Chemical Science

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: H. M. Devi, A. Badaya, A. maity, S. B. Saikia, R. Venkatramani and R. Swaminathan, *Chem. Sci.*, 2026, DOI: 10.1039/D5SC09293K.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemical-science

Label-Free Optical Detection of Protein Acetylation using *UV-Vis* Charge Transfer Spectra

Himanshi Maniram Devi [a#], *Apoorva Badaya*[b#], *Arijit Maity* [b], *Simangka Bor Saikia* [a], *Ravindra Venkatramani** [b] *and Rajaram Swaminathan** [a]

[a]*Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India*

[b]*Department of Chemical Sciences, Tata Institute of Fundamental Research, Homi Bhabha Road, Colaba, Mumbai 400005, India*

[#]*Equal Contributions*

**Corresponding Authors*

ravi.venkatramani@tifr.res.in *rsw@iitg.ac.in*

Abstract

We report the first spectroscopic detection of protein acetylation in solution using UV-Vis Protein Charge transfer spectra (ProCharTS). Acetylation is an important post-translational modification (PTM) that modulates diverse cellular processes, yet its detection relies on antibody-based cost-intensive and/or destructive techniques such as mass spectrometry. ProCharTS exploits the reduction of lysine charge post-acetylation to offer an alternate label-free, easily accessible, and cost-effective option to detect the PTM. Using two charge-rich proteins α_3C and α_3W, we demonstrate that the ProCharTS extinction coefficient between 370-800 nm monotonically decreases with increasing degree of chemical acetylation in the proteins. Complementary spectroscopic analysis and molecular dynamics (MD) simulations indicate that the ProCharTS signatures arise independent of secondary structure changes although tertiary interactions weaken, post-acetylation. Using a new computational MD and Time-Dependent Density Functional Theory (TDDFT) approach to simulate ProCharTS of whole proteins from their known 3D structure, we assign the observed PTM induced decrease in intensity to changes in the size, composition and spatial distribution of charged amino acid residue clusters. Our joint experimental-computational approach enables us to detect five or more acetylation events per protein with significant scope for further improvements in sensitivity. More broadly, this study presents a new optical mode (ProCharTS^{PTM}) exploiting charge transfer transitions to probe/track charged residue modifications in protein solutions.

Keywords: Charge Transfer Spectra, Protein Acetylation, UV-Visible spectroscopy, Time-Dependent Density Functional Theory (TDDFT), Photo-induced charge transfer, Post-translational Modification.

View Article Online
DOI: 10.1039/D5SC09293K

1. Introduction

Post-translational modifications (PTMs) are an important class of cellular mechanisms which expand the functional diversity, complexity and heterogeneity of the proteome by covalently modifying constituent amino acid residues in proteins[1,2]. There are over 400 different types of PTMs that can influence protein function[3–5]. PTMs can significantly transform the physiochemical properties of proteins including membrane binding propensities[6], site reactivity[7,8] and accessibility[9], and protein stiffness[10]. Acetylation is an important PTM in prokaryotes to eukaryotes which can be achieved by both enzymatic and non-enzymatic routes[11]. Acetylation adds an acetyl group (CH_3CO) to the protein that removes the positive charge from the amino group of Lys or N-terminus, making it neutral. Histone acetyltransferases (HATs) [12,13] are responsible for enzymatic acetylation of proteins; alternatively, acetic anhydride[14], aspirin[15], and acetyl CoA [16] are the major reagents used in chemical acetylation. Extensive research over the past 50 years have revealed the role of acetylation in regulating many cellular processes including transcription, metabolic states, cell cycle, nuclear transport, and actin nucleation [17–20] . Acetylation plays a role in regulating protein-protein interactions [21,22], in cancer biology [23,24], modulating enzyme activity [25] and transport of essential molecules [26–28]. Detection of acetylation in a protein is thus important for understanding many physiological and pathological conditions.

At present, Lys acetylation can be confirmed by several techniques such as western blotting[29], immunoprecipitation[30], mass spectrometry[31], radiolabelling[32], NMR[33], IR spectroscopy[34], and Flow cytometry[35]. While Western blot and ELISA[36] have both high sensitivity and accuracy, these assays involve multiple washing steps and expensive antibodies making them laborious, time-consuming and costly. Also, antibody cross-reactivity and false positive cases are limitations[37]. Mass spectrometry is a definitive tool for detecting PTMs[38], as it provides a quantitative and accurate change in the mass of the protein, upon addition of an external group to the parent protein. However, expensive instrumentation and limitations maintaining labile PTMs during sample preparation and need for careful separation techniques [39] pose challenges. The detection of sub-stoichiometric PTMs is especially difficult when the majority of the protein molecules may be unmodified[40]. Radiolabelled detection based upon [^{14}C] or [^{3}H]-acetyl-CoA enzymatic labeling and autoradiography analysis after separation with electrophoresis is another option[37]. These techniques for detection of acetylation in proteins described above are costly and require complex sample processing or labelling to be carried out, potentially perturbing the natural state of the proteins. Therefore, there is a need for techniques capable of detecting protein PTMs unambiguously in their native solution state without using external labels and requiring minimal sample preparation. In this context, optical spectroscopy may offer a solution. Interestingly, distinct spectral signatures for acetylation of both the N-terminus and the side chain amino groups of Lys have been reported experimentally using IR and Raman spectroscopy[41]. However, these signatures have not been tested for detecting/tracking acetylation in proteins. Post-translational modification of the amino group in Lys by acetylation alters the charge on Lys headgroup from +1 to zero. This radical change in the charge of Lys headgroup upon acetylation presents another approach that can be explored

View Article Online
DOI: 10.1039/D3SC09293K

to detect the PTM. Here we propose a UV-Vis spectroscopic approach to detect PTMs like acetylation which alter the charged state of protein amino acid residues.

Our approach is based on a novel UV-vis absorption observed in aqueous solutions of Lys-HCl[42] and charged amino acid rich proteins like α_3C, human serum albumin and calf thymus histones[43,44] This new spectra has been assigned to photo-induced charge transfer (PICT) transitions involving anionic and cationic head groups of all charged amino acid residues (including those that become charged with PTMs like phosphorylation) and the protein backbone. [44–46] The link between charged amino acids and ProCharTS has been further confirmed in recent studies of Vázquez and co-workers which showed that single alpha helical (SAH) peptides comprising solely of $(Lys)_4(Glu)_4$ or $(Arg)_4(Glu)_4$ repeats exhibit a broad absorption profile extending beyond 400 nm[47] Interestingly, the stability of Lys-Glu SAH peptides increases with the number of repeats not due to formation of salt bridges as might be expected intuitively, but rather due to the increasing number of possible charged states for the individual residues [48]**.** Our own studies have shown that within protein folds the sidechains and neutral backbones of Lys and Glu can act as electronic donor-acceptor pairs in the absence of salt-bridges to produce broad UV-vis ProCharTS profiles. Based on the charge complementarity of the amino acids and their separations, five types of inter-residue and intra-residue PICT transitions were shown to arise with diverse donor-acceptor separations ranging from 3-10Å [46]. The demonstrated sensitivity of the underlying PICT transition intensities and spectral range to the charge and clustering of amino acid residues makes ProCharTS potentially useful in identifying conformational states of proteins, their interactions, and PTMs which alter these properties. Indeed, ProCharTS absorption has been demonstrated as an easy *in vitro* technique to track protein aggregation[49,50], protein unfolding[51] and viral capsid assembly[52]. Beyond absorption, inter- or intra-residue PICT creates electron and hole pairs which can either undergo charge separation or recombination. The separated charges derived from ProCharTS excitations can be potentially harnessed for biochemical reactions as in the case of the light harvesting chlorophyll special pairs during photosynthesis[53]. On the other hand charge recombination leads to excitation dependent ProCharTS luminescence which has already been demonstrated to be useful in tracking protein aggregation [54] and in photosensitizing applications [47].

In this paper, we hypothesize that the ProCharTS absorption profile of proteins should be sensitive to the progressive acetylation of Lys rich proteins. To validate this hypothesis, we consider two small (~67 residues), folded (3-helix bundle), synthetic proteins, α_3C and α_3W. These well characterized proteins, each of which possesses a high content of charged amino acid residues (~54% of the sequence), serve as ideal models for studying the effects of acetylation on ProCharTS. We carried out systematic and progressive chemical acetylation of the proteins using acetic anhydride and employed a variety of experimental and computational methods to characterize acetylation and analyze its impact on protein structure and dynamics. Mass spectrometry (MALDI-ToF) confirmed the number of acetyl groups added to residues as a function of increasing concentration of acetic anhydride, besides confirming the heterogeneity in the acetylation levels of the sample protein. CD spectroscopy and the fluorescence of the tryptophan in α_3W probed changes in protein structure post-acetylation. Additionally, we developed a computational framework based on classical molecular dynamics (MD)

View Article Online
DOI: 10.1039/D5SC09293K

simulations and time-dependent density functional theory (TDDFT) calculations, to predict the ProCharTS profile of whole proteins from their 3D structure. Here absorption profiles are simulated from the spatiotemporal convolution of contributions from all charged amino acid residue chromophores within the proteins, providing residue level understanding of spectral changes between acetylated and unacetylated proteins and facilitating their deconvolution.

Our results demonstrate that ProCharTS and its associated luminescence are both sensitive to the presence of charge among the side chains of amino acid residues and their spatial proximity. Acetylation, which eliminates the positive charge of Lys, disrupts the size, composition and spatial proximities among charge clusters, leading to notable changes in absorbance and luminescence. Specifically, we find a monotonic quenching of the absorption spectra as a function of acetic anhydride concentration, thereby validating our hypothesis. The simulation and deconvolution of ProCharTS through computational methods allowed for the identification of oppositely charged dimers as critical chromophores which significantly contributed to the spectra changes brought about by progressive acetylation. Our study demonstrates an accessible, label-free, and non-invasive optical mode to detect acetylation in protein solutions. More generally, the results presented here highlight the potential of ProCharTS to detect any PTMs that alter the charge environment of proteins and conceptually advance our understanding of how specific modifications influence protein electronic spectra.

2. Methods:

2.1 Experimental

Materials: Trifluoroacetic acid (T6508); Sinapic acid (85429); β Mercaptoethanol (63689); Ammonium persulfate (A3678); Phenylmethylsulfonyl fluoride (P7626); $CaCl_2$ (C8106); 8-Anilinonapthalene-1-sulfonic acid (A3125); Folin-Ciocalteu reagent (F9252); Acetic anhydride (320102); N-acetyl tryptophan amide (A6501) were procured from Sigma Aldrich, Bengaluru, India. Luria Agar (M557); Luria Broth (M1245); Ampicillin sodium salt (TC021); Isopropyl β-D-1-thiogalactopyranoside (I5502); HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid); Sodium Chloride (GRM853); Sodium hydroxide (GRM467); Tris (Hydroxy methyl amino methane) (93315); Imidazole (GRM1864); MgCl2 were obtained from HiMedia Laboratories, India. The PD-10 column (17-0851-01) was purchased from GE Healthcare. Nuvia IMAC Ni-Charged Resin (780-0800) was purchased from Bio-Rad. All items purchased are of analytical grade with >98% purity.

2.1.1 Purification of α_3C and α_3W

Recombinant α_3C and its tryptophan (Trp) mutant form α_3W were expressed in *E. coli* strain BL21-DE3 and purified by Ni-NTA column chromatography, as reported earlier[55]. The α_3C amino acid sequence is shown in **Figure 1A.** Protein purity was determined by SDS-PAGE (ESI **Figure F1**) and additionally confirmed by MALDI-ToF (ESI **Figure F2)**.

2.1.2 Acetylation reaction

Sample preparation: 20 μM α_3C and α_3W were prepared in 100 mM HEPES buffer at pH 8. We also considered NaH_2PO_4 buffer, where the protein was less stable and found to precipitate. The concentrations of α_3C and α_3W were estimated using the Lowry method[56]. Here, protein absorbance with Lowry reagents (Reagent I (NaOH, Na_2CO_3, $KNaC_4H_4O_6.H_2O$, $CuSO_4.5H_2O$)

and Reagent II (Folin's)) was recorded at 550 nm and the ProCharTS contamination was eliminated by subtracting the protein absorbance in the absence of the reagents (ESI **Table T1**).

Reaction condition for chemical acetylation of α_3C and α_3W: In vitro chemical acetylation was carried out using acetic anhydride as described earlier [57] with slight modifications. Briefly, 20 µM α_3C and α_3W were titrated with varying concentrations (0.05, 0.1, 0.2, 2 and 20 mM) of acetic anhydride. The reaction vessel was kept at 4 ˚C with continuous stirring during addition of acetic anhydride in small aliquots and the pH of the reaction mixture was maintained between 7.5 and 8.5 by adding 0.5 N NaOH as required. The reaction was performed in HEPES buffer for 30 minutes. Subsequently the dialysis and centrifugation of the prepared samples were performed to remove the residual acetic anhydride, byproducts and any large particles which can contribute to the scattering (details are mentioned in ESI **Section M1.1.1** and spectral analysis shown in ESI **Figure F3**).

2.1.3 Confirmation of acetylation by Mass Spectrometry

Mass spectrometry was used to determine the molecular weight of the proteins. Protein samples (in deionised water) were mixed with a matrix formed by a saturated solution of sinapic acid in the TA-30 solution (0.1% TFA and Acetonitrile in a 7:3 ratio). Protein samples dissolved in matrix were analyzed in linear mode by flex analysis and flex control software from Bruker Daltonics, Germany. The native (unacetylated) proteins have *m/z* values of 7464.2 Da (α_3C) and 7545.2 Da (α_3W) as shown in ESI **Figure F2**. The addition of an acetyl group to the protein results in a 42 Da mass shift in the mass spectrum. For samples containing a heterogenous mixture of populations with varying degrees of acetylation, the mass spectra exhibit a broad distribution with multiple peaks. The number of residues acetylated in each population can be obtained from the *m/z* value of each acetylated species peak as: $(Acetylated_{m/z} - Native_{m/z})/42$

2.1.4 UV-Visible absorption

Absorption spectra of all native and acetylated protein samples were acquired using a double-beam UV-Vis Cary-100 spectrophotometer (Agilent Technology). Sample concentrations (~9-10 µM, for more details refer ESI **Table T1**) were determined spectrophotometrically, as mentioned above, using the calculated extinction coefficient and applying Beer-Lambert Law ($A = \varepsilon cl$, where A is Absorbance, ε is the extinction coefficient in $M^{-1}cm^{-1}$, l is the path length in cm and c is the concentration in M). Spectra of all protein samples were collected between 250 to 800 nm at room temperature (~22 °C) using a 10 mm path length quartz cuvette (Hellma: Z600210). Each spectrum was scanned three times with 1 nm bandwidth and 600 nm/min speed and subsequently averaged. The deionised water used in the samples as a solvent was used as a reference for baseline corrections during spectral measurements.

2.1.5 Steady State Luminescence

ProCharTS luminescence: Steady-state luminescence emission spectra were collected for all

native and acetylated proteins in deionised water at several excitation wavelengths: 280, 310, 340, 355, 370 and 410 nm (slit width: 2 nm) and emission spectra were recorded in the range 300-500, 330-550, 360-600, 375-650,390-700 and 430-750 nm respectively (slit width: 15 nm). Excitation at 280 nm was not carried out for α_3W due to interference from emission of indole in Trp which has a higher quantum yield (~0.14) compared to ProCharTS luminescence (~0.01). For acquiring the luminescence spectra of α_3C and α_3W, the protein concentration was kept at ~9-10 μM, and recorded spectra were corrected manually for Raman peaks (ESI **Table T2**) by subtracting the emission spectra in the appropriate range from blank samples devoid of protein.

Quantum Yield calculation: Quantum yields (QY) for native and acetylated proteins (α_3C and α_3W) were estimated from experiments in deionised water by exciting samples at 280 and 355 nm at room temperature (25 °C) Details on the standard protocols followed [58] are mentioned in ESI **Section M1.1.2**.

ANS fluorescence: Protein-ANS binding was monitored by fluorescence emission. The ANS stock sample was prepared in deionised water with concentrations determined from extinction coefficient values at 350 nm (4900 $M^{-1}cm^{-1}$). The stock sample was added to the protein samples (~9-10 μM native/acetylated forms) in deionised water to achieve a final ANS concentration of 10 μM. All samples were excited at 380 nm and emission spectra were collected from 400 to 700 nm using 2 nm excitation and 15 nm emission slit widths. The ProCharTS luminescence contribution was also subtracted manually from all the samples to avoid ProCharTS contamination All the spectra were recorded in triplicates with Fluoromax-4 (Horiba Scientific, USA) at 25 ^{o}C.

Trp fluorescence: Intrinsic fluorescence of Trp was measured using a spectrofluorometer (Fluoromax-4, Horiba, USA). Native and acetylated α_3W samples (~9-10 μM) were prepared in deionised water and excited at 280 nm. All readings were measured in an optical glass cuvette of path length 10 mm (Hellma: Z802875). Fluorescence data were acquired as S1/R1 signal with a wavelength increment of 1 nm and an integration time of 0.1 s. Emission spectra of α_3W (300 – 500 nm) were collected with 2 nm excitation and 15 nm emission slit width. All spectra were averaged over five scans and manually corrected for Raman peaks (ESI **Table T2**). The emission spectrum of NATA was recorded as a reference for calibrating the monochromator. Steady-state anisotropy of Trp in native and acetylated α_3W (~9-10 μM) was carried out in deionised water using the same instrument (see ESI **Section M1.1.2** for more details).

2.1.6 Time-resolved fluorescence

The fluorescence intensity decay was recorded using a Time-correlated single-photon counting instrument equipped with MCP detection (Horiba Jobin, Model: Ultrafast-01-DD). An LED of 290 nm was used to excite the sample, and emission was collected at 327 and 345 nm for ~9-10 μM native and acetylated samples with 20 mM acetic anhydride, respectively[59]. All samples were measured in deionised water. Further details along with analysis are provided in ESI **Section M1.1.2**).

View Article Online
DOI: 10.1039/D5SC09293K

2.1.7 Circular Dichroism

Protein secondary structures were investigated using a Jasco J-1500 CD spectrometer, using a 2 mm pathlength rectangular quartz cuvette. The CD spectra were recorded between 190 to 260 nm (bandwidth 2 nm) for ~9-10 μM α_3C and α_3W (native and acetylated forms) in deionised water. Each reported individual spectrum is an average of 10 individual scans. The CD spectrum of deionised water was also recorded as a baseline and a blank. The secondary structure content was analyzed by the K2D3 online software [60].

2.1.8 Zeta potential measurements

The Zeta potential was recorded in Anton Paar Litesizer 500. Native α_3C and α_3W and their acetylated forms (~9-10 μM) were filtered through a sterilized 0.2 μm syringe filter before use. Samples were run in triplicate, and each sample was scanned 50 times and subsequently averaged. A disposable 900 μL Omega cell was used for the zeta potential calculation and deionised water was used as the reference solvent.

2.2 Computational

We designed a general computational workflow (ESI **Figure F4** and **Section M1.2**) based on MD simulations and TDDFT calculations to generate the ProCharTS profile of proteins from their 3D structure. We apply this protocol to native and acetylated forms of α_3C and α_3W as detailed below:

2.2.1 Protein Models

PDB id 2LXY and 1LQ7 [61]were selected as starting structures for α_3C (**Figure 1A-B**) and α_3W respectively. These proteins are rich in charged amino acid residues (17 Lys, 17 Glu, 2 arginine (Arg)) and contain either a single Cys (α_3C) or Trp (α_3W) at position 34. Acetylated variants of the proteins were generated by mutating Lys residues to acetyl-Lys (Aly) in the PDB structures (**Figure 1B**). A total of six α_3C models were generated: native, Ac1, Ac3, Ac5, Ac12 and Ac17 (n = number of acetylated Lys residues in Acn). These systems were chosen to match the dominant protein states obtained at different concentrations of acetic anhydride as verified by mass spectrometry (**Figure 1C-N**). The Lys residues in α_3C were progressively acetylated in decreasing order of their relative solvent accessible surface area (rSASA) value (ESI **Section M1.2.1** and **Figure F5A-F**). Further to test the sensitivity of acetylation sites on protein spectra, three variants for Ac12 were created with three different permutations of acetylated states for three specific residues: Ac12-α_3C-v1 (Lys10, Aly32, Aly39), Ac12-α_3C-v2 (Aly10, Lys32, Aly39), and Ac12-α_3C-v3 (Aly10, Aly32, Lys39). For α_3W we generated only the native and fully acetylated models (α_3W and Ac17-α_3W). To summarize a total of ten native and acetylated protein systems were generated for MD simulations. More details of the generation of the modelled structures are provided in ESI **Sections M1.2.1** and **M1.2.5**.

View Article Online
DOI: 10.1039/D5SC09293K

2.2.2 Molecular Dynamics Simulations

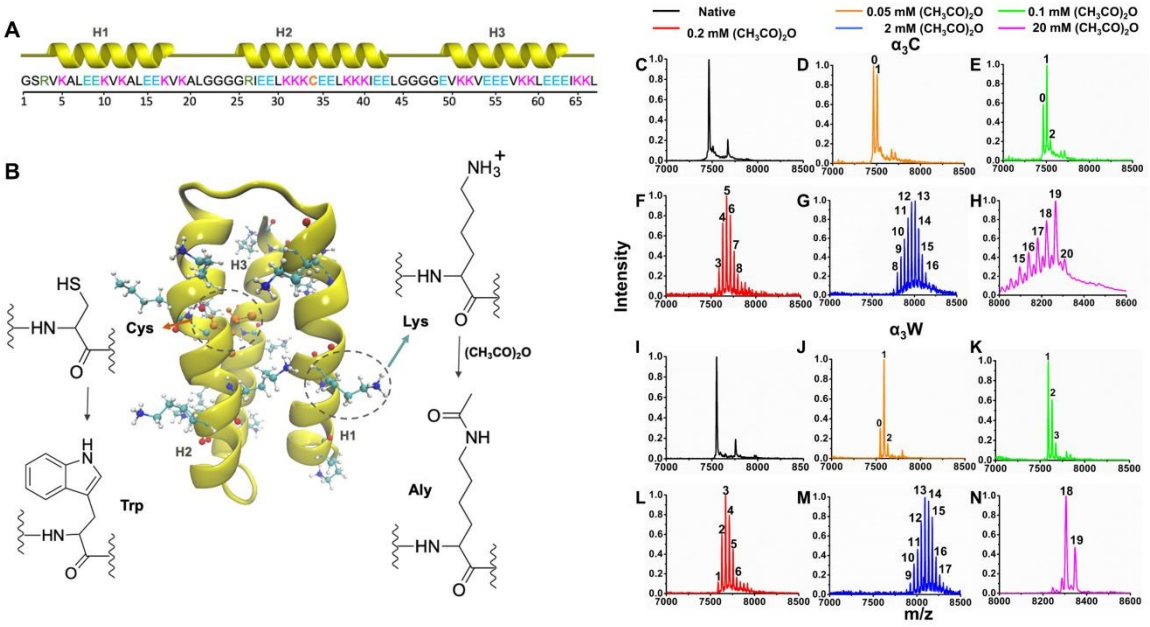

Figure 1: Structure of α_3C and α_3W proteins and their mass spectra under native (unacetylated) and acetylated conditions. (A) Amino acid sequence, helices (H1, H2 and H3) and (B) structure (PDB code: 2LXY) of α_3C protein. In the α_3W protein, the cysteine (Cys) residue at position 34 is mutated to tryptophan (Trp). In the acetylated (Ac-α_3C/Ac-α_3C) protein, lysine (Lys) residue is chemically modified to Acetyl-lysine (Aly); Mass spectra of native α_3C (C) and acetylated α_3C (D-H); native α_3W (I) and acetylated α_3W (J-N) reveal a gradual increase in mass with an increase in the concentration of acetic anhydride. Acetylated proteins display a clearly resolved distribution of mass peaks in each individual mass spectrum, reflecting the number of Lys residues acetylated in each protein species as indicated by the integer above.

Fully atomistic MD simulations on each of the ten protein models were carried out under periodic boundary conditions using Gromacs-5.1.7 [62–64] with the CHARMM36 force field [65]. Each model was solvated with explicit TIP3P water molecules [66] and neutralized by adding suitable number of Na^+ or Cl^- ions. During simulations, electrostatic interactions were calculated using the Particle Mesh Ewald (PME) algorithm [67] and non-bonded van der Waals interactions were described by the Lennard-Jones potential. The protein systems were subjected to energy minimization, 1 ns NVT (constant volume and temperature) and NPT (constant pressure and temperature) equilibrations. Here temperature and pressure were controlled using the velocity-rescale thermostat[68] and Parrinello-Rahman barostat [69] respectively. Following this, unrestrained 100 ns MD trajectories were generated under NPT conditions with a time step of 2 fs with the last 50 ns being considered as the production phase. The positions and the velocities of the atoms were saved every 2 ps during the production stage yielding a total of 25,000 snapshots for each system which were taken forward for stability analysis [70] and subsequent sampling of chromophores. Full details on the equilibrium protocol and stability analysis are provided in ESI **Sections M1.2.2** and **Sections M1.2.3**.

2.2.3 ProCharTS Chromophores

View Article Online
DOI: 10.1039/D5SC09293K

From each snapshot of the MD trajectories we selected amino acid residues clusters (C_i) comprising of i interacting charged residues and Trp (for α_3W). Here two amino acid residues are considered interacting if they lie within a specified cut-off distance (R_c = 6Å). A discussion on the choice of R_c is provided in ESI **Section M1.2.4**. The amino acid residues included in our clusters include Lys, Arg, Glu, and Trp. For the acetylated systems we considered clusters both with and without Aly. ProCharTS chromophores (clusters C_i) were extracted from 50 snapshots uniformly sampled from the 50 ns production MD trajectory using Visual Molecular Dynamics (VMD)[71]. For Ac3-α_3C, we sampled 25 snapshots from the first 25 ns equilibrated production trajectory segment based on stability analysis (ESI **Section M1.2.3**). Here we verified that the distributions of clusters C_i of different sizes (i = 1-10) over the 50 snapshots were representative of the distributions from the full set of 25,000 snapshots in native (α_3C, α_3W) and fully acetylated (Ac17-α_3C, Ac17-α_3W) protein trajectories (**Figure F7** and **Figure F8**). For each amino acid residue fragment in a cluster, the backbone of the adjacent residues was also extracted to preserve local geometry. For residue pairs that are non-adjacent in sequence, both the N–1 and N+1 backbone atomic positions were capped with hydrogens to create symmetric terminal methyl (CH_3) groups, thereby neutralizing dangling bonds and mimicking the extended backbone environment of the native protein[46]. For nearest-neighbor residue clusters, the peptide linkage between the residues was retained and only the N–1 backbone atomic position of the first residue and the N+1 backbone atomic position of the last residue were capped.

2.2.4 Computations of spectra

We calculated electronic transitions beyond 250 nm for each individual cluster using TDDFT with the OT-CAMB3LYP exchange correlation functional [46] and 6-31++G(d) basis set in Gaussian 09 [72]. An SCF convergence criteria of 10^{-4} was employed which was adequate to produce reliable spectra (ESI **Figure F30**). The spectral range covered in our calculations spans 250-800 nm, over which the total number of TDDFT transitions calculated varies from 11-150 depending on the cluster size (ESI **Figure F9)**. The absorption spectra of the j^{th} cluster with i ($1 \leq i \leq 10$) residues (C_{ij}) as a function of incident light wavelength λ is given by a weighted sum of all TDDFT transitions calculated for the cluster:

$$\varepsilon(C_{ij},\lambda) = \sum_k 1.3 \times 10^8 \times \left(\frac{f_{ijk}}{\sigma}\right) \times \exp\left(\frac{-\left[\frac{1}{\lambda} - \frac{1}{\lambda_{ijk}}\right]}{\sigma}\right)$$

where f_{ijk} is the oscillator strength of k^{th} transition with wavelength λ_{ijk} for the cluster C_{ij} and σ is the uniform broadening parameter (0.4 eV) used for all transitions. Since the clusters, by definition, are electronically uncoupled (ESI **Section M1.2.4**), the spectra of protein in each snapshot is obtained by adding the spectra of individual clusters as:

$$\epsilon_{snapshot}(\lambda) = \sum_{i,j} \varepsilon(C_{ij},\lambda)$$

View Article Online
DOI: 10.1039/D5SC09293K

where, the indices *i* and *j* run over the number of residues in the clusters and number of clusters in the snapshot respectively. Finally, the protein spectral profile is obtained by averaging the spectra over all *N* snapshots for each system:

$$\epsilon(\lambda) = \frac{1}{N}\sum_{k=1}^{N} \epsilon^{k}_{snapshot}(\lambda)$$

As discussed previously, we use *N*=50 for all protein systems except for Ac3-α_3C where *N*=25.

3. Results

3.1 Chemical acetylation of α_3C and α_3W and confirmation using mass spectrometry

We incubated α_3C and α_3W in acetic anhydride solutions and used mass spectrometry to quantify the number of acetylations in the two proteins as a function of acetic anhydride concentration. The mass of native α_3C and α_3W (**Figure 1C, I)**, as determined by MALDI-ToF, is consistent with the expected protein mass based on sequence. Both proteins contain 18 amino groups (17 ε-groups of Lys and 1 N-terminal) which can be potentially acetylated. Since a single acetyl group adds 42 Da to the parent protein the increment in the protein mass relative to the native protein reveals the number of acetylated residues. The MALDI-ToF data for the proteins in the presence of acetic anhydride in **Figure 1**, shows multiple *m/z* peaks which correspond to a distribution of protein species with different number of modified residues (molecular mass values for the peaks are given in ESI **Figure F11-F12**). We observe two peaks in the mass spectrum for α_3C in 0.05 mM acetic anhydride (**Figure 1D)** corresponding to the native unacetylated protein (~7,464 Da) and to a population with one residue modified (~7,506 Da)**.** As the concentration of acetic anhydride is increased to 0.1 mM, three peaks are observed that reveal a decrease in the native population and acetylated protein populations with one and two modified residues (**Figure 1E)**. At higher acetic anhydride concentrations, a broad distribution of peaks is observed (**Figure 1F-G**) indicating the presence of multiple protein species. While species with 3-7 modified residues (5 acetylations for the dominant species) are observed at 0.2 mM acetic anhydride, species with 8-16 modified residues (12/13 acetylations for the largest populations) are seen for 2 mM acetic anhydride. Finally, to achieve hyperacetylated protein, 20 mM acetic anhydride was used, which modifies 17-20 residues (**Figure 1H**). Since there are only 18 amino groups in α_3C, one or two residues other than Lys appear to be modified. Acetic anhydride has also been reported to target the hydroxyl groups of Ser, Thr, and tyrosine (Tyr), as well as the sulfur group of Cys [73,74]. In α_3C a Ser and a Cys residue are present that are presumably targeted by acetic anhydride, resulting in their acetylation. Between these two residues, the Cys sidechain (*RSH*) is intrinsically more nucleophilic than the Ser sidechain (*ROH*). However, because acetylation occurs within the folded protein rather than in free solution, steric effects and relative accessibility to acetic anhydride would also be important factors. A comparison of the relative SASA (rSASA) of the Ser and Cys sidechains using MD trajectories of α_3C before and after acetylation reveals that Ser remains moderately exposed to

View Article Online
DOI: 10.1039/D3SC09293K

solvent but exhibits a significantly broader thermal distribution of accessibilities upon acetylation (rSASA of 35.7±8.0 % in native protein versus 36.6±25.3 % in Ac17-α_3C). On the other hand, the relative SASA for Cys changes from (0.1±0.1) % in native protein to (0.8±0.7) % upon full Lys acetylation. Clearly Cys is deeply buried and is not very accessible to the solvent either before or after acetylation in MD simulations. Nevertheless, there is a relative increase in accessibility for the thiol (– SH) group post-acetylation of the 17 Lys residues. Here, the reactivity and relative accessibility of the -ROH and -SH groups have opposing influences on the acetylation potential of Ser and Cys. Based on the present analysis, it is therefore not possible to determine which of these residues is more likely to be acetylated within the folded protein environment.

Similarly, α_3W was progressively chemically acetylated by adding 0.05, 0.1, 0.2, 2, or 20 mM acetic anhydride. With 0.05 mM acetic anhydride, species with one modified residue is predominantly generated along with a small population of native species and an even smaller population with two modified residues (**Figure 1J)**. Increasing the acetylating agent concentration to 0.1 mM results in three *m/z* peaks which show 1-3 modified residues with the predominant species showing a single acetyl modification (**Figure 1K)**. When the acetic anhydride concentration is increased further, the distribution of modified mass peaks progressively changes, first exhibiting four prominent peaks with a range of 3-6 modified residues (0.2 mM) and then showing seven peaks with 10-16 modified residues (2 mM) **(Figure 1L-M)**. Finally, with a 20 mM concentration of acetic anhydride, α_3W species with 18 and 19 modified residues were generated. Here, 19 modifications in the protein can only be explained if a residue other than Lys is modified. In α_3W, the Cys is absent; however, a Ser residue is present.

3.2 Acetylation does not significantly alter the secondary structure of α_3C and α_3W but leads to weakening of tertiary interactions.

Circular dichroism (CD) spectroscopy (190-260 nm) was used to investigate changes in the secondary structure of α_3C and α_3W arising due to the reaction with acetic anhydride. Native α_3C and α_3W spectra (**Figure 2A-B**) show two negative bands at 208 and 222 nm and a positive band at 195 nm, which indicates the presence of alpha-helical structure in the proteins. The spectra does not change significantly for either protein after treatment with acetic anhydride (0.05-20 mM). Specifically, the CD spectra for α_3C and α_3W at low concentrations of acetic anhydride (0.05 and 0.1 mM) nearly overlap with that of the native protein. Even at higher concentrations (0.2, 2 and 20 mM) of acetic anhydride, the spectra show only minor changes indicating a slight increase in the helical content for both proteins (**Figure 2A, B).** An analysis of the CD spectra using the K2D3 web server [60] shows a consistent 95% helicity for α_3C and α_3W independent of acetic anhydride treatment (ESI **Table T4)**. Dictionary of Secondary Structure Prediction (DSSP) analyses on production MD trajectories also show that the coil, bend, turns, and alpha-helix elements of the protein are preserved following acetylation for both α_3C and α_3W (**Figure 2C**) and are presumably acetylated by acetic anhydride, as previously reported[73].

View Article Online
DOI: 10.1039/D5SC09293K

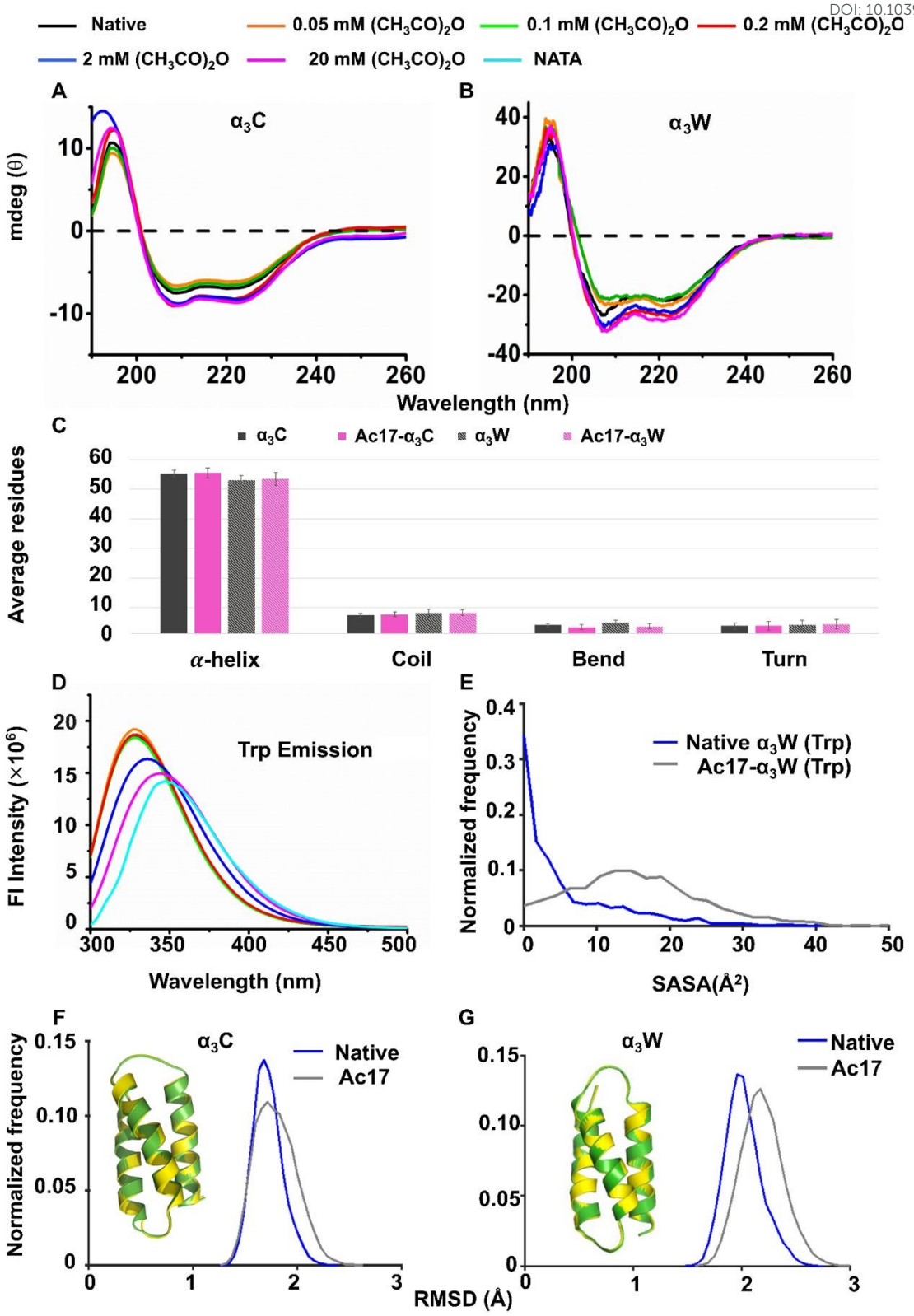

Figure 2: Global and local changes in protein structure post-acetylation. Far-UV CD spectra of ~9-10 μM α_3C (A) and α_3W (B) at 25 °C in deionised water, incubated with different concentrations of acetic anhydride for 30 minutes at 4 °C. (C) Secondary structure content of native and acetylated proteins from 50 ns production MD trajectories obtained using the DSSP module in Gromacs. (D) Steady-state Trp fluorescence emission spectra of native and acetylated α_3W excited at 280 nm; emission was

collected in the range 300 -500 nm with 10 μM NATA (as a control) at room temperature in deionised water using 2 nm excitation and 15 nm emission slit width. (E) Normalized frequency distribution of Trp SASA in native and acetylated α_3W proteins obtained from snapshots derived from 50 ns production MD trajectories. Normalized frequency distribution of backbone atom RMSDs of native and fully acetylated α_3C (panel F) and α_3W (panel G) relative to their initial conformation from 50 ns MD trajectories. For reference, the backbone RMSD between the initial conformations (inset figures: yellow = α_3C/ α_3W and green Ac17-α_3C/Ac17-α_3W) is ~ 0.3 Å.

We next used the fluorescence intensity of Trp34 in α_3W to probe potential local changes in protein structure after acetylation. NATA (a derivative of Trp) which showed emission maxima at 347 nm, was used as a standard control. Trp34 was excited at 280 nm and its peak emission intensity was observed at 327 nm in the native protein. The emission maximum remained fixed for lower concentrations (0.05 and 0.1 mM) of acetic anhydride but exhibited a systematic red shift at higher concentrations to 329 nm (0.2 mM), 336 nm (2 mM) and 345 nm (20 mM), accompanied by decrease in intensity (**Figure 2D)**. This data indicates that indole ring of Trp34 in native α_3W is buried in the hydrophobic core and as more residues get acetylated it gradually gets more exposed to the polar aqueous environment. We note that relative to the control NATA, the emission peaks for all systems are blue-shifted, indicating that even Ac17-α_3W does not have a fully exposed Trp. Additionally, the integrated area of fluorescence emission was calculated for native and acetylated α_3W. This analysis indicated that acetylation does not exhibit a significant change in total fluorescence emission relative to native protein **(**ESI **Figure F13)**. Fluorescence anisotropy and time-resolved data provide further insights into changes in Trp34 indole rotational motion [75] and local environment induced by acetylation. The steady state fluorescence anisotropy of indole chromophore in native α_3W was measured to be 0.083 and is sensitive to acetic anhydride concentrations (ESI **Figure F14** and **Table T5**). The Trp34 anisotropy is comparable to the native protein at lower concentrations (0.05,0.1, 0.2 mM) but decreases significantly at higher acetic anhydride concentrations (2 and 20 mM) where between 10-16 acetyl groups are added to the protein (**Figure 1M**). The large number of modifications leads to a significant exposure of Trp34 to the solvent, which should increase its rotational mobility and decrease fluorescence anisotropy. Note that the Trp34 anisotropy in α_3W is always higher than the anisotropy of NATA, as the former is linked to a polypeptide chain, unlike the latter, which can freely tumble. The anisotropy data are summarized in **Table S5**. Trp fluorescence intensity decay for native and acetylated α_3W (20 mM acetic anhydride) both show two lifetimes, which are assigned to different Trp rotamers (ESI **Figure F15 and Table T6**). The native protein exhibits a mean lifetime that is slightly lower than NATA (ESI **Table T6**). The amplitude values of native and fully acetylated (24-hour dialysis) samples are different, which reflects the exposure of Trp to a different environment as compared to the native protein, in agreement with the steady-state fluorescence data. The mean fluorescence lifetime of native (2.4 ns) and fully acetylated (24 hours dialysis) sample (2.5 ns) are nearly same, consistent with nearly equal integrated areas of their emission spectra (ESI **Figure F13).**

Computational SASA analysis (**Figure 2E**) for the Trp residue in α_3W and Ac17-α_3W MD trajectories reveal that acetylation indeed increases Trp solvent exposure, as shown by the broader distribution and a shift toward higher SASA values for Ac17-α_3W relative to native protein. An analysis of spatially proximal residues further reveals that two Lys residues (19 and

38) are located within 5.5 Å of Trp34 in ~50% frames in the α_3W production trajectory. Notably, Lys38 forms a stabilizing salt bridge with Glu35 in ~66% frames in native protein which is disrupted in Ac17- α_3W and weakens the tertiary fold near Trp. Our analysis reveals more global changes created by acetylation as well. The protein backbone for both α_3C and α_3W proteins tends to get slightly more flexible upon acetylation as seen by the distribution of RMSD (backbone atoms) of structures in the production MD trajectories (**Figure 2F** and **2G**). The distributions tend to get a bit broader and with a slight peak shift for both α_3C and α_3W upon acetylation. Finally, the radius of gyration (R_g) value increases systematically as we increase the number of Aly residues (ESI **Table T8** and **Figure F16**) in both α_3C and α_3W. indicative of a more expanded structure post-acetylation. Taken together the computational results support the hypothesis of a loosely packed tertiary structure in α_3W created by acetylation.

3.3 Acetylation makes the surface of α_3C and α_3W more polar and negatively charged

The sequence of native α_3C and α_3W is rich in polar, charged amino acid residues. However, the number of positive and negatively charged groups are balanced (17 Lys, 17 Glu, and 2 Arg) leading to a relatively small net charge of +2. Acetylation of Lys by acetic anhydride treatment is expected to significantly change the surface charge and polarity of the proteins which we probe experimentally and computationally in this subsection. First we use ANS, a hydrophobic probe which can bind to solvent-accessible clusters of non-polar atoms [76]. Specifically, acetylated proteins (in 0.05-20 mM acetic anhydride concentrations) were incubated with ANS, and the fluorescence intensity of the probe was measured. In the native α_3C, the ANS emission maximum appears at 490 nm, more intense than and blue shifted from the free ANS maximum at 525 nm, which confirms the binding of ANS to the hydrophobic core of the protein. With increasing acetylation, the emission maxima remain almost fixed with a slight blue shift (484nm) at acetic anhydride concentrations of 0.2 and 2 mM (**Figure 3A**). However, the emission peak intensity decreases with increasing acetic anhydride concentrations (≥ 0.2 mM), approaching that of free ANS. In α_3W, the emission maximum in native protein (479 nm) and its acetylated forms for acetic anhydride concentrations ranging 0.05-20 mM) are blue-shifted relative to native protein. The blue shift and reduction in the peak intensity above a threshold concentration (0.1 mM) of acetic anhydride were evident, while the emission spectra approach that of free ANS at 20 mM acetic anhydride concentration (**Figure 3B**). These observations suggest that subjecting the protein to acetic anhydride at concentrations above a threshold reduces ANS binding in acetylated α_3C and α_3W.

We employed further experimental and computational measurements to understand the mechanisms by which acetylation reduces ANS binding to the proteins. In addition to anilino-naphthalene a hydrophobic group, ANS also possesses a charged sulfonate group [77,78] which can bind to the positively charged Lys amino moiety. As the concentration of acetic anhydride increases, more Lys residues are derivatized with the acetyl group, making the surface more negative and creating unfavorable electrostatics for the ANS sulfonate ions. An examination of the overall charge of native and acetylated α_3C/α_3W by measuring their zeta potentials [79] reveals

View Article Online
DOI: 10.1039/D5SC09293K

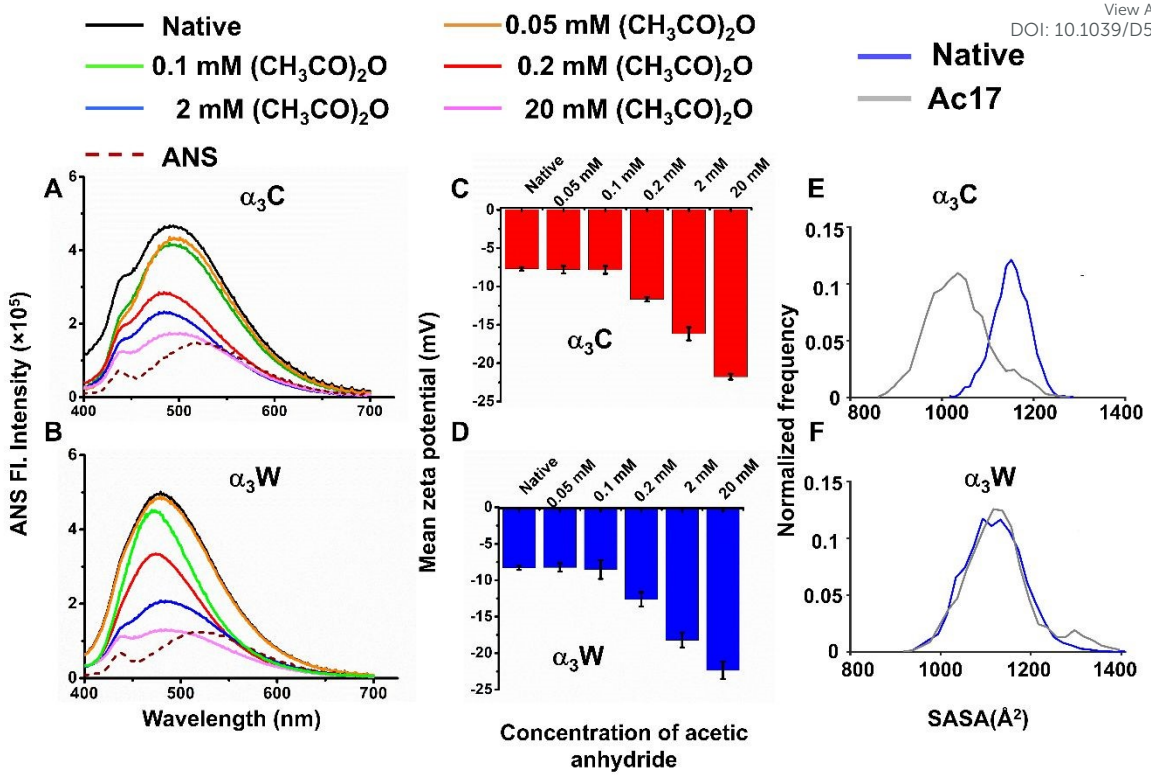

Figure 3: Illustrating the effect of acetylation on protein hydrophobicity and surface charge. Steady-state ANS Fluorescence emission spectra of ~9-10 µM native and different degrees of acetylated α_3C (A) and α_3W (B) were recorded in deionised water after incubating all the samples in ANS at room temperature for 10 minutes for the binding reaction. All the spectra were excited at 380 nm, and emissions were collected from 400 – 700 nm with excitation slit width 2 nm and emission slit width 15 nm. Zeta potentials of native α_3C (C); and α_3W (D); compared with different degrees of acetylated samples to depict the change in the surface charge of proteins after acetylation. (E) and (F) Normalized frequency distribution of solvent-accessible surface area (SASA) for the hydrophobic residues (Ala, Gly, Val, Ile, Leu, Phe, Met) in native and acetylated α_3C and α_3W proteins, respectively.

that the acetylated proteins have more negative values as compared to the native form. In α_3C, addition of 0.2, 2, or 20 mM acetic anhydride changes the mean zeta potential from -8.1±0.7 (native protein) to -11.5±1.3, -16.1±0.9 and -21.4±1.5, respectively (**Figure 3C**). Similarly, in α_3W (mean zeta potential of -8.3±0.2), after acetic anhydride treatment at the same three concentrations, the zeta potentials shift to -12.6±0.98, -18.2±1.0 and -22.3±1.1, respectively. However, at lower concentrations (0.05 and 0.1 mM), the values of mean zeta potential closely match those of native proteins with α_3C showing -7.8 ± 0.5 and -7.9 ± 0.5 mV and α_3W showing -8.2 ± 0.5 and -8.5 ± 1.2, respectively (**Figure 3D**). This data confirms that the protein surface becomes more negative post-acetylation. Further, SASA analysis on native and acetylated proteins shows that the hydrophobic residues (alanine (Ala), glycine (Gly), valine (Val), isoleucine (Ile), leucine (Leu), phenylalanine (Phe), and methionine (Met)) are also less exposed to the solvent post-acetylation (**Figure 3E-F**). The SASA distribution for hydrophobic residues in native α_3C (blue line) is shifted to higher values compared to that of acetylated α_3C (gray line). While the corresponding distribution for the native α_3W (blue line) is also shifted to higher SASA values relative to the acetylated α_3W (gray line), the shift is smaller as compared to α_3C.

View Article Online
DOI: 10.1039/D3SC09293K

These observations correlate with the higher threshold concentrations of acetic anhydride needed for α_3W to reduce ANS binding in experiments. Further, the redder ANS emission maximum position (463 nm) in α_3W compared to α_3C (446 nm) when acetylated with 20 mM acetic anhydride also indicates that ANS resides in less non-polar regions of α_3W. In summary, for both proteins (α_3C and α_3W), acetylation appears to reduce the solvent accessibility of hydrophobic residues and increase the surface negative charge leading to the observed weaker binding of ANS.

3.4 *The ProCharTS UV-Vis absorption of α_3C and α_3W proteins can track their progressive chemical acetylation*

The experimental absorption spectra (250-800 nm) of native and acetylated α_3C and α_3W in different concentrations of acetic anhydride (0.05, 0.1, 0.2, 2 and 20 mM) are compared in **Figure 4**. The broad absorption profile (black curve in **Figure 4A**) is characteristic of native α_3C and has been previously assigned by us to PICT in closely clustered charged Lys and Glu residues within the protein fold. [44] [46]. As the two proteins are exposed to increasing concentrations of acetic anhydride, they undergo progressive acetylation of the Lys residues (ESI **Figure F17**) which in turn alters the ProCharTS profiles (**Figure 4A-D)**. While α_3C is devoid of any aromatic amino acid residues, α_3W has a single Trp that shows strong absorption from 280 – 295 nm which overrides changes in this wavelength range observed after acetylation (**Figure 4A-B**). On the other hand, the ProCharTS from 370-800 nm for α_3W is free of any contributions from Trp and matches that from α_3C. In both proteins, the ProCharTS responds to acetylation only above a threshold acetic anhydride concentration of 0.2 mM, essentially overlapping with the spectra of native protein at lower (0.05 and 0.1 mM) concentrations. Beyond this threshold, the ProCharTS intensity across all wavelengths, monotonically drops with increasing concentrations of acetic anhydride for both α_3C (**Figure 4C**) and α_3W (**Figure 4D**). The threshold acetic anhydride concentration (0.2 mM) required for ProCharTS to register a change suggests a sensitivity limit of 3-7 Lys residue modifications. Note that at shorter wavelengths (300-400 nm), the experimental spectra for α_3C shows well resolved shifts (**Figure 4A-C**) even at lower concentration of acetic anhydride (0.05-0.1 mM) where only a single Lys site is acetylated in the protein. However, in α_3W, this sensitivity at bluer wavelengths is reduced due to the strong absorption features of Trp (**Figure 4B**).

Our computational data (**Figure 4E-G**) confirms that the decrease in absorption intensity for both α_3C and Ac17- α_3W originates from the addition of the acetyl group to the Lys sidechains in the proteins that contribute strongly to the ProCharTS intensity. As shown in **Subsection 3.3** the overall positive charge of the protein is progressively reduced with increasing acetic anhydride concentrations. Our mass spectrometry results (**Figure 1D-E**) show that at low acetic anhydride concentrations (0.05 and 0.1 mM) only a single Lys residue is modified, and this species co-exists with a significant population of native species. Our calculations show a pronounced decrease in ProCharTS intensity for Ac5- α_3C Ac12- α_3C, Ac17-α_3C (**Figure 4E**) and Ac17-α_3W relative to native protein (**Figure 4F**) consistent with experimental data (**Figure 4C-D**) at higher acetic anhydride concentrations (≥ 2 mM). The insensitivity of the spectra above 400 nm at lower concentrations is also captured by the computed spectra. In **Figure 4F**

View Article Online
DOI: 10.1039/D5SC09293K

we compare the weighted average spectra (orange and green data) from native α_3C and Ac1-α_3C with weights corresponding to the *m/z* peak ratios for the protein with 0.05- and 1-mM acetic anhydride (**Figure 1D-E**). The weighted average spectra nearly overlap with that from the native α_3C reproducing the observed insensitivity of ProCharTS at lower concentrations (**Figure 4C**). Our calculations predict that the spectra from Ac1-α_3C (**Figure 4G**) and Ac3-alone (ESI **Figure F18**) appear to be slightly more intense than native α_3C. We are not able to comment on whether this increase is significant to resolve the species with the present set of calculations and experiments. The spectral changes induced by acetylation can be better understood by the UV-Vis response of smaller charged residue clusters (ESI **Figure F19**) extracted from the Ac17-α_3C MD trajectory. We find that ProCharTS in native protein clusters above 370 nm is primarily composed of SS-CT transitions between oppositely charged headgroups of Lys and Glu (ESI **Figure F20 A-B**). Upon acetylation of Lys, these low energy transitions are replaced with higher energy BS-CT transitions between the charged carboxylate and backbone of Glu (ESI **Figure F20C**). Note that the computed spectra in **Figures 4G-F** are not expected to quantitatively reproduce the experimental data in **Figures 4C-D** since the former is calculated for a single species of protein with a distinct acetylated state, whereas the latter arises from the contribution of multiple species. A detailed analysis of these considerations is presented in the discussion section.

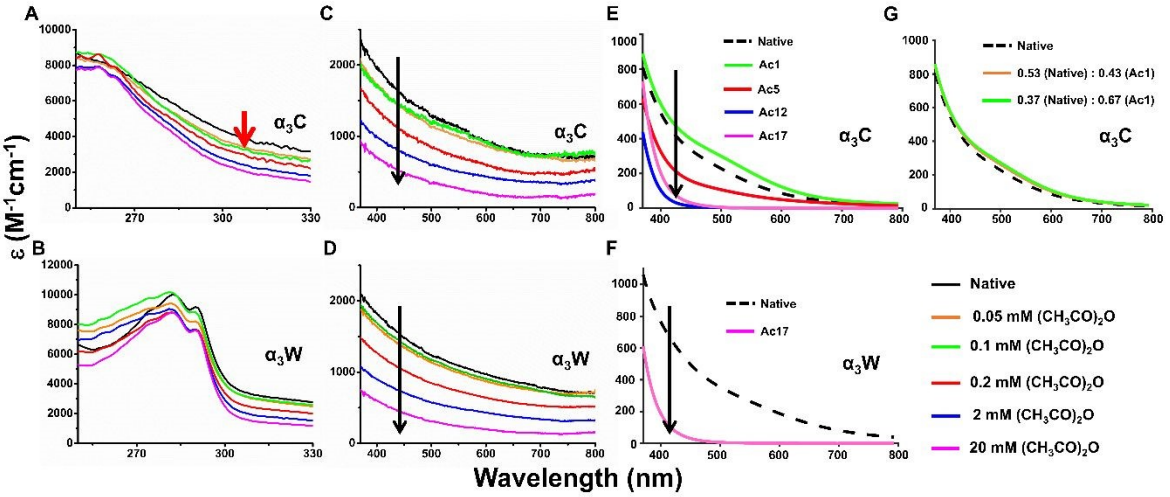

Figure 4: UV-Visible absorption spectra track progressive acetylation in α_3C and α_3W. Experimental (A-D) absorption spectra between 250-330 nm (A-B) and 370-800 nm (C-D) of native/acetylated α_3C and α_3W. All acetylation reactions are performed with 20 µM protein in 100 mM HEPES buffer at pH 8 and room temperature and with varying concentrations of acetic anhydride; 0 (native), 0.05, 0.1, 0.2, 2 and 20 mM. The absorption spectra of native/acetylated samples (~9-10 µM protein) were recorded in deionised water. In α_3C, the absorption is sensitive to low levels of acetylation between 300-400 nm (red arrow in panel A and short wavelength range in panel C). The presence of Trp masks this sensitivity at low acetic anhydride concentrations (B-D). Computed absorption spectra (370-800 nm) of native/ acetylated proteins (E-F). Computed spectra for either the most abundant or one of the highly populated species (see **Figure 1C-N**) present in the experiments at different concentrations of acetic anhydride (different colors in panels E, F). Panel G compares the computed α_3C native protein spectra with that obtained from a weighted average spectrum from α_3C and Ac1-α_3C as per experimental conditions (**Figure 1D-E**) with 0.05 mM (orange) and 0.1 mM (green) acetic anhydride.

We observe that the computed ProCharTS profile for Ac-α_3C and Ac-α_3W is not significantly altered if Aly is included along with charged residues in the clusters (ESI **Figures F21A-B**). This indicates that the neutral residue has a minimal impact on ProCharTS above 250 nm. This was further confirmed by computing the absorption spectra of small clusters (dimer/hexamer) of charged residues with and without Aly (ESI **Figure F19)**. The computed spectra of Aly-Glu dimer (blue line) are nearly identical to that of isolated Glu taken from the same dimer. The minute differences seen are probably due to electrostatic effects coming from Aly which are absent for the Glu monomer. In the case of hexamer, the cluster upon acetylation reduces to a Glu dimer and isolated Glu monomer (with R_c=6Å). In ESI **Figure F19D**, cluster spectra with and without the Aly show differences below 450 nm as the latter misses the electrostatic influence of three Aly residues. While the red-most transition, BS-CT from negatively charged carboxylate headgroup to the backbone of Glu remains the same independent of inclusion of Aly (ESI **Figure F20C-D**), the transition is redshifted in the presence of Aly which further confirms that electrostatic influence of the neutral residue. Nevertheless, the effect of acetylation in reducing the spectra is fully captured independent of whether the Aly is included or not. These calculations further strengthen our model of including only charged residue clusters to study the effect of acetylation on ProCharTS.

3.5 ProCharTS luminescence is not an effective probe of chemical acetylation in α_3C and α_3W

Luminescence can arise when charges separated in the excited state subsequent to charge transfer in charge-rich proteins undergo charge recombination. ProCharTS luminescence intensities have previously been shown to correlate with population of proximal charges within the protein[54]. We therefore examined the possibility of using ProCharTS luminescence features to track acetylation in α_3C and α_3W. Since ProCharTS arises from the absorption of multiple chromophores (charged amino acid residue clusters) over a broad wavelength range, the resultant emission is excitation dependent. We therefore excited the native and acetylated forms of the proteins at multiple wavelengths and recorded the emission as a function of wavelength (**Figure 5).** The ProCharTS native form luminescence of α_3C and α_3W is retained at low acetic anhydride concentrations (0.05 and 0.1 mM), and shows reduction in intensity as acetylation increased (for acetic anhydride concentrations ≥ 0.2 mM) at all the excitation wavelengths. However, the decrease is not systematic with increasing acetic anhydride concentrations with trends highly sensitive to the excitation wavelength.

To determine the primary cause behind the decrease in the luminescence intensity after acetylation, we calculated the integrated quantum yield (QY) of both native and acetylated forms of α_3C/α_3W relative to reference 9,10-diphenylanthracene (DPA) and NATA excited at 355 and 280 nm, respectively. In α_3C the luminescence from 280 nm excitation exhibits a significant reduction in the quantum yield as the concentration of acetic anhydride increases beyond 0.2 mM but is nearly similar to that of the native protein at lower concentrations (ESI **Figure F22**). This suggests that the charge recombination is lowered post-acetylation with higher concentrations of acetic anhydride. However, the luminescence QY with 355 nm excitation in both the proteins is insensitive to the degree of acetylation (ESI **Table T7**), which

View Article Online
DOI: 10.1039/D5SC09293K

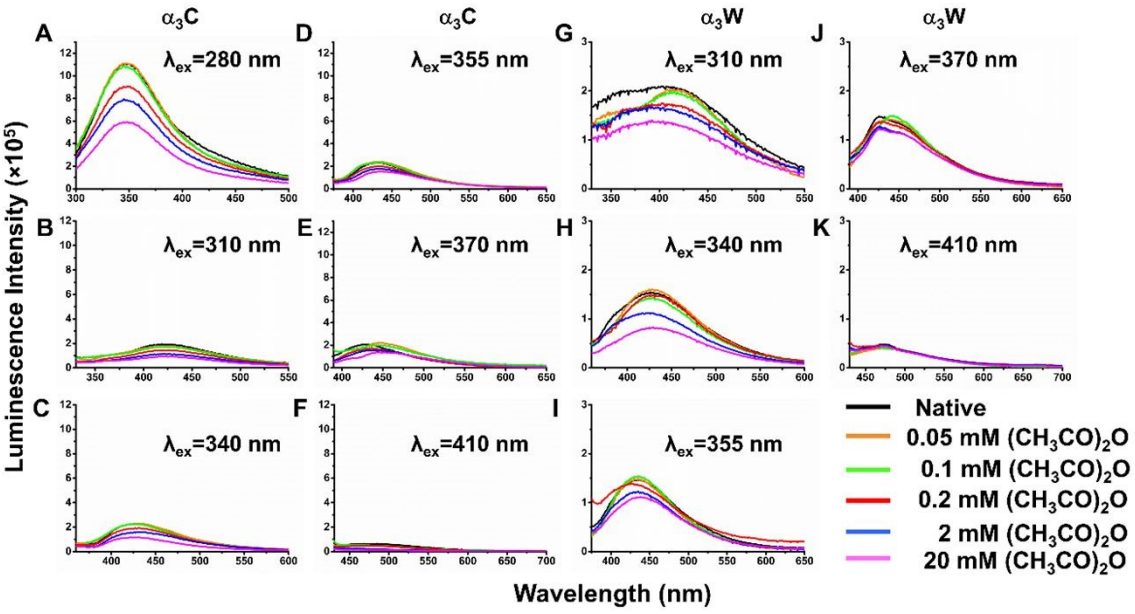

Figure 5: Emission spectra of acetylated and native α_3C and α_3W. (A-F) Luminescence emission spectra of ~9-10 µM native and acetylated α_3C (D-I) and α_3W (J-K) were recorded in deionised water at room temperature. Spectra were obtained by exciting the proteins at 280 nm (only α_3C) 310, 340, 355, 370 and 410 nm with a slit width of 2 nm, and their emission profiles recorded between 330-550, 360-600, 375-650 and 430-700 nm respectively with an emission slit width of 15 nm.

indicates that the reduction in luminescence cannot be attributed to the changes in QY. Rather, differences in protein absorption convoluted with different radiative and non-radiative relaxation mechanisms of the excitations lead to the complex changes in the luminescence of α_3C and α_3W post-acetylation. Furthermore, the sensitivity of the emission spectra is the same as that for the absorption profiles, both not responsive at low concentrations ($\leq$ 0.1 mM) of acetic anhydride. Based on these considerations, we conclude that ProCharTS luminescence does not directly shadow charge neutralizing modifications of residues and is therefore not as effective as absorption at tracking PTMs such as acetylation in proteins.

4. Discussion

A direct label-free optical spectroscopic detection of PTMs offers many advantages in comparison to conventional techniques such as mass spectrometry or antibody-based detection. Absorption and emission spectroscopy are some of the most accessible characterization techniques in protein biophysics and biochemistry labs around the world. Optical techniques are least perturbative and can report on functional assays in solution phase[49,51,80]. Additionally, extensions to *in-vivo* measurements and more advanced non-linear (ultrafast) spectroscopies is also possible[81,82]. Recently there is growing interest in using spectroscopy to probe the collective optical properties of amino acids in supramolecular architectures and utilize them in sensing and therapeutic applications[83]. However, studies in this direction have been restricted to aromatic amino acids (Phe, Tyr, Trp). From the vantage point of protein biophysics and biochemistry, it would therefore be very useful to extend optical spectroscopy to detect PTMs

View Article Online
DOI: 10.1039/D5SC09293K

in protein solutions by establishing distinct chemical signatures of modified and unmodified residues. In this direction, recently, vibrational markers were proposed for acetylated and non-acetylated forms of Lys using Raman and FTIR spectroscopy of powdered samples of the amino acids in combination with DFT calculations[41]. However, these vibrational fingerprints and their sensitivity to number of PTMs remain to be tested in solution phase. Here, we show for the first time that emergent charge transfer transitions from collective excitations of charged amino acid residues provide distinctive signatures in electronic absorption spectroscopy to detect PTMs in solution phase. We have demonstrated that ProCharTS absorption profiles of α_3C and α_3W show a monotonic decrease in intensities between 370-800 nm with progressive chemical acetylation of the proteins in acetic anhydride solutions. Mass spectrometry (**Figure 1C-N**) confirms that between 1-19 acetyl groups are added to both α_3C and α_3W which are assigned to modifications of 17 Lys residues, the protein N-terminus, and the single Ser residue in the proteins. Additionally, a weak signal for an extra acetyl group in α_3C is assigned to the buried Cys residue in the protein. Since the ProCharTS profiles are broad and featureless, we developed a computational framework to simulate and deconvolute the spectral profile in terms of the underlying charged amino acid residue chromophores. This provides us with new insights to develop optical signatures of acetylation in terms of PICTs arising within the protein fold as discussed below.

An analysis of the computed spectra of α_3C/α_3W show that the addition of acetyl groups to Lys residues significantly alters the clustering among charged amino acid residues to influence the spectra. For instance, as illustrated in ESI **Figures F7-8,** acetylation leads to an increase in the number of charged residue monomers in Ac17-α_3C compared to α_3C, accompanied by a corresponding decrease in the size of larger clusters, particularly tetramers to heptamers. For other acetylated systems, we find a wide range of clusters all of which are uniformly populated with charged residues (ESI **Figure F23A-E(i)**) when 12 or less residues are acetylated. However, as the degree of acetylation progresses and charged Lys residues are neutralized, we see a gradual increase in the population of monomers and dimers at the expense of the larger clusters (ESI **Figure F23A-E(ii)** and **Table T10-11**). Eventually, when all Lys residues are acetylated, we find most of the residues (~85%) in monomeric form for Ac17-α_3C (ESI **Figure F23F(i-ii**)). We find that the decrease in spectral intensity is correlated with the total number of monomers, dimers and trimers (ESI **Figure F28)** linking the decrease in the ProCharTS intensity of spectra post-acetylation to the reduction in the clustering propensities of oppositely charged Lys and Glu amino acid residues.

In previous computational studies we had shown that the excess charge on the sidechains of amino acid residues like Lys and Glu creates a polarization of frontier filled orbitals to create an electronic donor-bridge-acceptor (D-B-A) molecular architecture[45]. As a result, while monomeric Lys or Glu residues absorb in the deep UV, they show facile backbone-sidechain (BS) PICT transitions. More interestingly, when these D-B-A residues (Lys and Glu) electrostatically interact through their charged headgroups, both the nature of the PICT and their absorption wavelengths change. An oppositely charged Lys-Glu dimer which is separated by 5-6 Å shows 6 different inter/intra-residue PICTs including visible sidechain to sidechain (SS) PICT from the Glu carboxylate group to the Lys amino group[46]. ESI **Figure F29** shows that the decrease in spectral intensity is strongly correlated with number of Lys acetylated and

View Article Online
DOI: 10.1039/D5SC09293K

decrease in Lys-Glu dimers which are separated by more than 4 Å. Thus, the intensity decrease is a direct effect of Lys charge neutralization which is well accordance with the previous finding[44]. Thus, based on these results we hypothesize that the classic ProCharTS profile of α_3C in the near UV and visible arises from the contribution of oppositely charged amino acid residue dimers and higher order clusters. Below we carry out a deconvolution of the simulated spectra of native α_3C to validate this hypothesis which allows us to connect the spectral differences post-acetylation to the underlying changes in PICT processes.

To fully understand the contributions of charged amino acid residue clusters to ProCharTS, we first decomposed the computed spectral profiles of native and acetylated α_3C into contributions from individual snapshots from the MD production trajectories. **Figure 6A and 6B** showcase the final averaged spectra (Black line) of native α_3C and Ac17-α_3C against a backdrop of contributions (different colours) arising from 50 snapshots from the production runs. Similar analysis for other acetylated systems are shown in ESI **Figure F24**. Interestingly, while the averaged ProCharTS profile for native α_3C is featureless, 46 % of the MD snapshots show a rich array of features including clearly defined peaks ranging from the UV to the visible (**Figures 6A** and ESI **Figure F25**). Strikingly, in fully acetylated Ac17-α_3C all features disappear and the resultant weak ProCharTS absorption is comprised of featureless contributions from individual snapshots (**Figure 6B**). Further analysis reveals that oppositely charged dimers contribute significantly to the spectra exhibiting peak features (featured spectra). For instance, the spectral profile from the snapshot indicated by the black arrow in **Figure 6A**, arises predominantly from dimer and trimer contributions between 350-450nm and solely from dimers above 450 nm (**Figure 6C**). In this specific snapshot, there are a total of five oppositely charged dimers present but only two of them with separation distances between 5-6 Å exhibit SS-PICT transitions which contribute to the spectra beyond 450nm (**Figure 6D**) and mixture of SS-PICT and BS-PICT transitions which contribute below 450 nm. Extending this analysis to the full set of MD snapshots we find that isolated oppositely charged dimers predominantly contribute to the native α_3C absorption profile above 450 nm in more than 50% of the snapshots exhibiting featured spectra (**Figure 6E**). In remaining snapshots which exhibit featureless spectra, dimers and trimers both predominantly contribute (**Figure 6F**). Similar analysis for acetylated α_3C systems are also shown in ESI **Figures F26-27**. The ratio of snapshots showing featured versus those showing featureless spectra decreases as α_3C is progressively acetylated and falls rapidly when 5 or more Lys residues are acetylated.

Interestingly, while the averaged absorption spectrum decreases with increasing acetylation, it does not exhibit any peak features in any of these systems despite possessing multiple snapshots with featured spectra. When the degree of acetylation is low (Ac1- and Ac3-α_3C), oppositely charged dimers still dominate the spectral intensity above 450 nm and produce a slight enhancement in the spectra relative to native α_3C (**Figure 4G** and ESI **Figure F17**). Only when five or more Lys residues are acetylated, do we observe a significant decrease in the number of snapshots showing featured spectra. Further even for the featureless spectra, the tails beyond 500 nm reduce with increasing acetylation. We find that featureless spectra in all systems arise predominantly from isolated Glu monomers and either neutral (Lys-Glu pairs forming a salt-bridge) or negatively charged Glu dimers (ESI **Figures F26-27**). These DBA systems lack SS-PICT transitions and show only BS-PICT or backbone-backbone (BB) PICT [44,46]This is also

View Article Online
DOI: 10.1039/D5SC09293K

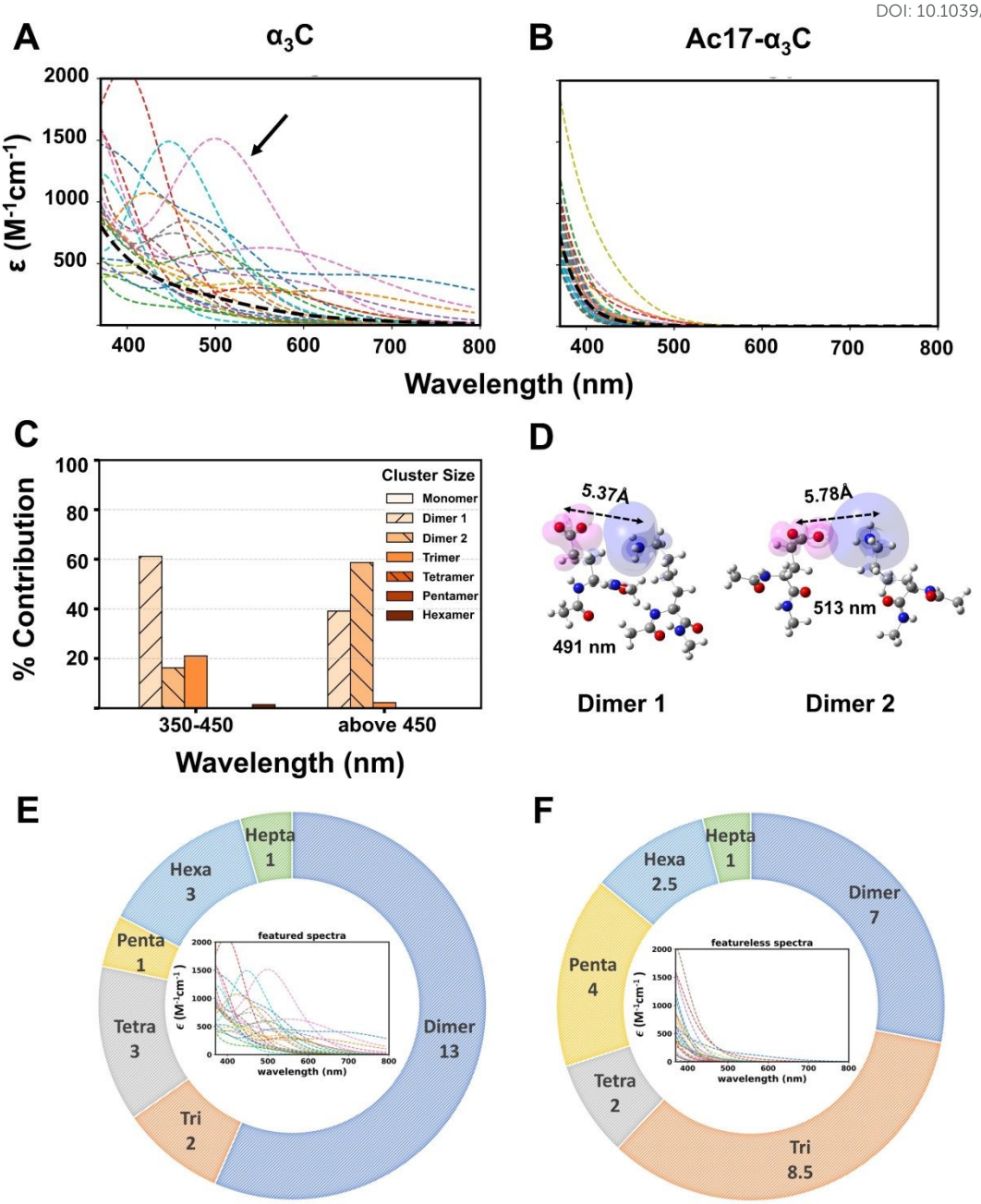

Figure 6: Spectral deconvolution reveals SS-PICT in oppositely charged dimers as key markers of the degree of acetylation. Simulated spectra (black dotted line) of (A) α_3C, and (B) Ac17-α_3C. The spectrum for each system is obtained by averaging over 50 snapshots (coloured dotted lines) from the 50 ns production MD runs. The spectra from individual snapshots are classified as featured or featureless based on whether they show peaks or not (C) Contribution of clusters to the intensity of a representative snapshot exhibiting featured spectra (black arrow in panel A) between 350-450nm and above 450 nm. (D) The two dominant dimers, Lys32:Glu28 (Dimer 1) and Lys40:Glu49 (Dimer 2) and their corresponding PICT transitions showing hole (pink) and electron (blue) density. These transitions produce the peak features in the profile of the representative snapshot (black arrow in panel A). Contributions of clusters to the set of featured (E) and featureless (F) spectra for native α_3C (panel A) in terms of the number of snapshots in which each cluster dominates the spectral intensity (above 450nm (E) and 350-450nm (F)). Two frames had zero transitions above 350nm and were discarded in this analysis.

contributing to the spectra below 450 nm. Taken together, the analysis here also shows that long range (5-6Å) SS-PICT transitions in oppositely charged dimers, are responsible for the extended absorption feature above 450 nm for the native protein. With increasing acetylation, these long-range PICT transitions are turned off due to the reduction in population of oppositely charged residue pairs and manifests in the observed drop in ProCharTS intensity in the visible. Given the broad features of the spatiotemporally averaged ProCharTS profile (**Figure 4**) and experimental limitations in selectively acetylating specific residues during chemical acetylation assay which results in an additional averaging over multiple species (in terms of acetylation sites), it is not possible to directly verify the transient clusters predicted here by computations. However, our sensitivity analysis of the Ac12-α_3C (ESI **Section M1.2.5** and **Figure F18**) indicates that the ProCharTS profile may be much more sensitive to PTMs at certain sites in relation to others, something which can be inferred from computational screening and then verified by experiments which can carry out selective single point perturbations.

Charge neutralization of Lys (acetylation or any such PTMs) can create a number of changes, both local and global, in the protein structure which impact the spectra to differing extents. Some changes are universal, such as changes in the composition and size distribution of charged amino acid clusters, tertiary interactions, and changes in the overall charge of the protein, Others such as changes in the protein secondary structure, conformational changes, or changes in function (e.g enzymatic activity) are very protein specific. The sensitivity of the protein spectral changes to PTMs will also be protein specific depending on whether these effects act in a concerted or opposing mode to alter the ProCharTS activity. Both computational and experimental techniques are required to understand the interplay of these effects and deconvolute them. In the present manuscript we have attempted to do precisely this for our model systems as noted below. For our model protein systems, the CD, Trp spectra, and MD simulations (**Figure 2**) indicate no major changes in the secondary structure of the protein or the overall protein fold which leads us to the conclusion that the associated spectral changes do not arise from factors such as backbone conformational changes. On the other hand, these techniques do suggest a weakening of tertiary interactions which leads to a systematic increase in R_g and can be attributed to changes in the surface charges and polarity of the protein as indicated by ANS binding, zeta potential measurements (**Figure 3**). Further, Trp emission and SASA calculations (**Figure 2D-E**) on α_3W also indicate local unfolding of the structure exposing the Trp to solvent post-acetylation is attributed from simulations to disruption of a Lys38-Glu35 salt bridge in the vicinity due to charge neutralization. To more clearly show the impact of these changes on the spectra we examined the distribution of cluster sizes as a function of acetylation which clearly indicates a reordering of cluster sizes upon acetylation. Scatter plots (ESI **Figure F28**) between spectral intensity and number of clusters of smaller size (monomers, dimers and trimers) show that the systematic decrease in the intensity of spectra is correlated to the reduction in the clustering propensities of oppositely charged Lys and Glu amino acids. The drop in spectral intensity correlates even more strongly (ESI **Figure F29** and **Tables 8-9**) with protein R_g, the number of Lys acetylated or the number of Lys-Glu dimers which are separated by more than 4 Å. Essentially, changes in cluster size reordering or reductions in the number of oppositely charged Lys-Glu dimers or even the changes in R_g can

View Article Online
DOI: 10.1039/D3SC09293K

all be traced back to charge neutralization (acetylation). As demonstrated here, the combination of computational and experimental analysis can be used to dissect the nature of changes induced by the PTMs which manifest in spectral changes.

Our experimental studies and computational analysis identify a new optical mode employing charge transfer transitions to follow acetylation or indeed any PTM that alters the charged state of residues. We term the new detection mode ProCharTS^{PTM} which has the potential to be developed into a cheap and viable alternative to techniques such as Mass Spectrometry and antibody-based assays. The strategy rests on the enhanced sensitivity of ProCharTS to charge altering PTMs is due to charge transfer (CT) character of the underlying transitions which show significant shifts and changes in intensity in response to perturbations in amino acid charge and their clustering[44]. In the present study, we note that experimental ProCharTS spectra show noticeable sensitivity (**Figure 4A-B**) to detecting lower levels of acetylation (at 0.05 and 0.1 mM acetic anhydride) in the 300-400 nm window. Mass spectrometry data indicates that only 0 and 1 Lys acetylated species are present at these low concentrations of acetic anhydride. Based on the data from α_3W, the strong absorbance of tryptophan can mask this sensitivity and could possibly be deconvoluted with the aid of computations as noted above. In order to assess the applicability of ProCharTS to biologically relevant proteins, we carried out chemical acetylation studies on two proteins, the GTPase K-RAS which regulates cellular response and the Histone H2A which regulates gene expression, where acetylation has been shown to play important roles[84,85]. Both proteins exhibit a clear systematic monotonic decrease in ProCharTS absorbance with increasing concentrations of acetic anhydride (ESI **Figure F31**). Additionally, both proteins have a lower predominance of charged amino acid residues (~30-40% of the sequence) and a higher molecular weight compared to α_3C and α_3W (ESI **Table T12**). These results show that ProCharTS^{PTM} can also be used to detect modifications in biologically relevant proteins, even when the charge content is low.

Despite these promising results, there are limitations on the ProCharTS^{PTM} measurements and analysis which need to be addressed in order to develop the technique further. We note that there are quantitative differences between the measured and computed spectra (**Figure 4**), particularly in terms of intensity, which is lower for the latter and the sharper drop in the ProCharTS tail for the computational data. Further, the computed data cannot resolve the drop in ProCharTS intensity when more than 12 Lys are acetylated. These differences and limitations can be attributed to the diversity of protein species contributing to the experimental spectra and approximations used in our computational analysis. The mass spectrometry data reveal (**Figure 1C-N**), that the experimental ProCharTS signal arises from multiple protein species with different numbers of acetylated residues. Additionally, even for a fixed number of acetylation, there are multiple protein species present. For instance, in Ac1-α_3C there are 18 possible amino groups which can be acetylated. While all modifications sites are not equally accessible, a solvent accessibility analysis (ESI **Figure F5A**) indicates that the number of degenerate possibilities is still large. Our computational examination of single amino acid permutations in the Ac12-α_3W system shows that the choice of amino acid site acetylated can impact the ProCharTS profile above 400 nm significantly. Thus, specific acetylation sites could potentially be detected with suitable combinations of experiments and computations. For instance, experiments could be improved to selectively acetylate Lys sites or generate samples with one

View Article Online
DOI: 10.1039/D3SC09293K

or a few species at most. Alternatively, the computations could be made more efficient to sample all possible permutations and combinations of PTM sites present in the sample. The prohibitive computational costs of a full detailed MD plus TDDFT framework presented can be mitigated by identifying geometric parameters/feature spaces to generate spectra using AI/ML approaches. Reducing the cost of the computations will also help improve the quality of simulated spectra. For instance, to generate the spectra in **Figure 4** at a reasonable cost, the clustering cut-off parameter R_c was set to 6 Å to restrict cluster sizes in ProCharTS simulations to decamers (**Subsection 2.2.4**). Increasing the value of R_c will allow larger clusters which should produce more accurate ProCharTS profiles. Further, our calculations do not include non-charged amino acid residues in the spectra simulations which contribute to the spectra below 300 nm. Finally, even for charged amino acid residue clusters computational cost limits the maximum number of calculated transitions to 120 per cluster covering a spectral range of 250-800 nm. These factors lead to a poor description of the spectra in UV range (up to 380 nm) and can be improved in the future. Moving forward, we suggest the following roadmap to develop ProCharTS^{PTM} into a potent tool for biochemistry and biomedicine:

Stage 1 (Accelerating Computations and Standardizing Experimental Protocols): Reducing the computational time to simulate the spectra of ProCharTS (currently two weeks on a high-performance computing cluster) to less than a day. This can be accomplished by developing either AI/ML strategies or a library of precomputed cluster spectra. On the experimental side, creating standard and systematic protocols for sample preparation (including enzyme catalyzed PTMs) which isolate specific ProCharTS active species along with their spectral response to modifications.

Stage 2 (Spectral Deconvolution and Experimental Validation): Development of computational strategies to deconvolute the experimental signals arising from multiple proteins present in a sample based on individual spectral fingerprints along with experimental validation

Stage 3 (Database of ProCharTS^{PTM} Fingerprints): Creation of a database of ProCharTS active proteins along with their spectral fingerprints. A comprehensive in-silico scan on protein structures can be carried out first followed by experimental validation and spectral characterization. Such a database can be regularly updated to include newly discovered ProCharTS^{PTM} fingerprints.

These steps lay the foundation to develop ProCharTS^{PTM} into a quantitative spectroscopic method for detecting charge-altering PTMs such as acetylation in biological samples.

5. Conclusion

Our results show for the first time that charge transfer transitions can serve as markers to track and study PTMs in charge rich proteins. Here, we use chemical acetylation of two model charge rich proteins α_3C and α_3W to develop an approach (ProCharTS^{PTM}) which combines simple UV-Vis absorption spectroscopy with computational analysis to follow the progressive modifications of their charged residues. Extensive spectroscopic and biochemical characterization enables us to map ProCharTS intensity changes to the number of modified

View Article Online
DOI: 10.1039/D5SC09293K

residues in the proteins. The analysis reveals that the ProCharTS profile obtained with different concentrations of acetic anhydride arises from heterogeneous protein populations with varying degrees of acetylation. Joint experiments and MD computational analysis show that while acetylation does not change the secondary structure of the proteins, their tertiary fold is weakened by acetylation. ANS binding studies, zeta potential measurements, and computational analysis indicate that acetylation makes the surface of the protein more polar and negatively charged. Further hydrophobic pockets become less accessible for ANS binding with increasing extent of acetylation. The ProCharTS emission profiles for the proteins as a function of acetylation are complex showing considerable sensitivity to excitation and emission wavelengths. As such, they are not conducive to tracking PTM modifications in proteins. We believe that our study sets up an exciting new paradigm for tracking biologically relevant PTMs among charge rich proteins, a biomedically relevant class that includes genome regulating proteins such as histones; intrinsically disordered proteins implicated in neuropathological disorders like tau, tumour suppressor protein p53 and gene-regulating transcription factor modifications such as TFIIE, TFIIF; proteins that are activated or inactivated by PTM in complex networks of signaling pathways.

Author Contributions

Conceptualization: RV and RS; Data curation: HMD, AB, AM, and SB; Formal Analysis: HMD, AB, and AM; Funding Acquisition: RV and RS; Investigation: HMD, AB, AM, and SB Methodology: HMD, AB, AM, RV and RS; Project Administration: RV and RS; Resources: HMD and AM; Software: AB and AM; Supervision: RV and RS; Validation: HMD, AB, AM; Visualization: HMD, AB, and AM; Writing-original draft: HMD and AB Writing-review and editing: All Authors.

Conflicts of Interest

There are no conflicts to declare.

Data Availability

All methodology and supporting data are provided in the main manuscript and supporting information (ESI) file. Additional data for this article, including MD trajectories and analysis scripts to generate the figures are available at the Zenodo repository at http://doi.org/10.5281/zenodo.18893836.

Acknowledgements

HMD, SB and RS thank the Department of Biotechnology, Government of India, for providing financial support to the North East Centre for Biological Science and Healthcare Engineering (NECBH) through project number BT/NER/143/SP44675/2023. The authors acknowledge the NECBH and Central Instrument Facility (CIF) at IIT Guwahati for providing access to their instruments. AB, AM, and RV acknowledge the funding support from the Department of

View Article Online
DOI: 10.1039/D3SC09293K

Atomic Energy (DAE), Government of India, under Project Identification No. 1303/10/2023-R&D-II-DAE/TIFR-17248 RTI 4015. HMD, SB, and RS thank Prof. Bhubaneswar Mandal, Ms. Subhashree Sahu (Department of Chemistry, IIT-Guwahati), Prof. Sachin Kumar (Department of Biosciences and Bioengineering, IIT-Guwahati) and Dr. Devyani Haldar (CDFD Hyderabad) for providing Acetic anhydride, for assisting with lifetime measurements, for gifting K-RAS protein, and for gifting Histone H2A DNA construct respectively. The authors thank Prof. Jyotishman Dasgupta (Department of Chemical Sciences, TIFR-Mumbai) for fruitful discussions on the electronic spectral profiles.

References

1 C. T. Walsh, S. Garneau-Tsodikova and G. J. Gatto, *Angew Chem Int Ed*, 2005, **44**, 7342–7372.
2 R. Aebersold, J. N. Agar, I. J. Amster, M. S. Baker, B. Zhang, et al., *Nat Chem Biol*, 2018, **14**, 206–214.
3 S. Ramazi and J. Zahiri, *Database*, 2021, **2021**, baab012.
4 B. Macek, K. Forchhammer, J. Hardouin, E. Weber-Ban, C. Grangeasse and I. Mijakovic, *Nat Rev Microbiol*, 2019, **17**, 651–664.
5 D. G. Christensen, X. Xie, N. Basisty, J. Byrnes, S. McSweeney, B. Schilling and A. J. Wolfe, *Front. Microbiol.*, 2019, **10**, 1604.
6 R. Bell, R. J. Thrush, M. Castellana-Cruz, M. Oeller, R. Staats, A. Nene, P. Flagmeier, C. K. Xu, S. Satapathy, C. Galvagnion, M. R. Wilson, C. M. Dobson, J. R. Kumita and M. Vendruscolo, *Biochemistry*, 2022, **61**, 1743–1756.
7 K. W. Barber and J. Rinehart, *Nat Chem Biol*, 2018, **14**, 188–192.
8 Á. Marín-Hernández, J. S. Rodríguez-Zavala, R. Jasso-Chávez, E. Saavedra and R. Moreno-Sánchez, *J of Cellular Biochemistry*, 2022, **123**, 701–718.
9 C. N. I. Pang, A. Hayen and M. R. Wilkins, *Journal of Proteome Research*, 2007, **6**, 1833–1845.
10 P. Sorlier, C. Viton and A. Domard, *Biomacromolecules*, 2002, **3**, 1336–1342.
11 M. Liu, L. Guo, Y. Fu, M. Huo, Q. Qi and G. Zhao, *Biotechnology Advances*, 2021, **53**, 107842.
12 E. L. Gershey, G. Vidali and V. G. Allfrey, *Journal of Biological Chemistry*, 1968, **243**, 5018–5022.
13 S. N. Khan and A. U. Khan, *Clinica Chimica Acta*, 2010, **411**, 1401–1411.
14 J. F. Riordan and B. L. Vallee, in *Methods in Enzymology*, Elsevier, 1967, vol. 11, pp. 565–570.
15 K. R. Bridges, G. J. Schmidt, M. Jensen, A. Cerami and H. F. Bunn, *Journal of Clinical Investigation*, 1975, **56**, 201–207.
16 M.-M. Wang, D. You and B.-C. Ye, *Scientific Reports*, 2017, **7**, 14790.
17 E. Verdin and M. Ott, *Nat Rev Mol Cell Biol*, 2015, **16**, 258–264.
18 M. D. Shahbazian and M. Grunstein, *Annu. Rev. Biochem.*, 2007, **76**, 75–100.
19 S. Zhao, W. Xu, W. Jiang, W. Yu, Y. Lin, T. Zhang, J. Yao, L. Zhou, Y. Zeng, H. Li, Y. Li, J. Shi, W. An, S. M. Hancock, F. He, L. Qin, J. Chin, P. Yang, X. Chen, Q. Lei, Y. Xiong and K.-L. Guan, *Science*, 2010, **327**, 1000–1004.
20 C. M. VanDrisse and J. C. Escalante-Semerena, *Annu. Rev. Microbiol.*, 2019, **73**, 111–132.
21 N. Šoštarić and V. Van Noort, *PLoS Comput Biol*, 2021, **17**, e1008988.
22 S. Deota, S. Rathnachalam, K. Namrata, M. Boob, A. Fulzele, S. Radhika, S. Ganguli, C. Balaji, S. Kaypee, K. K. Vishwakarma, T. K. Kundu, R. Bhandari, A. Gonzalez De Peredo, M. Mishra, R. Venkatramani and U. Kolthur-Seetharam, *Journal of Molecular Biology*, 2019, **431**, 2127–2142.
23 U. Mahlknecht and D. Hoelzer, *Mol Med*, 2000, **6**, 623–644.
24 X. Lu, L. Wang, C. Yu, D. Yu and G. Yu, *Front. Cell. Neurosci.*, 2015, **9**, 226.
25 M. S. Liyasova, L. M. Schopfer and O. Lockridge, *Biochemical Pharmacology*, 2010, **79**, 784–791.
26 I. Petitpas, C. E. Petersen, C.-E. Ha, A. A. Bhattacharya, P. A. Zunszain, J. Ghuman, N. V Bhagavan and S. Curry, *Proceedings of the National Academy of Sciences of the United States of America*, 2003, **100**, 6440–6445.

27 I. Petitpas, A. A. Bhattacharya, S. Twine, M. East and S. Curry, *Journal of Biological Chemistry*, 2001, **276**, 22804–22809.
28 P. A. Zunszain, J. Ghuman, A. F. McDonagh and S. Curry, *Journal of Molecular Biology*, 2008, **381**, 394–406.
29 C. Krawic, M. W. Luczak and A. Zhitkovich, *Chemical research in toxicology*, 2024, **37**, 1588–1597.
30 F. Middleton-Davis, A. Davis and K. Middleton, *PloS one*, 2022, **17**, e0268887.
31 K. Zhang, S. Tian and E. Fan, *Analyst*, 2013, **138**, 1628.
32 P. Li, Y. Han, Y. Li, R. Zhu, H. Wang, Z. Nie and S. Yao, *Analytical and Bioanalytical Chemistry*, 2016, **408**, 2659–2668.
33 S. Liokatis, A. Dose, D. Schwarzer and P. Selenko, *Journal of the American Chemical Society*, 2010, **132**, 14704–14705.
34 M. Kasaai, *Carbohydrate Polymers*, 2008, **71**, 497–508.
35 S. Ronzoni, M. Faretta, M. Ballarini, P. G. Pelicci and S. Minucci, *Cytometry Part A*, 2005, **66**, 52–61.
36 E. Hesping, T. S. Skinner-Adams, G. M. Fisher, T. Kurz and K. T. Andrews, *International journal for parasitology. Drugs and drug resistance*, 2020, **14**, 249–256.
37 I. Diallo, M. Seve, V. Cunin, F. Minassian, J.-F. Poisson, S. Michelland and S. Bourgoin-Voillard, *Expert Review of Proteomics*, 2019, **16**, 139–159.
38 L. A. Murray, A. N. Combs, P. Rekapalli and I. M. Cristea, in *Methods in Enzymology*, Elsevier, 2019, vol. 626, pp. 587–620.
39 L. E. Smith and A. Rogowska-Wrzesinska, *Essays in Biochemistry*, 2020, **64**, 135–153.
40 A. R. Farley and A. J. Link, in *Methods in Enzymology*, Academic Press Inc., 2009, vol. 463, pp. 725–763.
41 G. Yao and Q. Huang, *Spectrochimica acta. Part A, Molecular and biomolecular spectroscopy*, 2022, **278**, 121371.
42 L. Homchaudhuri and R. Swaminathan, *Chemistry Letters*, 2001, **30**, 844–845.
43 L. Homchaudhuri and R. Swaminathan, *Bulletin of the Chemical Society of Japan*, 2004, **77**, 765–769.
44 S. Prasad, I. Mandal, S. Singh, A. Paul, B. Mandal, R. Venkatramani and R. Swaminathan, *Chemical Science*, 2017, **8**, 5416–5433.
45 I. Mandal, S. Paul and R. Venkatramani, *Faraday Discuss.*, 2018, **207**, 115–135.
46 I. Mandal, S. Manna and R. Venkatramani, *Journal of Physical Chemistry B*, 2019, **123**, 10967–10979.
47 C. González-González, R. Lopez-Blanco, J. A. González-Vera, S. D'Ingiullo, D. Bouzada, M. Melle-Franco, A. Orte and M. E. Vázquez, *Cell Reports Physical Science*, 2025, **6**, 102631.
48 M. J. Fossat, A. E. Posey and R. V. Pappu, *ChemPhysChem*, 2023, **24**, e202200746.
49 S. E. Alom, S. Kalita, A. H. Kawa, B. Mandal and R. Swaminathan, *Analytica Chimica Acta*, 2024, **1297**, 342374.
50 Mohd. Z. Ansari, A. Kumar, D. Ahari, A. Priyadarshi, P. Lolla, R. Bhandari and R. Swaminathan, *Faraday Discuss.*, 2018, **207**, 91–113.
51 A. Priyadarshi, H. M. Devi and R. Swaminathan, *Biochemistry*, 2023, **62**, 1643–1658.
52 S. E. Alom, K. Swaminathan, V. Nuzelu, A. Singh, H. De Rocquigny and R. Swaminathan, *Biomacromolecules*, 2024, **25**, 6425–6438.
53 N. M. Ennist, S. Wang, M. A. Kennedy, M. Curti, G. A. Sutherland, C. Vasilev, R. L. Redler, V. Maffeis, S. Shareef, A. V. Sica, A. S. Hua, A. P. Deshmukh, A. P. Moyer, D. R. Hicks, A. Z. Swartz, R. A. Cacho, N. Novy, A. K. Bera, A. Kang, B. Sankaran, M. P. Johnson, A. Phadkule, M. Reppert, D. Ekiert, G. Bhabha, L. Stewart, J. R. Caram, B. L. Stoddard, E. Romero, C. N. Hunter and D. Baker, *Nat Chem Biol*, 2024, **20**, 906–915.
54 A. Kumar, D. Ahari, A. Priyadarshi, M. Ziauddin Ansari and R. Swaminathan, *Journal of Physical Chemistry B*, 2020, **124**, 2731–2746.

55 C. Tommos, K. G. Valentine, M. C. Martínez-Rivera, L. Liang and V. R. Moorman, *Biochemistry*, 2013, **52**, 1409–1418.
56 J. H. Waterborg, in *The Protein Protocols Handbook*, ed. J. M. Walker, Humana Press, Totowa, NJ, 2009, pp. 7–10.
57 A. A. Ansari, S. A. Kidwai and A. Salahuddin, *Journal of Biological Chemistry*, 1975, **250**, 1625–1632.
58 J. R. Lakowicz, Ed., *Principles of Fluorescence Spectroscopy*, Springer US, Boston, MA, 2006.
59 S. Sahu, T. Debnath and K. Sahu, *Journal of Physical Chemistry Letters*, 2024, **15**, 3677–3682.
60 C. Louis-Jeune, M. A. Andrade-Navarro and C. Perez-Iratxeta, *Proteins*, 2012, **80**, 374–381.
61 H. M. Berman, *Nucleic Acids Research*, 2000, **28**, 235–242.
62 S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. Van Der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845–854.
63 B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
64 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J Comput Chem*, 2005, **26**, 1701–1718.
65 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell, *Journal of Computational Chemistry*, 2010, **31**, 671–690.
66 P. Mark and L. Nilsson, *Journal of Physical Chemistry A*, 2001, **105**, 9954–9960.
67 T. Darden, D. York and L. Pedersen, *The Journal of Chemical Physics*, 1993, **98**, 10089–10092.
68 G. Bussi, D. Donadio and M. Parrinello, *The Journal of Chemical Physics*, 2007, **126**, 014101.
69 M. Parrinello and A. Rahman, *Journal of Applied Physics*, 1981, **52**, 7182–7190.
70 S. Paul, S. R. K. Ainavarapu and R. Venkatramani, *J. Phys. Chem. B*, 2020, **124**, 4247–4262.
71 W. Humphrey, A. Dalke and K. Schulten, *Journal of molecular graphics*, 1996, **14**, 27-28,33-38.
72 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, D. J. Fox, et al., 2013.
73 K. J. Barber and J. J. Warthesen, *J. Agric. Food Chem.*, 1982, **30**, 930–934.
74 H. R. Ball and S. E. Winn, *Poultry Science*, 1982, **61**, 1041–1046.
75 J. R. Lakowicz, B. P. Maliwal, H. Cherek and A. Balter, *Biochemistry*, 1983, **22**, 1741–1752.
76 L. Stryer, *Journal of Molecular Biology*, 1965, **13**, 482–495.
77 D. Matulis and R. Lovrien, *Biophysical Journal*, 1998, **74**, 422–429.
78 M. Cardamone and N. K. Puri, *Biochemical Journal*, 1992, **282**, 589–593.
79 S. Bhattacharjee, *Journal of Controlled Release*, 2016, **235**, 337–351.
80 A. Priyadarshi, S. B. Saikia and R. Swaminathan, *J. Phys. Chem. B*, 2024, **128**, 9656–9668.
81 G. R. Fleming and G. D. Scholes, *Quarterly reviews of biophysics*, 2024, **57**, e11.
82 J. Pansieri, V. Josserand, S.-J. Lee, A. Rongier, D. Imbert, M. M. Sallanon, E. Kövari, T. G. Dane, C. Vendrely, O. Chaix-Pluchery, M. Guidetti, J. Vollaire, A. Fertin, Y. Usson, P. Rannou, J.-L. Coll, C. Marquette and V. Forge, *Nat. Photonics*, 2019, **13**, 473–479.
83 A. P. Kalra, S. Biswas, I. Mulrain, M. Wang, J. A. Tuszynski and G. D. Scholes, *J. Phys. Chem. Lett.*, 2023, **14**, 5891–5900.
84 M. H. Yang, S. Nickerson, E. T. Kim, C. Liot, G. Laurent, R. Spang, M. R. Philips, Y. Shan, D. E. Shaw, D. Bar-Sagi, M. C. Haigis and K. M. Haigis, *Proc. Natl. Acad. Sci. U.S.A.*, 2012, **109**, 10843–10848.
85 T. Nguyen, D. Sridaran, S. Chouhan, C. Weimholt, A. Wilson, J. Luo, T. Li, J. Koomen, B. Fang, N. Putluri, A. Sreekumar, F. Y. Feng, K. Mahajan and N. P. Mahajan, *Nat Commun*, 2023, **14**, 3357.

View Article Online
DOI: 10.1039/D5SC09293K

Data Availability Statement for the manuscript ""*Label-Free Optical Detection of Protein Acetylation using UV-Vis Charge Transfer Spectra*" by Devi et. al."

All methodology and supporting data are provided in the main manuscript and supporting information (ESI) file. Additional data for this article, including MD trajectories and analysis scripts to generate the figures are available at the Zenodo repository at http://doi.org/10.5281/zenodo.18893836.