Biophysical Journal
Article

Biophysical Society

Temperature-Dependent Interactions Explain Normal and Inverted Solubility in a γD-Crystallin Mutant

Amir R. Khan,[1] Susan James,[2] Michelle K. Quinn,[2] Irem Altan,[3] Patrick Charbonneau,[3] and Jennifer J. McManus[2,*]

[1]School of Biochemistry and Immunology, Trinity College Dublin, Dublin, Ireland; [2]Department of Chemistry, Maynooth University, Maynooth, Ireland; and [3]Department of Chemistry, Duke University, Durham, North Carolina

ABSTRACT Protein crystal production is a major bottleneck in the structural characterization of proteins. To advance beyond large-scale screening, rational strategies for protein crystallization are crucial. Understanding how chemical anisotropy (or patchiness) of the protein surface, due to the variety of amino-acid side chains in contact with solvent, contributes to protein-protein contact formation in the crystal lattice is a major obstacle to predicting and optimizing crystallization. The relative scarcity of sophisticated theoretical models that include sufficient detail to link collective behavior, captured in protein phase diagrams, and molecular-level details, determined from high-resolution structural information, is a further barrier. Here, we present two crystal structures for the P23T + R36S mutant of γD-crystallin, each with opposite solubility behavior: one melts when heated, the other when cooled. When combined with the protein phase diagram and a tailored patchy particle model, we show that a single temperature-dependent interaction is sufficient to stabilize the inverted solubility crystal. This contact, at the P23T substitution site, relates to a genetic cataract and reveals at a molecular level the origin of the lowered and retrograde solubility of the protein. Our results show that the approach employed here may present a productive strategy for the rationalization of protein crystallization.

SIGNIFICANCE Understanding, controlling, and modeling complex protein-protein interactions is key to directing protein assembly. Using a soft matter physics approach and combining protein phase diagrams, two new, to our knowledge, high-resolution protein crystal structures, and a custom patchy particle model, we present a significant advance. The protein forms two crystals, one that melts when heated and another that melts when cooled. Our work identifies the microscopic origin of a human cataract and explains solubility inversion for proteins. In short, this work produces a major advance in our understanding of the nature and impact of the complex anisotropic (or patchy) interactions between proteins.

INTRODUCTION

The rationalization of protein crystallization remains a major obstacle to efficient structure determination—a requirement to understand the molecular basis for many diseases and to pinpoint targets for new drug development (1). Sampling hundreds (or sometimes thousands) of solution conditions (i.e., mixtures of different buffers, salts, and precipitants) is often the most productive strategy to identify lead conditions for protein crystallization. Even when coupled with rational design strategies such as surface entropy reduction (2), this approach can be time-consuming and costly because screening methods often fail to produce crystalline material or diffraction-quality crystals. Protein phase diagrams that map how a given protein behaves across sets of solution conditions dramatically improve the success of the process and narrow the screening required for producing diffraction-quality crystals but have only been measured for a small number of proteins ((3) and references therein). These reference studies have identified key challenges in guiding and improving protein crystallization.

An excellent such reference is human γD-crystallin (HGD), a major structural protein found in the eye lens. HGD is unusually stable in the eye lens in mixtures with α- and β-crystallins, often over a whole lifetime (4). Its

Submitted April 30, 2019, and accepted for publication July 8, 2019.

*Correspondence: jennifer.mcmanus@mu.ie

Amir R. Khan and Susan James contributed equally to this work.

Editor: Jill Trewhella.

https://doi.org/10.1016/j.bpj.2019.07.019

phase behavior is otherwise generally similar to that of a large group of important globular proteins that includes hemoglobin (5), immunoglobulins (6), lysozyme (7), and thaumatin (8). These phase diagrams are defined by net attractive short-range interactions that result in liquid-liquid phase separation and crystallization. Although native HGD itself does not readily form crystals, several of its genetic-cataract-related single-amino-acid substitutions do so easily, without any major structural changes (9–14). The P23T substitution—which is a naturally occurring mutation associated with congenital cataracts—however, has unusual phase behavior in that its aggregates have inverted solubility, i.e., they melt as temperature is decreased (15,16). As a result, the protein is insoluble at physiological concentration and temperature, leading to eye-lens opacity. In the related P23V mutant, both aggregates and crystals are observed, both also with inverted solubility (15,16). Yet, crystallization of the P23T mutant under physiological conditions has remained elusive. Numerous structural and biophysical studies, including x-ray structures at pH 4.5 and NMR solution studies, have failed to unambiguously identify major structural changes in the P23T mutant, and hence a full explanation for its anomalous inverted solubility is still unknown (17–20).

Physicochemical insights into protein phase behavior—both normal and anomalous—are often gleaned from colloidal science. Simple colloidal models do capture key features of protein phase diagrams, such as their metastable critical point (21). However, protein phase diagrams cannot be completely rationalized without including some level of anisotropy, in terms of the directional contacts between proteins in solution or within a crystal lattice (22–27) or of shape anisotropy (28). This anisotropy gives rise to rich protein phase diagrams and is more widely exploited for the controlled assembly of biological and biomimetic materials (29). It has even been proposed that these types of interactions are important in controlling liquid-liquid phase separation in cells (30), with important implications in understanding stress responses, RNA processing, and gene expression. However, understanding and predicting anisotropic protein-protein interactions ab initio is not yet possible because of the extreme heterogeneity of amino-acid side chains on the protein surface. Although measurements indicative of net protein-protein interactions such as the osmotic second virial coefficient, B_{22}, or the diffusivity constant, k_D, can provide some insight, they reflect the averaged pair interactions between proteins. These parameters are typically insufficient to trace back the specific, directional protein-protein interactions that control the dramatic (and often unpredictable) changes in protein assembly upon mutagenesis (14,16). Enhanced numerical models that capture the details of anisotropic protein-protein interactions may allow for the prediction of protein phase diagrams and hence optimal crystallization conditions (17,21–28,31–34). To identify the microscopic origin of inverted solubility, however, we need high-resolution structural information detailing the underlying anisotropic interactions, using, for instance, crystal structures of the protein of interest.

To design a P23T mutant that crystallizes at pH 7, we focused our interest on HGD mutant structures that do not form specific protein-protein contacts near proline 23. One such mutant, R36S, readily crystallizes by forming a crystal lattice contact at position 36. By combining the R36S and P23T substitutions, we reasoned that crystals of the double mutant would display inverted solubility based on a comparison of the phase diagrams for the single-mutant proteins, thus providing insights into the mechanism for the P23T mutant retrograde solubility. Remarkably, the double mutant P23T + R36S formed two distinct crystals forms—one with normal solubility and one with inverted solubility (35). Although inverted solubility in proteins has been previously observed, a protein that forms two distinct crystal lattices, each with opposite temperature dependence of the solubility line, had not, and therefore, this double mutant offers a rare opportunity to access the microscopic origin of solubility inversion, which we now probe further.

Here, we report the x-ray structures of the two crystal forms of the P23T + R36S mutant of HGD. We find that the two are polymorphs with different unit cells and crystal contacts and that it is possible to interchange between them solely by varying the solution temperature. In the inverted solubility crystal, a lattice contact involving the cataract-associated Thr23 residue is formed. This is a new contact with the same binding energy determined from a statistical mechanics analysis of the chemical potentials of the solubility lines in earlier work (16). We have used both the phase diagram for P23T + R36S and crystal structures to design a custom patchy particle model that incorporates specific contacts formed in the crystal lattice. We find that when temperature-dependent patchy interactions are included, the temperature dependence of the solubility lines for both crystal lattices can be reproduced by simulations performed using the custom model. Specifically, we show that a change to the contact that contains the 23rd residue in the inverted solubility crystal is sufficient to cause inverted solubility. This contact becomes engaged as temperature increases, stabilizing the inverted solubility crystal phase and thus revealing the molecular origin of the inverted solubility for P23T.

MATERIALS AND METHODS

Preparation and characterization of double mutant

The double mutant was created, expressed, and purified as described previously (35). Sodium dodecyl sulfate-polyacrylamide gel electrophoresis and size-exclusion high performance liquid chromatography were used to confirm protein purity at >98%. The intact molecular weight for the mutant protein was analyzed by electrospray ionization mass spectrometry (Finger

Prints Proteomics Facility, College of Life Sciences, University of Dundee, Dundee, UK), which confirmed a molecular mass of 20,541 ± 1 Da for the P23T + R36S mutant.

Crystallization and data collection

The crystals from the P23TR36S double mutant of HGD protein were obtained and grown in capillaries in 100 mM sodium phosphate buffer (pH 7) in the presence of 20 mM dithiothreitol. The solution concentration of protein was in the range of 1–2 mg/mL, and there was no additional precipitant in the solution. Crystals of the double mutant with inverted solubility (DBI) were grown at 310 K, whereas crystals with normal solubility (DBN) grew at 277 K. Both crystal types formed within a few hours of incubation at the relevant temperature. Crystals were harvested from capillaries and mixed with 25% glycerol, flash-cooled in liquid nitrogen, and subjected to x-ray diffraction. Data sets from two crystals, one grown at 310 K and the second grown at 277 K, were collected at the PX2 beamline at Le Soleil Synchrotron (Saint-Aubin, France) on an ADSC Q315 detector (Area Detector Systems Corporation, Poway, CA).

Solubility measurements

Protein solutions were prepared initially by diafiltration against 100 mM sodium phosphate buffer (pH 7.0) using Ultracel 10 KDa ultrafiltration disks (Merck Millipore, Tullagreen, Ireland). Protein concentrations for the double mutant was measured by ultraviolet absorbance using the extinction coefficient value of 2.09 mg^{-1} mL cm^{-1} after filtration through 0.22 μm Millex-GV Millipore (Merck Millipore) syringe-driven filters. When required, protein solutions were further concentrated by ultrafiltration using Amicon Ultra-4 centrifugal filter units (Merck Millipore) and the protein concentration reestablished by ultraviolet absorbance.

Data processing and structure solution

The structure of DBN was solved using the model that contains the R36S single-site mutation in HGD (Protein Data Bank (PDB): 2G98) (36). The program Phaser (37) provided the starting model, which was improved through cycles of manual model-building using Coot (38) and Phenix refinement (39). The structure of DBI was solved using the high-resolution 1.25 Å structure of wild-type HGD (PDB: 1HK0 (13)). The obtained structure was further refined using the same refinement procedure as for DBN. Statistics from the data collection and refinement strategies are detailed in Table 1. Crystal contacts determined from the structural analysis were used to determine the patch-patch interactions for the phase diagram, as described below.

TABLE 1 Data Collection and Refinement Statistics

	DBI	DBN
Resolution range (Å)	44.01–1.197 (1.24–1.197)	48.16–2.20 (2.277–2.20)
Wavelength (Å)	0.9786	0.9786
Space group	P 1 21 1	P 21 21 21
Unit cell (Å)	44.02, 31.70, 52.50, 90, 91.29, 90	54.04, 82.10, 106.25, 90, 90, 90
Total reflections	89,029 (8455)	164,792
Unique reflections	45,768 (4504)	24,656 (2391)
Multiplicity	1.9 (1.9)	6.7
Completeness (%)	99.46 (99.03)	99.63 (98.03)
Mean I/σ(I)	8.93 (1.22)	9.2 (1.4)
Wilson B-factor ($Å^2$)	10.73	38.69
R-merge	0.03956 (0.5534)	0.129
R-meas	0.05595 (0.7827)	0.139
R-pim	0.03956 (0.5534)	0.053
Mn(I) half-set CC(1/2)	0.998 (0.39)	0.997 (0.71)
Reflections used for R-free	1999 (190)	1233 (119)
R-work	0.1493 (0.2662)	0.2304 (0.2993)
R-free	0.1794 (0.2835)	0.2658 (0.3614)
Number of nonhydrogen atoms	1687	2888
Macromolecules	1479	2776
Solvent	208	112
Protein residues	173	341
RMS bonds (Å)	0.013	0.025
RMS angles (°)	1.21	1.49
Ramachandran favored (%)	98.83	94.93
Ramachandran allowed (%)	1.17	4.18
Ramachandran outliers (%)	0.00	0.9
Rotamer outliers (%)	1.85	0.69
Average B-factor	14.98	56.9
Macromolecules	13.30	57.26
Solvent	26.91	47.89
Chain A (DBN): 1–81	–	40.16
Chain A (DBN): 82–173	–	40.81
Chain B (DBN): 1–81	–	51.48
Chain B (DBN): 82–173	–	101.38[a]

Statistics for the highest-resolution shell are shown in parentheses.
[a]B-factors for the C-terminal domain of molecule B reveal significant domain flexibility.

Description of the model

Because transitions between the two crystal forms occur upon temperature change, we consider the phase behavior of the double mutant using a patchy particle model with temperature-dependent patches. This choice accounts for the associated change in bonding free energy (16).

Model definition

Proteins are modeled as patchy particles with interactions adapted from the Kern-Frenkel model (40). Hard spheres with a diameter σ, chosen as the largest center of mass distance between protein-protein crystal contacts, interact with directional, attractive patches of range $\lambda_{\alpha\beta}\sigma$ for each protein-protein contact $\alpha\beta$. The patch interaction potential,

$$u(r_{ij}, \Omega_i, \Omega_j) = u_{\mathrm{HS}} + \sum_{\alpha,\beta}^{n} u_{\alpha\beta}(r_{ij}, \Omega_i, \Omega_j),$$

thus includes a factorized attractive contribution, $u_{\alpha\beta} = v_{\alpha\beta}(r_{ij})f_{\alpha\beta}(\Omega_i, \Omega_j)$, that depends on interparticle distance, r_{ij}, and particle orientations, Ω_i and Ω_j. Its orientational component is

$$f_{\alpha\beta} = \begin{cases} 1, & \theta_{\alpha,ij} \le \delta_\alpha \text{ and } \theta_{\beta,ij} \le \delta_\beta \\ 0, & \text{otherwise} \end{cases} \times \begin{cases} 1, & \psi_{\alpha\beta,ij} \in [\phi_{\alpha\beta} - \Delta\phi_{\alpha\beta}, \phi_{\alpha\beta} + \Delta\phi_{\alpha\beta}] \\ 0, & \text{otherwise} \end{cases},$$

where the first term ensures that patch vectors face each other (Fig. S1). The second restricts the torsion between the two particles (Fig. S2). Its radial component is a square-well potential

$$v_{\alpha\beta}(r_{ij}) = \begin{cases} -\varepsilon_{\alpha\beta}(T), & \sigma < r_{ij} < \lambda_{\alpha\beta}\sigma \\ 0, & \text{otherwise} \end{cases},$$

where $\varepsilon_{\alpha\beta}$ is constant if the patch is not temperature-dependent and otherwise has a modulated interaction $\varepsilon_{\alpha\beta}(T) = (\bar{\varepsilon}_{\alpha\beta}/2)(1 + \tanh(T - T_a/\tau_{\alpha\beta}))$ (41), which becomes deactivated below temperature T_a over a rate set by $\tau_{\alpha\beta}$, thus capturing the change in free energy upon increasing the temperature. Model parameters were determined from all-atom molecular dynamics simulations of all patches of DBI and DBN (see Supporting Materials and Methods for methodological details and model parameters).

Phase diagram determination

The phase diagram of the schematic model was obtained by specialized Monte Carlo simulations: 1) the reference crystal free energies were obtained by integrating from an ideal Einstein crystal using the Frenkel-Ladd method (42); 2) the chemical potentials of the crystal phases as a function of temperature were obtained by thermodynamic integration along isobars from the reference in 1; and 3) fluid free energies were approximated using the second virial coefficient, B_{22}, because of the inefficiency of traditional Monte Carlo sampling at low densities (Supporting Materials and Methods). Coexistence points between the fluid and crystal phases were determined from the intersection of chemical potential curves, and coexistence lines were then traced out using a Gibbs-Duhem integration scheme (43,44).

RESULTS AND DISCUSSION

The equilibrium phase diagram for P23T + R36S is shown in Fig. 1 (35). Two different crystal types are observed, distinguished by the temperature dependence of their respective solubility lines; one with normal solubility (DBN), which melts as temperature increases, and a second with inverted solubility (DBI), which forms at higher temperatures and melts as temperature is lowered. The solubility lines intersect at ~303 K, where both crystal forms coexist. Remarkably, the two crystals form under

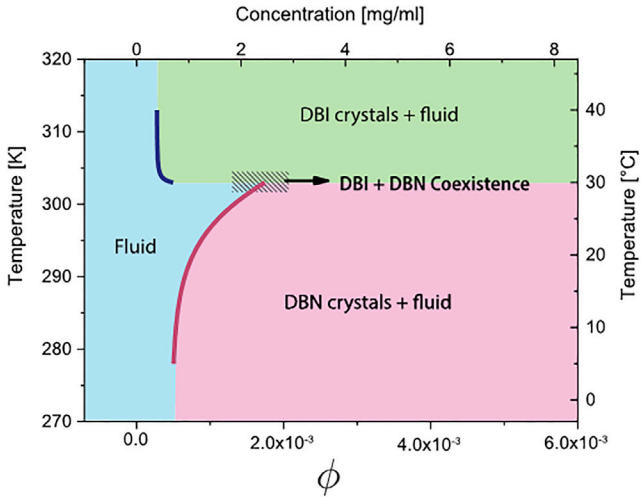

FIGURE 1 Experimental phase diagram for P23T + R36S mutant of HGD, indicating the equilibrium phase boundaries for the two crystals formed and their respective fluid phases (*solid lines*: data taken from 35). The volume fraction (ϕ) is calculated as $\phi = c \times v_{sp}$, where c is the concentration of protein in mg/mL and v_{sp} is the partial specific volume = 7.1 × 10^{-4} mg/mL (16). Coexistence of the two crystals is observed at the temperature at which the phase boundaries overlap (~303 K).

near-physiological conditions of temperature, pH, and salt, unlike the previously determined structure of P23T (pH 4.6, PEG4K) (20).

The crystals have different morphologies: DBN crystals are rod-shaped, and DBI crystals are rhombic. The proteins remain in their fully folded globular state across the temperatures probed in this work. We further showed in previous work that this mutant protein displays no significant change to its secondary structure relative to native HGD (35). Our determination of the structures for these two crystal forms by x-ray crystallography confirms this finding for these polymorphs.

The DBI (PDB: 6ETC) and DBN (PDB: 6ETA) structures consist of paired homologous domains that each adopt a Greek key motif. DBI crystallized as a monomer at high resolution (1.2 Å), whereas DBN crystallized with two molecules in the asymmetric unit at medium resolution (2.2 Å).

The overall structures of DBN and DBI are otherwise highly conserved, with a root-mean-square deviation of 0.44 Å for the main chain atoms (residues 1–173; superposition of molecule A of DBN onto DBI). The side chain of Thr23 in DBI is involved in a crystal contact, within which it forms a hydrogen bond to the backbone of Gly128 in a symmetry-related molecule (Fig. 2 *A*). This interaction is unique to DBI because Thr23 is not involved in lattice contacts of DBN. These polar interactions unambiguously demonstrate that the pathogenic P23T mutation enables direct interactions in the crystal lattice. By contrast, Ser36 is not involved in any direct contact in the DBI crystal (Fig. 2 *B*). It only contributes a hydrogen bond within a DBN crystal contact (Fig. 2 *C*).

If we are to relate our findings to the P23T single mutant, it is important to ponder whether the P23T + R36S mutant is a good model for it. The R36S contact is not activated in DBI, suggesting that it does not influence the structure of the DBI crystal to any significant extent, and the DBN crystal has the same structure and lattice contacts as the R36S single mutant. Because P23T and R36S reside on opposite sides of the N-terminal domain, we expect the structural and energetic influence of the two to be uncoupled. The molecular interactions at the 36 locus are also distinct. In the structures of DBN and the R36S single mutant (PDB: 2G98) (12), Ser36 forms a hydrogen bond with Asn24 with a symmetry mate in the lattice. In contrast, DBI forms an intramolecular hydrogen bond with Asp62, which in turn ion-pairs with Arg140 in a symmetry-related molecule. Thus, Ser36 in DBI orients Asp62 lattice interactions, which is distinct from DBN.

There is a significant degree of flexibility in the C-terminal domain of DBN (molecule B), which likely explains why only a medium resolution structure could be obtained, as evidenced by the associated B-factors (Table 1). Other HGD mutants with medium resolution structures, namely the P23T (PDB: 4JGF) (20) and R36S (PDB: 4JGF) (12) single mutants, display comparable flexibility in the

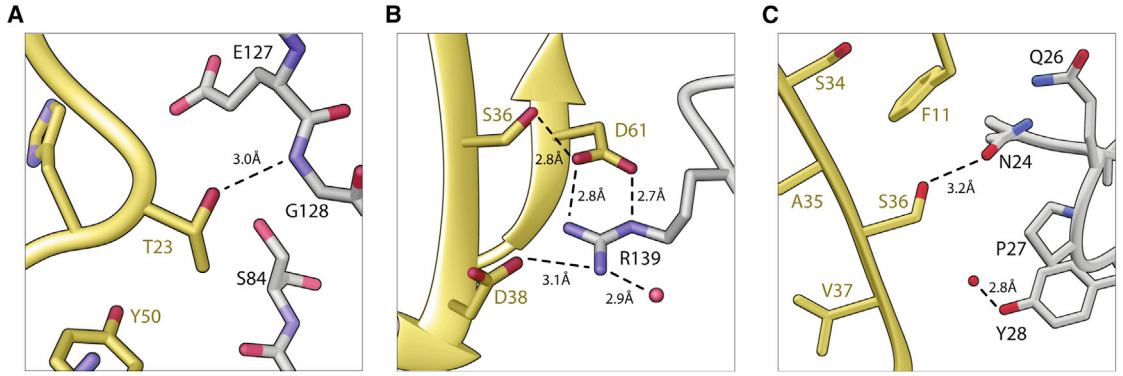

FIGURE 2 (*A*) Interactions between Thr23 and the crystal lattice. The side chain of Thr23 (*gold*) forms a hydrogen bond with the backbone of NH of Gly128, indicating the close contacts between the P23T locus and a symmetry-related molecule. (*B*) Crystal contacts near the R36S locus of DBI are shown. Ser36 does not make direct contacts in the crystal lattice. However, the side of Ser36 hydrogen bonds with Asp61, which additionally forms a salt bridge to Arg139 in a symmetry-related molecule. (*C*) Crystal contacts of DBN involve R36S. In this case, there is a hydrogen bond between Ser36 and Asn24.

C-terminal domain. By contrast, the corresponding domain in the DBI crystal is more rigid. A stabilizing lattice contact is formed between Ser173 (Oγ) and a symmetry-related Gly157 (O), which is associated with the higher-resolution structure (Fig. 3 *A*). Strikingly, the C-terminal carboxylate forms an ion pair with Arg141 from the same symmetry mate. There are also nonpolar interactions between Phe172 and Gln67 from a second symmetry-related molecule, indicative of the intimate associations between the C-termini of DBI in the crystal lattice. It is not clear, however, whether this flexibility is a result of a lack of a stabilizing crystal contact or whether, conversely, it precludes contact formation.

The formation of a hydrogen bond between Thr23 and the backbone of Gly129 in DBI suggests a molecular basis for a change in protein-protein interactions in the region of position 23 in the mutant protein. The change in the net binding energy between native HGD and the P23T single-mutant protein, calculated from the solubility data (16), corresponds to $\sim$2.4 k_BT, which is indeed the strength of a typical hydrogen bond. Beyond this observation, there is no obvious structural basis for the inverted temperature dependence of the solubility line. Therefore, we employed a modeling strategy based on custom patchy particle colloidal models to investigate the microscopic origins of the inverted solubility of the double mutant.

The model describes proteins as having a hard, spherical core with directional, short-ranged attractive patches representing crystal contacts derived from the crystal structures (see Supporting Materials and Methods). DBI and DBN are modeled with five patches each, as determined from their crystal contacts, which we assume recapitulate the relevant physical chemistry for crystal formation (Fig. 4; see Supporting Materials and Methods for details, including the amino acids involved in the different contacts). Despite its very crude description of protein-protein interactions, such models can recapitulate the characteristic topology of protein phase diagrams. Because solubility inversion necessarily implies some degree of temperature dependence for the patch interactions (35), we first consider deactivating the contact that contains the 23rd residue, where the new crystal contact is formed, around a temperature T_a with

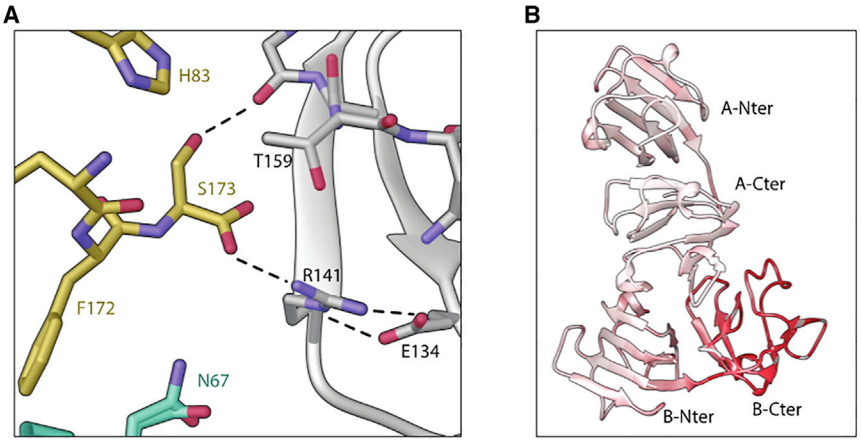

FIGURE 3 (*A*) Interactions at the C-terminus of DBI (*gold*). DBI is green, and symmetry-related molecules are gray and teal, respectively. (*B*) Flexibility of the C-terminal domain of DBN is shown. The two ribbon models (*A* and *B*) in the asymmetric unit are annotated by gradient colors of backbone B-factors. Regions with high values which denote flexibility are red, and ordered regions are white.

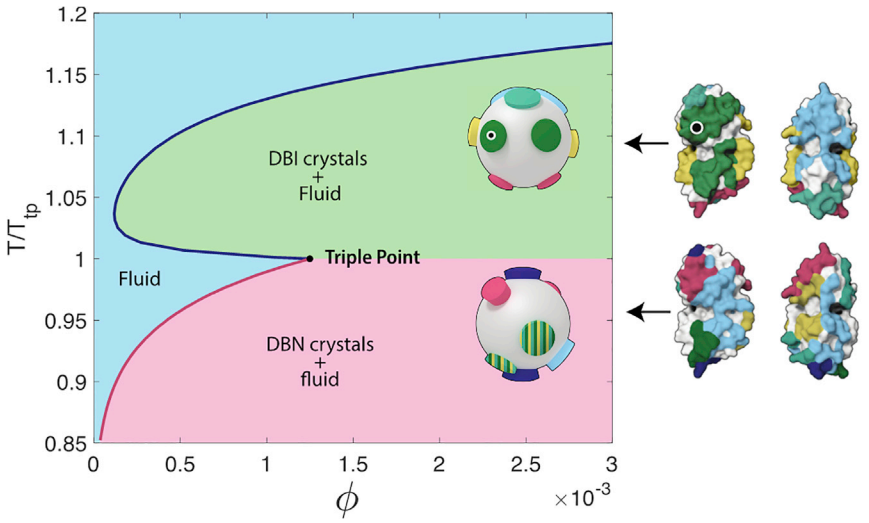

FIGURE 4 Crystal structures of DBI and DBN (*right*) are used to devise patchy particle models (*center*). Patches derive from crystal contacts, and each of them is represented here by a different color (see Table S1). A black dot specifically denotes the mutated 23rd residue. The resulting DBI (*red*) and DBN (*blue*) solubility lines obtained by deactivating the contact containing the 23rd residue below temperature T_a intersect at the nearby triple-point temperature, T_{tp}, in near-quantitative agreement with experimental observations.

rate τ, set by the experimentally observed inverted solubility temperature and density ranges, respectively.

Simulations of this model with specialized Monte Carlo methods determined the equilibrium phase diagram shown in Fig. 4. Upon cooling, the model solubility line for DBI crystals reaches a minimal volume fraction, $\phi \sim 10^{-4}$, before exhibiting an inverted solubility regime, all in remarkable agreement with experimental observations. The DBN solubility line, which shows normal solubility, intersects with that of the DBI crystal around $\phi \sim 10^{-3}$, forming a triple point. Experimental results are also suggestive of a triple point for comparable densities, but the flatness of the DBI solubility line in this regime precludes its accurate determination. This model allows us to speculate about the phase behavior of other double mutants that could be designed similarly, i.e., R36S + P23S and R36S + P23V, knowing that the single mutants P23S and P23V also exhibit inverted solubility. Strengthening the patch containing the 23rd residue in the model would push the DBI solubility line to lower ϕ, which suggests that the putative (inverted solubility) crystals of R36S + P23S and R36S + P23V may have higher solubilities than DBI. Such behavior is consistent with the binding energy estimates in (16,35) and therefore presents a new, to our knowledge, mechanism for the inverted solubility of proteins. How common this mechanism is compared to other proposals, however, remains to be determined.

Note that although similarly deactivating a larger set of DBI patches can also reproduce the observed experimental phase behavior, no microscopic basis exists for these changes, and doing so to more than a couple of patches melts the crystal before solubility inversion can be observed. Prior experimental observations suggest that a change to surface hydrophobicity using either small molecule dyes (31,45) or by mutagenesis at position 23 (46,47) may give rise to entropic gain upon crystallization and could explain the lowered solubility of the mutant protein. The functional form of the temperature-dependent patch energy in our model may suggest that additional flexibility in amino-acid side chains with increasing temperature in the solution phase may be more likely than a hydrophobic patch effect, but this possiblity cannot be excluded. However, because there is no experimental evidence for local unfolding or structural changes, we should not exclude the possibility that inverted solubility could have some other microscopic origin that has not yet been considered.

CONCLUSIONS

The rational design of a double mutant based on phase diagrams of single-mutant proteins has allowed us to produce two crystal forms of the P23T + R36S mutant of HGD that are polymorphs with different unit cells and distinct crystal contacts. The use of a single amino-acid substitution (R36S), previously shown to increase the crystallization propensity of HGD and to be unrelated to the mutant under consideration (P23T), is not standard but could provide an alternative design strategy to assist large-scale crystallization screening. The crystal displaying inverted solubility (DBI) forms a hydrogen bond at position 23, which distinguishes it from other γ-crystallin structures. We employed crystallographic data for both crystals, which made further investigation of the microscopic origin of inverted solubility and greater understanding of the solution behavior of the P23T single mutant. By considering a patchy particle model parameterized for this particular system, the phase diagram for the double-mutant protein was reproduced by simulations. A single temperature-dependent contact, specifically the contact that includes the P23T mutation, is sufficient to explain the crystallization behavior for the protein. Activation of the patch that contains this mutation was found to stabilize the inverted solubility crystal. This overall analysis illustrates that although noncovalent protein-protein interactions are far from trivial and thus challenging to predict, the

combination model and experimental phase diagrams could be a productive approach to rationalize and provide support for future crystallization studies.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2019.07.019.

AUTHOR CONTRIBUTIONS

A.R.K., S.J., M.K.Q., I.A., P.C., and J.J.M. performed experiments and analyzed data. A.R.K., I.A., P.C., and J.J.M. wrote the manuscript.

ACKNOWLEDGMENTS

The authors thank Nicolette Lubsen for permission to use the HGD plasmid DNA. I.A. and P.C. thank Diana Fusco for supporting discussions. The authors thank beamline scientists at Proxima 2 (Soleil Synchrotron) for their help in data collection. Data and scripts used in this work have been archived and can be accessed at https://doi.org/10.7924/r49w0dx6s.

This work was supported in part by Science Foundation Ireland Grant 11/RFP.1/PHY/3165, the Irish Research Council, and Science Foundation Ireland Stokes Lectureship to J.J.M. I.A. and P.C. acknowledge support from National Science Foundation Grant No. DMR-1749374. Computational parts of this work used the Extreme Science and Engineering Discovery Environment, which is supported by National Science Foundation Grant No. ACI-1548562, as well as the Open Science Grid and the Duke Compute Cluster. A.R.K. was supported by Science Foundation Ireland (grant number 12/IA/1239).

SUPPORTING CITATIONS

References (48–55) can be found in the Supporting Material.

REFERENCES

1. McPherson, A. 1990. Current approaches to macromolecular crystallization. *Eur. J. Biochem.* 189:1–23.
2. Goldschmidt, L., D. R. Cooper, …, D. Eisenberg. 2007. Toward rational protein crystallization: a web server for the design of crystallizable protein variants. *Protein Sci.* 16:1569–1576.
3. McManus, J. J., P. Charbonneau, …, N. Asherie. 2016. The physics of protein self-assembly. *Curr. Opin. Coll. Int. Sci.* 22:73–79.
4. Zigler, J. S., Jr. 1994. Lens proteins. *In* Principles and Practice of Ophthalmology. E. M. Alberts and F. A. Jacobiec, eds. Saunders, pp. 97–113.
5. Galkin, O., K. Chen, …, P. G. Vekilov. 2002. Liquid-liquid separation in solutions of normal and sickle cell hemoglobin. *Proc. Natl. Acad. Sci. USA.* 99:8479–8483.
6. Wang, Y., A. Lomakin, …, G. B. Benedek. 2013. Phase transitions in human IgG solutions. *J. Chem. Phys.* 139:121904.
7. Muschol, M., and F. Rosenberger. 1997. Liquid-liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *J. Chem. Phys.* 107:1953–1962.
8. Asherie, N., C. Ginsberg, …, S. Knafo. 2008. Solubility of thaumatin. *Cryst. Growth Des.* 8:1815–1817.
9. Hejtmancik, J. F. 1998. The genetics of cataract: our vision becomes clearer. *Am. J. Hum. Genet.* 62:520–525.
10. Stephan, D. A., E. Gillanders, …, M. J. Brownstein. 1999. Progressive juvenile-onset punctate cataracts caused by mutation of the gammaD-crystallin gene. *Proc. Natl. Acad. Sci. USA.* 96:1008–1012.
11. Pande, A., J. Pande, …, G. B. Benedek. 2000. Molecular basis of a progressive juvenile-onset hereditary cataract. *Proc. Natl. Acad. Sci. USA.* 97:1993–1998.
12. Pande, A., J. Pande, …, G. B. Benedek. 2001. Crystal cataracts: human genetic cataract caused by protein crystallization. *Proc. Natl. Acad. Sci. USA.* 98:6116–6120.
13. Basak, A., O. Bateman, …, J. Pande. 2003. High-resolution X-ray crystal structures of human gammaD crystallin (1.25 A) and the R58H mutant (1.15 A) associated with aculeiform cataract. *J. Mol. Biol.* 328:1137–1147.
14. Banerjee, P. R., A. Pande, …, J. Pande. 2011. Cataract-associated mutant E107A of human gammaD-crystallin shows increased attraction to α-crystallin and enhanced light scattering. *Proc. Natl. Acad. Sci. USA.* 108:574–579.
15. Pande, A., O. Annunziata, …, J. Pande. 2005. Decrease in protein solubility and cataract formation caused by the Pro23 to Thr mutation in human γ D-crystallin. *Biochemistry.* 44:2491–2500.
16. McManus, J. J., A. Lomakin, …, G. B. Benedek. 2007. Altered phase diagram due to a single point mutation in human gammaD-crystallin. *Proc. Natl. Acad. Sci. USA.* 104:16856–16861.
17. Banerjee, P. R., S. S. Puttamadappa, …, J. Pande. 2011. Increased hydrophobicity and decreased backbone flexibility explain the lower solubility of a cataract-linked mutant of γD-crystallin. *J. Mol. Biol.* 412:647–659.
18. Jung, J., I. J. Byeon, …, A. M. Gronenborn. 2009. The structure of the cataract-causing P23T mutant of human gammaD-crystallin exhibits distinctive local conformational and dynamic changes. *Biochemistry.* 48:2597–2609.
19. Boatz, J. C., M. J. Whitley, …, P. C. A. van der Wel. 2017. Cataract-associated P23T γD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH. *Nat. Commun.* 8:15137.
20. Ji, F., L. M. Koharudin, …, A. M. Gronenborn. 2013. Crystal structure of the cataract-causing P23T γD-crystallin mutant. *Proteins.* 81:1493–1498.
21. Asherie, N., A. Lomakin, and G. B. Benedek. 1996. Phase diagram of colloidal solutions. *Phys. Rev. Lett.* 77:4832–4835.
22. Lomakin, A., N. Asherie, and G. B. Benedek. 1999. Aeolotopic interactions of globular proteins. *Proc. Natl. Acad. Sci. USA.* 96:9465–9468.
23. Dorsaz, N., L. Fillion, …, D. Frenkel. 2012. Spiers Memorial lecture: effect of interaction specificity on the phase behaviour of patchy particles. *Faraday Discuss.* 159:9–21.
24. Sear, R. P. 1999. Phase behavior of a simple model of globular proteins. *J. Chem. Phys.* 11:4800–4806.
25. Bianchi, E., P. Tartaglia, …, F. Sciortino. 2008. Theoretical and numerical study of the phase diagram of patchy colloids: ordered and disordered patch arrangements. *J. Chem. Phys.* 128:144504–144510.
26. Curtis, R. A., H. W. Blanch, and J. M. Prausnitz. 2001. Calculation of phase diagrams for aqueous protein solutions. *J. Phys. Chem. B.* 105:2445–2452.
27. Quang, L. J., S. I. Sandler, and A. M. Lenhoff. 2014. Anisotropic contributions to protein-protein interactions. *J. Chem. Theory Comput.* 10:835–845.
28. Sun, G., Y. Wang, …, S. V. Buldyrev. 2016. The phase behavior study of human antibody solution using multi-scale modeling. *J. Chem. Phys.* 145:194901.
29. Sarikaya, M., C. Tamerler, …, F. Baneyx. 2003. Molecular biomimetics: nanotechnology through biology. *Nat. Mater.* 2:577–585.
30. Nguemaha, V., and H. X. Zhou. 2018. Liquid-liquid phase separation of patchy particles illuminates diverse effects of regulatory components on protein droplet formation. *Sci. Rep.* 8:6728.

31. Quinn, M. K., N. Gnan, …, J. J. McManus. 2015. How fluorescent labelling alters the solution behaviour of proteins. *Phys. Chem. Chem. Phys.* 17:31177–31187.
32. Fusco, D., J. J. Headd, …, P. Charbonneau. 2014. Characterizing protein crystal contacts and their role in crystallization: rubredoxin as a case study. *Soft Matter.* 10:290–302.
33. Fusco, D., and P. Charbonneau. 2016. Soft matter perspective on protein crystal assembly. *Colloids Surf. B Biointerfaces.* 137:22–31.
34. Fusco, D., and P. Charbonneau. 2013. Crystallization of asymmetric patchy models for globular proteins in solution. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 88:012721.
35. James, S., M. K. Quinn, and J. J. McManus. 2015. The self assembly of proteins; probing patchy protein interactions. *Phys. Chem. Chem. Phys.* 17:5413–5420.
36. Kmoch, S., J. Brynda, …, M. Elleder. 2000. Link between a novel human gammaD-crystallin allele and a unique cataract phenotype explained by protein crystallography. *Hum. Mol. Genet.* 9:1779–1786.
37. McCoy, A. J., R. W. Grosse-Kunstleve, …, R. J. Read. 2007. Phaser crystallographic software. *J. Appl. Cryst.* 40:658–674.
38. Emsley, P., B. Lohkamp, …, K. Cowtan. 2010. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66:486–501.
39. Adams, P. D., P. V. Afonine, …, P. H. Zwart. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66:213–221.
40. Kern, N., and D. Frenkel. 2003. Fluid–fluid coexistence in colloidal systems with short-ranged strongly directional attraction. *J. Chem. Phys.* 118:9882–9889.
41. de Las Heras, D., and M. M. da Gama. 2016. Temperature (de)activated patchy colloidal particles. *J. Phys. Condens. Matter.* 28:244008.
42. Frenkel, D., and A. J. C. L. Ladd. 1984. New Monte Carlo method to compute the free energy of arbitrary solids. Application to the fcc and hcp phases of hard spheres. *J. Chem. Phys.* 81:3188–3193.
43. Kofke, D. A. 1993. Direct evaluation of phase coexistence by molecular simulation via integration along the saturation line. *J. Chem. Phys.* 98:4149–4162.
44. Kofke, D. A. 1993. Gibbs-Duhem integration: a new method for direct evaluation of phase coexistence by molecular simulation. *Mol. Phys.* 78:1331–1336.
45. Quinn, M. K., S. James, and J. J. McManus. 2019. Chemical modification alters protein-protein interactions and can lead to lower protein solubility. *J. Phys. Chem. B.* 123:4373–4379.
46. Pande, A., K. S. Ghosh, …, J. Pande. 2010. Increase in surface hydrophobicity of the cataract-associated P23T mutant of human gammaD-crystallin is responsible for its dramatically lower, retrograde solubility. *Biochemistry.* 49:6122–6129.
47. Migas, U. M., M. K. Quinn, and J. J. McManus. 2017. Protein self-assembly following in situ expression in artificial and mammalian cells. *Integr. Biol (Camb).* 9:444–450.
48. Krissinel, E., and K. Henrick. 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372:774–797.
49. Berendsen, H. D., van der Spoel, and R. van Drunen. 1994. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91:43–56.
50. Hub, J. S. H., B. L. de Groot, and D. van der Spoel. 2010. g_wham—A free weighted histogram analysis implementation including robust error and autocorrelation estimates. *J. Chem. Theory Comput.* 6:3713–3720.
51. Ponder, J., and D. Case. 2003. Force fields for protein simulations. *Adv. Protein Chem.* 66:27–85.
52. Horn, H. W., W. C. Swope, …, T. Head-Gordon. 2004. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* 120:9665–9678.
53. Evans, D. J., and B. L. Holian. 1985. The Nose-Hoover thermostat. *J. Chem. Phys.* 83:4069–4074.
54. Essmann, U., L. Perera, …, L. G. Pedersen. 1995. A smooth particle mesh Ewald method. *J. Chem. Phys.* 103:8577–8593.
55. Hess, B., H. Bekker, …, J. G. E. M. Fraaije. 1998. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* 18:1463–1472.