# Chemical Modification Alters Protein−Protein Interactions and Can Lead to Lower Protein Solubility

Michelle K. Quinn, Susan James, and Jennifer J. McManus*

Department of Chemistry, Maynooth University, Maynooth, Co. Kildare, Ireland

ABSTRACT: The chemical modification of proteins is at the frontier of developments in biological imaging and biopharmaceutics. With the advent of more sensitive and higher resolution imaging techniques, researchers increasingly rely on the functionalization of proteins to enable these techniques to capture cellular processes. For biopharmaceutical therapies, chemically modified proteins, for example, antibody-drug conjugates (ADCs) offer the possibility of more tailored treatments for the disease with lower toxicities than traditional small molecule therapies. However, relatively little consideration is paid to how chemical modifications impact protein−protein interactions and solution stability. Using human γD-crystallin as a model, we demonstrate that chemical modification of the protein surface alters protein−protein interactions, which can result in lower solubility depending on the chemical nature of the modifier and the position on the protein where the modification is made. Understanding these effects is essential to ensure that modifying proteins effectively occurs with minimum self-association and that studies carried out using labeled proteins accurately reflect those of unmodified proteins.
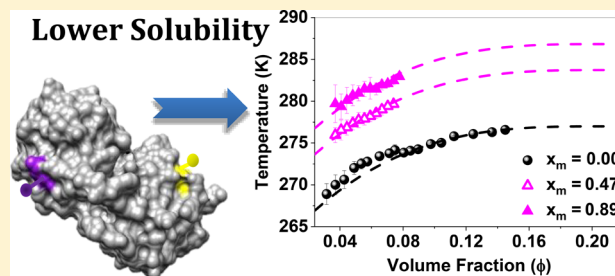
INTRODUCTION

Of the 22 biologically active proteinogenic amino acids, 15 are subject to diversification via posttranslational modification (PTM; there are no known modifications of Leu, Ile, Val, Ala, Phe, Sec, and Pyl).[1] In almost all cases, naturally occurring PTMs are crucial for the function and structure of proteins.[2−4]

The artificial chemical modification of a protein surface (which will be referred to here simply as "chemical modification") was first reported in 1904 and has, since then, been a mainstay in the study of proteins.[5] Covalent posttranslational modifications (PTMs) are chemical modifications of proteins that occur either in vivo or in vitro after their biosynthesis and serve to change the structure or function of the protein by means of covalent attachment of a functional group or protein (such as ubiquitination), proteolytic (or less commonly, autocatalytic) cleavage, or degradation.[1,6] In addition to ubiquitination, common naturally occurring PTMs include acetylation, acylation, ADP-ribosylation, lipidation, amidation, γ-carboxylation, β-hydroxylation, disulfide bond formation, methylation, glycosylation, phosphorylation, proteolytic processing, and sulfation.[2,7,8] Given the diversity of structure and function imposed on proteins as a result of naturally occurring PTMs, the ability to chemically modify proteins by artificial means is highly attractive.[6] The chemical modification of a protein surface is an important mechanism for altering protein function and is commonly employed for protein tracking using fluorophores in biological systems and is in widespread use in the biopharmaceutical industry to alter or enhance the function of protein biotherapeutics using either natural or synthetic covalent PTMs (e.g., via glycosylation and PEGylation).[7,9,10] PEGylation of a protein has also been successfully achieved via noncovalent modification of a fucose-binding lectin with a fucose-capped PEG.[11] More recently, chemical modification has been employed in the development of antibody-drug conjugates (ADCs) for the targeted treatment of malignant tumor cells.[12−15] These chemical modifications have led to the development of a diverse range of proteins with functionalities beyond their initial capabilities. As a result, understanding, quantifying, and predicting how modifications made to the surface of a protein affect its behavior and stability are required.

Lysine (Lys, K) and cysteine (Cys, C) are the amino acids most often used for chemical modification[13,16] owing to their inherent nucleophilic character.[6,9,17] Lysine is one of the most abundant amino acids in proteins, and there are many established protocols to preferentially chemically modify the amine of its side chain, which is typically the most nucleophilic amine on a protein surface.[16−19] Preferential modification of Lys requires the use of "harder" nucleophiles such as activated esters, sulfonyl chlorides, and isothiocyanates or modification can be achieved through reductive alkylation using aldehydes in the presence of cyanoborohydride.[9] The abundance of Lys can lead to heterogeneous mixtures of the modified protein, which may be useful in some applications, where multiple conjugates are required but undesirable in others such as the development of some biopharmaceutics where a more homogenous mixture is preferred.[9] Cysteine is regarded as the most important naturally occurring amino acid for chemical modification because in addition to its highly nucleophilic nature (it is much more nucleophilic than Lys),

it has a low abundance (<2%), which allows for chemical modification with a greater degree of specificity.[9,16,19,20] There are a number of electrophiles available for the specific labeling of Cys residues such as α-halocarbonyls, iodoacetamides, and maleimides, with maleimides prevailing as the most commercialized and widely used derivative for Cys modification due to ease of use and straightforward synthesis.[9] Indeed, a selenocysteine residue, which is more nucleophilic than Cys, can be inserted into a protein for highly specific modification also using maleimide derivatives.[21]

Other amino acids can be chemically modified but can be more synthetically challenging to achieve.[16] Tyrosine (Tyr) and tryptophan (Trp) are rarely found on the surface of a protein (due to their hydrophobic nature), which has made them the subject of focus as sites for chemical modification because single-site modification is advantageous[6,22] The modification of Tyr and Trp is generally performed using water-soluble transition metal catalysis.[23] Trp has the lowest natural abundance of all the canonical amino acids, which can allow for Trp modification specificity provided the protein being modified has no surface-bound Tyr residues. Chemical modification of histidine (His), arginine (Arg), glutamic acid (Glu), and aspartic acid (Asp) is also possible as each has a nucleophilic residue that can be used as a site for modification but their use is much rarer than modifications of Lys, Cys, Tyr, and Trp.[17] His is a target for nucleophilic attack and is highly selective toward vinyl sulfone, which allows for the site-specific labeling with molecules containing a vinyl sulfone group as an acetylating agent.[24] While Arg has a similar abundance to Lys, it has a lower tendency to be found on the surface of the protein, which makes it more favorable for site-specific modification.[25] Specific modification of Arg involves the formation of a pyrimidine ring using two out of the three nitrogen atoms found in its side chain.[26] Although it is highly selective, modification of Arg requires a 14 day reaction time, which may make it an unsuitable choice of site for some studies that require specificity and a short reaction time. Modification of the carboxylate groups in Glu and Asp involves amidation using ethylenediamine.[27] However, ethylenediamine indiscriminately labels both Glu and Asp, which make these sites unfavorable, as sites for modification. Methionine (Met) is the only other sulfur-containing proteinogenic amino acid. Though it is similar to Cys in terms of its natural abundance, it does not share the same popularity as a site of modification due to its weak nucleophilicity.[28] However, redox-based reagents are emerging as a strategy for chemoselective Met modification.[28] An extensive review of synthetic routes for selective proteinogenic amino acid modification has been recently published.[29] In addition to modification of proteinogenic amino acids, there has been recent interest in the insertion of noncanonical amino acids for use as sites for modification to enhance site selectivity.[30,31]

The surface of a protein is inherently anisotropic as a result of the varying amino acid side chains on the protein surface.[32] The interactions of amino acid side chains determine the specific biological, chemical, and technological functions of each protein.[32,33] Proteins in solution can undergo a number of different phase transitions to form condensed phases and states: reversible or irreversible solid–liquid phase separation leading to the formation of amorphous or ordered solids or less frequently, liquid–liquid phase separation (LLPS), also known as coacervation, whereby two metastable liquid phases are formed, one protein rich and one protein poor.[34] LLPS is also observed in some colloidal systems and is characteristic of short-range (less than a quarter of the diameter of the particle) attractive interactions.[34,35] Understanding the fundamental mechanisms, which determine protein phase behavior, is a key requirement in the biopharmaceutical industry.[36] Protein phase diagrams are used to predict the physical state of a protein for a given set of conditions across a range of volume fractions (e.g., temperature, pressure, salt concentration, and pH)[35,37,38] and are a good measure of the strength of protein interactions.[37] Changes in the protein environment, such as varying ionic strength or pH,[33] the introduction of another protein type,[39,40] the introduction of a mutation,[41,42] or a chemical modification of the protein surface[43] can change its phase diagram. Because protein solution stability is intrinsically linked to the strength of its interaction potential, phase diagrams are a useful way of understanding, measuring, and predicting the effect modifications to the surface of a protein have on its behavior in solution.

The lens crystallins are a family of proteins that undergo LLPS at physiological pH. These highly soluble proteins are essential for vision and are found in the nucleus and cortex of the eye lens and remain soluble, in many cases over a lifetime to maintain lens transparency.[44] Human γD-crystallin (HGD) is soluble at concentrations in excess of 400 mg mL^{-1}.[45,46] HGD, which has a near-neutral pI (7.1), exhibits short-ranged attractive interactions and therefore can undergo LLPS at physiological pH. Cataract is the leading cause of blindness worldwide and results in a loss of transparency due to increased light scattering within the eye lens.[44,47] Congenital (or childhood) cataracts arising from single amino acid mutations in the sequence of HGD have been studied extensively.[41,48−56] These point mutations lead to a decrease in lens protein solubility and lens transparency as a result of increased light scattering and can lead to blindness.

Surprisingly, given the prevalence of Lys, which can account for as much as 10% of the amino acids in a typical protein,[18] HGD contains only a single Lys residue located at position 2. Despite it having six unpaired Cys residues, only one, at position 110, is surface-exposed.[41,57] The surface exposure of a single Lys and Cys residues in HGD coupled with its stability at high concentrations (>400 mg mL^{-1}) and its ability to undergo LLPS make HGD an ideal model protein to assess how chemical modification of the protein surface impacts on its solution behavior. Because we have explicit control over the number of labeling sites on the protein surface, any changes in the temperature at which LLPS occurs are indicative of changes in the net attractive protein–protein interactions on the protein surface due to the introduction of further chemical anisotropy to the protein surface by covalent attachment of a specific molecule. We have already demonstrated using a combination of experiments and coarse grained simulations[43] that fluorescently labeling HGD using small molecule fluorophores increases the temperature at which LLPS occurs (T_{ph}) due to an increase in net attraction in the system. Small molecule fluorophores are significantly larger and inherently more hydrophobic in nature (given their aromatic structures) than any naturally occurring amino acid, and we hypothezised based on the experimental data that the decrease in protein stability was due to the addition of a large attractive patch on the protein surface due to the hydrophobicity of the fluorophore. To test this hypothesis, we have labeled the protein at the same amino acid residues with a hydrophilic tag, biotin. If the change in solution behavior observed in the

presence of a fluorophore was due to its hydrophobicity, then modification of the same amino acids with a hydrophilic molecule should restore (or eliminate) any decrease in solubility observed. Hence, based on the nature of the molecule used to adapt a protein and the position it is located, chemical modification can be used to modulate protein−protein interactions where this is required (for example in protein crystallization) or decrease the impact of a necessary but problematic modification (e.g., by increasing hydrophilicity or changing its position on the protein) to improve solubility.

■ EXPERIMENTAL SECTION

Preparation of Materials. Analytical grade glacial acetic acid, sodium hydroxide (NaOH), sodium chloride, sodium dihydrogen orthophosphate dihydrate, dibasic sodium phosphate heptahydrate, hydrochloric acid (HCl), sodium azide, LB Agar, LB Broth, isopropyl β-D-1-thiogalactopyranoside (IPTG), magnesium sulfate heptahydrate, ethylenediaminetetraacetic acid (EDTA), and dithiothreitol (DTT) were purchased from Fisher Scientific (UK). Lysozyme and Tris HCl were purchased from Merck Millipore (Ireland). Protease inhibitor cocktail tablets were purchased from Roche Diagnostics (USA). Buffers were prepared using ultrapure (Milli-Q) water, and pH values were adjusted using concentrated NaOH or HCl. Buffers were filtered through a 0.45 μm MillexHV syringe filter or nylon membrane filter (Merck Millipore, Ireland) before use. Amicon Ultra 4 mL centrifugal filters were used for buffer exchange (Merck Millipore, Ireland).

Expression, Purification, and Characterization of HGD. HGD was prepared by recombinant methods as previously described.[41] HGD protein purity was confirmed at >98% by SDS-PAGE and size exclusion HPLC. The molecular weight of HGD was confirmed to be 20,608 ± 1 Da by intact molecular weight analysis using electrospray ionization mass spectrometry, carried out in "FingerPrints" Proteomics Facility at the University of Dundee, Scotland, UK, and is in agreement with previously published data.[41] Protein concentration in mg mL^{-1} was determined spectroscopically using a mass extinction coefficient equal to 2.09 mg^{-1} mL cm^{-1} for HGD.[39] The volume fraction, ϕ, was calculated using the expression $\phi = c\nu$, where c is the protein concentration in mg mL^{-1}, and ν is the partial specific value in mL mg^{-1} equal to 7.1 ± 0.1 × 10^{-4} for the γ crystallins.[58]

Chemical Modification of Lys-2 and Cys-110 Amino Acid Residues. Amine modification of Lys-2 in HGD was carried out after purification using EZ-Link NHS-PEG$_4$-Biotin (Pierce Biotechnology, USA) as per supplied instructions. Thiol modification of Cys-110 was carried out after purification using EZ-Link Maleimide-PEG$_2$-Biotin (Pierce Biotechnology, USA). Interprotein disulfide bonds were reduced using 2-mercaptoethylamine-HCl (Pierce Biotechnology, USA) as per supplied instructions before modification. Biotinylated proteins were quantified using a Pierce Biotin Quantitation Kit (Thermo Scientific, USA) as per supplied instructions. Fluorescence measurements were performed using a Molecular Devices M2e spectrofluorometer, and spectral data were replotted and analyzed using OriginPro 2016 data analysis software.

Circular Dichroism Spectroscopy. Circular dichroism (CD) spectra were recorded at near- and far-UV wavelengths for samples of unmodified HGD, amine modified HGD (x_m =

0.86), and thiol modified HGD (x_m = 0.88) at a concentration of 0.1 mg mL^{-1} in 0.1 M sodium phosphate buffer at pH 7. Measurements were performed using a JASCO J-810 spectropolarimeter (Institute of Molecular, Cell, and Systems Biology, College of Medical, Veterinary, and Life Sciences, University of Glasgow, Scotland, UK). Spectral data were plotted using OriginPro 2016 data analysis software.

Liquid−Liquid Phase Separation Measurements. Liquid−liquid coexistence curves were measured using a method previously outlined[59] for unmodified HGD, mixtures of unmodified HGD and amine modified HGD, and mixtures of unmodified HGD and thiol modified HGD in 0.1 M sodium phosphate buffer at pH 7. A Perkin Elmer Lambda 35 UV−vis spectrophotometer and UV Win Lab − Timedrive Lambda 35 associated software were used to measure the percentage transmission of light at λ = 600 nm through a Hellma quartz cuvette with a pathlength of 1 cm. A Thermo Scientific K10 water bath attached to a Thermo Scientific D10 temperature control was used for temperature regulation. The system temperature was monitored using an Omega HH509R thermocouple. OriginPro 2016 was used to construct phase diagrams for each mixture. LLPS was induced for a mixture of unmodified and thiol modified HGD (x_m = 0.89) in a Thermo Scientific Heraeus Multifuge 3SR+ temperature controlled centrifuge using a method described previously to determine modified protein partitioning.[60] The volume fraction and fraction of biotinylated protein for each phase were quantified as before using a Biotin Quantitation Kit.

■ RESULTS AND DISCUSSION

Two forms of chemically modified HGD were used to investigate the phase behavior of mixtures of modified and unmodified protein: HGD chemically modified with PEGylated biotin at the Lys-2 position and at the Cys-110 position. The sites used for chemical modification (Lys-2 and Cys-110) are shown in Figure 1.
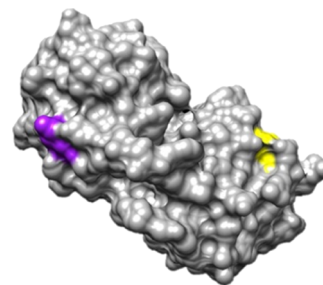

Figure 1. Topological rendering of HGD showing the locations of modified sites, Lys-2 (purple) and Cys-110 (yellow).

The liquid−liquid phase separation (LLPS) temperature (T_{ph}) for each modified protein in mixtures with native protein was used to quantify the change in net interaction potential between proteins for mixtures of modified and unmodified proteins.

Chemical modification of Lys-2 and Cys-110 using PEGylated biotin. PEGylated biotin was used to chemically modify sites 2 and 110 of HGD. Biotin is a vitamin (also known as vitamin H, vitamin B7, or coenzyme R) with a valeric acid side chain that allows derivatization of this moiety for enhanced functionality. It is a relatively small molecule with a molecular weight of 244.3 g mol^{-1} compared to the molecular

weight of HGD, which is approximately 21 kDa. While biotin itself is water-soluble, its derivatization with a PEG linker is intended to enhance the hydrophilicity of the biotin molecule by increasing the solvation shell around the molecule.

Separate chemical modifications of HGD were carried out to create two forms of modified protein: modification of HGD at the Lys-2 position and modification of a free thiol at the Cys-110 position. T_{ph} was measured up to a volume fraction of ~0.14 for mixtures of Lys-2 modified HGD across two modified protein compositions. Separately, T_{ph} was also measured up to a volume fraction of ~0.07 for two different Cys-110 modified protein compositions. The liquid−liquid coexistence curves for these mixtures are shown in Figure 2.
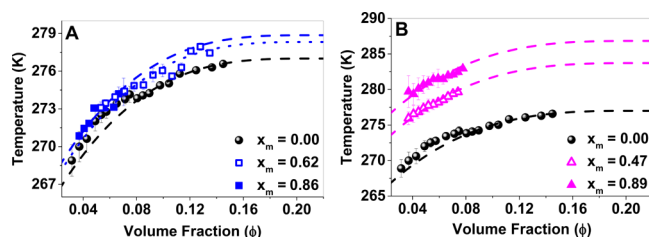

Figure 2. Liquid−liquid coexistence curves for unmodified HGD (black) and (A) protein mixtures in aqueous solution of unmodified and amine modified (blue) and (B) thiol modified (pink) HGD with increasing modified protein compositions (x_m). Panel A indicates liquid−liquid coexistence curves for protein mixtures in aqueous solution of unmodified HGD and HGD amine modified with PEG_4-Biotin. Panel B shows liquid−liquid coexistence curves for unmodified HGD and binary protein mixtures in aqueous solution of unmodified HGD and HGD thiol modified using PEG_2-Biotin.

The lower concentration limit for the cysteine-modified protein was necessary because modification at this site reduced solubility and precipitation of protein begins to occur at higher volume fractions.

The LLPS data for unmodified HGD and each of the mixtures with modified protein were fitted using eq 1[58]

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

where ϕ and ϕ_c are the volume fraction and critical volume fraction, respectively (where the partial specific volume, ν, of HGD and its modified forms is taken to be $7.1 \pm 0.1 \times 10^{-4}$ mL mg^{-1}), A is a parameter related to the width of the coexistence curve, T and T_c are the phase separation temperature and critical temperature, respectively, and β is an exponent term for the three-dimensional Ising model equal to 0.325. The critical volume fraction, ϕ_c, was taken to equal 0.21, a value consistent with ϕ_c for all gamma crystallins.[61] The estimated values for T_c, ΔT_c, and A are given in Table 1 for each composition of chemically modified protein.

The change in T_{ph} after modifying the thiol at site 110 is significantly greater than modifying the primary amine at the site 2. The site-specific differences on the effect of a modification is consistent with our previous work, which also indicated that HGD chemically modified at Lys-2 had less impact than labeling at Cys-110.[43] When fluorescently labeled protein was present in mixtures with unlabeled HGD at protein compositions between 0.0001 and 0.01 (partial data is shown in Figure 3), an increase in T_{ph} with an increase in the proportion of modified protein was observed. The proportion

Table 1. Values for Estimated Changes in Critical Temperature to the Nearest Kelvin, K (ΔT_c) with Respect to Unmodified HGD Using eq 1 for Each Binary Protein Mixture in Aqueous Solution

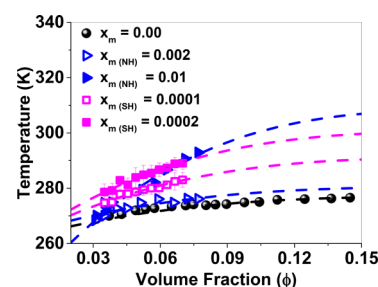| modified protein	T_c (K)	ΔT_c (K)	A
unmodified HGD ($x_m = 0.00$)	277		2.6
HGD amine modified with PEG_4-Biotin ($x_m = 0.62$)	278	1	2.6
HGD amine modified with PEG_4-Biotin ($x_m = 0.86$)	279	2	2.6
HGD thiol modified with PEG_2-Biotin ($x_m = 0.47$)	283	6	2.6
HGD thiol modified with PEG_2-Biotin ($x_m = 0.89$)	287	10	2.6

Figure 3. Liquid−liquid coexistence curves for HGD (black) and HGD chemically modified via fluorescent labeling of a primary amine (blue) and thiol (pink).

of labeled protein that was required to increase T_{ph} was substantially lower for protein modified at the Cys position (by a factor of ~50).

While there is very little change in T_{ph} for HGD modified with PEGylated biotin at the amine position (i.e., ΔT_c is within the maximum experimental error in measurements of T_{ph}), there is a substantial increase in T_{ph} with increasing proportions of HGD modified with PEGylated biotin at the thiol position. The range of the protein−protein interactions is related to the width of the liquid−liquid coexistence curve where a decrease in the coexistence curve width is indicative of an increase in the interaction range of the protein.[61−63] However, there is no change in the width of the coexistence curve for HGD modified with PEGylated biotin (parameter A in eq 1, see Table 1). This suggests that the increased net attraction as indicated by the increase in T_{ph} is not due to an increase in the range of the attraction, unlike the observations for proteins modified with a fluorescent tag in our earlier work.[43] Figure 3 shows partial coexistence curves for HGD in mixtures with fluorescently labeled HGD modified at the amine position (NH) and thiol position (SH) where an increase in T_{ph} was accompanied by a decrease in curve width. For biotinylated protein, we need to add a significantly higher fraction of modified protein required to induce a change in T_{ph} of similar magnitude ($x_m = 0.0001$ for fluorescently labeled HGD vs $x_m = 0.89$ for HGD chemically modified with PEGylated biotin). Hence, there are notable differences between this and the case of fluorescently labeled proteins.

Electrostatic interprotein interactions between charged groups on the protein surface contribute to its solubility.[64] The pI of HGD (pI = 7.1) is almost at physiological pH and is very close to the working pH used in the data presented here. Lysine (along with histidine and arginine) is one of three positively charged amino acid residues at neutral pH (pH 7). The modification of Lys-2 by covalent attachment of the

PEGylated biotin thus removes one positive charge from the protein surface and results in a protein with a net negative charge. This change in charge may be a contributing factor as to why the increase in net attraction (as indicated by the lower change in T_{ph}) is not as substantial as the increase in T_{ph} for HGD modified at Cys-110 with PEGylated biotin. Because the structures of the small molecule fluorescent dyes previously used for the specific labeling of HGD are proprietary, it is not possible to comment on how this affected the electrostatic contribution of the net interaction potential. Fluorescein isothiocyanate, FITC, which has a partial negative charge at neutral pH, can modify both primary and secondary amines. As a result, HGD labeled with FITC will carry a net negative charge, which should contribute to its solution stability. A mixture of HGD and HGD fluorescently labeled with FITC ($x_L = 0.0002$) led to an increase in a T_{ph} of 20 K. This suggests that the hydrophobicity rather than electrostatics contributed more to the net interaction potential in that case.

Chemically Modified HGD Partitioning. A mixture of unmodified and modified HGD biotinylated at the Cys-110 position ($x_m = 0.89$) was cooled to below its T_{ph} to induce LLPS and centrifuged to macroscopically separate the dilute and dense liquid phases (Figure 4). The modified protein
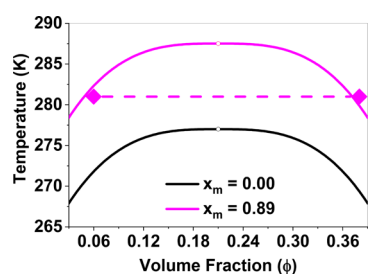

Figure 4. Liquid−liquid coexistence curves for unmodified HGD (black) and a mixture of unmodified HGD and HGD modified with biotin ($x_m = 0.89$). The diamonds indicate the volume fractions for the dilute phase ($\phi = 0.06$) and concentrated phase ($\phi = 0.38$) after liquid−liquid phase separation. The dashed line is a tie line and connects the pair of volume fractions after phase separation. The composition of biotinylated protein was 0.80 in the dilute phase and 0.87 in the dense phase.

composition was determined for each phase following the complete separation of the dense and dilute phases and found to be 0.80 in the dilute phase and 0.87 in the dense phase. The preferential partitioning of the more attractive protein (i.e., the modified HGD) into the dense protein phase is consistent with previously published data using fluorescently labeled HGD where the majority of the chemically modified protein partitioned into the dense protein phase.[39,43]

Circular Dichroism. A structural or conformational change as a result of unfolding can cause an increase in the net attraction of a protein due to the exposure of hydrophobic residues. To determine if this was the case for both the amine and thiol modified proteins, near-and far-ultraviolet (UV) circular dichroism (CD) spectroscopies were used to assess the impact of chemically modifying HGD with PEGylated biotin on its respective tertiary and secondary structures (Figure 5).

The spectral data indicate that there was no significant change in the tertiary and secondary structures of the modified protein relative to the unmodified protein thus that a structural or conformational change is not responsible for the net increase in attraction of the modified protein, indicated by the
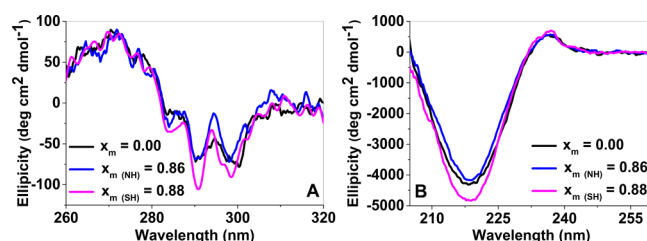

Figure 5. (A) Near-UV and (B) far-UV circular dichroism spectra showing unmodified HGD (black), amine modified HGD (blue), and thiol modified HGD (pink).

increase in T_{ph}. This data is consistent with our previously published CD data for HGD chemically modified with small molecule fluorescent labels.[43]

■ CONCLUSIONS

HGD was chemically modified at two different amino acid residues (Lys-2 and Cys-110) using PEGylated biotin. These modified proteins were then used in mixtures with unmodified HGD to assess what impact, if any, the presence of the chemically modified protein had on solution behavior relative to solutions containing unmodified protein alone. Both mixtures containing a chemically modified protein (i.e. Lys modified HGD and Cys modified HGD) showed an increase in the net attraction of the system (as indicated by an increase in T_{ph}) relative to the unmodified protein, albeit to different extents; mixtures containing HGD amine modified at Lys-2 with modified protein compositions of $x_m = 0.62$ and 0.86 both showed a slight increase in T_{ph} ($\leq \sim 2$ K), while mixtures containing HGD thiol modified at Cys-110 with modified protein compositions of $x_m = 0.47$ and 0.89 showed a more significant increase ($\leq \sim 10$ K). These results are qualitatively consistent with our previously published data using HGD fluorescently labeled with a small molecule fluorescent label, which showed that the protein modified at the Cys-110 position resulted in a much greater increase in T_{ph} than the protein modified at Lys-2. In this study, the width of the liquid−liquid coexistence curve measured for each protein mixture containing chemically modified and unmodified protein was unchanged. This is in contrast with our previous study using small molecule fluorescent dyes, where a decrease in curve width was observed, which indicates an increase in the range of the attraction. Complete phase separation of a sample containing thiol modified HGD indicated that, as expected, the majority of the chemically modified protein (that had a higher net attraction than the unmodified protein) preferentially partitioned into the dense protein phase. This observation is consistent with previous studies using HGD and other proteins with a higher net attraction. Near- and far-UV CD data indicated that the increase in T_{ph} of solutions containing the chemically modified proteins is not as a result of a significant structural change in either the secondary or tertiary structure of the modified protein. These results are also consistent with our previously published data using small molecule fluorescent dyes.

Our results show that the effect of modifying the surface of a protein on its solution behavior is not only highly dependent on the physicochemical properties of the modifier used but also on the site of the modification. This study has direct industrial applicability given the widespread use of surface

modifications, especially the use of Lys and Cys modifications, in biotherapeutic innovation and development.

AUTHOR INFORMATION

Corresponding Author
*E-mail: jennifer.mcmanus@mu.ie.

ORCID
Jennifer J. McManus: 0000-0003-3630-5108

Notes
The authors declare no competing financial interest.

REFERENCES

(1) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J., Jr. Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications. *Angew. Chem.* **2005**, *44*, 7342−7372.

(2) Whitford, D. *Proteins: Structure and Function*; West Sussex, John Wiley & Sons Ltd.: 2005.

(3) Gunnoo, S. B.; Madder, A. Chemical Protein Modification Through Cysteine. *ChemBioChem* **2016**, *17*, 529−553.

(4) Li, Q.; Shortreed, M. R.; Wenger, C. D.; Frey, B. L.; Schaffer, L. V.; Scalf, M.; Smith, L. M. Global Post-Translational Modification Discovery. *J. Proteome Res.* **2017**, *16*, 1383−1390.

(5) Stephanopoulos, N.; Francis, M. B. Choosing an Effective Protein Bioconjugation Strategy. *Nat. Chem. Biol.* **2011**, *7*, 876−884.

(6) Baslé, E.; Joubert, N.; Pucheault, M. Protein Chemical Modification on Endogenous Amino Acids. *Chem. Biol.* **2010**, *17*, 213−227.

(7) Walsh, G.; Jefferis, R. Post-translational Modifications in the Context of Therapeutic Proteins. *Nat. Biotechnol.* **2006**, *24*, 1241−1252.

(8) Jiang, H.; Zhang, X.; Chen, X.; Aramsangtienchai, P.; Tong, Z.; Lin, H. Protein Lipidation: Occurrence, Mechanisms, Biological Functions, and Enabling Technologies. *Chem. Rev.* **2018**, *118*, 919−988.

(9) Spicer, C. D.; Davis, B. G. Selective Chemical Protein Modification. *Nat. Commun.* **2014**, *5*, 4740.

(10) Rosen, C. B.; Francis, M. B. Targeting the N Terminus for Site-Selective Protein Modification. *Nat. Chem. Biol.* **2017**, *13*, 697−705.

(11) Antonik, P. M.; Eissa, A. M.; Round, A. R.; Cameron, N. R.; Crowley, P. B. Noncovalent PEGylation via Lectin-Glycopolymer Interactions. *Biomacromolecules* **2016**, *17*, 2719−2725.

(12) Agarwal, P.; Bertozzi, C. R. Site-Specific Antibody−Drug Conjugates: The Nexus of Bioorthogonal Chemistry, Protein Engineering, and Drug Development. *Bioconjugate Chem.* **2015**, *26*, 176−192.

(13) Polakis, P. Antibody Drug Conjugates for Cancer Therapy. *Pharmacol. Rev.* **2016**, *68*, 3−19.

(14) Akkapeddi, P.; Azizi, S.-A.; Freedy, A. M.; Cal, P. M. S. D.; Gois, P. M. P.; Bernardes, G. J. L. Construction of Homogeneous Antibody-Drug Conjugates Using Site-Selective Protein Chemistry. *Chem. Sci.* **2016**, *7*, 2954−2963.

(15) Nanna, A. R.; Li, X.; Walseng, E.; Pedzisa, L.; Goydel, R. S.; Hymel, D.; Burke, T. R., Jr.; Roush, W. R.; Rader, C. Harnessing a Catalytic Lysine Residue for the One-Step Preparation of Homogeneous Antibody-Drug Conjugates. *Nat. Commun.* **2017**, *8*, 1112.

(16) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angew. Chem., Int. Ed.* **2009**, *48*, 6974−6998.

(17) Boutureira, O.; Bernardes, G. J. L. Advances in Chemical Protein Modification. *Chem. Rev.* **2015**, *115*, 2174−2195.

(18) Turecek, P. L.; Bossard, M. J.; Schoetens, F.; Ivens, I. A. PEGylation of Biopharmaceuticals: A Review of Chemistry and Nonclinical Safety Information of Approved Drugs. *J. Pharm. Sci.* **2016**, *105*, 460−475.

(19) Matos, M. J.; Oliveira, B. L.; Martínez-Sáez, N.; Guerreiro, A.; Cal, P. M. S. D.; Bertoldo, J.; Maneiro, M.; Perkins, E.; Howard, J.; Deery, M. J.; et al. Chemo- and Regioselective Lysine Modification on Native Proteins. *J. Am. Chem. Soc.* **2018**, *140*, 4004−4017.

(20) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. Chemical Modification of Proteins at Cysteine: Opportunities in Chemistry and Biology. *Chem. − Asian J.* **2009**, *4*, 630−640.

(21) Hofer, T.; Thomas, J. D.; Burke, T. R., Jr.; Rader, C. An Engineered Selenocysteine Defines a Unique Class of Antibody Derivatives. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 12451−12456.

(22) Joshi, N. S.; Whitaker, L. R.; Francis, M. B. A Three-Component Mannich-Type Reaction for Selective Tyrosine Bioconjugation. *J. Am. Chem. Soc.* **2004**, *126*, 15942−15943.

(23) Antos, J. M.; Francis, M. B. Transition Metal Catalyzed Methods for Site-Selective Protein Modification. *Curr. Opin. Chem. Biol.* **2006**, *10*, 253−262.

(24) del Castillo, T.; Marales-Sanfrutos, J.; Santoyo-González, F.; Magez, S.; Lopez-Jaramillo, F. J.; Garcia-Salcedo, J. A. Monovinyl Sulfone β-Cyclodextrin. A Flexible Drug Carrier System. *ChemMedChem* **2014**, *9*, 383−389.

(25) Gauthier, M. A.; Klok, H.-A. Peptide/Protein−Polymer Conjugates: Synthetic Strategies and Design Concepts. *Chem. Commun.* **2008**, 2591−2611.

(26) Oya, T.; Hattori, N.; Mizuno, Y.; Miyata, S.; Maeda, S.; Osawa, T.; Uchida, K. Methylglyoxal Modification of Protein. Chemical and Immunochemical Characterization of Methylglyoxal-Arginine Adducts. *J. Biol. Chem.* **1999**, *274*, 18492−18502.

(27) Díaz-Rodríguez, A.; Davis, B. G. Chemical Modification in the Creation of Novel Biocatalysts. *Curr. Opin. Chem. Biol.* **2011**, *15*, 211−219.

(28) Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Toste, F. D.; et al. Redox-Based Reagents for Chemoselective Methionine Bioconjugation. *Science* **2017**, *355*, 597−602.

(29) deGruyter, J. N.; Malins, L. R.; Baran, P. S. Residue-Specific Peptide Modification: A Chemist's Guide. *Biochemistry* **2017**, *56*, 3863−3873.

(30) Lang, K.; Chin, J. W. Cellular Incorporation of Unnatural Amino Acids and Bioorthogonal Labeling of Proteins. *Chem. Rev.* **2014**, *114*, 4764−4806.

(31) Freedy, A. M.; Matos, M. J.; Boutureira, O.; Corzana, F.; Guerreiro, A.; Akkapeddi, P.; Somovilla, V. J.; Rodrigues, T.; Nicholls, K.; Xie, B.; et al. Chemoselective Installation of Amine Bonds on Proteins through Aza-Michael Ligation. *J. Am. Chem. Soc.* **2017**, *139*, 18365−18375.

(32) Lomakin, A.; Asherie, N.; Benedek, G. B. Aeolotopic Interactions of Globular Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9465−9468.

(33) Broide, M. L.; Tominc, T. M.; Saxowsky, M. D. Using Phase Transitions to Investigate the Effect of Salts on Protein Interactions. *Phys. Rev. E* **1996**, *53*, 6325−6335.

(34) Asherie, N.; Lomakin, A.; Benedek, G. B. Phase Diagram of Colloidal Solutions. *Phys. Rev. Lett.* **1996**, *77*, 4832−4835.

(35) Möller, J.; Grobelny, S.; Schulze, J.; Bieder, S.; Steffen, A.; Erlkamp, M.; Paulus, M.; Tolan, M.; Winter, R. Reentrant Liquid-Liquid Phase Separation in Protein Solutions at Elevated Hydrostatic Pressures. *Phys. Rev. Lett.* **2014**, *112*, No. 028101.

(36) Amin, S.; Barnett, G. V.; Pathak, J. A.; Roberts, C. J.; Sarangapani, P. S. Protein Aggregation, Particle Formation, Character-

ization & Rheology. *Curr. Opin. Colloid Interface Sci.* **2014**, *19*, 438–449.

(37) Taratuta, V. G.; Holschbach, A.; Thurston, G. M.; Blankschtein, D.; Benedek, G. B. Liquid-Liquid Phase Separation of Aqueous Lysozyme Solutions: Effects of pH and Salt Identity. *J. Phys. Chem.* **1990**, *94*, 2140–2144.

(38) Muschol, M.; Rosenberger, F. Liquid–Liquid Phase Separation in Supersaturated Lysozyme Solutions and Associated Precipitate Formation/Crystallization. *J. Chem. Phys.* **1997**, *107*, 1953–1962.

(39) Wang, Y.; Lomakin, A.; McManus, J. J.; Ogun, O.; Benedek, G. B. Phase Behavior of Mixtures of Human Lens Proteins Gamma D and Beta B1. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 13282–13287.

(40) Liu, C.; Lomakin, A.; Thurston, G. M.; Hayden, D.; Pande, A.; Pande, J.; Ogun, O.; Asherie, N.; Benedek, G. B. Phase Separation in Multicomponent Aqueous-Protein Solutions. *J. Phys. Chem.* **1995**, *99*, 454–461.

(41) Pande, A.; Pande, J.; Asherie, N.; Lomakin, A.; Ogun, O.; King, J. A.; Lubsen, N. H.; Walton, D.; Benedek, G. B. Molecular Basis of a Progressive Juvenile-Onset Hereditary Cataract. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 1993–1998.

(42) James, S.; Quinn, M. K.; McManus, J. J. The Self Assembly of Proteins; Probing Patchy Protein Interactions. *Phys. Chem. Chem. Phys.* **2015**, *17*, 5413–5420.

(43) Quinn, M. K.; Gnan, N.; James, S.; Ninarello, A.; Sciortino, F.; Zaccarelli, E.; McManus, J. J. How Fluorescent Labelling Alters the Solution Behaviour of Proteins. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31177–31187.

(44) Sharma, K. K.; Santhoshkumar, P. Lens Aging: Effects of Crystallins. *Biochim. Biophys. Acta, Gen. Subj.* **2009**, *1790*, 1095–1108.

(45) Ji, F.; Jung, J.; Koharudin, L. M. I.; Gronenborn, A. M. The Human W42R γD-Crystallin Mutant Structure Provides a Link between Congenital and Age-related Cataracts. *J. Biol. Chem.* **2013**, *288*, 99–109.

(46) Vendra, V. P. R.; Khan, I.; Chandani, S.; Muniyandi, A.; Balasubramanian, D. Gamma Crystallins of the Human Eye Lens. *Biochim. Biophys. Acta, Gen. Subj.* **2016**, *1860*, 333–343.

(47) Bloemendal, H.; de Jong, W.; Jaenicke, R.; Lubsen, N. H.; Slingsby, C.; Tardieu, A. Ageing and Vision: Structure, Stability and Function of Lens Crystallins. *Prog. Biophys. Mol. Biol.* **2004**, *86*, 407–485.

(48) Kmoch, S.; Brynda, J.; Asfaw, B.; Bezouška, K.; Novák, P.; Řezáčová, P.; Ondrová, L.; Filipec, M.; Sedláček, J.; Elleder, M. Link Between a Novel Human γD-crystallin Allele and a Unique Cataract Phenotype Explained by Protein Crystallography. *Hum. Mol. Genet.* **2000**, *9*, 1779–1786.

(49) Pande, A.; Pande, J.; Asherie, N.; Lomakin, A.; Ogun, O.; King, J.; Benedek, G. B. Crystal Cataracts: Human Genetic Cataract Caused by Protein Crystallization. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 6116–6120.

(50) Basak, A.; Bateman, O.; Slingsby, C.; Pande, A.; Asherie, N.; Ogun, O.; Benedek, G. B.; Pande, J. High-Resolution X-ray Crystal Structures of Human γD Crystallin (1.25 Å) and the R58H Mutant (1.15 Å) Associated with Aculeiform Cataract. *J. Mol. Biol.* **2003**, *328*, 1137–1147.

(51) Héon, E.; Priston, M.; Schorderet, D. F.; Billingsley, G. D.; Girard, P. O.; Lubsen, N.; Munier, F. L. The γ-Crystallins and Human Cataracts: A Puzzle Made Clearer. *Am. J. Hum. Genet.* **1999**, *65*, 1261–1267.

(52) Stephan, D. A.; Gillanders, E.; Vanderveen, D.; Freas-Lutz, D.; Wistow, G.; Baxenvanis, A. D.; Robbins, C. M.; Van Auken, A.; Quesenberry, M. I.; Bailey-Wilson, J.; et al. Progressive Juvenile-Onset Punctate Cataracts Caused by Mutation of the γD-Crystallin Gene. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 1008–1012.

(53) Plotnikova, O. V.; Kondrashov, F. A.; Vlasov, P. K.; Grigorenko, A. P.; Ginter, E. K.; Rogaev, E. I. Conversion and Compensatory Evolution of the γ-Crystallin Genes and Identification of a Cataractogenic Mutation That Reverses the Sequence of the Human CRYGD Gene to an Ancestral State. *Am. J. Hum. Genet.* **2007**, *81*, 32–43.

(54) Santhiya, S. T.; Manohar, M. S.; Rawlley, D.; Vijayalakshmi, P.; Namperumalsamy, P.; Gopinath, P. M.; Löster, J.; Graw, J. Novel Mutations in the γ-Crystallin Genes Cause Autosomal Dominant Congenital Cataracts. *J. Med. Genet.* **2002**, *39*, 352–358.

(55) Nandrot, E.; Slingsby, C.; Basak, A.; Cherif-Chefchaouni, M.; Benazzouz, B.; Hajaji, Y.; Boutayeb, S.; Gribouval, O.; Arbogast, L.; Berraho, A.; et al. Gamma-D Crystallin Gene (CRYGD) Mutation Causes Autosomal Dominant Congenital Cerulean Cataracts. *J. Med. Genet.* **2003**, *40*, 262–267.

(56) Banerjee, P. R.; Pande, A.; Patrosz, J.; Thurston, G. M.; Pande, J. Cataract-Associated Mutant E107A of Human γD-Crystallin shows Increased Attraction to α-crystallin and Enhanced Light Scattering. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 574–579.

(57) Srikanthan, D.; Bateman, O. A.; Purkiss, A. G.; Slingsby, C. Sulfur in Human Crystallins. *Exp. Eye Res.* **2004**, *79*, 823–831.

(58) Broide, M. L.; Berland, C. R.; Pande, J.; Ogun, O. O.; Benedek, G. B. Binary-Liquid Phase Separation of Lens Protein Solutions. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 5660–5664.

(59) Asherie, N. Protein Crystallization and Phase Diagrams. *Methods* **2004**, *34*, 266–272.

(60) Thomson, J. A.; Schurtenberger, P.; Thurston, G. M.; Benedek, G. B. Binary Liquid Phase Separation and Critical Phenomena in a Protein/Water Solution. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 7079–7083.

(61) Liu, C.; Asherie, N.; Lomakin, A.; Pande, J.; Ogun, O.; Benedek, G. B. Phase Separation in Aqueous Solutions of Lens Gamma-Crystallins: Special Role of Gamma S. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 377–382.

(62) Chen, Q.; Vekilov, P. G.; Nagel, R. L.; Hirsch, R. E. Liquid-Liquid Phase Separation in Hemoglobins: Distinct Aggregation Mechanisms of the β6 Mutants. *Biophys. J.* **2004**, *86*, 1702–1712.

(63) Lomakin, A.; Asherie, N.; Benedek, G. B. Monte Carlo Study of Phase Separation in Aqueous Protein Solutions. *J. Chem. Phys.* **1996**, *104*, 1646–1656.

(64) Cooper, A. *Biophysical Chemistry*; 2 Ed.; RSC Publishing: Cambridge UK, 2011.