagged at 25 °C.

To analyze and understand the temperature dependent pattern of tagging, the crystal structure of TadA was used as a guide to probe solvent accessibility. Further, to corroborate the observations, hydrophobicity score of the residues around the three cysteine residues was also analyzed and plotted (Fig. 5B). It was observed that C63 was most solvent exposed (Fig. 3C) and is also most hydrophilic (Fig. 5B) and hence absence of this cysteine leads to largest decrease in fluorescence intensity. Analysis shows that C109 which also tags efficiently at 25 °C, is present on a loop and the hydrophobicity profile shows that it lies at a cusp of a hydrophobic section followed by a hydrophilic stretch (Fig. 5B). Since L has aromatic hydrophobic rings they likely pack against residues P84, F122 and L113 that line the C109 pocket (Fig. 3D), allowing L to sit comfortably near C109. C137, on the other hand, is in a helix that is packed in a neutral, moderately hydrophobic environment,

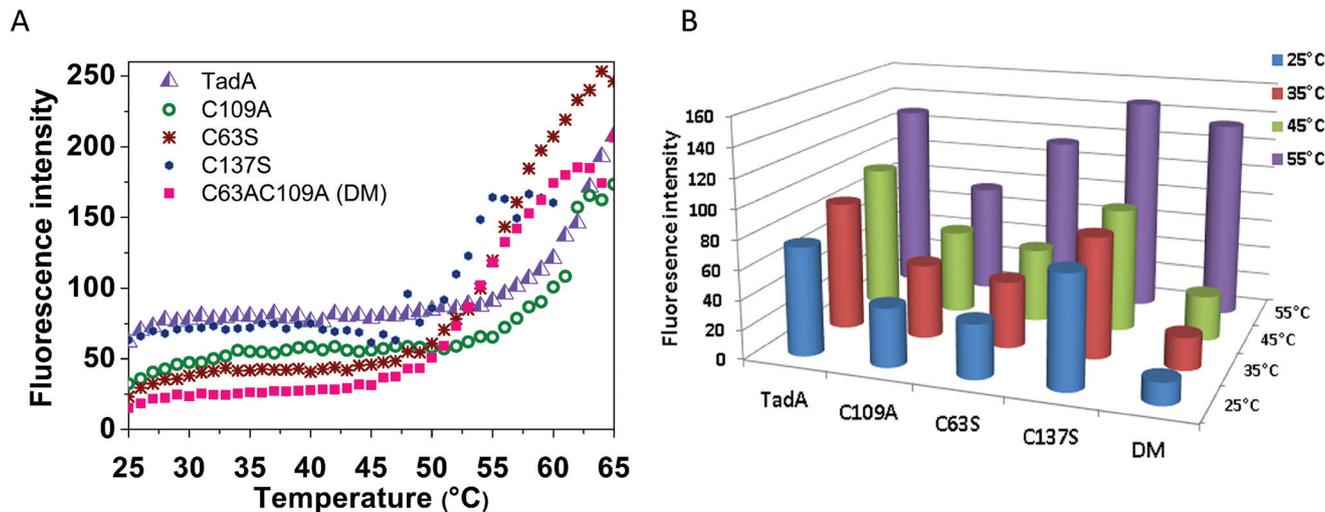


Fig. 4 Analysis of thermal unfolding of TadA with probe L. (A) Thermal unfolding profile of TadA and its mutants obtained using probe L, $\lambda_{\text{ex}} = 445$ nm. (B) Bar diagram showing the average fluorescence emission intensity of TadA and its mutants at four representative temperatures of 25, 35, 45 and 55 °C, the standard deviations are shown in Fig. S7.†

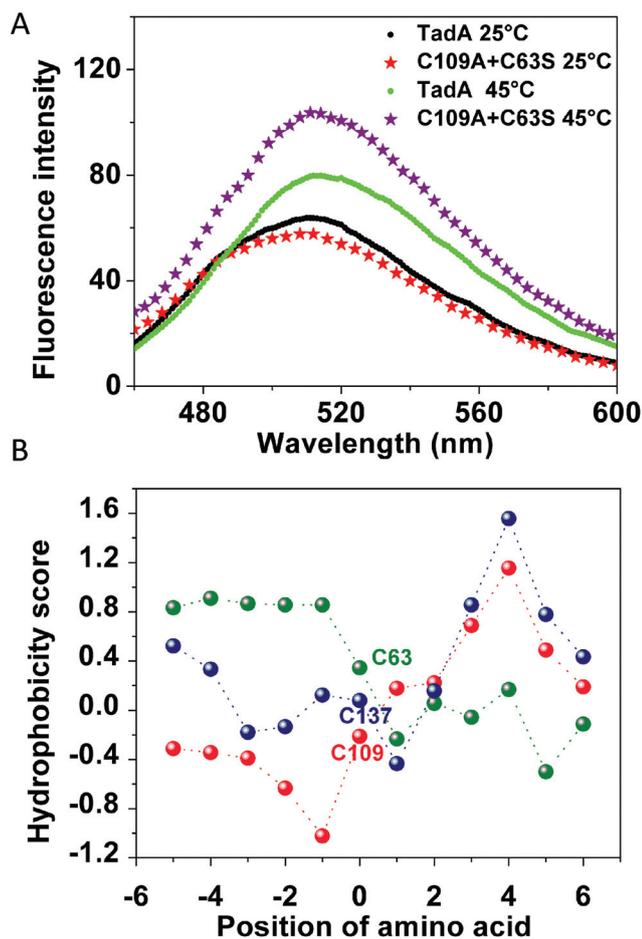


Fig. 5 Biophysical analysis of TadA and its cysteine mutants. (A) Additive fluorescence emission spectra of C109A and C63S mutants overlaid with the fluorescence emission spectra of native TadA at 25 and 45 °C. (B) Hydrophobicity plot of a stretch of 20 residues of TadA with the corresponding cysteine at the 0th position. Numbers to the left and right of zero denote the residues preceding and succeeding the corresponding cysteine.

sandwiched between two helices with almost no space for the probe to enter. The compact cavity that harbours C137 allows it to interact with methionine, threonine, and isoleucine residues that shield it from the outside (Fig. 3E). Hence, the probe is unable to penetrate here and no tagging is observed.

In the double mutant C63A C109A, since the two prominently tagged cysteines are rendered silent, the sole contributor to fluorescence is C137. Analysis of the fluorescence profile of the double mutant supports the above finding that C137 is not significantly tagged at 25 °C (Fig. 4B). The temperature dependent tagging profile shows that C137 tags at 45 °C. At this temperature, the additive spectra of C63S and C109A no longer overlaps with the original native TadA spectra but is observed to add up to a higher value (Fig. 5A). Thus at 45 °C the helix harboring C137 becomes accessible. This observation points to the temperature where structurally compact regions start becoming exposed. Earlier reports have shown that the helix harbouring C137 is important for the presentation of the RNA to the protein and important for activity (Fig. S8†).³² This shows that functionally important regions form more compact structures and are less accessible to the outside environment. An advantage of the probe is that it reports early changes in unfolding, however, a drawback is that spectroscopic quantification is hindered at higher temperatures, in the regime where aggregation takes place.

Viscosity and polarity effects on probe L

To further explore whether L can sense changes in viscosity or polarity, the fluorescence emission spectra of L in glycerol and dioxane were probed. It was established that even 70% glycerol resulted in no switch-ON fluorescence of L (Fig. S9A†). However, fluorescence was induced in the presence of dioxane (Fig. S9B†), a non-polar solvent. Thus L was unaffected by alterations in viscosity but was sensitive to the changes in

polarity. This corroborates with our observation where at temperatures, beyond 50–55 °C, when the hydrophobic core of the protein becomes accessible, the probe shows a sharp enhancement of fluorescence (Fig. 4B). Proteins with higher instability are expected to aggregate earlier, providing more hydrophobic hotspots, which subsequently enhance the fluorescence of L.

Conclusions

To conclude, switch-ON fluorescent responses of a chemodosimetric molecular probe L was used here to specifically tag various proteins (with a free sulfhydryl group in the Cys moiety) in their native state. It demonstrated efficient tagging for accessible cysteine residues at room temperature, even in absence of an excess probe. Moreover, the molecular properties of L make it an excellent reporter to monitor temperature dependent changes in protein structure. This probe therefore has great potential to be employed further for detailed protein unfolding pathway studies.

Experimental

The proteins were cloned, as described in the ESI (page S1†), followed by expression and purification as follows.

Protein purification

KsgA and PurL. Both proteins were expressed and purified in a similar way as described below for TadA. The detailed protocol and buffer concentrations are as described in Bhujbalrao *et al.*²⁷ and Tanwar *et al.*³⁰

TadA and its mutants. An overnight culture of Rosetta cells transformed with TadA plasmid (or the mutant) was inoculated into 1.2 liters of Luria Bertani broth containing kanamycin (35 µg ml⁻¹) and chloramphenicol (30 µg ml⁻¹) and incubated at 37 °C, 220 rpm. The cells were induced in the mid-log phase (O.D 600 nm ≈ 0.6) with 0.6 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) and grown at 25 °C. The cells were harvested after 6 h by centrifugation (4500 rpm, 20 min) and stored at -80 °C. Cell pellets were thawed quickly and suspended in buffer A (binding buffer): 50 mM HEPES, 500 mM NaCl, 3 mM imidazole, pH 7.5. Cells were then lysed by sonication and the lysate was centrifuged at 14 000 rpm, 4 °C for 1 h. The supernatant was mixed with Ni-NTA resin which was equilibrated with buffer A and placed on a rocker for 1 h. Thereafter the beads were centrifuged at 1200 rpm and loaded onto a column wherein the washing step was carried out with buffer B (wash buffer): 50 mM HEPES, 500 mM NaCl, 30 mM imidazole, pH 7.5. Gradient elution of the protein was performed with buffer containing: 50 mM HEPES, 500 mM NaCl, 100 mM imidazole, followed by buffer with 150 mM imidazole. Fractions containing TadA were pooled and desalted in desalting columns equilibrated with desalting buffer (25 mM HEPES, 300 mM NaCl, 3% glycerol, pH 7.5). TadA was finally eluted with desalting buffer. The pooled fractions were concen-

trated in Corning concentrators and checked for purity by performing SDS-PAGE. The concentration of TadA and its mutants were calculated using absorbance at 280 nm with a molar extinction coefficient of 10 220 L mol⁻¹ cm⁻¹. For crystallization studies, TadA was purified in the same way, except that Tris was used instead of HEPES and 5% glycerol was included in the final desalting buffer.

Synthesis of L

As outlined in Ashoka *et al.*²⁶ Stock solutions of 1 mM L were prepared in dimethylsulfoxide (DMSO).

Crystallization and structure determination of native TadA

The crystallization screens (Hampton research crystal screen) of TadA (around 8 mg ml⁻¹) were set up by mixing 1 µl protein and 1 µl of reservoir solution using the hanging drop vapour diffusion method. Crystal trays were kept in a temperature-controlled cabinet at 20 °C and were monitored regularly. Thin elongated bunch of needle shaped crystals was observed within one week in several conditions. One of the conditions (8% tacsimate pH 6.0, 20% PEG 3350) was optimized by preparing its grid wherein pH of 10% tacsimate was varied from 6.0 to 7.5 and the concentration of PEG 3350 was varied from 18 to 28%. Well isolated diffraction quality crystals were seen in a few of these grid conditions in 15–20 days. Crystals were cryo-protected with a solution composed of mother liquor and 20% glycerol and then flash cooled in liquid nitrogen. The diffraction data for native TadA was collected at the micro-focus beamline (MX2) of the Australian synchrotron. iMOSFLM program was used for indexing, integrating the intensities.³³ Then the intensities were scaled using the program SCALA from CCP4i software.³⁴ The structure was solved at a resolution of 2.3 Å. The crystal (space group *P*₂₁₂₁₂₁; unit-cell parameters, *a*-38.53, *b*-41.25, *c*-202.08, *α*, *β*, *γ* = 90°) contains two molecules (a dimer) per asymmetric unit with a solvent content of 44%. Molecular replacement method was employed to obtain phases using the Phaser program.³⁴ The template model structure of *Staphylococcus aureus* TadA (PDB ID-2B3J)³² was used to obtain phases since it shared a sequence identity of 60% with TadA. The resulting model was further built-in COOT³⁵ and refined in the REFMAC5³⁶ and phenix-refine program.³⁷ Initially, rigid body refinement was performed to fit the model into the experimental data. Hereafter, restrained refinement was carried out to further refine the structure. The data and refinement statistics are listed in Table 1.

Fluorescence measurements

Fluorescence emission spectra of equimolar concentration (20 µM) of L and KsgA (or KsgA C119A) in 25 mM HEPES, 300 mM NaCl, pH 7.4 buffer were recorded on Cary Varian spectrofluorimeter in the wavelength range of 460–600 nm, with an excitation wavelength of 445 nm. For KsgA and KsgA C119A, the samples were prepared by mixing L with the protein and pre-incubating at 37 °C for 20 min before recording the spectra. Excitation and emission slit widths were maintained at 5 nm. The same instrument parameters and buffer composition were maintained for TadA and PurL. For thermo-

profile experiments of L with Tada/PurL, the temperature was ramped at the rate of 1 °C min⁻¹. The proteins PurL (5 μM), Tada and its mutants (20 μM) were pre-incubated with the probe L at 25 °C for 15 minutes and the initial fluorescence spectra were then recorded. For PurL, further fluorescence spectra were recorded at an interval of 5 °C with a hold time of 5 minutes after the desired temperature was attained. In the case of Tada (or mutants) thermoprofile, subsequent fluorescence spectra were recorded at an interval of 1 °C with a hold time of 2 minutes after the target temperature was achieved. Spectra of appropriate controls *i.e.* only buffer and only probe were also recorded. The fluorescence intensity values of at least four data sets were taken for analysis.

Thermoprofile using circular dichroism

Thermal melting curves of PurL and Tada (5 μM) in 25 mM phosphate buffer, pH 7.4 were recorded on Jasco 815 CD spectrophotometer. CD mdeg value at 222 nm was recorded as a function of temperature (25 °C–80 °C). The temperature was ramped at the rate of 1 °C min⁻¹ with a hold time of 120 s.

The unfolding curves obtained from fluorescence (using probe L) and CD were normalized to the apparent unfolded fraction f_U (relative unfolding), using the following equation

$f_U = (Y_{obs} - Y_N)/(Y_u - Y_N)$, where Y_{obs} , Y_u and Y_N is the fluorescence emission intensity (at 515 nm) or the ellipticity (measured at 222 nm) at the observed temperature, temperature of the completely unfolded (denatured) and native state of the protein respectively.

Determination of solvent accessibility and hydrophathy score

The solvent accessibility of cysteine residues was estimated from Pymol. The hydrophathy score of these residues was calculated using ProtScale-ExPasy tool (Kyte and Doolittle method).³⁸ A sequence of 20 residues were chosen with nine and ten residues preceding and succeeding the corresponding cysteine and a window size of 9 was selected.

Author contributions

Jessy Mariam: Data curation, investigation, formal analysis, writing-original draft preparation, visualization; Anila Hoskere Ashoka: investigation; Vandana Gaded: investigation, formal analysis, writing-original draft preparation; Firoj Ali: investigation; Harshada Malvi: validation; Amitava Das: conceptualization, methodology, project administration, resources, supervision, writing-review and editing; Ruchi Anand: conceptualization, methodology, project administration, resources, supervision, writing-original draft preparation, writing-review and editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

JM is grateful to Department of Biotechnology, Govt. of India for the DBT-RA fellowship (2015–2018). AD acknowledges SERB (EMR/2016/001850 & JCB/2017/000004) and DBT India (BT/PR22251/NNT/28/1274/2017) grants for financial support. RA gratefully acknowledges funding from National Women Bioscientists Award 2017 & 2018, Department of Biotechnology, Govt. of India (BT/HRD/NWBA/39/03/2018-19). We acknowledge Mr Subhankar Sahu for his support during the revision stage of the manuscript.

Notes and references

- L. Xue, I. A. Karpenko, J. Hiblot and K. Johnsson, *Nat. Chem. Biol.*, 2015, **11**, 917–923.
- L. Xue, E. Prifti and K. Johnsson, *J. Am. Chem. Soc.*, 2016, **138**, 5258–5261.
- H. S. Jung, X. Chen, J. S. Kim and J. Yoon, *Chem. Soc. Rev.*, 2013, **42**, 6019–6031.
- Y. Kim, S. O. Ho, N. R. Gassman, Y. Korlann, E. V. Landorf, F. R. Collart and S. Weiss, *Bioconjugate Chem.*, 2008, **19**, 786–791.
- M. Chilamari, N. Kalra, S. Shukla and V. Rai, *Chem. Commun.*, 2018, **54**, 7302–7305.
- B. Krishnan, A. Szymanska and L. M. Gierasch, *Chem. Biol. Drug Des.*, 2007, **69**, 31–40.
- S. Lata, M. Gavutis, R. Tampé and J. Piehler, *J. Am. Chem. Soc.*, 2006, **128**, 2365–2372.
- S. Frillingos, M. Sahin-Tóth, J. Wu and H. R. Kaback, *FASEB J.*, 1998, **12**, 1281–1299.
- M. J. Patel, G. Yilmaz, L. Bhatia, E. E. Biswas-Fiss and S. B. Biswas, *MethodsX*, 2018, **5**, 419–430.
- M. C. Puljung and W. N. Zagotta, *Curr. Protoc. Protein Sci.*, 2012, ch. 14, Unit 14.14–Unit 14.14.
- J. C. Lee and P. M. Horowitz, *J. Biol. Chem.*, 1992, **267**, 2502–2506.
- A. M. Saxena, J. B. Udgaonkar and G. Krishnamoorthy, *J. Mol. Biol.*, 2006, **359**, 174–189.
- F. Ali, A. H. Ashoka, N. Taye, R. G. Gonnade, S. Chattopadhyay and A. Das, *Chem. Commun.*, 2015, **51**, 16932–16935.
- E. Branigan, C. Pliotas, G. Hagelueken and J. H. Naismith, *Nat. Protoc.*, 2013, **8**, 2090–2097.
- J. J. Ferrie, N. Ieda, C. M. Haney, C. R. Walters, I. Sungwienwong, J. Yoon and E. J. Petersson, *Chem. Commun.*, 2017, **53**, 11072–11075.
- Z. Guo, S. Nam, S. Park and J. Yoon, *Chem. Sci.*, 2012, **3**, 2760–2765.
- D. Lee, G. Kim, J. Yin and J. Yoon, *Chem. Commun.*, 2015, **51**, 6518–6520.
- Y. H. Lee, W. X. Ren, J. Han, K. Sunwoo, J.-Y. Lim, J.-H. Kim and J. S. Kim, *Chem. Commun.*, 2015, **51**, 14401–14404.

- 19 U. G. Reddy, H. Agarwalla, N. Taye, S. Ghorai, S. Chattopadhyay and A. Das, *Chem. Commun.*, 2014, **50**, 9899–9902.
- 20 C. Han, H. Yang, M. Chen, Q. Su, W. Feng and F. Li, *ACS Appl. Mater. Interfaces*, 2015, **7**, 27968–27975.
- 21 C. Y. Kim, H. J. Kang, S. J. Chung, H.-K. Kim, S.-Y. Na and H.-J. Kim, *Anal. Chem.*, 2016, **88**, 7178–7182.
- 22 A. H. Ashoka, F. Ali, S. Kushwaha, N. Taye, S. Chattopadhyay and A. Das, *Anal. Chem.*, 2016, **88**, 12161–12168.
- 23 S. Chatterjee, S. Ghosh, S. Mishra, K. D. Saha, B. Banerji and K. Chattopadhyay, *Biochemistry*, 2019, **58**, 1109–1119.
- 24 M. Z. Chen, N. S. Moily, J. L. Bridgford, R. J. Wood, M. Radwan, T. A. Smith, Z. Song, B. Z. Tang, L. Tilley, X. Xu, G. E. Reid, M. A. Pouladi, Y. Hong and D. M. Hatters, *Nat. Commun.*, 2017, **8**, 474.
- 25 S. Zhang, M. Liu, L. Y. F. Tan, Q. Hong, Z. L. Pow, T. C. Owyong, S. Ding, W. W. H. Wong and Y. Hong, *Chem. – Asian J.*, 2019, **14**, 904–909.
- 26 A. H. Ashoka, U. G. Reddy, F. Ali, N. Taye, S. Chattopadhyay and A. Das, *Chem. Commun.*, 2015, **51**, 15592–15595.
- 27 R. Bhujbalrao and R. Anand, *J. Am. Chem. Soc.*, 2019, **141**, 1425–1429.
- 28 R. Anand, A. A. Hoskins, J. Stubbe and S. E. Ealick, *Biochemistry*, 2004, **43**, 10328–10342.
- 29 N. Sharma, N. Ahalawat, P. Sandhu, E. Strauss, J. Mondal and R. Anand, *Sci. Adv.*, 2020, **6**, eaay7919.
- 30 A. S. Tanwar, V. D. Goyal, D. Choudhary, S. Panjekar and R. Anand, *PLoS One*, 2013, **8**, e77781–e77781.
- 31 J. Wolf, A. P. Gerber and W. Keller, *EMBO J.*, 2002, **21**, 3841–3851.
- 32 H. C. Losey, A. J. Ruthenburg and G. L. Verdine, *Nat. Struct. Mol. Biol.*, 2006, **13**, 153–159.
- 33 T. G. Battye, L. Kontogiannis, O. Johnson, H. R. Powell and A. G. Leslie, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 271–281.
- 34 C. C. Project, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1994, **50**, 760–763.
- 35 P. Emsley and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126–2132.
- 36 G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1997, **53**, 240–255.
- 37 D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams, *Acta Crystallogr., Sect. D: Struct. Biol.*, 2019, **75**, 861–877.
- 38 J. Kyte and R. F. Doolittle, *J. Mol. Biol.*, 1982, **157**, 105–132.