



Cite this: *Phys. Chem. Chem. Phys.*,  
2015, **17**, 31177

## How fluorescent labelling alters the solution behaviour of proteins

M. K. Quinn,<sup>a</sup> N. Gnan,<sup>b</sup> S. James,<sup>a</sup> A. Ninarello,<sup>c</sup> F. Sciortino,<sup>bc</sup> E. Zaccarelli<sup>bc</sup> and J. J. McManus<sup>\*a</sup>

A complete understanding of the role of molecular anisotropy in directing the self assembly of colloids and proteins remains a challenge for soft matter science and biophysics. For proteins in particular, the complexity of the surface at a molecular level poses a challenge for any theoretical and numerical description. A soft matter approach, based on patchy models, has been useful in describing protein phase behaviour. In this work we examine how chemical modification of the protein surface, by addition of a fluorophore, affects the physical properties of protein solutions. By using a carefully controlled experimental protein model (human gamma-D crystallin) and numerical simulations, we demonstrate that protein solution behaviour defined by anisotropic surface effects can be captured by a custom patchy particle model. In particular, the chemical modification is found to be equivalent to the addition of a large hydrophobic surface patch with a large attractive potential energy well, which produces a significant increase in the temperature at which liquid–liquid phase separation occurs, even for very low fractions of fluorescently labelled proteins. These results are therefore directly relevant to all applications based on the use of fluorescent labelling by chemical modification, which have become increasingly important in the understanding of biological processes and biophysical interactions.

Received 29th July 2015,  
Accepted 19th October 2015

DOI: 10.1039/c5cp04463d

[www.rsc.org/pccp](http://www.rsc.org/pccp)

### Introduction

The ability to harness the versatility of anisotropic colloids to build new materials is of significant interest to soft matter scientists.<sup>1,2</sup> Advances in synthesis have allowed these materials to be exploited as biomaterials and/or photonic devices.<sup>1,3</sup> Concurrently, development of patchy particle models has occurred in which isotropic interaction potentials are replaced by those incorporating anisotropy, by specifically defining regions on the particle surface with a number of fixed (attractive) potentials.<sup>4–6</sup> Here, there is a direct analogy to be drawn with proteins. Proteins can also be considered as patchy particles, in which the diversity in the chemical surface, due to variations in the surface exposed amino acids, creates anisotropy at a molecular level.<sup>7</sup> Indeed, it is widely acknowledged that protein phase behaviour can only be fully explained by accounting for these anisotropic effects.<sup>8,9</sup>

The modification of proteins by chemical methods is a versatile mechanism to label a protein for detection, or to modify its behaviour, for example by (PEG)ylation, which increases the number of potential variations (and therefore functions) of a

particular protein.<sup>10–13</sup> It is currently exploited for several applications, ranging from the adaptation of therapeutic biomolecules to improve stability and/or function to the production of novel, functional materials and bottom-up approaches to synthetic biology.<sup>10,14,15</sup> One of the most useful ways that chemical modification is already exploited is for the fluorescent labelling of proteins, which allows direct visualisation of biomolecules *in situ*, with increasing resolution due to the ongoing development of specialised imaging techniques.<sup>16,17</sup> Indeed, exciting advances in both microscopy and single-molecule detection are currently used for investigating protein distribution, translocation and interactions both *in vitro* and *in vivo*.<sup>18</sup> Covalently attached fluorescent dyes have also been used as a means to screen for the formation of protein crystals.<sup>19</sup> Thus, the use of fluorescent labelling of proteins plays an important role for the exploration and understanding of the mechanisms involved in many biological problems, such as intra/inter-cellular communication, genomics, unravelling the origins of pathologies associated with protein condensation diseases and the mechanisms that govern protein-self association during the production of proteins as therapies.<sup>12,18,20,21</sup> To fluorescently label proteins, there are currently three main established strategies; small organic molecules that are covalently attached to the biomolecule of interest, fusion proteins in which an inherently fluorescent protein is co-expressed (or fused) with the protein of interest or the use of quantum dots (or small molecule fluorescent dyes),

<sup>a</sup> Department of Chemistry, Maynooth University, Maynooth, Co. Kildare, Ireland.  
E-mail: [jennifer.mcmanus@nuim.ie](mailto:jennifer.mcmanus@nuim.ie)

<sup>b</sup> CNR-ISC UOS Sapienza, Piazzale A. Moro 2, 00186 Roma, Italy

<sup>c</sup> Dipartimento di Fisica, "Sapienza" Università di Roma, Piazzale A. Moro 2, 00186, Roma, Italy

that are functionalised to form a biotin–avidin pair with the biomolecule under study.<sup>22–24</sup> Of these three routes, covalent attachment of an organic fluorophore is considered to be the most versatile method of labelling for fluorescence imaging and sensing of biological specimens.<sup>22</sup>

While the pace of advances in fluorescence imaging and single molecule analysis increases rapidly, little consideration has been given to the impact that fluorescent labelling of proteins has on protein solution behaviour and if these need to be considered when interpreting fluorescence data. The solution behaviour of a protein is very sensitive to changes in its environment.<sup>25</sup> Protein solution behaviour is also sometimes dramatically altered by (comparatively minor) changes to the protein surface, for example by mutagenesis, specific binding of ions or by changes in pH and this can have consequences for the pathogenesis of several known protein condensation diseases.<sup>26–28</sup> Hence, one would also expect a significant impact due to the presence of a fluorescent label. Some studies have discussed the role that fluorescent labelling has on specific protein characteristics; bovine serum albumin (BSA) labelled with fluorescein isothiocyanate (FITC) exhibited a larger diffusion coefficient and higher levels of irreversibly adsorbed proteins at the oil–water interface.<sup>29</sup> A chemically modified green fluorescent protein (GFP) exhibited a more acidic pI than the unlabelled protein.<sup>30</sup> Several different small molecule fluorophores have also been investigated to determine which had least effect on cellular function and adherence of leucocytes to endothelial cells.<sup>31</sup> A further study has shown that monomerically labelled and mixtures of unlabelled and labelled Gag (a viral protein) assemble into morphologically indistinguishable clusters at a lipid membrane.<sup>32</sup>

It is therefore reasonable to ask how and to what extent fluorescent labelling of proteins alters the solution behaviour and if this has consequences for understanding protein interactions within the cellular environment and more widely in interpreting analytical data which requires proteins to be labelled. To do this directly and unambiguously, we examine variations in the phase behaviour of protein solutions upon addition of labelled proteins. Phase diagrams are used to characterize the interactions between macromolecules, including colloidal particles and proteins.<sup>8,25,33,34</sup> These studies have shown that the effective inter-protein interaction potential results from the complex interplay of both attractive and repulsive contributions.<sup>35</sup> Proteins are inherently anisotropic, since each surface exposed amino acid contributes to varying degrees to the overall net-interaction potential.<sup>9</sup> This anisotropy has been shown to have a direct influence on protein phase behaviour.<sup>26,33,36,37</sup> Liquid–liquid phase separation (LLPS) occurs for proteins interacting *via* short-ranged net attractive forces.<sup>25,33,34</sup> Indeed, the temperature at which LLPS occurs is sensitive to the strength of net attraction between proteins in solution and can be used to probe the relative changes in protein inter-particle interactions as a result of altered protein chemistry, changing solution conditions or the presence of a second protein type.<sup>38,39</sup>

In this study, we experimentally assess the phase diagram for human  $\gamma$ D-crystallin (HGD), which is a structural protein in the eye lens that undergoes LLPS due to short-ranged attractive

interactions which dominate its behaviour at physiological pH (at the pI of the protein).<sup>9,37,39,40</sup> The native protein is remarkably stable *in vivo*.<sup>41</sup> Several point mutations in HGD have been linked to congenital cataract formation due to significantly decreased solubility of the protein at body temperature leading to increased light scattering and, eventually, blindness.<sup>37,40–44</sup> Therefore, LLPS of HGD is a readily quantifiable phase transition through which the effect that the addition of a fluorescent dye, by covalent attachment of an organic fluorophore, can be probed. Additionally, the surface chemistry of HGD ensures that the both the position and number of small molecules conjugated to the protein can be controlled.

We will quantify the changes in phase behaviour of HGD upon addition of a small amount (less than 1% total protein) of fluorescently labelled proteins. We will demonstrate that, even when the fraction of labelled proteins is very low, the chemical modification introduced by small molecule tagging leads to a collective increase in net attraction between proteins in solution. To account for these surprising results, we propose the hypothesis that the addition of a fluorescent molecule to a protein is the equivalent of adding a large attractive hydrophobic patch to the surface of a protein. By comparison of experimental results with numerical investigations of a specifically devised patchy particle model, we will provide a description that accounts for the increased tendency of protein solutions containing labelled proteins to phase separate. We will also show how fluorescent labelling is a simple yet effective way of demonstrating the directionality of protein–protein interactions by chemical modification (at two different amino acid positions) of HGD. Our results are therefore relevant for all current uses of fluorescent labelling of proteins, clearly indicating that the enhanced protein–protein interactions need to be taken into account for a correct comparison between labelled and unlabelled conditions.

## Methods

### Experiments

#### Sample preparation and experimental techniques

*Preparation of materials.* Analytical grade sodium hydroxide, glacial acetic acid, sodium azide, sodium dihydrogen orthophosphate dihydrate, dibasic sodium phosphate heptahydrate, sodium chloride, sodium borate, dithiothreitol (DTT), hydrochloric acid, LB Broth, LB agar, magnesium sulfate heptahydrate, and ethylenediaminetetraacetic acid (EDTA) were purchased from Fisher Scientific (Dublin, Ireland). Tris hydrochloride was purchased from Merck Millipore (Cork, Ireland). Dimethylformamide (DMF) was purchased from Romil (Cambridge, UK) and used without further purification. All buffers were prepared using Milli Q water. All buffers were filtered through a 0.45  $\mu\text{m}$  MillexHV syringe filter or a 0.45  $\mu\text{m}$  nylon membrane filter (Millipore, Cork, Ireland) prior to use. Amicon Ultra 4 ml centrifugal filters were used for buffer exchange (Millipore, Cork, Ireland). Protein concentration was determined using a mass extinction coefficient value equal to 2.09  $\text{mg}^{-1} \text{ml cm}^{-1}$ .

**Expression, purification and characterisation of protein.** HGD was prepared by recombinant methods as previously described.<sup>45</sup> HGD purity was confirmed at >98% by SDS-PAGE and size exclusion HPLC. The molecular weight of HGD was confirmed to be  $20\,608 \pm 1$  Da by intact molecular weight analysis using electrospray ionisation mass spectrometry, carried out in 'FingerPrints' Proteomics Facility at the University of Dundee, Scotland, UK and is in agreement with previously published data.<sup>45</sup> The volume fraction is calculated using the expression  $c = \phi\nu$  where  $c$  is the concentration in  $\text{mg ml}^{-1}$ ,  $\phi$  is the volume fraction and  $\nu$  is the partial specific volume equal to  $7.1 \pm 0.1 \times 10^{-4} \text{ ml mg}^{-1}$  for the  $\gamma$  crystallins.<sup>46</sup>

**Modification of Lys-2 and Cys-110 amino acid residues.** Amine modification of Lys-2 in HGD was carried out using an AnaTag HiLyte 405 kit, purchased from AnaSpec (Freemont, CA, USA), as per supplied instructions *via* amine modification after purification. Absorbance values to calculate labelling efficiency were determined using a Perkin Elmer Lambda 35 UV/Vis spectrophotometer and UV Win Lab – Scan Lambda 35 software. Thiol modification of Cys-110 in HGD using DyLight 405, Maleimide (Pierce Biotechnology, Rockford, IL, USA) was performed as per manufacturer's instructions after protein purification. Labelling efficiency was determined spectroscopically as for amine modification. Fluorescently labelled proteins were exchanged with 0.1 M sodium phosphate buffer at pH 7 prior to use.

**Modification with FITC.** HGD was labelled using FITC (Pierce Biotechnology, Rockford, IL, USA) as per supplied instructions after purification. Conjugation efficiency was determined spectroscopically as per manufacturer's instructions after extensive washing using a Perkin Elmer Lambda 35 UV/Vis spectrophotometer and UV Win Lab – Scan Lambda 35 software.

**Circular dichroism spectra.** Circular dichroism (CD) spectra were attained for samples of unlabelled HGD, amine modified HGD and thiol modified HGD at a concentration of  $0.1 \text{ mg ml}^{-1}$  in 0.1 M sodium phosphate buffer at pH 7 using near and far UV wavelengths at the Institute of Molecular, Cell, and Systems Biology, College of Medical, Veterinary, and Life Sciences, University of Glasgow, Scotland, UK using a JASCO J-810 spectropolarimeter. OriginPro 9.1 was used to construct the spectra.

**Phase separation measurements.** Liquid–liquid coexistence curves were measured for unlabelled HGD, mixtures of unlabelled HGD and amine modified HGD, and mixtures of unlabelled HGD and thiol modified HGD using a method outlined previously.<sup>25</sup> A Perkin Elmer Lambda 35 UV/Vis spectrophotometer and UV Win Lab – Timedrive Lambda 35 associated software was used to measure the percentage transmission of light ( $\lambda = 600 \text{ nm}$ ). A Thermo Scientific K10 water bath attached to a Thermo Scientific D10 temperature control was used for temperature regulation. The temperature of the system was monitored using an Omega HH509R thermocouple. OriginPro 9.1 was used to construct phase diagrams.

**Fluorescence measurements.** A Molecular Devices Spectra Max M2e plate reader and SoftMax Pro 6.2.1. software were used to spectroscopically measure fluorescence intensity.

## Model and simulation details

**Model for the unlabelled proteins.** The unlabelled HGD proteins (U-type particles) are modelled based on the work of H. Liu *et al.* where Kern–Frenkel patchy particles are complemented with a square-well (SW) attraction.<sup>47,48</sup> This model reproduces the correct width of the experimental coexistence curve of bovine gamma-crystallin. Thus we represent the unlabelled proteins as hard spheres of diameter  $\sigma_m$  interacting *via* a SW attraction of width  $\delta_{\text{SW}}$  and depth  $\varepsilon_{\text{SW}}$

$$V_{\text{SW}}(|\vec{r}_{ij}|) = \begin{cases} \infty & \text{if } |\vec{r}_{ij}| < \sigma_m, \\ -\varepsilon_{\text{SW}} & \text{if } \sigma_m \leq |\vec{r}_{ij}| \leq \sigma_m + \delta_{\text{SW}}, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

where  $|\vec{r}_{ij}|$  is the vector between the centres of particles  $i$  and  $j$ . In addition the surface of the particles is decorated by four randomly-located attractive sites (patches) which do not overlap. The interaction potential between patches is the product of a radial contribution modulated by an angular function

$$V_{ij}^{\text{U}} = V_{\text{SW,U}}(|\vec{r}_{ij}|)G(\hat{r}_{ij}, \hat{r}_{i\alpha}, \hat{r}_{j\beta}) \quad (2)$$

where  $\hat{r}_{i\alpha}$  and  $\hat{r}_{j\beta}$  are the unit vectors from the centre of particle  $i(j)$  to the centre of the  $\alpha(\beta)$  patch on the surface and  $V_{\text{SW,U}}$  is a square well potential of width  $\delta_{\text{U,U}}$  and depth  $\varepsilon_{\text{U,U}}$ . The function  $G$  modulates the potential and depends on the reciprocal orientation of two particles

$$G(\hat{r}_{ij}, \hat{r}_{i\alpha}, \hat{r}_{j\beta}) = \begin{cases} 1 & \text{if } \begin{cases} \hat{r}_{ij} \cdot \hat{r}_{i\alpha} > \cos(\theta_{\text{max}})_{\text{U}}, \\ -\hat{r}_{ij} \cdot \hat{r}_{j\beta} > \cos(\theta_{\text{max}})_{\text{U}}, \end{cases} \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

$\sigma_m$  and  $\varepsilon_{\text{SW}}$  are chosen as the units of length and energy respectively. The temperature  $T$  is measured in units of  $\varepsilon_{\text{SW}}$ . Parameters for the unlabelled patches (U) are shown in Table 1. This is a modified version of the model used in ref. 47 where the patches were not randomly located, but arranged in a tetrahedral geometry on the surface of the particle. This model captures the experimental value of the critical packing fraction  $\phi_c = 0.21$  for HGD. In addition the model correctly reproduces the experimental width of the liquid–liquid coexistence curve.

**Model for the labelled proteins.** Since no minimal models have been proposed to describe the addition of a fluorescent dye to a protein, we have accounted for the presence of a fluorescent dye by adding a fifth patch to the model of the unlabelled protein. Hence a labelled HGD (L-type particles) is the same as an unlabelled particle except for the presence of an additional patch on the particle surface. U-type and L-type particles are sketched in Fig. 1. Accounting for the differences in chemistry

**Table 1** Parameters of the interaction potential for the model of unlabelled proteins as in ref. 47

$\delta_{\text{U,U}}$	$\varepsilon_{\text{U,U}}$	$\cos(\theta_{\text{max}})_{\text{U}}$	$\delta_{\text{SW}}$	$\varepsilon_{\text{SW}}$
0.05	5	0.95	0.5	1.0

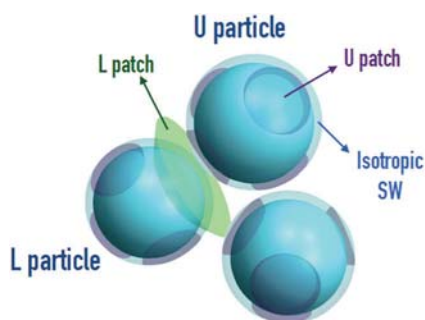


Fig. 1 Patchy particles used for modelling HGD proteins. Particles with four patches (U-patches) are unlabelled proteins (U-type), while the particle with green (wider) patch (L-patch) corresponds to a fluorescently labelled protein (L-type).

between an aromatic amino acid (*i.e.* Phenylalanine, Tryptophan, Tyrosine) and a typical fluorescent dye, we create the additional patch (L-patch) such that it is wider and more attractive than the patches on the unlabelled protein. This assumption is reasonable on the basis that the majority of standard fluorescent molecules have molecular weights that are significantly larger than any of the aromatic amino acid side chains on the protein surface and are also significantly more hydrophobic due to the presence of several aromatic groups (Fig. 2). Using these considerations, we estimate the size of the patch to be 3–4 larger than a hydrophobic amino acid.

The label is chemically attached to the surface of the protein rather than being a composite part of the structure and therefore has the potential to extend beyond the surface of the molecule, thereby allowing contact with more than one other protein. Hence we allow the L-patch to form more than one bond with the U-patches of both U-type and L-type particles (Fig. 1). A second characteristic of the model is that the L-patch on a particle does not directly interact with a L-patch on another L-type particle. If this happens it would cause the L-type particles to associate in small stable clusters since the L–L bond would be the most energetically favourable. This possibility was excluded experimentally.

**Evaluation of the critical points.** To accurately evaluate the critical point and the liquid–liquid coexistence line of the unlabelled particles we perform umbrella sampling grand-canonical Monte Carlo simulations (US-GCMC) and histogram reweighting.<sup>49,50</sup> Thus we identify the correct order parameter  $M$  that, at the critical point, has a probability distribution  $P(M)$  that follows the one typical of the Ising universality class.<sup>51–54</sup> This parameter  $M = \rho + su$  is the linear combination of the

number density  $\rho$  and the energy density  $u$  through the field-mixing parameter  $s$ .<sup>51,52,55</sup>

To locate the critical point we tune the temperature  $T$  and the chemical potential  $\mu$  of the system until the numerical joint distribution  $P(N, E; T, \mu)$  projected over  $N$  shows the double-peaked shape typical of the Ising probability distribution  $P(M)$ ; this is the signature of the presence of large density fluctuations in the system in the vicinity of the critical point. We then implement a fitting procedure based on the histogram reweighting technique that allows us to transform the joint distribution  $P(N, E; T, \mu)$ , evaluated at  $T$  and  $\mu$ , into  $P(N, E; T', \mu')$  at  $T'$  and  $\mu'$  and to extract the best values of  $T'$ ,  $\mu'$  and  $s$  for which the numerical  $P(M)$  is the closest to the Ising distribution.<sup>50,51</sup> The reweighting technique is also used to evaluate the liquid–liquid coexistence line; at temperatures lower than  $T_c$  the liquid–liquid coexistence densities are obtained by imposing the equality of the areas below the two peaks of the reweighted  $P(N, E; T, \mu)$  projected over  $N$  and centred in  $N_c = \rho_c L^3$  where  $\rho_c$  is the critical density and  $L$  is the length of the simulation box.

The procedure to evaluate the critical point in the binary mixtures is similar to that employed for the pure system except for the fact that a different order parameter needs to be considered. In particular we choose  $M = \rho_U + \alpha\rho_L$ , where  $\rho_{U(L)}$  is the number density of the species U(L) and  $\alpha$  is a mixing parameter.<sup>55</sup> Hence the numerical joint distribution  $P(N_U, N_L, \mu_U, \mu_L)$  will depend only on the number of particles of the two species and the associated chemical potentials. Note that since the joint distribution does not depend on the energy,  $T$  is a parameter that cannot be varied to locate the critical points. This means that for each  $\mu_U$  and  $\mu_L$  we search for the critical temperature  $T_c$  of a mixture of a given concentration  $x_{L(U)}$  of species U and L, mapping in this way  $T_c$ . We then select the specific  $T_c$  corresponding to the desired concentration of L-type particles, *e.g.*  $x_L = 0.01$  to  $0.02$ , to compare with the experimental values of concentrations of labelled proteins with amine modification.

## Results and discussion

HGD has two surface-exposed primary amines (the terminal amine and  $\epsilon$ -amine of lysine in position 2). Control of pH during conjugation ensures that only the free amine of the lysine is modified by labelling. HGD has 6 unpaired cysteine residues but only one is surface-exposed (at position 110).<sup>45</sup> Therefore, using commercially available labelling methodologies, it is only possible to label HGD in one position on the protein in each case (*i.e.* for a particular label type only a single molecule of dye is conjugated to the protein) (Fig. 3). This allows us to explore both the chemical and directional anisotropy imposed by the addition of the fluorophore.

Two commonly used fluorescent dyes were selected initially; HiLyte Fluor 405 which labels the lysine in position 2 by formation of a carboxamide bond between the reduced  $\epsilon$ -lysine group of the lysine and the succinimidyl ester of the dye and DyLight Maleimide 405, used to label the cysteine residue at

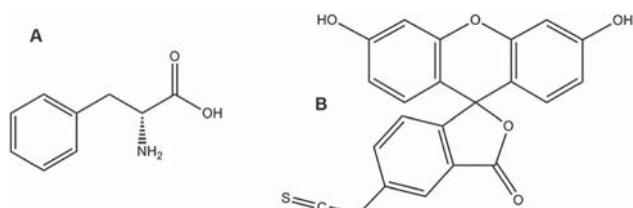


Fig. 2 Structure of Phenylalanine (A) and FITC (B).



Fig. 3 Topological rendering of HGD indicating each site used for fluorescent labelling. Lys-2 is shown in green and Cys-110 in red.

position 110 *via* the formation of a thioether bond between the reduced thiol group of the cysteine and the maleimide group of the dye.

For many applications, labelled protein is used at low labelling densities in solution in combination with unlabelled protein. The high sensitivity of fluorescence methods in general requires low quantities of labelled protein. Using lower amounts of labelled protein may also be used as a strategy to mitigate any potential impact labelling may have on the structure or behaviour of a protein. Typically labelled proteins ranging from  $x_L = 0.0001$  to  $x_L = 0.01$  (where  $x_L$  is the fraction of labelled proteins for a combined protein volume fraction  $x_T = 1$ ) are used. Therefore, protein labelled by the methods described above, were mixed in known quantities with unlabelled HGD and the temperature at which liquid–liquid phase separation occurred was measured at each composition (Fig. 4a and b).

The temperature at which the onset of liquid–liquid phase separation occurs is referred to as the phase separation temperature ( $T_{ph}$ ) and is determined by taking the average of the  $T_{cloud}$  (defined as the temperature at which the solution clouds upon cooling (at 50% transmission)), and  $T_{clear}$  (defined as the temperature at which the solution clears upon reheating) from the experimental measurements. We observe significant increases in the liquid–liquid phase separation temperatures ( $T_{ph}$ ) for the mixtures in both cases at compositions with very low fractions of labelled protein (*i.e.* there is an increase in the temperature at which the liquid–liquid coexistence curve is measured). To a first approximation, we can use the single-component description of the liquid–liquid coexistence curve to describe the experimental data even for the mixtures (eqn (4))

$$\left(\frac{\phi - \phi_c}{\phi_c}\right) = A \left(1 - \frac{T}{T_c}\right)^\beta \quad (4)$$

where  $\phi$  and  $\phi_c$  are the volume fraction of the sample and critical volume fractions respectively;  $A$  is a parameter that determines the width of the liquid–liquid coexistence curve;  $T$  and  $T_c$  are the sample temperature and critical temperature respectively (in Kelvin) and  $\beta$  is an exponent term for the three dimensional Ising model equal to 0.325.<sup>46</sup> The degree to which  $T_{ph}$  changes is, however, different for the two positions at which the protein was labelled. For the protein mixture containing protein modified with HiLyte 405 at the amine 2 position, it was possible to add up to  $x_L = 0.01$  of labelled protein before significant precipitation of protein occurred.

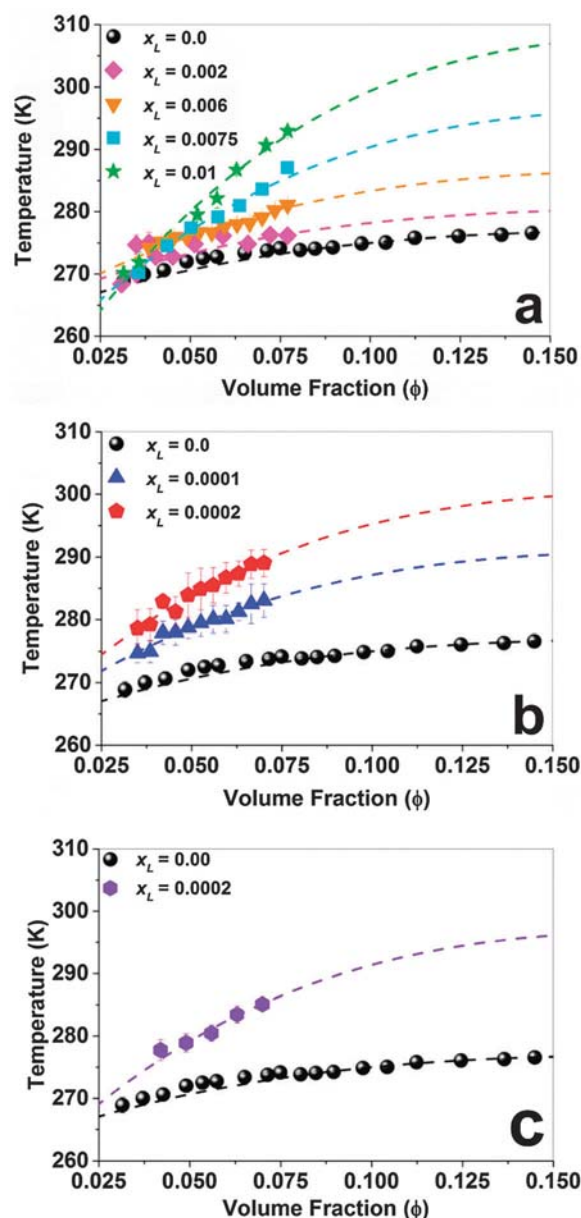


Fig. 4 Liquid–liquid coexistence curves for unlabelled HGD and binary protein mixtures in aqueous solution of unlabelled and amine modified (a and c) and thiol modified (b) HGD with increasing labelled protein fractions ( $x_L$ ). Panel A indicates binary protein mixtures in aqueous solution of unlabelled HGD and HGD amine modified with HiLyte Fluor 405. Panel B shows unlabelled HGD and binary protein mixtures in aqueous solution of unlabelled HGD and HGD thiol modified using DyLight 405 Maleimide. Panel C indicates the liquid–liquid coexistence curves for unlabelled HGD and a binary protein mixture in aqueous solution of unlabelled HGD and HGD modified using FITC (assuming on average 1 dye molecule per labelled protein).

At protein compositions containing up to  $x_L = 0.01$  the protein remained in solution up to a total protein volume fraction of  $\sim 0.07$ . The aggregation that did occur at  $x_L = 0.01$  was due to the formation of reversible non-covalent aggregates above this total protein volume fraction. It was possible to re-disperse these aggregates upon dilution. For the protein

modified at the cysteine 110 position, it was only possible to add labelled protein fractions,  $x_L \approx 0.0002$  before precipitation occurred. We also observe a change in the shape of the liquid–liquid coexistence curve with increasing fractions of labelled protein in each case. It was not possible to obtain further points on the liquid–liquid coexistence curve experimentally, since further concentration of protein within the ranges that we have measured resulted in precipitation (due to the net increase in attraction).

To ensure that these observations were not related only to the specific chemistry of the dyes that we selected, we also labelled protein with FITC, which is a widely used fluorescent dye (for which the chemical structure is known). In this case, we have not specifically labelled a particular amino acid (since FITC will covalently attach to both primary and secondary amines). Again, the FITC labelled proteins were used in mixtures with unlabelled HGD. A significant change in the liquid–liquid coexistence curve is also observed in this case (Fig. 4c). Therefore, the addition of the labelled proteins (and therefore a strong hydrophobic patch) alters the solution behaviour, the extent of which is governed by both the specific chemistry of the dye (*i.e.* its hydrophobicity, and therefore the strength of the attraction) and also by the position on the protein surface in which the label is placed.

The critical point for several gamma crystallins has been measured previously and occurs at  $\phi_c = 0.21$  (300 mg ml<sup>-1</sup>) and 277 K (4 °C).<sup>56–58</sup> To estimate the critical temperature for each of the protein mixtures, we fit each liquid–liquid coexistence curve using eqn (4), assuming that the critical volume fraction ( $\phi_c$ ) did not change with the addition of labelled protein. Values for  $T_c$  (the critical temperature) and  $A$  (related to the width of the liquid–liquid coexistence curve) are shown in Table 2. The width of the liquid–liquid coexistence curve is determined by the range of the inter-protein interaction, with a decrease in width indicative of an increase in the range of the interaction.<sup>57,59</sup> The decreasing width of the liquid–liquid coexistence curve observed indicates that the range and strength of the attractive interaction is increasing. This is consistent with our view that fluorescent labelling is the equivalent of adding a hydrophobic patch that increases inter-protein attraction.

Furthermore, both the molecular weight and number of hydrophobic groups in the fluorescent dye are significantly greater than for a hydrophobic amino acid (Fig. 2). Therefore, we expect that the strength of the attraction at this patch to be higher than other attractive patches on the protein surface.

**Table 2** Estimated changes in critical temperature ( $\Delta T_c$ ) for each protein mixture with different labelled protein fractions ( $x_L$ ) at a total protein volume fraction of 0.21

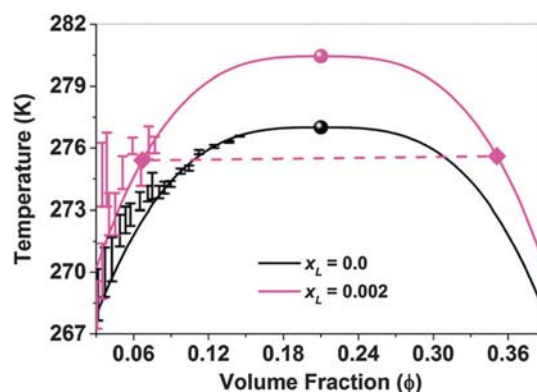
Labelled protein	$T_c$ (K)	$\Delta T_c$ (K)	$A$
Unlabelled HGD	277	0	2.6
HiLyte Fluor 405 labelled amine, $x_L = 0.0002$	280	3	2.5
HiLyte Fluor 405 labelled amine, $x_L = 0.006$	287	10	2.2
HiLyte Fluor 405 labelled amine, $x_L = 0.0075$	297	20	1.8
HiLyte Fluor 405 labelled amine, $x_L = 0.01$	308	31	1.6
DyLight 405 Maleimide labelled thiol, $x_L = 0.0001$	291	14	2.1
DyLight 405 Maleimide labelled thiol, $x_L = 0.0002$	301	24	1.9
FITC labelled HGD, $x_L = 0.0002$	297	20	1.9

These results clearly show that upon increasing labelled protein fractions a significant increase in net-attraction of the system is observed.

The above results are based on the strong assumption that, within the volume fraction range that we have probed, the protein remains monomeric, the critical volume fraction does not change and there is no significant change to protein structure. We performed SE-HPLC (size exclusion high pressure liquid chromatography) on the protein mixtures to determine if it was possible to detect even small amounts of protein aggregates in our samples. None were detected, even at high total protein volume fraction ( $\phi = 0.07$ ). The presence of protein aggregates has been previously shown to increase  $T_{ph}$ , but we find no evidence that this is the source of the increase in  $T_{ph}$  in our case.<sup>60,61</sup>

To probe this further, we induced LLPS in a sample and allowed the concentrated and dilute phases to fully phase separate.<sup>62</sup> We then measured both the total protein volume fraction and the fluorescence intensity in each of the two phases (Fig. 5). The protein volume fractions measured in each phase are consistent with a critical volume fraction that does not change upon the addition of labelled protein (to within the experimentally measurable range). Furthermore, the measurements clearly indicate that the density of labelled protein in the concentrated phase is higher than in the dilute phase and therefore that the fluorescently labelled protein preferentially partitions into the concentrated phase after phase separation. In this case before LLPS, labelled protein is present at  $x_L = 0.002$ . After phase separation the labelled protein fraction in the dilute phase is  $x_L = 0.0012$  and in the concentrated phase is  $x_L = 0.0025$ .

We also performed circular dichroism spectroscopy to determine if structural change to the protein occurs after labelling (Fig. 6). A structural change as a result of unfolding can lead to an increase in net attraction due to exposure of hydrophobic amino acid residues from the protein interior. To establish if this was the case for our protein mixtures in aqueous solution,



**Fig. 5** Liquid–liquid coexistence curves for unlabelled HGD and for a protein mixture containing  $x_L = 0.002$  amine modified HGD (with HiLyte Fluor 405). The diamonds indicate the volume fractions for the dilute phase ( $\phi = 0.067$ ) and concentrated phase ( $\phi = 0.351$ ) after liquid–liquid phase separation. The dashed line is a tie-line and connects the pair of volume fractions after phase separation. The measurements are consistent in both cases with a critical volume fraction  $\phi_c = 0.21$ .

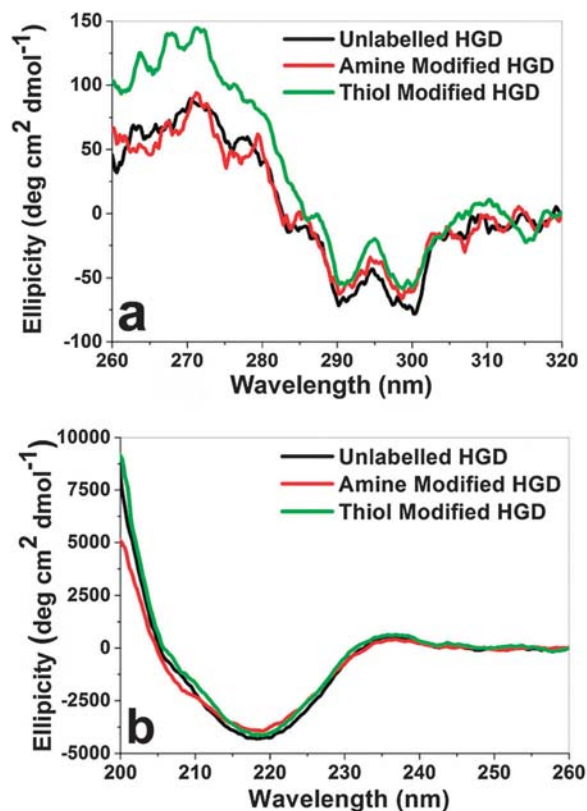


Fig. 6 Near (a) and far (b) UV spectra showing unlabelled HGD (black), amine modified HGD (green) and thiol modified HGD (red). While there is no change in the secondary structure of the protein for either modification, there is a slight perturbation in the near UV spectral data in the region corresponding to phenylalanine (255 nm to 270 nm) for the amine modified protein.<sup>65</sup> There is only 1 phenylalanine in close proximity to lysine in position 2 which is the phenylalanine in position 11. However, the CD spectrum indicates that this change in structure is very local and does not significantly impact the protein structure. There is no measured change in the tertiary structure for the thiol modified protein.

we used far and near circular dichroism (CD) spectroscopy for comparative analysis of secondary and tertiary structure between the unlabelled protein and the protein labelled at both positions. No significant change in protein structure is observed.

We have so far demonstrated that fluorescent labelling of HGD in position 2 *via* an amine modification and at position 110 *via* a thiol modification leads to an increase in the net attraction between proteins in solution. This increase is unrelated to the formation of protein aggregates, although the increase in net attraction does eventually lead to aggregation. When aggregation does occur, it involves a significant volume of the protein material (not only the labelled protein) and appears to arise due to the collective increase in net attraction (rather than by nucleation *via* the labelled protein). The increase in attraction arises from a small amount of labelled protein in solution and not as a result of a significant change to the protein structure. While the effect is not limited to a single type of small molecule label, the extent of the impact is related both to the specific chemistry of the dye and the position on the protein in which the modification is made. While it is not overly

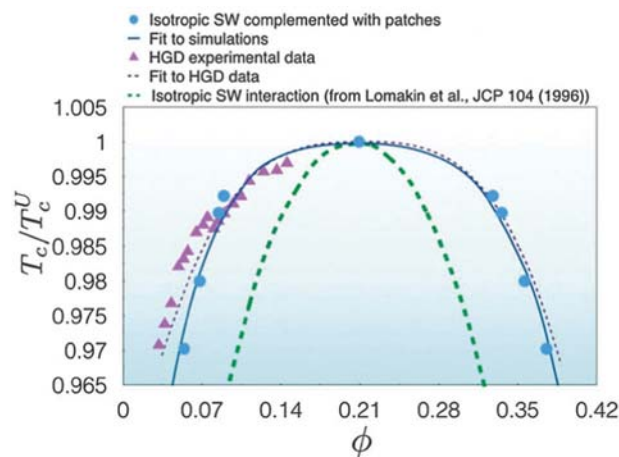


Fig. 7 Comparison between the experimental liquid–liquid coexistence curve of unlabelled HGD protein and different numerical models. Circles are results for the Kern–Frenkel patchy model with four patches randomly distributed on the surface and complemented with a SW attraction. The thick dashed line is the liquid–liquid coexistence curve of a simple SW model from ref. 61: the isotropic attraction captures the position of the critical volume fraction but doesn't quantitatively describe the amplitude of the liquid–liquid coexistence curve. Solid lines are fits to the critical scaling to describe the liquid–liquid coexistence curve.  $T_c$  is the critical temperature. For the chosen model we find  $T_c = 0.8184$ .

surprising that the addition of a fluorescent (hydrophobic) molecule has some impact on solution behaviour, the very low labelled protein fractions at which a significant effect occurs is somewhat unexpected.

To gain a microscopic understanding of the primary mechanism responsible for the shift in  $T_c$  for LLPS when a small fraction of protein is labelled, we numerically study a simple patchy model which incorporates the essential elements required to describe the experimentally observed behaviour. Patchy particle models with implicit solvent have been used to describe the phase behaviour of rubredoxin and lysozyme.<sup>47,61,64</sup> To show that the model used in this work (eqn (1) and (2) and parameters reported in Table 1) correctly captures the value of the critical volume fraction as well as the experimental width of the liquid–liquid coexistence curve of HGD we report a comparison between numerical and experimental data in Fig. 7.

To model unlabelled–labelled interactions we have performed a study in which we vary the parameters characterising the L-patch: attraction strength  $\epsilon_{L,U}$ , angular width  $(\theta_{\max})_L$  and width of the patch  $\delta_{L,U}$ , aiming to achieve an increase of  $T_c$  in agreement with experimental results. In Table 3 we summarise all the different combinations of parameters for the L-patch and the corresponding values of  $\phi_c$  and  $T_c$  with respect to the critical temperature of the unlabelled system ( $T_c/T_c^U$ ). The resulting fits to the Ising probability distribution for estimating the critical point are shown in Fig. 8 for unlabelled particles and mixtures with  $x_L \sim 0.02$ .

Most of the simulations have been carried out with a wide angular width compared to that of the U-patches, in order to allow for multiple bonding. The bonding volume, *i.e.* the volume available to form bonds with other particles depends

**Table 3** Parameters of the L-patch and corresponding critical temperatures for different fractions of labelled protein

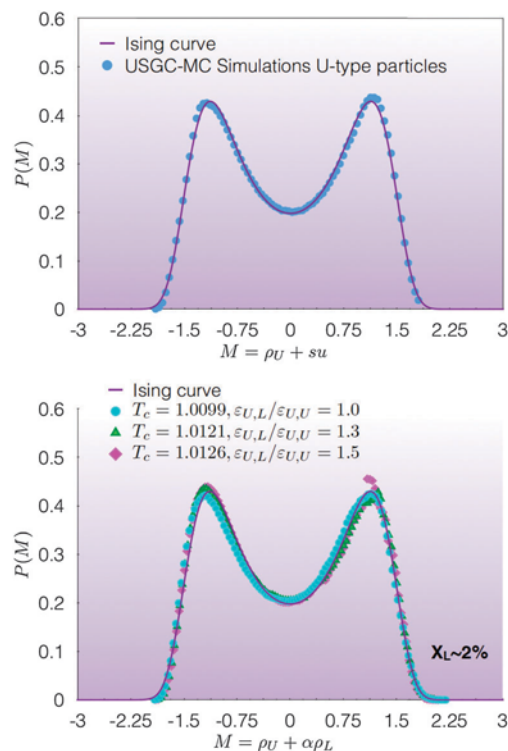
$\varepsilon_{U,L}/\varepsilon_{U,U}$	$(\theta_{\max})_L$	$\delta_{U,L}$	$x_L$ (%)	$T_c/T_c^U$	$\phi_c$
1.0	0.723	0.175	1.12	1.0056	0.214
	0.723	0.175	1.57	1.0081	0.215
	0.723	0.175	2.14	1.0099	0.214
	0.723	0.175	2.90	1.0130	0.214
1.1	0.723	0.175	1.13	1.0062	0.215
	0.723	0.175	1.48	1.0081	0.214
	0.723	0.175	2.16	1.0105	0.214
	0.723	0.175	3.10	1.0142	0.214
1.2	0.723	0.175	1.97	1.0105	0.217
	0.723	0.175	1.95	1.0105	0.217
	0.723	0.175	2.61	1.0105	0.213
	0.723	0.175	1.62	1.0105	0.217
	0.723	0.175	2.05	1.0111	0.217
	0.723	0.175	2.60	1.0142	0.216
	0.723	0.175	4.16	1.0203	0.217
	0.723	0.175	4.16	1.0203	0.217
1.5	0.723	0.175	1.66	1.0105	0.217
	0.723	0.175	2.23	1.0126	0.219
	0.723	0.175	2.51	1.0139	0.219

on the angular width and on the interaction width. Whenever the condition

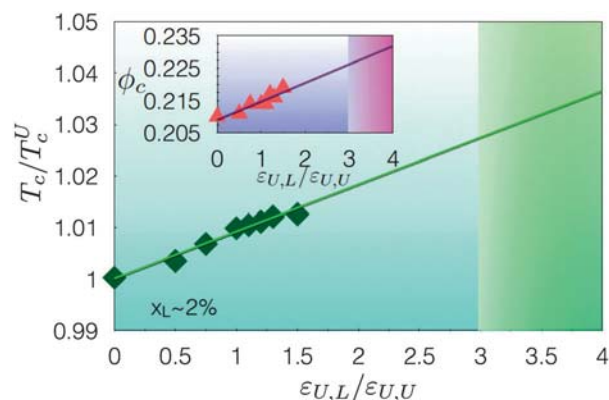
$$\sin(\theta_{\max})_L > \frac{1}{2(1 + \delta_{U,L})} \quad (5)$$

is satisfied, the patch can be involved in more than one bond. We note that, since the bonding angle is already very wide, the increase of the interaction width does not contribute to a significant increase of the critical temperature. On the other end, a further increase of the bonding angle would create overlaps with the other patches of the same particle. Finally, if the L-patch is too attractive, bonds with other proteins becomes irreversible and it is not possible to use grand-canonical simulations to estimate the critical point of the system.

In our simulations this occurs for bonding energy  $\varepsilon_{U,L}/\varepsilon_{U,U} > 1.5$  (when the bonding angle of the L-patch is set to  $(\theta_{\max})_L = 0.723$  and the width of the patch is  $\delta_{U,L} = 0.175$ ). Hence we cannot directly probe the range of bonding energies directly relevant to experiments, where the energetic contribution brought by the fluorescent molecule is estimated to be 3–4 times larger than the typical protein–protein attraction. As a consequence, to allow for a comparison with experiments we need to resort to extrapolation of the critical temperature to higher bonding energies. For instance, focusing on the case in which the composition of the labelled proteins is  $x_L \sim 0.02$ , the extrapolation values of the bonding energies comparable to the energetic contribution brought by the fluorescent molecule in the experimental system is shown in Fig. 9. The addition of a very small percentage of labelled proteins, as little as  $x_L = 0.02$  produces a significant increase in the critical temperature of the mixtures, in qualitative agreement with experimental results. We also see a shift of the critical volume fraction to larger values when compared to the unlabelled system. However a crude extrapolation suggests that this would shift to  $\phi_c = 0.23$  when  $\varepsilon_{U,L}/\varepsilon_{U,U} \sim 4$ .



**Fig. 8** Top Panel: Best fit of numerical data (full circles) to the Ising order parameter distribution  $P(M = \rho_U + s u)$  (solid line) of monodisperse system of unlabelled (U-type) particles. The mixing parameter is  $s = 0.028$ . The simulation box is  $L = 5$ . Bottom Panel: the same as before but for the binary mixtures with  $x_L \sim 0.02$ , when the interaction strength of the L-patch ( $\varepsilon_{U,L}/\varepsilon_{U,U}$ ) is varied. The order parameter is given by  $M = \rho_U + \alpha \rho_L$ . For all the three curves  $\alpha = 0.001$ . The simulation box is  $L = 5$ . The angular width of the L-patch is set to  $(\theta_{\max})_L = 0.723$  and the width  $\delta_{U,L} = 0.175$ .



**Fig. 9** Evolution of the critical temperature of the mixture of labelled and unlabelled particles (normalised to the critical temperature of the unlabelled system  $T_c^U$ ) as a function of the bond energy between a U-patch and a L-patch  $\varepsilon_{U,L}$ . The fraction of labelled particles is  $x_L \sim 0.02$ . The darker area in the plot corresponds to the range of interaction energies comparable to experiments. Inset: critical volume fraction  $\phi_c$  as a function of  $\varepsilon_{U,L}$ .

HGD in mixtures with the amine modified protein at  $x_L = 0.002$ , for which we have an experimental point on the descending arm of the liquid–liquid coexistence line, the experimental data is consistent with a similar critical volume



fraction ( $\phi_c = 0.21$ ) as for the unlabelled case, although an increase to higher values cannot be excluded since the fitting function (eqn (4)) strictly holds only for one-component systems. Finally, we have also monitored the relative concentration of labelled proteins in the two phases of the liquid–liquid phase separated mixtures, finding that they are present only in the dense phase. This is due to the larger strength of the fluorescent label attraction which forms longer-lived bonds with respect to unlabelled proteins making it favourable for them to be in the concentrated phase. This is consistent with the experimental results where a 1 : 2 ratio of labelled protein is found between the dilute and concentrated phases (from an initial  $x_L = 0.002$  of labelled protein) (Fig. 5).

## Conclusion

We have demonstrated that small molecule fluorescent labelling of protein increases the net attraction in protein solutions at low labelled protein fractions. The temperature at which liquid–liquid phase separation for HGD occurs increases when using fluorescent dyes with different chemistries and is also affected by positioning these dyes in different sites on the protein. By careful control of the labelling conditions, it was possible to ensure that each labelled protein contained only 1 dye molecule and hence the ratio of labelled to unlabelled protein could be unambiguously calculated. It appears that both the specific chemistry of the dye molecule and the position at which it is located on the surface of the protein are important in determining the extent to which the net attraction increases. To understand the microscopic basis for these observations, we have numerically assessed a simple patchy particle model in which unlabelled proteins are represented by a 4-patch particle complemented with an isotropic SW potential and labelled proteins by the addition of a fifth patch to this model. Performing Monte Carlo simulations for mixtures of these particles at labelled protein compositions within the range of the experimental values, we find a good agreement between the model and experiments. The net increase in attraction in the system is derived from an increase in the total number of bonds due to the addition of a large, hydrophobic region on the protein surface, which occurs upon labelling. We note that while a simple model can quantitatively describe the liquid–liquid phase coexistence of the unlabelled HGD solutions, it does not fully describe the chemical modification induced by the fluorescent labelling in terms of the interaction potential. Our attempt to model this by the presence of an additional patch only qualitatively captures the experimental behaviour. We have tested different potential models (*e.g.* with and without the presence of the isotropic SW) as well as different combinations of parameters of the fifth patch but we have not been able to reach quantitatively the same change in the critical temperatures. However, the fact that the critical volume fraction does not change significantly is well reproduced by the present model. These findings suggest that more fundamental work is needed to understand how to model correctly the fluorescent label in a quantitative way. There may of course be other surface effects with the addition of dye

that we have not yet considered. For example the fluorescent dye, when present alongside other hydrophobic amino acids on the surface, could increase the effective size of the additional attractive site on the protein surface. These possibilities will be explored in future work.

The current work however provides important observations. While anecdotally we know that fluorescent labelling can alter the solubility of the protein (and this has led to a number of new dyes being developed to offset these effects), we explain for the first time that the origin of the decrease in solubility is not due to the formation of aggregates in itself (although this is a downstream consequence) or a significant change in protein structure, but as a result of a net-increase in attraction due to an increase in the total number of attractive sites in the system. We have also demonstrated that these effects can be significant, even at sites on the protein which may be considered biologically inert. While the cysteine at position 110 for HGD is important in maintaining its solubility, the amine in position 2 has no known biological significance and there are no known mutations at this position for this protein which result in a change in protein solution behaviour.

With the increasing dependence on fluorescence methods to understand protein behaviour, this study provides for the first time a quantitative assessment of the effects that these established methodologies can have on the stability of protein solutions. Thus, while it is clear that the use of fluorescence by chemical modification requires careful control and may best be limited to the lowest possible amounts of labelled protein, it also suggests that the interpretation of results involving even small amounts of fluorescently labelled protein must account for these important effects. More broadly, these results may suggest a more general mechanism by which anisotropic protein–protein interactions in protein mixtures determine solution behaviour.

## Acknowledgements

JJMCM, MKQ and SJ gratefully acknowledge funding from Science Foundation Ireland grant 11/RFP.1/PHY/3165, Stokes Lectureship to JJMCM and also from the Irish Research Council. NG and EZ acknowledge support from MIUR Futuro in Ricerca ANISOFT project (RBFR125H0M). NG and EZ thank L. Rovigatti for useful discussions. Molecular graphics and analyses (Fig. 1) were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualisation, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

## References

- 1 E. Bianchi, R. Blaak and C. N. Likos, *Phys. Chem. Chem. Phys.*, 2011, **13**, 6397–6410.
- 2 E. Bianchi, J. Largo, P. Tartaglia, E. Zaccarelli and F. Sciortino, *Phys. Rev. Lett.*, 2006, **97**(16), 168301.
- 3 Q. Chen, S. C. Bae and S. Granick, *Nature*, 2011, **469**, 381–384.

- 4 S. C. Glotzer and M. J. Solomon, *Nat. Mater.*, 2007, **6**, 557–562.
- 5 G. Pellicane, G. Smith and L. Sarkisov, *Phys. Rev. Lett.*, 2008, **101**(24), 248102.
- 6 D. Fusco and P. Charbonneau, *Colloids Surf., B*, 2015, DOI: 10.1016/j.colsurfb.2015.07.023.
- 7 H. Rezvantlaba and S. Shojaei-Zadeh, *Phys. Chem. Chem. Phys.*, 2014, **16**, 8283.
- 8 R. P. Sear, *J. Chem. Phys.*, 1999, **111**(10), 4800–4806.
- 9 A. Lomakin, N. Asherie and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**(17), 9465–9468.
- 10 C. D. Spicer and B. G. Davis, *Nat. Commun.*, 2014, **5**, 4740.
- 11 O. Boutoureira and G. J. L. Bernardes, *Chem. Rev.*, 2015, **115**, 2174–2195.
- 12 H. Sahoo, *RSC Adv.*, 2012, **2**, 7017–7029.
- 13 F. M. Veronese, *Biomaterials*, 2001, **22**, 405–417.
- 14 G. Walsh and R. Jefferis, *Nat. Biotechnol.*, 2006, **24**(10), 1241–1252.
- 15 S. I. van Kasteren, H. B. Kramer, H. H. Jensen, S. J. Campbell, J. Kirkpatrick, N. J. Oldham, D. C. Anthony and B. G. Davis, *Nature*, 2007, **446**, 1105–1109.
- 16 F. H. Kasten, *Fluorescent and luminescent probes for biological activity*, W.T. Mason, London, 2nd edn, 1999, pp. 17–39.
- 17 B. Huang, M. Bates and X. Zhuang, *Annu. Rev. Biochem.*, 2009, **78**, 993–1016.
- 18 F. Hillger, D. Nettels, S. Dorsch and B. Schuler, *J. Fluoresc.*, 2007, **17**(6), 759–765.
- 19 M. Pusey, J. Barcena, M. Morris, A. Singhal, Q. Yuan and J. Ng, *Acta Crystallogr.*, 2015, **F71**, 806–814.
- 20 J. Zhang, R. E. Campbell, A. Y. Ting and R. Y. Tsien, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**(12), 906–918.
- 21 A. Hawe, M. Sutter and W. Jiskoot, *Pharm. Res.*, 2008, **25**(7), 1487–1499.
- 22 Q. Zheng, M. F. Juette, S. Jockusch, M. R. Wasserman, Z. Zhou, R. B. Altman and S. C. Blancharda, *Chem. Soc. Rev.*, 2014, **43**(4), 1044–1056.
- 23 R. Y. Tsien, *Annu. Rev. Biochem.*, 1998, **67**, 509–544.
- 24 U. Resch-Genger, M. Grabolle, S. Cavaliere-Jaricot, R. Nitschke and T. Nann, *Nat. Methods*, 2008, **5**(9), 763–775.
- 25 N. Asherie, *Methods*, 2004, **34**(3), 266–272.
- 26 S. James, M. K. Quinn and J. J. McManus, *Phys. Chem. Chem. Phys.*, 2015, **17**(7), 5413–5420.
- 27 W. Li, B. A. Persson, M. Morin, M. A. Behrens, M. Lund and M. Zackrisson Oskolkova, *J. Phys. Chem. B*, 2015, **119**(2), 503–508.
- 28 F. Roosen-Runge, F. Zhang, F. Schreiber and R. Roth, *Sci. Rep.*, 2014, **4**, 7016.
- 29 A. Gajraj and R. Y. Ofoli, *Langmuir*, 2000, **16**(21), 8085–8094.
- 30 D. P. Richards, C. Stathakis, R. Polakowski, H. Ahmadzadeh and N. J. Dovichi, *J. Chromatogr. A*, 1999, **853**(1–2), 21–25.
- 31 L. S. De Clerck, C. H. Bridts, A. M. Mertens, M. M. Moens and W. J. Stevens, *J. Immunol. Methods*, 1994, **172**(1), 115–124.
- 32 J. Gunzenhäuser, R. Wyss and S. Manley, *PLoS One*, 2014, **9**(12), e115095.
- 33 N. Asherie, A. Lomakin and G. B. Benedek, *Phys. Rev. Lett.*, 1996, **77**(23), 4832–4835.
- 34 J. Möller, S. Grobelny, J. Schulze, S. Bieder, A. Steffen, M. Erilkamp, M. Paulus, M. Tolan and R. Winter, *Phys. Rev. Lett.*, 2014, **112**(2), 028101.
- 35 S. James and J. J. McManus, *J. Phys. Chem. B*, 2012, **116**(34), 10182–10188.
- 36 F. Bonneté, S. Finet and A. Tardieu, *J. Cryst. Growth*, 1999, **96**(2–4), 403–414.
- 37 J. J. McManus, A. Lomakin, O. Ogun, A. Pande, M. Basan, J. Pande and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**(43), 16856–16861.
- 38 M. L. Broide, T. M. Tominc and M. D. Saxowsky, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1996, **53**(6), 6325–6335.
- 39 Y. Wang, A. Lomakin, J. J. McManus, O. Ogun and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**(30), 13282–13287.
- 40 P. R. Banerjee, A. Pande, J. Patrosz, G. M. Thurston and J. Pande, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(2), 574–579.
- 41 U. P. Andley, *Prog. Retinal Eye Res.*, 2007, **26**(1), 78–98.
- 42 J. F. Hejtmancik, *Semin. Cell Dev. Biol.*, 2008, **19**(2), 134–149.
- 43 E. Héon, M. Priston, D. F. Schorderet, G. D. Billingsley, P. O. Girard, N. Lubsen and F. L. Munier, *Am. J. Hum. Genet.*, 1999, **65**(5), 1261–1267.
- 44 P. Das, J. A. King and R. Zhou, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(26), 10514–10519.
- 45 A. Pande, J. Pande, N. Asherie, A. Lomakin, O. Ogun, J. A. King, N. H. Lubsen, D. Walton and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**(5), 1993–1998.
- 46 M. L. Broide, C. R. Berland, J. Pande, O. O. Ogun and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**(13), 5660–5664.
- 47 H. Liu, S. K. Kumar and F. Sciortino, *J. Chem. Phys.*, 2007, **127**, 084902.
- 48 N. Kern and D. Frenkel, *J. Chem. Phys.*, 2003, **118**, 9882.
- 49 P. Virnau and M. Mueller, *J. Chem. Phys.*, 2004, **120**, 10925.
- 50 A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.*, 1988, **61**, 2635.
- 51 N. B. Wilding, *J. Phys.: Condens. Matter*, 1997, **9**, 585–612.
- 52 F. Romano, P. Tartaglia and F. Sciortino, *J. Phys.: Condens. Matter*, 2007, **19**, 322101.
- 53 D. Nicolaides and A. D. Bruce, *J. Phys. A: Math. Gen.*, 1988, **21**, 233–234.
- 54 M. M. Tsy-pin and H. W. J. Blöte, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2000, **62**, 73–76.
- 55 L. Rovigatti, D. de las Heras, J. M. Tavares, M. M. Telo da Gama and F. Sciortino, *J. Chem. Phys.*, 2013, **138**(16), 164904.
- 56 C. Liu, N. Asherie, A. Lomakin, J. Pande, O. Ogun and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**(1), 377–382.
- 57 A. Lomakin, N. Asherie and G. B. Benedek, *J. Chem. Phys.*, 1996, **104**(4), 1646–1656.
- 58 Y. Wang, A. Lomakin, J. J. McManus, O. Ogun and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**(30), 13282–13287.

- 59 Q. Chen, P. G. Vekilov, R. L. Nagel and R. E. Hirsch, *Biophys. J.*, 2004, **86**(3), 1702–1712.
- 60 J. Pande, A. Lomakin, B. Fine, O. Ogun, I. Sokolinski and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 1995, **92**(4), 1067–1071.
- 61 N. Asherie, J. Pande, A. Lomakin, O. Ogun, S. R. Hanson, J. B. Smith and G. B. Benedek, *Biophys. Chem.*, 1998, **75**(3), 213–227.
- 62 J. A. Thomson, P. Schurtenberger, G. M. Thurston and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**(20), 7079–7083.
- 63 S. M. Kelly, T. J. Jess and N. C. Price, *Biochim. Biophys. Acta*, 2005, **1751**(2), 119–139.
- 64 D. Fusco, J. J. Headd, A. De Simone, J. Wang and P. Charbonneau, *Soft Matter*, 2014, **10**(2), 290–302.