Tardigrade CAHS Proteins Act as Molecular Swiss Army Knives to Mediate Desiccation Tolerance Through Multiple Mechanisms

Cherie Hesgrove[1], Kenny H. Nguyen[1], Sourav Biswas[1], Charles A. Childs[1], Shraddha KC[1], Bryan X. Medina[2], Vladimir Alvarado[2], Feng Yu[3], Shahar Sukenik[3,4], Marco Malferrari[5], Francesco Francia[6], Giovanni Venturoli[6,7], Erik W. Martin[8], Alex S. Holehouse[9,10], Thomas C. Boothby[1]

[1] Molecular Biology Department, University of Wyoming, Laramie, WY, USA
[2] Chemical Engineering Department, University of Wyoming, Laramie, WY USA
[3] Quantitative Systems Biology Program, University of California, Merced, CA, USA
[4] Department of Chemistry and Biochemistry, University of California, Merced, CA, USA
[5] Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna, via Selmi 2, I-40126 Bologna, Italy
[6] Laboratorio di Biochimica e Biofisica Molecolare, Dipartimento di Farmacia e Biotecnologie, FaBiT, Università di Bologna, via Irnerio 42, I-40126 Bologna, Italy
[7] Consorzio Nazionale Interuniversitario per le Scienze Fisiche della Materia (CNISM), c/o Dipartimento di Fisica e Astronomia (DIFA), Università di Bologna, via Irnerio 46, I-40126 Bologna, Italy
[8] Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN, USA
[9] Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO, USA
[10] Center for Science and Engineering of Living Systems (CSELS), Washington University in St. Louis, St. Louis, MO, USA

Abstract

Tardigrades, also known as water bears, make up a phylum of small but extremely hardy animals, renowned for their ability to survive extreme stresses, including desiccation. How tardigrades survive desiccation is one of the enduring mysteries of animal physiology. Here we show that CAHS D, an intrinsically disordered protein belonging to a unique family of proteins possessed only by tardigrades, undergoes a liquid-to-gel phase transition in a concentration dependent manner. Unlike other gelling proteins, such as gelatin, our data support a mechanism in which gel formation of CAHS D is driven by intermolecular β-β interactions. We find that gel formation corresponds with strong coordination of water, and slowing of water diffusion. The degree of water coordination correlates with the ability of CAHS D to protect lactate dehydrogenase from unfolding when dried. This implies that the mechanism for unfolding protection could be attributed to a combination of hydration and slowed molecular motion. Conversely, rapid diffusion leading to efficient molecular shielding appears to be the predominant mechanism preventing protein aggregation. Our study demonstrates that distinct mechanisms are required for holistic protection during desiccation, and that protectants, such as CAHS D, can act as molecular 'Swiss Army Knives' capable of providing protection through several different mechanisms simultaneously.

Introduction

Anhydrobiosis, the ability to survive near-complete water loss, is an intriguing trait found in all kingdoms of life.[1] Extreme drying can impart a number of stresses on biological systems,[2–4] with protein dysfunctions being a major set of common perturbations.[2–4] The two prevalent, and non-mutually exclusive, forms of protein dysfunction during desiccation are protein unfolding and protein aggregation.[2]

Many desiccation-tolerant organisms protect their cells from drying induced damage by accumulating non-reducing sugars, such as sucrose[5] or trehalose.[6–10] The enrichment of disaccharides was long thought to be a universal feature of desiccation tolerance. However, several robustly anhydrobiotic organisms, such as tardigrades and rotifers, do not accumulate high levels of sugars during drying.[11–13] Instead, these animals use a diverse array of intrinsically disordered proteins (IDPs) to provide adaptive protection against desiccation.[2,3,12,14–17]

One example of stress tolerant IDPs are Cytoplasmic Abundant Heat Soluble (CAHS) proteins, which are employed by tardigrades to survive desiccation.[2,3,12,17] A model CAHS protein, CAHS D, is required for anhydrobiosis and provides desiccation protection when heterologously expressed in yeast and bacteria.[12] *In vitro,* CAHS D can protect lactate dehydrogenase (LDH) from denaturation when subjected to desiccation and rehydration.[3,12] However, a holistic molecular understanding of how CAHS proteins confer desiccation tolerance remains unknown.

A general mechanism proposed to explain desiccation tolerance is the vitrification hypothesis.[18] This hypothesis hinges on slowed molecular motion reducing the frequency and speed of damaging processes, such as protein unfolding. In the early stages of drying, protection is proposed to occur through inducing high viscosity in the system, slowing diffusion and molecular motion. Once in the vitrified solid state, molecular motion is slowed so dramatically that biological processes are essentially stopped, preventing further degradation of the system. While it has been shown that tardigrades and their CAHS proteins form vitrified solids, vitrification is not mutually exclusive with other potential mechanisms of desiccation tolerance (see Results for further discussion).[2,3,12,19,20]

Here we set out to understand the mechanisms that underlie desiccation protection of client proteins by CAHS D. We present evidence that CAHS D can undergo a sol-gel transition in a concentration- and temperature-reversible manner. To understand how gelation impacts desiccation tolerance, we combined rational sequence design with a suite of complementary structural and biophysical techniques. Unexpectedly, we found that the mechanisms underlying protein stabilization do not correlate with the inhibition of protein aggregation, revealing that protection against these two major forms of protein dysfunction are mechanistically distinct. We also find that CAHS D's interactions with water, as measured through T_2 relaxation, was a strong predictor of unfolding protection. Our findings shed light not only on the fundamental biology underlying tardigrade anhydrobiosis and the function of IDPs during desiccation, but also provides avenues for pursuing applications such as the engineering of stress tolerant crops, and the stabilization of temperature sensitive therapeutics in a dry state.

Results

CAHS D undergoes a sol-gel transition

CAHS D (Uniprot: P0CU50) is a highly charged 227-residue protein that is predicted to be fully disordered.[2,12,21] In the course of its expression and purification, it was observed that CAHS D undergoes a sol-gel phase transition,[22] transitioning from a liquid into a solid gel state (Fig. 1a). CAHS D gelation is concentration dependent: solutions below ~10 g/L (0.4 mM) remain diffuse, solutions 10 g/L − 15 g/L are increasingly viscous, and above ~15 g/L (0.6 mM), form robust gels (Fig. 1a).

Gelation is reversible through heating (Fig. 1a) and dilution (Fig. 1b). Reversibility suggests that gelation is driven by non-covalent physical crosslinks, as opposed to chemical

crosslinks which would yield an irreversible gel.[23] The thermal dependence of the sol-gel transition suggests that favorable enthalpy drives gelation, rather than the hydrophobic effect, since hydrophobic interactions are stabilized by increasing temperature.[24–26] This is reinforced by the reversibility seen through resolvation (Fig. 1b), which would not lead to dissolution of hydrophobic interactions. Thus, gelation is driven by non-hydrophobic, non-covalent interactions such as hydrogen bonding, polar interactions, or charge interactions.[25]

High-resolution imaging reveals that CAHS D gels form reticular networks (Fig. 1c). A fine meshwork of CAHS D fibers is interspersed with large pores (Fig. 1c, S1a bottom left). This topology is similar to gels formed by gelatin, (Fig. 1c) and morphologically distinct from crystalline solids formed by lysozyme (Fig. 1c). The macromolecular architecture of CAHS D gels is reminiscent of that formed by synthetic polymers, in which relatively sparse physical crosslinks underlie the network connectivity.[27] This implies that specific regions of CAHS D may drive intermolecular interactions.

Gelation Slows Diffusion, Vitrification Immobilizes Proteins

Slowed molecular motion through high viscosity and eventual vitrification is a cornerstone hypothesis in desiccation tolerance.[18,28] To determine how molecular motion is impacted in CAHS D gels, we assessed the diffusion of water in solutions of varying CAHS D concentration using low-field time-domain nuclear magnetic resonance spectroscopy (TD-NMR). This method is commonly used to determine the diffusion of water and oils within liquid, gel, and even solid materials.[29,30]

CAHS D dramatically slowed the diffusion of water, both below and above its gelation point (15 g/L, 0.6 mM) (Fig. 1d). Gelatin, which forms gels through entwined triple helices, slowed diffusion more than CAHS D. In contrast, lysozyme, a non-gelling protein, slowed diffusion much less than either gelling proteins (Fig. 1d).

To probe the degree to which hydration influences immobilization within the vitrified gel matrix, we performed two experiments on dried and rehydrated gels; hydrogen deuterium exchange (HDX) of CAHS D amides, and kinetics of a model protein embedded in the gel matrix. HDX experiments can distinguish between a tightly packed, conformationally restricted matrix (slow exchange), and a loosely packed pliable matrix (rapid exchange).We found that both increased HDX rates (Fig. 1e) and client protein kinetics (S1 c,d) were strongly correlated with humidity, demonstrating rehydration-dependent immobilization of client proteins and CAHS D when vitrified (for further detail, see supplemental text).

In summary, biochemical, imaging, and biophysical assessment of CAHS D demonstrate that this protective protein undergoes a sol-gel transition, likely driven by assembly through non-covalent, non-hydrophobic interactions, and that the dynamics of the gel and embedded client proteins is dramatically influenced by the level of residual moisture in the system. The strong responsiveness of CAHS D to hydration and its ability to slow water diffusion has implications for the molecular mechanisms of client protein protection.

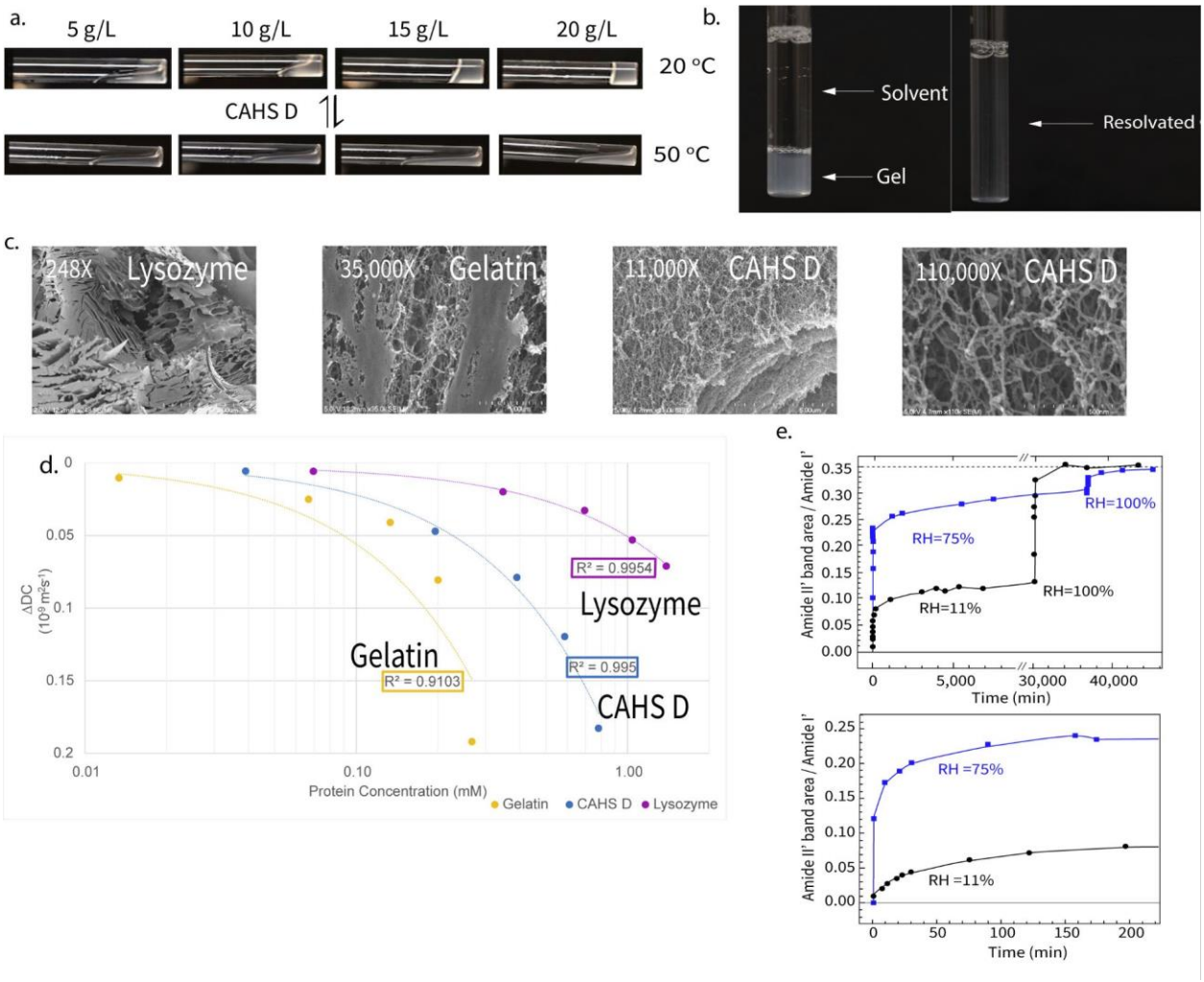


Fig. 1: Gel properties of CAHS D. a) Reversible temperature and concentration dependent gelation of CAHS D. CAHS D at 20˚C (top panel) shows gelation beginning at 15 g/L (0.6 mM), while at 50˚C gelation is not observed in any concentration. Once cooled, gel formation recurs at 15 g/L. b) Dilution of 20 g/L (0.8 mM) CAHS D gel in 20mM Tris buffer results in resolvation. c) SEM images of lysozyme (248X), gelatin (35,000X), and CAHS D(11,000X and 110,000X). Images show the reticular nature of CAHS D gel structure is similar to that of gelatin. All SEM imaging was performed with proteins at 50 g/L. d) Δ Diffusion coefficients for lysozyme (pink), CAHS D (blue), and gelatin (gold). Gelled proteins show more dramatic slowing of ΔDC than non-gelling lysozyme, in a concentration dependent manner. e) Kinetics of amide HDX in CAHS D glassy matrices. Left panel: CAHS vitrified gels at uniform initial hydration level were exposed to a D_2O atmosphere at RH=11% (black symbols) or RH=75% (blue symbols). Subsequently both were transitioned to pure D_2O atmosphere (RH=100%), which demonstrated that both dried gels showed similar HDX when fully saturated (supplemental text for more details). The dashed line represents the value of the amide II' band area normalized to the area of the amide I' band area. Right panel is an expansion of the left panel, showing the initial phase of the HDX kinetics.

Dumbbell-like Ensemble of CAHS D

CAHS proteins are highly disordered,[2,12] so standard structural studies are not feasible.[2,12] Instead, we performed all-atom Monte Carlo simulations to assess the predicted ensemble-state adopted by monomeric CAHS D proteins. Simulations revealed a dumbbell-like ensemble, with the N- and C-termini of the protein forming relatively collapsed regions that are held apart from one another by an extended and highly charged linker region (LR) (Fig. 2a, Mov. S1). Moreover,

meta-stable transient helices are observed throughout the linker region (LR), while transient β sheets are observed in the N- and C-terminal regions (Fig.2a, S2a,b).

To validate simulations, we performed Small Angle X-ray Scattering (SAXS) on monomeric CAHS D. The radius of gyration (R_g) – a measure of global protein dimensions – we obtained was 4.84 nm (simulation R_g = 5.1 nm), and the aligned scattering profiles obtained from simulation and experiment show good agreement (Fig. 2e, S2c). These dimensions are substantially larger than those expected if a protein of this size was either folded (R_g = 1.5 - 2.5 nm) or behaved as a Gaussian chain (R_g = 3.8 nm), yet smaller than a self-avoiding random coil (R_g = 6.5 nm) (Fig. 2f).

The expanded nature of CAHS D derives from the LR, which contains a high density of well-mixed oppositely charged residues (Fig. 2b, S2d), preventing its conformational collapse.[31,32] The transient helices formed in the LR have an amphipathic nature, (Fig 2a,b,c, Fig S2a) and are predicted to have a hydrophobic and a charged face[2] (Fig 2c). We note that the CAHS D linker is among the most well-mixed, high-charge sequences in the entire tardigrade proteome (Fig S2c); thus the extended nature of this sequence likely represents a functional, evolutionarily selected trait.

Circular dichroism (CD) spectroscopy confirms the largely disordered nature of full-length CAHS, with some propensity for α-helical and β-sheet formation (Fig. 2d). CD spectroscopy performed on truncation mutants containing only the LR or N-terminal region confirmed substantial helical content in the LR (~50%), and β-sheet content in the N-terminal region (~20%) (Fig 2d). We observed no residual structure by CD in the isolated C-terminal region, contrary to the predicted β-sheet content. This could be caused by the loss of sequence context in the truncated terminus, so we inserted three structurally disruptive prolines[33,34] into the predicted C-terminal β-sheets of the full-length protein (Fig S3a). CD on FL-Proline showed a ~50% reduction in β-sheet content relative to wildtype (Fig. 2d; Fig. S1b), confirming the β-sheet nature of the C-terminus.

Overall these data indicate that CAHS D exists in a dumbbell-like ensemble, which moves through conformational states consisting largely of β-sheeted termini held apart by an extended α-helical linker.

Gelation Driven by Terminal Interactions

A dumbbell-like protein with beta-sheeted termini would be an ideal candidate for gelation via the 'stickers and spacers' model.[35,36] In this model, discrete sites along a protein that contribute attractive intermolecular interactions are designated as stickers while non-interactive regions are spacers. Here, the N- and C-terminal regions of CAHS D can be considered stickers, while the LR is a spacer. When the spacer has a large effective solvation volume, like the expanded linker of CAHS D, phase separation is suppressed in favor of a sol-gel transition.[37] Moreover, we predict that intra-protein terminal interactions, which would suppress assembly through valence capping, are reduced by the separation enforced through the extended LR (Fig. 2c).[38] Based on our biophysical characterization of the monomeric protein, we hypothesized that gel formation of CAHS D occurs through inter-protein β-β interactions mediated between termini (Fig. 2g).

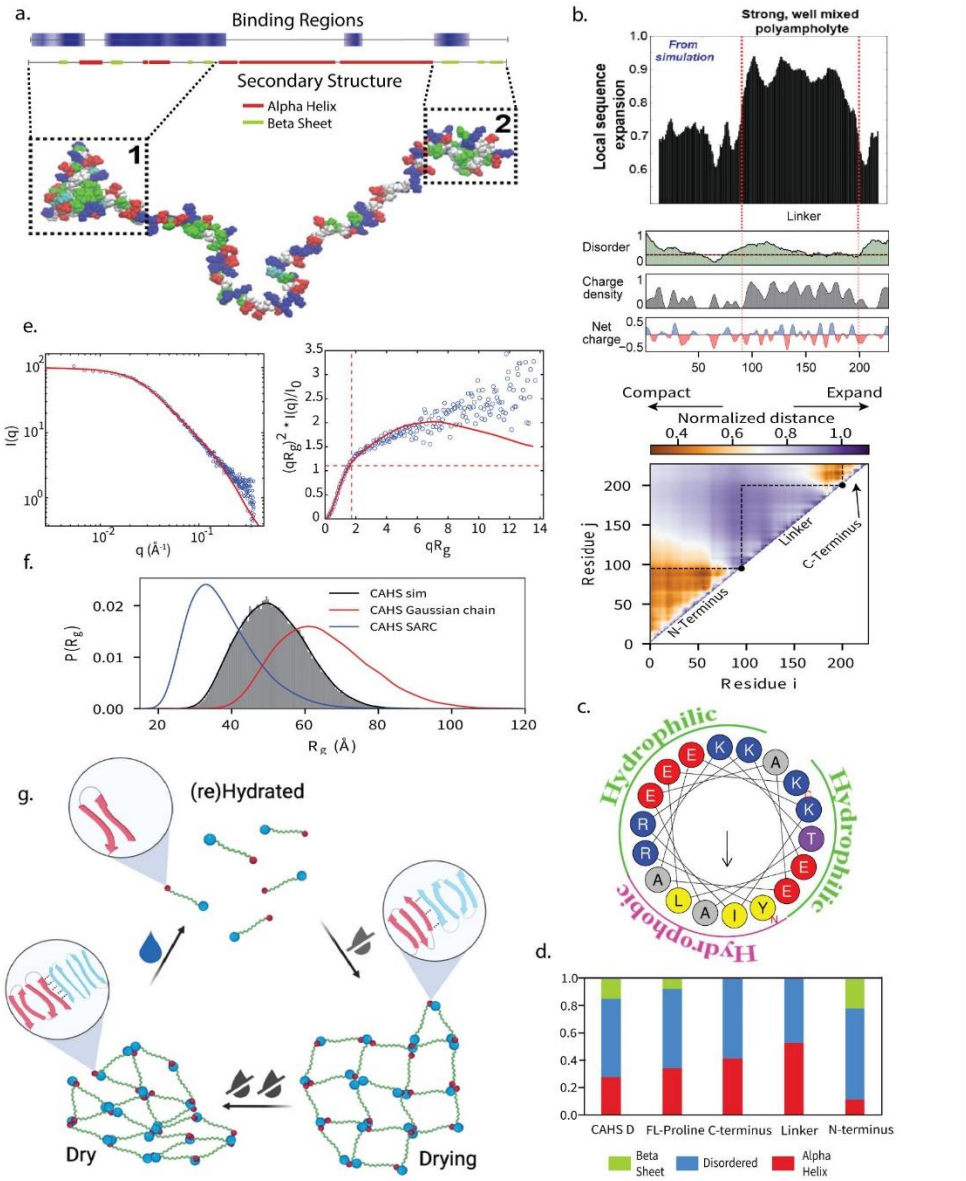


Fig 2. **CAHS D gelation is driven by the stickers and spacers model.** a) Bioinformatic predictions of secondary structure characteristics (top) and representative conformational global ensemble model (bottom) of CAHS D showing the extended central linker and the two terminal 'sticky' ends. b) Top panels illustrate the relationship between sequence expansion, disorder, charge density, and net charge of CAHS D. Heat map (bottom) shows normalized distance between residues as predicted in simulations. c) Helical wheel plot (HeliQuest, plothelix) shows the distribution of charged faces in a predicted linker helix. Yellow residues are hydrophobic, grey residues are non-polar, red residues are acidic and blue residues are basic. The central arrow indicates the direction of the hydrophobic moment. d) The ratio of α helix, disorder, and β sheet secondary structure propensity determined by CD spectroscopy for CAHS D, FL-Proline, C-terminus, Linker and N-terminus. e) Raw SAXS data for monomeric CAHS D protein (left) and the Kratky transformation (right). Experimental data (blue circles) and the form factor calculated from simulations (red lines) were normalized to the zero-angle scattering and overlaid. f) CAHS D (black line) radius of gyration determined from simulations, compared radii of a self-avoiding random coil (blue line) and a gaussian chain (red line), of equal linear size. g) Proposed mechanism of gelation for WT CAHS D. As water is removed from the system, monomers assemble through β-β interactions in the termini. These interactions are strengthened as drying progresses. Upon rehydration, CAHS D gel can easily disassemble as seen in Fig. 1c.

To test this, we generated a range of CAHS D variants (Fig. 3a) which disrupt the dumbbell-like ensemble, and thus should not gel. Consistent with our hypothesis that β-β interactions drive gelation, all variants lacking at least one termini resulted in a loss of gelation (N, LR, FL-Proline, NL1, CL1; Fig. 3a, Fig. S3a,b). Unexpectedly, variants that replaced one terminal region for another (NLN and CLC; Fig. 3a) also did not form gels under the conditions tested. These results show that heterotypic interactions between N- and C-termini are crucial for strong gel formation, implicating molecular recognition and specificity encoded by the termini. Interestingly, the 2X Linker variant, which maintained heterotypic termini but doubled the length of the linker, gelled rapidly at 5 g/L (0.1 mM), well below the gelling point of the wildtype protein at 15 g/L (0.6 mM) (Fig. 3a). Thus, the length of the linker may tune the gel point by determining the monomeric molecular volume, setting the overlap concentration, which is a key determinant of the gel point.[39,40]

Optimal Hydration for β-β Stability

Using Fourier-transform infrared spectroscopy (FTIR) we observed that β-sheet interactions are maximal in CAHS glasses at 95% relative humidity (RH), and lowest in fully hydrated gels and glasses at 11% RH (Fig. 3b). This raises the possibility that there is an optimal hydration level for stabilizing β-β contacts, which may relate to the need for higher stability while the matrix is undergoing drying or rehydration.

To confirm the role of β-β interactions in gelation, we assayed CAHS D solutions at ranging concentrations with thioflavin T (ThT). ThT is used as a fluorescent indicator of amyloid fibrils,[41,42] and can report on β-β interactions.[43–45] We observed increases in ThT fluorescence intensity as a function of CAHS D concentration, with the most dramatic increase between 5 – 10 g/L (0.2 – 0.4 mM), suggesting nucleation & assembly of CAHS D monomers prior to gelation at 15 g/L (0.6 mM) (Fig. 3c). Gelatin did not show concentration dependent changes in ThT fluorescence, and was not significantly different from buffer controls (Fig. 3c). ThT labeling of NLN and CLC variants showed a concentration-dependent increase in fluorescence, with more interactions shown in NLN than CLC (Fig 3c), consistent with the predicted degree of β-sheet content in each (Fig. 2a,d). FL-Proline β-β interactions were lesser than variants with two folded termini, as expected. These data suggest that β-β interactions increase parallel to gelation and that assembly events are occurring between monomers, prior to an observed system-wide sol-gel transition.

Together, our SAXS data and bioinformatic analyses show that the extended nature of the protein can be attributed to the LR, and that this extension helps to hold the termini of an individual protein apart. CD, FTIR, and ThT labeling data show that CAHS D gelation is mediated by inter-protein β-β contacts formed through the interaction of the termini, which are influenced by hydration.

Mechanistic Determinants of Protection

Protein gelation is uncommon in a biologically relevant context, and in these instances is often functional.[46,47] This led us to wonder if gelation of CAHS D is linked to protection. The enzyme lactate dehydrogenase (LDH) unfolds and becomes irreversibly non-functional,[12,17,48] but does not aggregate,[49,50] during desiccation and an LDH activity assay is commonly employed to measure protection against unfolding.[12,17,50] In contrast, citrate synthase (CS) forms non-functional aggregates after

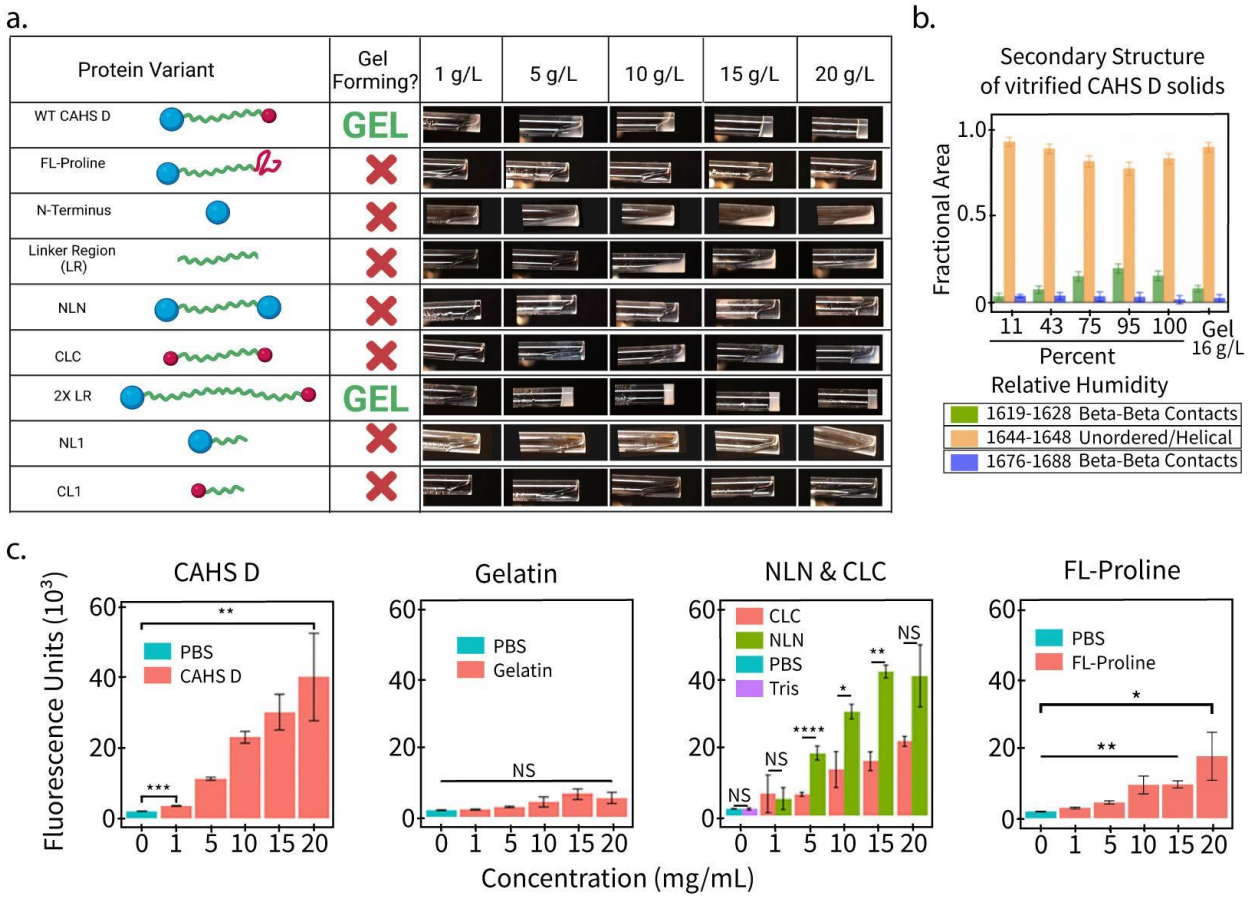


Fig. 3: Gelation of CAHS D relies on its dumbbell-like ensemble, and intermolecular β-β interactions. a) Schematic of the variant structure, the gel propensity of each listed concentration, and photos of the proteins in solution at indicated concentrations. Only CAHS D and the 2X Linker variant showed gelation. Future depictions of these constructs include a green "G" next to the cartoon to indicate their ability to form a gel. b) Relative secondary structure content of CAHS D as determined by FTIR analysis of the amide I' band in a glassy matrix at different hydration levels, and in the hydrated gel state at a concentration of 16 g/L (0.6 mM). The Gaussian sub-bands centered in the wavenumber interval (1644-1648 cm^{-1}) are attributed to unordered/helical regions, while the Gaussian components peaking in the wavenumber intervals (1619-1628 cm^{-1}) and (1676-1688 cm^{-1}) are indicative of interprotein β-sheet structures. c) Thioflavin T (ThT) fluorescence as a function of concentration for CAHS D, gelatin, NLN, CLC and FL-Proline. Error bars represent standard deviation. Significance determined using a Welch's t-test. All experiments presented used a minimum of 3 replicates. Asterisks represent significance relative to wild type CAHS D. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, NS is not significant.

successive rounds of dehydration and rehydration.[49] We tested all gelling and non-gelling variants (Fig. 3a) using both assays, to determine how gelation influences CAHS D protective capacity (Fig. S4, Fig. 4a&b), and were surprised to see that while gelling variants prevented unfolding best, they performed least well at preventing aggregation (Fig. S4, Fig. 4a&b), suggesting that different forms of protein dysfunction are prevented through distinct mechanisms.

Determinants of Aggregation Protection

The mechanism most often attributed to the prevention of desiccation-induced protein aggregation is the molecular shielding hypothesis.[49,51,52] This posits that protein aggregation can be prevented by protectants that act as disordered spacers, which impede interactions between

aggregation prone molecules.[49,51–53] Shielding proteins generally have some means of interacting with client proteins, although the interactions are weak.[49,51,52,54]

Results from our CS aggregation assay are consistent with this mechanism; the presence of extended LR with a single terminus emerged as the primary determinant of CS aggregation prevention (Fig. 4b). Meanwhile, the presence of two folded termini seemed antagonistic to aggregation protection, suggesting that higher order assembly, made possible by two termini, is detrimental to preventing aggregation.

To determine the role of diffusion in protection, we measured diffusion of water and compared this with aggregation protection (Fig. 4c,d). We found that neither aggregation protection nor diffusion were dependent on molecular weight or gelation. For example, FL-Proline (~25 kDa) is best at aggregation protection, does not form gels, and has much faster water diffusion than CAHS D (~25 kDa). On the other hand, NLN (~31 kDa) slowed diffusion nearly identically to CAHS D, yet did not form gels, and ranked 6th at protecting from aggregation (CAHS D was ranked 7th) (Fig. 4c,d).

We found that diffusion at 1mM normalized to the variant's molecular weight loosely followed an inverse trend, where faster diffusion trended with lower aggregation PD50s (Fig. 5d, S5f). In general, the higher linker content a variant contained, the better it performed at aggregation protection, so long as the variant also had a single folded terminus. Variants with a greater than 75% proportion of N-terminal content (NL1 and N-terminus), were outliers to this trend, both by having a high PD50 with fast diffusion (N-terminus), and by having a very low PD50 with slow diffusion (NL1). This is an interesting result that indicates that the N-terminus may in some way impact the relationship between diffusion and aggregation protection.

The negative relationship between diffusion and aggregation protection conflicts with key aspects of the vitrification hypothesis. However, if we consider aggregation protection to be primarily driven through molecular shielding, then this antagonism between anti-aggregation activity and diffusion can be understood more clearly. Slowed diffusion would dampen dynamic movement of protectants around clients, limiting client isolation, making the protectant a less effective aggregation shield. This suggests that rapid diffusion may be a key part of molecular shielding.

Determinants of Unfolding Protection

Gelation and diffusion also cannot fully explain protection against unfolding. For example, NLN did not gel, had diffusion identical to that of CAHS D, and was more protective against unfolding damage than CAHS D. Thus neither diffusion nor gelation alone can predict unfolding protection.

A potential driver of unfolding protection is the amount of extended LR present in a variant. This is true both in dual-termini variants, and in single-termini variants. An example is found by the higher protection afforded by CL1 (9.7 kDa Linker content) versus NL1 (2.9 kDa Linker content).

Monomer assembly, even if it does not ultimately lead to gelation, is critical for unfolding protection. All variants with two folded termini were either better or equivalent to CAHS D unfolding protection. This property, however, cannot predict protection displayed by variants with only one folded terminus.

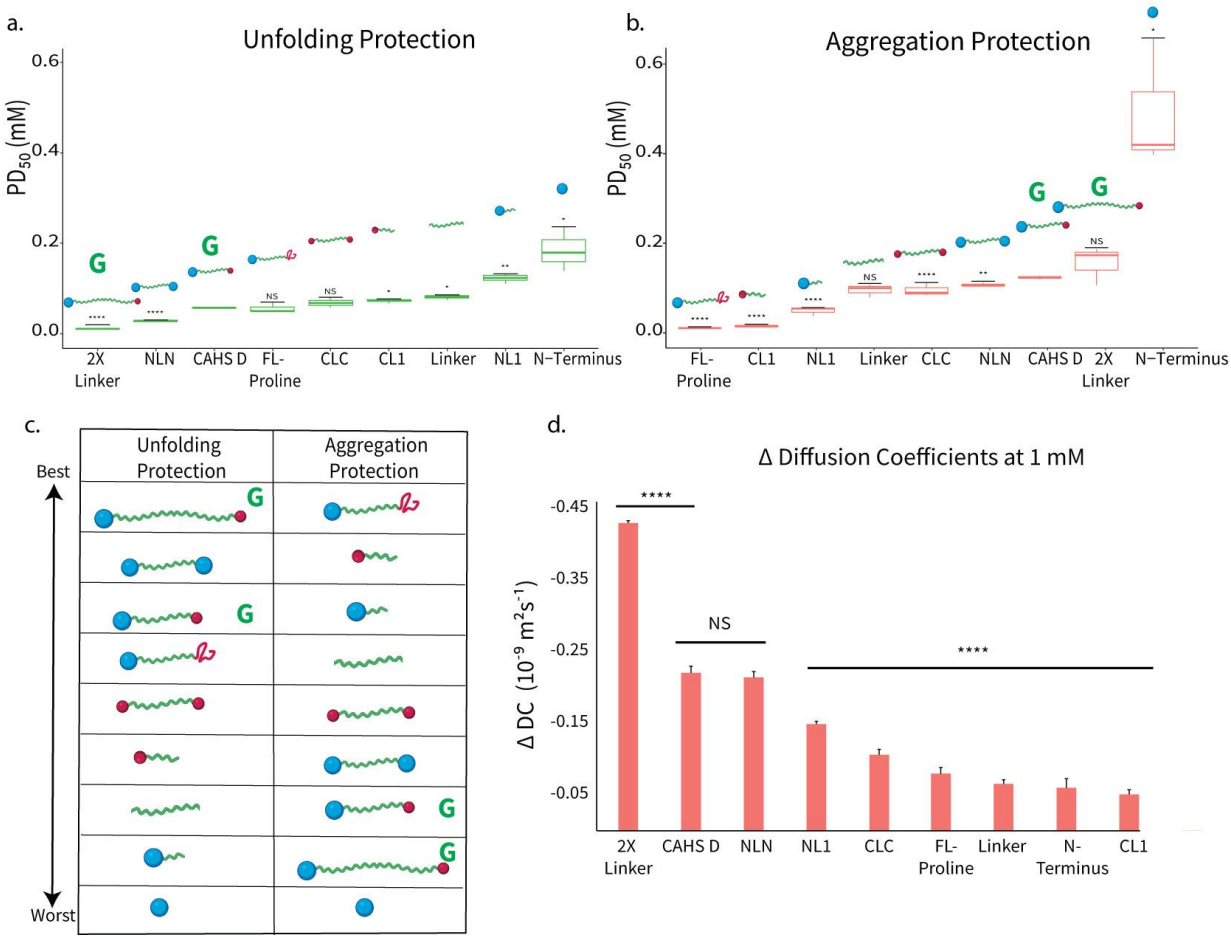


Fig. 4: Gelation promotes protection against protein unfolding, but not aggregation, during desiccation. Unfolding (a) and aggregation (b) PD50 values (shown left to right, in order from best to worst) of each variant. Lower PD50s correspond with better protective capability. Significance was determined for both (a) and (b) using a Welch's t-test, of variant's protection relative to that of wildtype. c) Ranking of variants in both the unfolding and aggregation assays, with the lowest PD50 (best protectant) at the top and the highest PD50 (worst protectant) at the bottom. Refer to (a) and (b) for information on relative protective ranking of each variant. Green G indicates that a variant is gel forming. d) ΔDiffusion Coefficients for all variants at 1 mM, calculated from linear fits for each variant's full concentration range (1-20 mg/mL). Error bars represent the standard error for the full concentration range for each variant. Significance was determined as a χ^2 analysis of the equality of linear regression coefficients for the linear fit of each variant, compared to that of CAHS D. (a-d) All experiments used a minimum of 3 replicates. Asterisks represent significance relative to the wildtype.*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, NS is not significant.

Taken together, our LDH and CS assays show that linker content is a driver of protection in both forms of protein dysfunction. However, the context of the linker ultimately determines its functionality. The properties that drive unfolding protection are distinct from those that dictate aggregation protection (Fig. 4c), again implying that the mechanisms underlying prevention of these unique forms of protein dysfunction are distinct.

Water coordination predicts unfolding, but not aggregation, protection

Since gelation, assembly, linker content, and diffusion do not fully account for the protective trends observed in our unfolding assay, we looked to other mechanistic hypotheses regarding desiccation protection to explain how this form of protection might be mediated.

The water entrapment hypothesis addresses how a protectant may compensate for the loss of a stabilizing hydrogen bond network (HBN) experienced by a client protein during desiccation stress. This hypothesis proposes that the protectant coordinates a layer of water between itself and the client protein, to maintain the client's HBN.[55]

To understand how CAHS D and its variants interact with water, TD-NMR was used to measure T_2 relaxation (Fig. 5a). T_2 relaxation yields information regarding the coordination of water molecules in a system.[56–61] Strongly coordinated water molecules are structurally organized, have reduced motion overall and less freedom of motion, and slower exchange within hydration layers.[62–67] Therefore, faster T_2 relaxation times correspond with coordinated or structured water molecules, and slower relaxations indicate less coordinated water, such as is found in bulk liquids (Fig. S5a) (see supplemental text for further details).

Among our CAHS D variants, we observed a remarkable diversity of water coordinating properties (Fig. 5a, S5b). Concentration-dependent water coordination was found, in order of strength, in 2X linker, CAHS D, CLC, NLN, and N-terminus (Fig. 5a). For a deeper understanding of the relationship between water coordination and protection, we plotted the PD50 for our LDH and CS assays as a function of the major T_2 peak midpoints normalized to the variant's MW (T_2/MW) at 1 g/L for all variants (Fig. 5b,c). This represents what is essentially water-coordinating ability per kDa of protein, and gives insight into the degree to which a protein can coordinate water when the overall size of each protein is taken into account. We were surprised to find a strong correlation between T_2/kDa and unfolding protection (Fig. 5c) which was not replicated in aggregation protection data (Fig. S5c).

This further reinforces the notion that CAHS D uses disparate mechanisms to protect from different forms of protein damage; molecular shielding for aggregation protection and vitrification and water coordination for protein unfolding. The relationship between T_2 relaxation and unfolding protection signals the relative importance of water coordinating capability in protecting a client protein's native fold, which in turn provides evidence for the importance of accounting for hydration when considering the proposed mechanisms for desiccation protection.

Discussion

In this work we set out to understand the drivers of CAHS D desiccation protection. We show that CAHS D forms reversible physical gels, an observation explained by a model in which attractive interactions encoded by the N- and C-terminal regions facilitate intermolecular self-assembly. In contrast, the large central linker region acts as an extended, highly charged spacer, reducing intramolecular interactions and setting the gel-point based on the overall protein dimensions.

From a functional perspective, neither gelation nor diffusion correlate directly with prevention of protein aggregation or unfolding. Protection from unfolding is influenced by structural features, but can be directly related to a variant's ability to coordinate water. Aggregation protection, on the other hand, seems to be favored in systems with more rapid diffusion that do not form higher order assemblies.

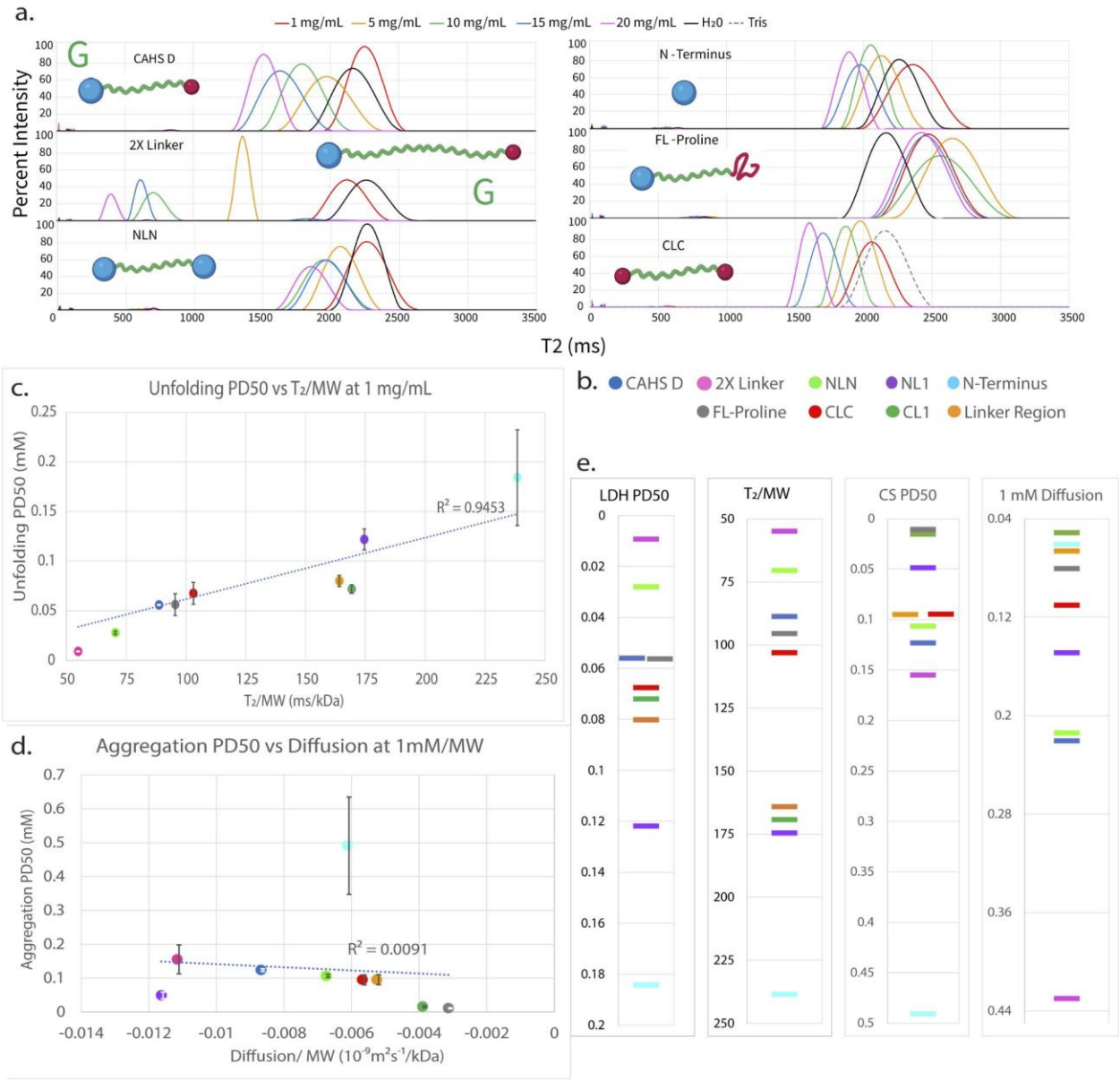


Fig. 5: Slowed diffusion and coordination of water underlie CAHS D's ability to promote protection against protein unfolding. a) T_2 distributions of variants WT CAHS D, 2X Linker, N-terminus, NLN, LR, and CLC. (All variants not shown here can be found in Fig. S5b). 2X linker, CAHS D, and NLN show the most shifted T_2 peak midpoints as a function of concentration. b) Legend for plots in (c-e). c)Unfolding PD50 as a function of water coordination per kilodalton for the 1 g/L. As shown by the linear trend of these plots, the water coordination per kilodalton is a strong indicator of unfolding protection. The relationship is shown in further detail in (e). d) Aggregation PD50 as a function of water diffusion at 1 mM per kilodalton. As shown, there is not a strong correlation between these properties. e) Single variable plots of Unfolding PD50, water coordination per kilodalton, Aggregation PD50, and ΔDC at 1mM. The top three variants that protect best against unfolding also show the slowest diffusion, however the relationship between ΔDC and unfolding protection breaks down after this point. Comparison of the Unfolding PD50 ranking and the T_2/kDa ranking shows a striking relationship. However, no such relationship is seen for aggregation protection (Fig S5c).

Based on our work, we propose that water entrapment, wherein the protectant traps a layer of hydrating water between itself and the client protein,[55,68] is the major mechanism driving CAHS D-mediated unfolding protection. Conversely, molecular shielding provides the best

explanation of the trends observed for aggregation protection. We propose that CAHS D functions as a molecular Swiss-Army knife, offering multiple protective capabilities though distinct mechanisms.

Dynamic Functionality for the Dynamic Process of Drying

Surviving desiccation is not about surviving a singular stress, but rather surviving a continuum of inter-related stresses.

During early drying, the need to prevent aggregation may outweigh the need to maintain hydration of most cellular proteins. At this early stage, monomeric or low-oligomeric CAHS D could be performing important shielding functionalities to mitigate a cascade of aggregation. Such an aggregation cascade has been implicated in proteostatic dysregulation.[69] Aggregation protection during early drying would help preserve functional proteins during a critical window when the cell must adapt to water loss.

As drying progresses and further water is lost, CAHS D concentration increases[12] and its primary function may transition from aggregation prevention to protection against unfolding. As small oligomers of CAHS D grow and combine, clients could become trapped between assembled CAHS D. This would enable CAHS D to hold coordinated water to the surface of the entrapped clients, hydrating them and maintaining their hydrogen bonding networks. These smaller assemblies would slow diffusion in the system more and more as they condense to form the gel fibers observed in SEM (Fig. 1c).

This fibrous network of gelled CAHS D could provide a bridge from the drying phase to the fully dried vitrified solid phase. Molecular motion in a CAHS D gel is slower than a liquid, but faster than a vitrified solid, and residual water can be coordinated and held to the surface of client proteins. This would mitigate the water loss experienced by client proteins, allowing them to retain their native fold for as long as possible. By the time dehydration overwhelms CAHS D's capacity to hold water, a vitrified solid will have formed. At this point, even though the client's native fold may no longer be stabilized by a hydrogen bond network; full unfolding *still* would not occur because of the degree to which molecular motion is reduced in the vitrified solid.

The final phase of surviving desiccation is to rehydrate and return to active biological functions. CAHS D gels easily solubilize after gelation, indicating that they can rehydrate, resolvate, and release entrapped, protected clients, likely an essential step in efficiently returning the rehydrating organism to activity.

Our work provides empirical evidence that multiple mechanisms contribute to desiccation tolerance, and that these mechanisms can be mediated by a single protectant molecule.

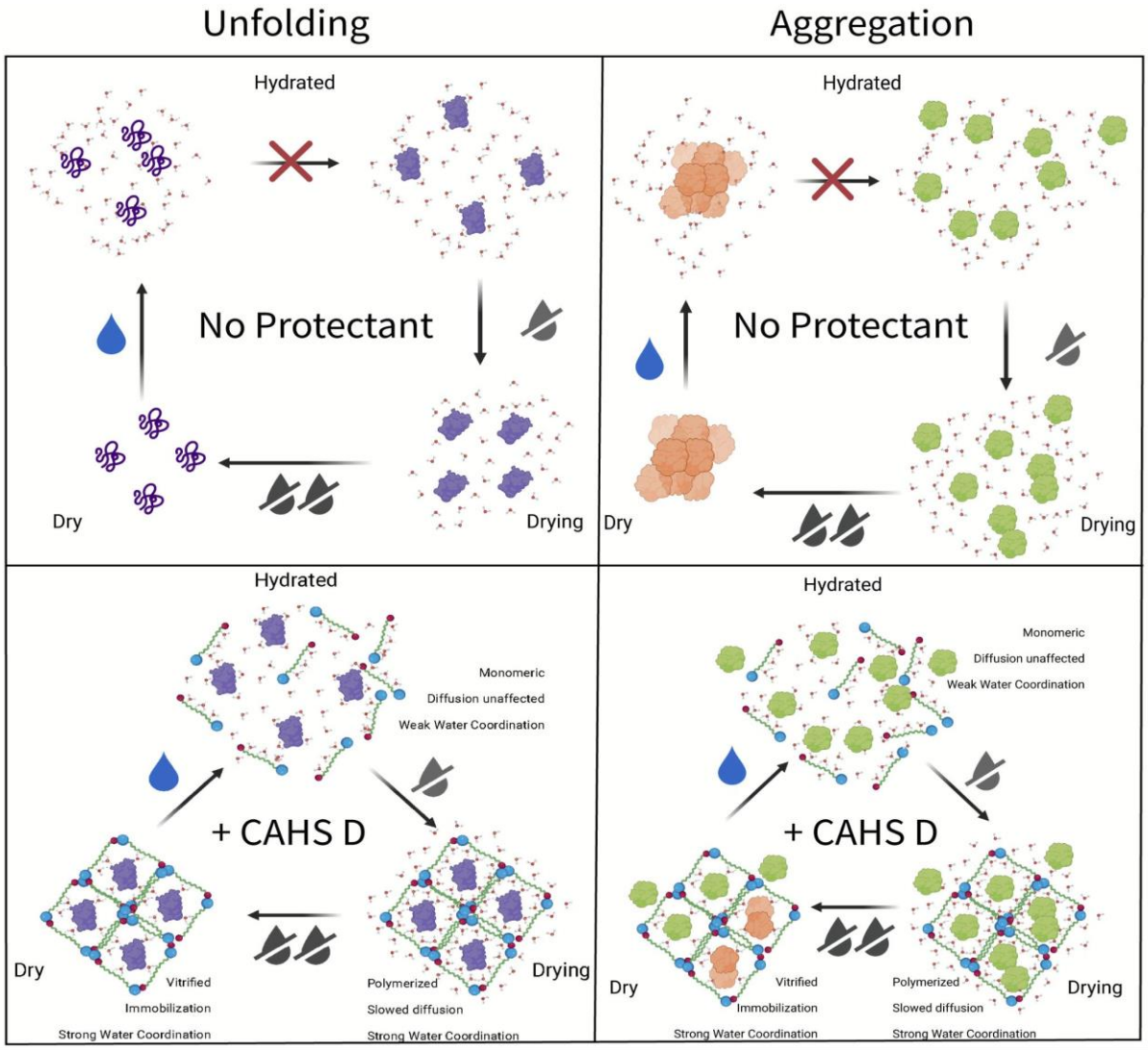


Fig. 6: Proposed mechanism of protection against desiccation induced protein aggregation and unfolding. Schematics depict the proposed mechanisms occurring within the meshwork of CAHS D fibers**.** The left panels show what occurs in the case of protein unfolding, while the right panels illustrate protein aggregation. Top panels display the system without protectants, while the lower panel of each shows the system with wildtype CAHS D. As water is lost, self-assembly begins in samples with CAHS D. Water is coordinated along the linker and to some degree the N-terminus of CAHS D. As samples enter a vitrified state, some water remains coordinated by the linker, but most water has been lost. Upon rehydration, the matrix readily disassembles and releases molecules from the gel structure.

Materials & Methods

Cloning

All variants and wild type CAHS D were cloned into the pET28b expression vector using Gibson assembly methods. Primers were designed using the NEBuilder tool (New England Biolabs, Ipswitch, MA), inserts were synthesized as gBlocks and purchased from Integrated DNA Technologies (Integrated DNA Technologies, Coralville, IA).

Protein Expression

Expression constructs were transformed in BL21 (DE3) E. coli (New England Biolabs), and plated on LB agar plates containing 50 µg/mL Kanamycin. At least 3 single colonies were chosen for each construct and tested for expression.

Large-scale expression was performed in 1 L LB/Kanamycin cultures, shaken at 37ºC (Innova S44i, Eppendorf, Hamburg, Germany) until an OD_{600} of 0.6, at which point expression was induced using 1 mM IPTG. Protein expression continued for four hours, after which cells were collected at 4000 g at 4°C for 30 minutes. Cell pellets were resuspended in 10 mL of resuspension buffer (20 mM tris, pH 7.5, 30 µL protease inhibitor [Sigma Aldrich, St. Louis, MO). Pellets were stored at -80ºC.

Protein Purification

Purification largely follows the methods in Piszkiewicz *et al.,* 2019.[17] Bacterial pellets were thawed and heat lysis was performed. Pellets were boiled for five minutes and allowed to cool for 10 minutes. All insoluble components were removed via centrifugation at 5,000 g at 10°C for 30 minutes. The supernatant was sterile filtered with 0.45 µm and 0.22 µm syringe filters (Foxx Life Sciences, Salem, NH). The filtered lysate was diluted 1:2 in purification buffer UA (8 M Urea, 50 mM sodium acetate [Acros Organics, Carlsbad, CA], pH 4). The protein was then purified using a cation exchange HiPrep SP HP 16/10 (Cytiva, Marlborough, MA) on the AKTA Pure 25 L (Cytiva), controlled using the UNICORN 7 Workstation pure-BP-exp (Cytiva). Variants were eluted using a gradient of 0-50% UB (8 M Urea, 50 mM sodium acetate, and 1 M NaCl, pH 4), over 20 column volumes.

Fractions were assessed by SDS-PAGE and pooled for dialysis in 3.5 kDa MWCO dialysis tubing (SpectraPor 3 Dialysis Membrane, Sigma Aldrich). For all variants except CLC, pooled fractions were dialyzed at 25°C for four hours against 2 M urea, 20 mM sodium phosphate at pH 7.0, then transferred to 20 mM sodium phosphate at pH 7 overnight. This was followed by six rounds of 4 hours each in Milli-Q water (18.2 MΩcm). Dialyzed samples were quantified fluorometrically (Qubit4 Fluorometer, Invitrogen, Waltham, MA), aliquoted in the quantity needed for each assay, lyophilized (FreeZone 6, Labconco, Kansas City, MO) for 48 hours, then stored at -20°C until use. CLC was dialyzed in 2 M urea, 20 mM Tris at pH 7 for four hours, followed by 6 rounds of 4 hours each in 20 mM Tris pH 7. CLC samples were quantified using Qubit4 fluorometer as described, concentrated using amicon spin-concentrators (Sigma-Aldrich) to the desired concentration and used immediately.

Visual gelation and heat/dilution gel resolubilization assay

Quantitated and lyophilized protein samples were transferred as powder into NMR tubes (Wilmad Lab Glass, Vineland, NJ)and resuspended in 500 µL of water to a final concentration of 5 g/L, 10 g/L, 15 g/L, and 20 g/L. Samples were left at room temp for 5 minutes to solubilize. If solubilization was not occurring (as determined visually), samples were moved to 55°C for intervals of 5 minutes until solubilized. If solubilization was not progressing at 55°C after 10 minutes of heating (as determined visually), then samples were transferred 95°C for 5 minute intervals until fully solubilized.

Solubilized proteins were transferred from heat blocks to the bench and left at ambient temperature for 1 hour. Tubes were then loaded horizontally into a clamp holder and photographed. Gelation was visually assessed by the degree of solidification or flow of the sample in the NMR tube.

After 1 hour at ambient temperature, proteins that had been found to form gels were transferred to a 55°C heat block for 1 hour. After an hour samples were returned to the photographic clamp holder and imaged immediately to confirm that gelation had been disrupted. Sample where placed upright on the bench at ambient temperature for 1 hour and allowed to reform gels and then imaged again as above.

A duplicate sample of 20 g/L CAHS D was prepared as described, and assayed for dilution resolubility. Buffer (Tris, 20 mM pH 7) was added to the gel to bring the final concentration of solvated CAHS D to 5 g/L, which is below the gelation point of the protein. Sample was photographed immediately after addition of buffer, vortexed for 5 seconds, and left to resolubilize. Sample completely dissolved within 30 minutes of buffer addition.

Helical Wheel Generation

Helical wheel plots were generated using the heliQuest sequence analysis module. CAHS D linker was used as the base template, with the α-helix option chosen as the helix type to model. A window size of 36 amino acids was used.

Scanning electron microscopy and critical point drying

Protein samples were heated to 95°C for 5 minutes and 50 µl of each sample were transferred to microscope coverslips. Samples were fixed in a 2.5% glutaraldehyde / 4% paraformaldehyde solution for 10 minutes. Samples were then dehydrated in an ethanol series going from 25%, 50%, 75%, 3x 100% with each incubation time being 10 minutes. Dehydrated samples were prepared for imaging using critical point drying (Tousimis Semidri PVT-3, SGC Equipment, Austin, TX) and supporter coating (108 Auto, Cressington Scientific Instruments, Watford, United Kingdom). Imaging was performed on a Hitachi S-47000 scanning electron microscope.

SAXS

All SAXS measurements were performed at the BioCAT (beamline 18ID at the Advanced Photon Source, Chicago, IL). SAXS measurements on monomeric CAHS were collected with in-line size exclusion chromatography (SEC-SAXS) coupled to the X-ray sample chamber to ensure the protein was monomeric. Concentrated protein samples were injected into a Superdex 200 increase column (Cytiva) pre-equilibrated in a buffer containing 20 mM Tris pH 7, 2 mM DTT, and 50 mM NaCl. Scattering intensity was recorded using a Pilatus3 1 M (Dectris) detector placed 3.5 m from the sample, providing a q-range from 0.004-0.4 $Å^{-1}$. One-second exposures were acquired every two seconds during the elution. Data were reduced at the beamline using the BioXTAS RAW 1.4.0 software.[70] The contribution of the buffer to the X-ray scattering curve was determined by averaging frames from the SEC eluent, which contained baseline levels of integrated X-ray scattering, UV absorbance, and conductance. Baseline frames were collected immediately pre-

and post-elution and averaged. Buffer subtraction, subsequent Guinier fits, and Kratky transformations were done using custom MATLAB (Mathworks, Portola Valley, CA) scripts.

CAHS samples were prepared for SAXS measurements by dissolving 5 mg/mL lyophilized CAHS protein into a buffer containing 20 mM Tris pH 7 and 50 mM NaCl. Samples were incubated at 60°C for 20 minutes to ensure the sample was completely dissolved. Samples were syringe filtered to remove any remaining undissolved protein before injecting 1 mL onto the Superdex 200 column.

SAXS data for CAHS gels were obtained by manually centering capillaries containing premade gels in the X-ray beam. Data was recorded as a series of thirty 0.2 second exposures, but only the first exposure was analyzed to minimize artifacts from X-ray damage. The final analyzed data was corrected for scattering from the empty capillary and a buffer containing capillary. CAHS gel containing samples were made by dissolving 100 mg/mL lyophilized protein in a buffer containing 20 mM Tris pH 7 and 50 mM NaCl. The sample was incubated for 20 minutes at 60°C to ensure the protein was completely dissolved. Double open-ended quartz capillaries with an internal diameter of 1.5 mm (Charles Supper) were used to make the samples. Dissolved protein was directly drawn into the capillary via a syringe. Concentration gradients were generated by layering the protein with buffer. Both ends of the capillary were then sealed with epoxy. Samples were allowed to cool for 5 hours prior to measurement. Data were collected along the concentration gradient by collecting data in 2 mm increments vertically along the capillary.

All data analysis was done using custom MATLAB (Mathworks) scripts. First, an effective concentration was calculated by assuming the maximum concentration was 100 mg/mL and scaling the remaining samples by the integrated intensity of the form factor. It should be noted that the actual concentration could be significantly less than 100 mg/mL in the maximum concentration sample. Data was fit to an equation containing three elements to describe the features apparent in the scattering data. The high-angle form factor was modeled using a Lorentzian-like function to extract the correlation length and an effective fractal dimension.

$$I(q) = \frac{A_1}{1 + (q\xi)^d} \qquad (1)$$

The correlation length is given by ξ and is related to the mesh size inside the fiber bundles seen in SEM images. The fractal dimension, d, is related to the density of the mesh.

No clear correlation length was observed in the smallest angle data, and thus a power law was used to account for this component. The exponent d, is related to the nature of the interface inside and outside of the bundles.

$$I(q) \sim A_2 * q^{-d} \qquad (2)$$

Finally, a Lorentzian peak was used to fit the diffraction peak that is apparent at higher concentrations. The width of the peak, B, appeared constant and was thus fixed so that the amplitude could be accurately extracted.

$$I(q) \sim \frac{A_3}{1 + \left|\frac{q-q_0}{B}\right|^2} \quad (3)$$

In all fit components, A_x is a scale factor.

CD Spectroscopy

Lyophilized protein constructs were weighed and dissolved in a 20 mM Tris-HCl (Fisher Bioreagents, Hampton, NH) buffer at pH 7.0. CD spectra were measured using a JASCO J-1500 CD spectrometer with 0.1 cm quartz cell (Starna Cells, Inc, Atascadero, CA) using a 0.1 nm step size, a bandwidth of 1 nm, and a scan speed of 200 nm/min. Each spectrum was measured 7 times and averaged to increase signal to noise ratio. The buffer control spectrum was subtracted from each protein spectrum. Each protein construct was measured at several concentrations to ensure there is no concentration-dependent change in CD spectra (Fig. S1b, lower panels).

The resulting spectra were analyzed using the lsq_linear function from the SciPy library. To do this, base spectra for α-helix, β-sheet, and random coil spectra (taken from [Perczel, Park, and Fasman 1992] and available in S1b, upper right panel) were linearly fit to match the experimental data set. The three fit coefficients were normalized to give the relative contribution of each base spectrum to the experimental spectrum.

Thioflavin T Assay

Proteins were dissolved in phosphate buffered saline, pH 7.2 (Sigma-Aldrich) was prepared in Dimethyl Sulfoxide (Sigma -Aldrich), and diluted to 20 µM in PBS for use in the assay. 25 µl of each solution were combined into a 96-well plate (Costar, Fisher Scientific, Hampton, NH) and incubated for 15 minutes at room temperature in the dark. Fluorescence was measured using a plate reader (Spark 10M, Tecan, Männedorf, Switzerland) with an excitation at 440 nm, emission was collected at 486 nm. CLC was suspended in 20 mM Tris buffer pH 7.5, the Thioflavin T was diluted in this buffer as well when using CLC.

Lactate Dehydrogenase Protection Assay

LDH desiccation protection assays were performed in triplicate as described previously.[12] Briefly, protectants were resuspended in a concentration range from 20 g/L to 0.1 g/L in 100 µL resuspension buffer (25 mM Tris, pH 7.0). Rabbit Muscle L-LDH (Sigma-Aldrich) was added to this at 0.1 g/L. Half each sample was stored at 4°C, and the other half was desiccated for 17 hours without heating (OFP400, Thermo Fisher Scientific, Waltham, MA). Following desiccation all samples were brought to a volume of 250 µL with water. The enzyme/protectant mixture was added 1:10 to assay buffer (100 mM Sodium Phosphate, 2 mM Sodium Pyruvate [Sigma-Aldrich], 1 mM NADH [Sigma-Aldrich], pH 6). Enzyme kinetics were measured by NAD+ absorbance at 340 nm, on the NanodropOne (Thermo Fisher Scientific). The protective capacity was calculated as a ratio of NAD+ absorbance in desiccated samples normalized to non-desiccated controls.

Citrate Synthase Protection Assay

The Citrate Synthase Kit (Sigma-Aldrich) was adapted for use in this assay.[49,53] All samples were prepared in triplicate, except desiccated negative control samples, which were prepared in

quadruplicate, so that the extra sample could be used for assessment of desiccation efficiency. Concentration of gelatin (Sigma-Aldrich) was determined based on an average mass of 150 kDa. Lyophilized variants were resuspended in either purified water (samples to be desiccated) or 1X assay buffer (control samples) to a concentration of 20 g/L, and diluted as necessary for lower concentrations. Citrate synthase (Sigma-Aldrich) was added at a ratio of 1:10 to resuspended protectants. Non-desiccated control samples were measured as described in the assay kit immediately following resuspension. Desiccated samples were subjected to 5-6 rounds of desiccation and rehydration (1 hour speedvac desiccation [Thermo Fisher Scientific] followed by resuspension in water). Following the 5th round of desiccation, a negative control sample was resuspended and assayed to determine if activity remained. If the negative control sample retained more than 10% activity, a 6th round of desiccation/rehydration was performed. After the final round of desiccation, samples were resuspended in 10 µL of cold 1X assay buffer. Samples were diluted 1:100 in the assay reaction mixture supplied, and all subsequent steps followed the kit instructions. The colorimetric reaction was measured for 90 seconds at 412 nm using the Spark 10M (Tecan).

TD-NMR sample preparation

Quantitated and lyophilized protein samples were transferred as powder into 10 mm TD-NMR tubes (Wilmad Lab Glass) and resuspended in 500 µL of water to a final concentration of 1 g/L, 5 g/L, 10 g/L, 15g/L, and 20 g/L. Samples were left at room temp for 5 minutes to solubilize. If solubilization was not occurring (as determined visually), samples were moved to 55˚C for intervals of 5 minutes until solubilized. If solubilization was not progressing at 55˚C after 10 minutes of heating (as determined visually), then samples were transferred 95˚C for 5 minutes intervals until fully solubilized. Samples were allowed to return to room temperature before being stored at 4˚C. Measurements were taken within 72 hours following solubilization.

Measurement of T_2 relaxation

NMR was performed using a Bruker mq20 minispec low-field nuclear magnetic resonance spectrophotometer, with a resonance frequency of 19.65 MHz. Samples were kept at 25˚C during measurements through the use of a chiller (F12-MA, Julabo USA Inc., Allentown, PA) circulating a constant-temperature coolant. T_2 free induction decays were measured using a Carr-Purcell-Meiboom-Grill (CPMG) pulse sequence with 8000 echoes, and an echo time of 1000 µs. Pulse separation of 1.5 ms, recycle delay of 3 ms, 32 scans was used for all samples. Gain was determined for each sample individually and ranged from 53-56. Conversion of the free induction decay to T_2 relaxation distribution was processed using the CONTIN ILT software provided by Bruker. Each variant was measured for the full concentration range, along with a buffer control (water or 20 mM Tris pH 7) in a single day.

Measurement of Diffusion Coefficients

The diffusion coefficient was determined by the pulsed field gradient spin echo (PFGSE) method using a gradient pulse of 0.5 ms, gradient pulse separation 7.51416 ms, 90°–180° pulse separation 7.51081 ms and 90° first gradient pulse of 1 ms. Each variant was measured for the full concentration range, along with a buffer control (water or 20 mM Tris pH 7) on the same day that T_2 relaxation data was collected.

In this work, diffusion coefficients of water molecules in protein solutions are represented relative to the diffusion coefficient of the water molecules in bulk solution, to emphasize the differences between proteins. This is calculated as: Buffer Diffusion minus Protein Diffusion, and is referred to as Δ Diffusion Coefficient (ΔDC).

Fourier Transform Infrared (FTIR) Measurements

FTIR absorption measurements of CAHS D gels, dehydrated glasses, and CAHS D glasses embedding RC were performed at room temperature using a Jasco Fourier transform 6100 spectrometer, equipped with a DLATGS detector. Spectra in the mid-IR range (7000–1000 cm^{-1}) were acquired using a standard high-intensity ceramic source. To determine the relative content in secondary structure of CAHS D, second and fourth derivative spectra in the amide I' region were calculated numerically and amide I' bands were decomposed into Gaussian sub-bands; the peak wavenumber of Gaussian sub-bands were allowed to be varied over narrow intervals centered at the values indicated by the peaks detected in the second and fourth derivative spectra (see Supplemental text for details).

The kinetics of the amide H/D exchange (HDX) in CAHS D glassy matrices has been followed by recording the onset of the amide II' band around 1450 cm^{-1}, approximately 100 cm^{-1} downshifted compared to the amide II mode in H_2O[71], during incubation with saturated D_2O solutions of LiCl or NaCl, providing a relative humidity of 11% or 75% respectively.

The residual water content of CAHS D-RC and trehalose-RC glassy matrices, equilibrated at RH=11-95%, was determined by FTIR spectroscopy from the area of the (ν_2+ν_3) combination band of water at 5155 cm^{-1}, using the absorption band of the RC at 802 nm as an internal standard.[72] To this end, optical measurements were extended to the NIR region (15000-2200 cm^{-1}) using a halogen lamp source, replacing the Ge/KBr with a Si/CaF2 beam splitter and the KRS-5 with a CaF_2 exit interferometer window. The water content, determined primarily as (H_2O/RC) molar ratio, has been expressed as the ratio (mass water/mass dry matrix). When evaluating the mass of the dry matrix, we included the CAHS protein, the RC and the detergent belt of the RC formed by 289 LDAO molecules[72] plus 14 molecules of free LDAO per RC.

Time resolved optical absorption spectroscopy

The RCs were isolated and purified from the photosynthetic bacterium *Rb. sphaeroides* R-26.[73] The kinetics of charge recombination between the primary photo oxidized electron donor (P^+) and the primary photo reduced acceptor (Q_A^-) were measured by time-resolved optical absorption spectroscopy, by recording the absorbance change at 422 nm following a 200 mJ light pulse (7 ns width) provided by a frequency doubled Nd:YAG laser, essentially as described.[74]

Proteome-wide bioinformatics

The tardigrade proteome (tg.default.maker.proteins.fasta), taken from https://github.com/sujaikumar/tardigrade was used as assembled previously.[75] The proteome file was pre-processed using protfasta (https://protfasta.readthedocs.io/), IDRs predicted with metapredict[76] and IDR kappa values calculated using localCIDER.[31,32] Metapredict identified 35,511 discrete IDRs distributed across 39,532 proteins in the tardigrade proteome.

Dimension estimates for LDH and CS

Radii of gyration for LDH and CS were calculated from the tetramer structure in PDB:1I10 and for the monomeric CS in PDB:3ENJ.[77,78]

All-atom simulations

All-atom simulations were run using the ABSINTH implicit solvent model and CAMPARI Monte Carlo simulation engine (https://campari.sourceforge.net/). ABSINTH simulations were performed with the ion parameters derived by Mao et al.[79] The combination of ABSINTH and CAMPARI has been used to generate experimentally-validated ensembles for a wide range of IDRs [80–82].

Simulations were performed for the full-length CAHS D protein starting from a randomly generated non-overlapping starting state. Monte Carlo simulations update the conformational state of the protein using moves that perturb backbone dihedrals, and sidechain dihedrals, and rigid-body coordinates of all components (including explicit ions).

Ten independent simulations were run for 150,000,000 steps each in a simulation droplet with a radius of 284 Å at 310 K. The combined ensemble generated consists of 69,500 conformations, with ensemble average properties computed across the entire ensemble where reported.

Given the size of the protein, reliability with respect to residue-specific structural propensities is likely to be limited, such that general trends should be taken as a qualitative assessment, as opposed to a quantitative description. Simulations were analyzed using soursop (https://github.com/holehouse-lab/soursop) and MDTraj [83].

Acknowledgements

This work was supported by DARPA award W911NF-19-2-0019 and an Institutional Development Award (IDeA) from NIH grant (P20GM103432) to T.C.B. NIH award R35GM137926 to S.S.; Financial support from MIUR of Italy (RFO2019) is gratefully acknowledged by M.M., F.F. and G.V. A fellowship to S.B. and this work were partially funded by Wyoming NASA EPSCoR, NASA Grant #80NSSC19M0061. Drs. Gary J. Pielak and Samantha Piszkiewicz are acknowledged and thanked for their early discussions and efforts, which were made possible with support from NIH award R01GM127291 to G.J.P. Lorena Rebecchi (University of Modena and Reggio Emilia, UNIMORE, Italy) is thanked for stimulating discussions and valuable advice.

Competing Interest

The authors declare no competing interest.

References

1. Crowe, J. H. & Clegg, J. S. *Anhydrobiosis*. (Hutchinson Ross Publishing Company, 1973).
2. Hesgrove, C. & Boothby, T. C. The biology of tardigrade disordered proteins in extreme stress tolerance. *Cell Commun. Signal.* **18**, 178 (2020).
3. Boothby, T. C. & Pielak, G. J. Intrinsically Disordered Proteins and Desiccation Tolerance: Elucidating Functional and Mechanistic Underpinnings of Anhydrobiosis. *Bioessays* **39**, (2017).
4. Boothby, T. C. Mechanisms and evolution of resistance to environmental extremes in animals. *Evodevo* **10**, 30 (2019).
5. Yathisha, N. S. *et al.* Vegetative desiccation tolerance in : biochemical and physiological responses. *Heliyon* **6**, e04948 (2020).
6. Erkut, C. *et al.* Trehalose renders the dauer larva of Caenorhabditis elegans resistant to extreme desiccation. *Curr. Biol.* **21**, 1331–1336 (2011).
7. Tapia, H., Young, L., Fox, D., Bertozzi, C. R. & Koshland, D. Increasing intracellular trehalose is sufficient to confer desiccation tolerance to Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6122–6127 (2015).
8. Tapia, H. & Koshland, D. E. Trehalose is a versatile and long-lived chaperone for desiccation tolerance. *Curr. Biol.* **24**, 2758–2766 (2014).
9. Mitsumasu, K. *et al.* Enzymatic control of anhydrobiosis-related accumulation of trehalose in the sleeping chironomid, Polypedilum vanderplanki. *FEBS J.* **277**, 4215–4228 (2010).
10. Laskowska, E. & Kuczyńska-Wiśnik, D. New insight into the mechanisms protecting bacteria during desiccation. *Curr. Genet.* **66**, 313–318 (2020).
11. Lapinski, J. & Tunnacliffe, A. Anhydrobiosis without trehalose in bdelloid rotifers. *FEBS Lett.* **553**, 387–390 (2003).
12. Boothby, T. C. *et al.* Tardigrades Use Intrinsically Disordered Proteins to Survive

Desiccation. *Mol. Cell* **65**, 975–984.e5 (2017).

13. Hengherr, S., Heyer, A. G., Köhler, H.-R. & Schill, R. O. Trehalose and anhydrobiosis in tardigrades--evidence for divergence in responses to dehydration. *FEBS J.* **275**, 281–288 (2008).
14. Tripathi, R., Boschetti, C., McGee, B. & Tunnacliffe, A. Trafficking of bdelloid rotifer late embryogenesis abundant proteins. *J. Exp. Biol.* **215**, 2786–2794 (2012).
15. Tripathi, R. *Functional Characterisation of LEA Proteins from Bdelloid Rotifers*. (2012).
16. Denekamp, N. Y., Reinhardt, R., Kube, M. & Lubzens, E. Late embryogenesis abundant (LEA) proteins in nondesiccated, encysted, and diapausing embryos of rotifers. *Biol. Reprod.* **82**, 714–724 (2010).
17. Piszkiewicz, S. *et al.* Protecting activity of desiccated enzymes. *Protein Sci.* **28**, 941–951 (2019).
18. Crowe, J. H., Carpenter, J. F. & Crowe, L. M. The role of vitrification in anhydrobiosis. *Annu. Rev. Physiol.* **60**, 73–103 (1998).
19. Boothby, T. C. Water content influences the vitrified properties of CAHS proteins. *Molecular cell* vol. 81 411–413 (2021).
20. Hengherr, S., Worland, M. R., Reuner, A., Brümmer, F. & Schill, R. O. High-temperature tolerance in anhydrobiotic tardigrades is limited by glass transition. *Physiol. Biochem. Zool.* **82**, 749–755 (2009).
21. Yamaguchi, A. *et al.* Two novel heat-soluble protein families abundantly expressed in an anhydrobiotic tardigrade. *PLoS One* **7**, e44209 (2012).
22. Sakai, T., Katashima, T., Matsushita, T. & Chung, U.-I. Sol-gel transition behavior near critical concentration and connectivity. *Polym. J.* **48**, 629–634 (2016).
23. Almdal, K., Dyre, J., Hvidt, S. & Kramer, O. Towards a phenomenological definition of the term ‘gel’. *Polymer Gels and Networks* vol. 1 5–17 (1993).
24. Dill, K. A., Alonso, D. O. V. & Hutchinson, K. Thermal stabilities of globular proteins.

Biochemistry vol. 28 5439–5449 (1989).

25. Dignon, G. L., Zheng, W., Kim, Y. C. & Mittal, J. Temperature-Controlled Liquid-Liquid Phase Separation of Disordered Proteins. *ACS Cent Sci* **5**, 821–830 (2019).
26. van Dijk, E., Hoogeveen, A. & Abeln, S. The hydrophobic temperature dependence of amino acids directly calculated from protein structures. *PLoS Comput. Biol.* **11**, e1004277 (2015).
27. McComb, D. W., Lengyel, J. & Barry Carter, C. Cryogenic transmission electron microscopy for materials research. *MRS Bulletin* vol. 44 924–928 (2019).
28. Buitink, J. & Leprince, O. Glass formation in plant anhydrobiotes: survival in the dry state. *Cryobiology* **48**, 215–228 (2004).
29. Alacik Develioglu, I., Ozel, B., Sahin, S. & Oztop, M. H. NMR Relaxometry and magnetic resonance imaging as tools to determine the emulsifying characteristics of quince seed powder in emulsions and hydrogels. *Int. J. Biol. Macromol.* **164**, 2051–2061 (2020).
30. Rondeau-Mouro, C. *et al.* Characterization of gluten-free bread crumb baked at atmospheric and reduced pressures using TD-NMR. *Magn. Reson. Chem.* **57**, 649–660 (2019).
31. Das, R. K. & Pappu, R. V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 13392–13397 (2013).
32. Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G. & Pappu, R. V. CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophys. J.* **112**, 16–21 (2017).
33. Williams, A. D. *et al.* Mapping abeta amyloid fibril secondary structure using scanning proline mutagenesis. *J. Mol. Biol.* **335**, 833–842 (2004).
34. Imai, K. & Mitaku, S. Mechanisms of secondary structure breakers in soluble proteins. *Biophysics* **1**, 55–65 (2005).

35. Choi, J.-M., Holehouse, A. S. & Pappu, R. V. Physical Principles Underlying the Complex Biology of Intracellular Phase Transitions. *Annu. Rev. Biophys.* **49**, 107–133 (2020).

36. Li, P. *et al.* Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 (2012).

37. Harmon, T. S., Holehouse, A. S., Rosen, M. K. & Pappu, R. V. Intrinsically disordered linkers determine the interplay between phase separation and gelation in multivalent proteins. *Elife* **6**, (2017).

38. Sanders, D. W. *et al.* Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. *Cell* **181**, 306–324.e28 (2020).

39. Colby, R. H. Structure and linear viscoelasticity of flexible polymer solutions: comparison of polyelectrolyte and neutral polymer solutions. *Rheol. Acta* **49**, 425–442 (2010).

40. Sakumichi, N., Yoshikawa, Y. & Sakai, T. Linear elasticity of polymer gels in terms of negative energy elasticity. *Polym. J.* (2021) doi:10.1038/s41428-021-00529-4.

41. Biancalana, M., Makabe, K., Koide, A. & Koide, S. Molecular mechanism of thioflavin-T binding to the surface of beta-rich peptide self-assemblies. *J. Mol. Biol.* **385**, 1052–1063 (2009).

42. Wu, C., Biancalana, M., Koide, S. & Shea, J.-E. Binding modes of thioflavin-T to the single-layer beta-sheet of the peptide self-assembly mimics. *J. Mol. Biol.* **394**, 627–633 (2009).

43. Ge, X., Yang, Y., Sun, Y., Cao, W. & Ding, F. Islet Amyloid Polypeptide Promotes Amyloid-Beta Aggregation by Binding-Induced Helix-Unfolding of the Amyloidogenic Core. *ACS Chem. Neurosci.* **9**, 967–975 (2018).

44. Namioka, S., Yoshida, N., Konno, H. & Makabe, K. Residue-Specific Binding Mechanisms of Thioflavin T to a Surface of Flat β-Sheets within a Peptide Self-Assembly Mimic. *Biochemistry* **59**, 2782–2787 (2020).

45. Peccati, F. *et al.* Binding of Thioflavin T and Related Probes to Polymorphic Models of Amyloid-β Fibrils. *J. Phys. Chem. B* **121**, 8926–8934 (2017).

46. Kesimer, M., Makhov, A. M., Griffith, J. D., Verdugo, P. & Sheehan, J. K. Unpacking a gel-forming mucin: a view of MUC5B organization after granular release. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **298**, L15–22 (2010).
47. Böni, L. J. *et al.* Hagfish slime exudate stabilization and its effect on slime formation and functionality. *Biol. Open* **6**, 1115–1122 (2017).
48. Rani, R. *Lactate Dehydrogenase (LDH): Biochemistry, Function and Clinical Significance*. (2019).
49. Chakrabortee, S. *et al.* Intrinsically disordered proteins as molecular shields. *Mol. Biosyst.* **8**, 210–219 (2012).
50. Popova, A. V., Rausch, S., Hundertmark, M., Gibon, Y. & Hincha, D. K. The intrinsically disordered protein LEA7 from Arabidopsis thaliana protects the isolated enzyme lactate dehydrogenase and enzymes in a soluble leaf proteome during freezing and drying. *Biochim. Biophys. Acta* **1854**, 1517–1525 (2015).
51. Hatanaka, R. *et al.* An abundant LEA protein in the anhydrobiotic midge, PvLEA4, acts as a molecular shield by limiting growth of aggregating protein particles. *Insect Biochem. Mol. Biol.* **43**, 1055–1067 (2013).
52. Furuki, T. *et al.* Group 3 LEA Protein Model Peptides Suppress Heat-Induced Lysozyme Aggregation. Elucidation of the Underlying Mechanism Using Coarse-Grained Molecular Simulations. *J. Phys. Chem. B* **124**, 2747–2759 (2020).
53. Goyal, K., Walton, L. J. & Tunnacliffe, A. LEA proteins prevent protein aggregation due to water stress. *Biochem. J* **388**, 151–157 (2005).
54. Ikeda, K. *et al.* Presence of intrinsically disordered proteins can inhibit the nucleation phase of amyloid fibril formation of Aβ(1–42) in amino acid sequence independent manner. *Scientific Reports* vol. 10 (2020).
55. Belton, P. S. & Gil, A. M. IR and Raman spectroscopic studies of the interaction of trehalose with hen egg white lysozyme. *Biopolymers* **34**, 957–961 (1994).

56. Käiväräinen, A. I., Sukhanova, G. & Goryunov, A. S. Changes in water properties in serum albumin solutions induced by alterations in protein flexibility. NMR studies. *Folia Biol.* **30**, 84–92 (1984).

57. Ghi, P. Y., Hill, D. J. T. & Whittaker, A. K. (1)H NMR study of the states of water in equilibrium poly(HEMA-co-THFMA) hydrogels. *Biomacromolecules* **3**, 991–997 (2002).

58. Qiao, S., Tian, Y., Song, P., He, K. & Song, S. Analysis and detection of decayed blueberry by low field nuclear magnetic resonance and imaging. *Postharvest Biol. Technol.* **156**, 110951 (2019).

59. Zheng, J.-M., Chin, W.-C., Khijniak, E., Khijniak, E., Jr & Pollack, G. H. Surfaces and interfacial water: evidence that hydrophilic surfaces have long-range impact. *Adv. Colloid Interface Sci.* **127**, 19–27 (2006).

60. Wang, H. *et al.* A fast and non-destructive LF-NMR and MRI method to discriminate adulterated shrimp. *J. Food Meas. Charact.* **12**, 1340–1349 (2018).

61. Hydration properties and proton exchange in aqueous sugar solutions studied by time domain nuclear magnetic resonance. *Food Chem.* **132**, 1644–1650 (2012).

62. Adhikari, A., Park, W.-W. & Kwon, O.-H. Hydrogen-Bond Dynamics and Energetics of Biological Water. *Chempluschem* **85**, 2657–2665 (2020).

63. Raschke, T. M. Water structure and interactions with protein surfaces. *Curr. Opin. Struct. Biol.* **16**, 152–159 (2006).

64. Pattni, V., Vasilevskaya, T., Thiel, W. & Heyden, M. Distinct Protein Hydration Water Species Defined by Spatially Resolved Spectra of Intermolecular Vibrations. *J. Phys. Chem. B* **121**, 7431–7442 (2017).

65. Lerbret, A. *et al.* How strongly does trehalose interact with lysozyme in the solid state? Insights from molecular dynamics simulation and inelastic neutron scattering. *J. Phys. Chem. B* **116**, 11103–11116 (2012).

66. Ahmed, M., Namboodiri, V., Singh, A. K. & Mondal, J. A. On the intermolecular vibrational

coupling, hydrogen bonding, and librational freedom of water in the hydration shell of mono- and bivalent anions. *J. Chem. Phys.* **141**, 164708 (2014).

67. Laage, D., Elsaesser, T. & Hynes, J. T. Water Dynamics in the Hydration Shells of Biomolecules. *Chem. Rev.* **117**, 10694–10725 (2017).
68. Shimizu, S. & Smith, D. J. Preferential hydration and the exclusion of cosolvents from protein surfaces. *J. Chem. Phys.* **121**, 1148–1154 (2004).
69. Klaips, C. L., Jayaraj, G. G. & Ulrich Hartl, F. Pathways of cellular proteostasis in aging and disease. *Journal of Cell Biology* vol. 217 51–63 (2018).
70. Hopkins, J. B., Gillilan, R. E. & Skou, S. : improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J. Appl. Crystallogr.* **50**, 1545–1553 (2017).
71. Barth, A. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta* **1767**, 1073–1101 (2007).
72. Malferrari, M., Francia, F. & Venturoli, G. Coupling between electron transfer and protein-solvent dynamics: FTIR and laser-flash spectroscopy studies in photosynthetic reaction center films at different hydration levels. *J. Phys. Chem. B* **115**, 14732–14750 (2011).
73. Gray, K. A., Farchaus, J. W., Wachtveitl, J., Breton, J. & Oesterhelt, D. Initial characterization of site-directed mutants of tyrosine M210 in the reaction centre of Rhodobacter sphaeroides. *EMBO J.* **9**, 2061–2070 (1990).
74. Francia, F. *et al.* Protein-matrix coupling/uncoupling in ‘dry’ systems of photosynthetic reaction center embedded in trehalose/sucrose: the origin of trehalose peculiarity. *J. Am. Chem. Soc.* **130**, 10240–10246 (2008).
75. Boothby, T. C. *et al.* Evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 15976–15981 (2015).
76. Emenecker, R. J., Griffith, D. & Holehouse, A. S. Metapredict: A fast, accurate, and easy-to-use predictor of consensus disorder and structure. *bioRxiv* (2021)

doi:10.1101/2021.05.30.446349.

77. Read, J. A., Winter, V. J., Eszes, C. M., Sessions, R. B. & Brady, R. L. Structural basis for altered activity of M- and H-isozyme forms of human lactate dehydrogenase. *Proteins* **43**, 175–185 (2001).

78. Larson, S. B., Day, J. S., Nguyen, C., Cudney, R. & McPherson, A. Structure of pig heart citrate synthase at 1.78 A resolution. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **65**, 430–434 (2009).

79. Mao, A. H. & Pappu, R. V. Crystal lattice properties fully determine short-range interaction parameters for alkali and halide ions. *J. Chem. Phys.* **137**, 064104 (2012).

80. Metskas, L. A. & Rhoades, E. Conformation and Dynamics of the Troponin I C-Terminal Domain: Combining Single-Molecule and Computational Approaches for a Disordered Protein Region. *J. Am. Chem. Soc.* **137**, 11962–11969 (2015).

81. Martin, E. W. *et al.* Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).

82. Martin, E. W. *et al.* Sequence Determinants of the Conformational Properties of an Intrinsically Disordered Protein Prior to and upon Multisite Phosphorylation. *J. Am. Chem. Soc.* **138**, 15323–15335 (2016).

83. McGibbon, R. T. *et al.* MDTraj: a modern, open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **109**, 1528–1532 (2015).

Supplemental Information

Mov. S1: Monte Carlo simulation of CAHS D. Movie showing the simulated ensemble structures of wildtype CAHS D.

File. S1: File containing the protein and nucleotide sequence for each protein variant used in this study.

Supplemental Text

Theory and application of TD-NMR Relaxometry for assessing water coordination

NMR measures the relaxation of nuclear spins at certain frequencies, which relate to the particular frequency of the molecular motion being measured (Horii, NMR Relaxation and Dynamics, 1998). Frequencies in the 10^3 -10^6 range are commonly used for ^{13}C measurements, while higher frequency motions, in the 10^9 -10^{12} range, can be used to measure motion of ^1H nuclei in water molecules. Time-Domain NMR (TD-NMR) uses a low field magnet to measure frequencies in the 10^9 -10^{12} range, enabling measurements of water dynamics. Spin-lattice relaxation (T_1), measures longitudinal spin of ^1H nuclei relative to the magnetic field, which is dependent on how water molecules interact with the magnetic field. Spin-spin relaxation (T_2) is the measurement of ^1H spins in a plane transverse to the magnetic field. This is a reflection of how water molecules interacting with other molecules influences their properties.[1,2] Both measures are useful for different applications, with T_2 being particularly useful for applications measuring slower water dynamics and water caging in porous structures.[3,4] To measure T_2, a sample is subjected to a magnetic pulse sequence[5,6] which induces the hydrogen nuclei of the sample to align their spins with the magnetic field. Once the magnetic pulse ends, the nuclei spins return to their unaligned, stochastic equilibrium state. The rate at which a sample's nuclei return to equilibrium in the transverse plane is measured as a rate of decay, called the Free Induction Decay (FID) for T_2, or the spin-spin FID. The FID is subjected to an inverse LaPlace transform, which can then be expressed as a frequency distribution of spin-spin decay times (Fig. S5a), called T_2 relaxation.[2,4]

T_2 relaxation is dependent on the state of matter of the sample being measured, because state of matter impacts the ability of molecules to interact with one another. Liquid molecules are free to interact with each other, and this interaction serves to hold the spins aligned for a longer period of time than in solids, where molecules cannot interact (Fig. S5a). To better understand the impact of state of matter on spin decay, consider the analogy of a sample's hydrogen nuclei as magnets; two magnets in close proximity will exert force on each other to maintain alignment, whereas two magnets that are held apart from one another cannot influence their respective orientations. Nuclei in the solid state are essentially independent, because of the distance between them and the rigidity of the solid state, and therefore their decay to equilibrium is rapid. Nuclei in the liquid state, on the other hand, can interact freely with other molecules in solution, and therefore their spin decay is slower.

The principles that govern decay rates with respect to a sample's state of matter can also be applied to the ability of a protein to interact with water. For a theoretical protein that is incapable of interacting with water, water would be blind this protein and thus would behave as if it were a pure liquid, with a similar T_2 relaxation profile. A highly hydrophobic protein would exclude water molecules from itself, increasing the local density of water molecules and thus increasing their interactions, which would further slow T_2 relaxation relative to pure water. Hydrophilic proteins coordinate water molecules through their dielectric field, which interacts with the dipole of water molecules, causing an ordering of water molecules into hydration shell(s) surrounding the protein. Being organized into hydration shells decreases the freedom of motion and interactions between different molecules of water, which would speed up T_2 relaxation. Thus, T_2 relaxation can inform us about the ability of a protein to coordinate or exclude water.

Characterization of tardigrade CAHS D by FTIR spectroscopy

Analysis of the amide I' band in CAHS gels and glasses at different contents of residual water

The amide I band centered around 1650 cm^{-1} is sensitive to the structure of the protein backbone; as a result a particular secondary structure absorbs predominantly in a specific range of the amide I region.[7] However the various secondary-structure components overlap leading to a rather broad and scarcely structured band and band-narrowing procedures are necessary to resolve the component bands.[8] The linewidth of the second and fourth derivative of a band is smaller than that of the original band, so that the minima and maxima of the second and fourth derivative spectra, respectively, can considerably help in evaluating the number and the positions of the overlapping spectral components.[9] On this basis, in order to gain information on the secondary structure of the protein, the amide I band is fitted to the sum of the component bands, as inferred by the derivative analysis. The resulting component bands are assigned to the specific secondary structures, and the fractional area of each component band is taken as the relative content of the corresponding secondary structures. We have used this approach with the aim of assessing CAHS secondary structure and its possible changes upon vitrification and dehydration.

Methods

All samples were prepared by dissolving the CAHS protein in D_2O to avoid the overlapping of the spectral contribution due to the H_2O bending mode (around 1640 cm^{-1})[10] with amide I band of the CAHS protein centered at approximately 1650 cm^{-1}. Deuteration causes a blue shift of the water bending mode by more than 400 cm^{-1}, while inducing only a small shift of the amide I band (amide I' band). Gel samples were obtained by dissolving the lyophilized CAHS protein at a concentration of 16 g/L in D_2O heated at about 50 °C and gently stirred for 5 minutes. A volume of 10 µL of the CAHS solution was deposited between two CaF_2 windows, separated by a 50 µm teflon spacer, and mounted into a Jasco MagCELL sample holder. When cooled at room temperature (25°C) CAHS solution forms a gel.

Glassy samples were obtained by depositing a volume of 38 µL of the heated CAHS solution (16 g/L) in D_2O on a 50 mm diameter CaF_2 window. Before gelling at room temperature, the drop spread out to form a layer of approximately 1 cm diameter. The sample was dried under N_2 flow for 5 minutes and subsequently the optical window on which the glass had formed was inserted into a specifically designed sample holder equipped with a second CaF_2 window to form a gas-tight cylindrical cavity.[11] Different hydration levels of the CAHS glass were obtained by an

isopiestic method, i.e. by equilibrating the sample with saturated salt solutions providing defined values of relative humidity, contained at the bottom of the gas-tight sample holder cavity. The following saturated solutions in D_2O were employed to obtain the desired relative humidity[12] at 297 K; KNO_3 (RH=95%), NaCl (RH=75%), K_2CO_3 (RH=43%), and LiCl (RH=11%).

FTIR absorption measurements were performed at room temperature with a Jasco Fourier transform 6100 spectrometer equipped with a DLATGS detector. The spectra were acquired with a resolution of 2 cm^{-1} in the whole mid-IR range (7,000 - 1,000 cm^{-1}) using a standard high-intensity ceramic source and a Ge/KBr beam splitter. All spectra were obtained by averaging 10^3 interferograms.

Spectral analysis of the amide I bands recorded in gel samples was performed after subtracting an appropriately normalized D_2O spectrum. In the case of glassy samples, equilibrated at all the relative humidity values tested, this procedure did not result in significant corrections in the amide band region, due to the much lower content of residual D_2O. When extracting the amide I' band from the spectrum, the background was approximated by a straight baseline drawn between the minima on either side of the band. Second and fourth derivative spectra in the amide I' region were calculated using the *i-signal* program (version 2.72) included in the SPECTRUM suite (http://terpconnect.umd.edu/~toh/spectrum/iSignal.html) written in MATLAB language. A Savitsky-Golay algorithm was employed to smooth the signals and calculate the derivatives. The smooth width was chosen by evaluating step-by-step the impact of increasing the smoothing on the calculation of the derivative spectrum, with the aim of optimizing the signal to noise ratio without losing spectral information. For all the amide I' bands analyzed, in gel and glassy samples, both the second and fourth derivative spectra (not shown) were consistently dominated by three minima and maxima, respectively, suggesting the presence of three spectral components. This is supported by the finding that the positions of each peak detected in the second and fourth derivative spectra for each amide I' band analyzed were very close (in most cases coincident within our spectral resolution). Over the whole data set (glassy matrices at different hydration and hydrated gels), the peak wavenumbers inferred from the second and fourth derivative analysis belonged to the intervals 1619-1628 cm^{-1}, 1644-1648 cm^{-1}, and 1676-1688 cm^{-1}. On this basis, amide I' bands were decomposed into three Gaussian components by using a locally developed least-squares minimization routine,[13] using a modified grid search algorithm.[14] Confidence intervals for the best-fit parameters were evaluated numerically, as previously detailed.[15] In the fitting procedure, the peak wavenumbers of the three Gaussian components were allowed to be varied over the intervals reported above, while the areas and the widths were treated as free, unconstrained parameters.

Results

Fig. S3c shows a series of spectra recorded in a CAHS glassy matrix, which has been initially dehydrated by equilibration at a relative humidity RH=11%, and subsequently rehydrated progressively by equilibration at the indicated relative humidity. At increasing relative humidity, the OD stretching band of D_2O, centered around 2500 cm^{-1},[10] increased in amplitude, indicating the increasing hydration of the CAHS matrix. During incubation at each relative humidity, spectra were acquired in sequence, and the increase of the band at 2500 cm^{-1} was monitored until a steady amplitude was reached, demonstrating the attainment of equilibrium at that specific relative humidity.

Fig. S3c shows the amide spectral region: the bands centered at about 1440 cm^{-1} and 1640 cm^{-1} are ascribed to the amide II' and amide I' modes of the protein, respectively.[7] The peak at 1575-1580 cm^{-1} is attributed to the contribution of amino acid residue side chains, and specifically to the carboxylate antisymmetric stretch mode $n_{as}(COO^-)$ from Asp and Glu and possibly to the Arg mode $n_s(CN_3H_5^+)$.[7,16] The absence of an amide II band around 1550 cm^{-1} indicates that essentially all amide protons (NH) have been exchanged for deuterium, as expected for an unordered protein structure.[17] Changes in the hydration level of the CAHS glassy matrix cause slight alterations in the amide spectral region. However, as described below, a detailed analysis of the amide I' band reveals subtle but systematic changes in the relative contribution of the secondary structure motives of the protein.

Fig. S3d top and bottom panels show the result of decomposition into Gaussian sub-bands, performed as described above, of the amide I' band recorded in the glassy CAHS matrix equilibrated at a relative humidity RH=11% (Fig. S3d top) and RH=95% (Fig. S3d bottom).

Both under strongly dehydrated (RH=11%) and hydrated (RH=95%) conditions the amide I' band is predominantly accounted for by a broad Gaussian band centered at 1644 cm^{-1}, indicating the presence of a dominating unstructured region in the protein.[7] The two Gaussian bands at 1621 - 1623 cm^{-1} and 1688 cm^{-1} can be assigned to intermolecular β-sheet structures.[8,18] Interestingly, following hydration of the glass, from RH=11% to RH=95%, the relative contribution of the sub-band centered around 1624 cm^{-1} increases significantly, at the expenses of the central band at approximately 1644 cm^{-1}. This behavior, as a function of the hydration level of the glassy matrix, appears to be systematic, as shown by Fig.3b, which presents the fractional area of the three Gaussian sub-bands obtained for the CAHS glass equilibrated at increasing values of relative humidity and for the gel sample.

The analysis summarized in Fig. 3b indicates that the CAHS protein is mostly (90%) unordered in the fully hydrated gel state, and largely retains this unstructured character upon drying. This behavior is at variance with what observed for other intrinsically disordered proteins, like Late Embryogenesis Abundant (LEA) proteins, which, although largely disordered under fully hydrated conditions, acquired α-helical structure upon drying.[19,20] Fig. 3b shows that the hydration state of the glass has a modest, but significant systematic influence on the secondary structure. Under the most dehydrated conditions (RH=11%) the unordered region is maximal, accounting for ~93% of the fractional area (FA), while interprotein β-sheet structures comprise only 7% of the FA. When the glass is progressively hydrated, the unstructured portion is coincidentally reduced (down to 77% FA at RH=95%) while interprotein β-contacts increase (up to 23% total β-sheet structures at RH=95%). Apparently, further hydration (RH=100%) slightly reverses this trend, and in the fully hydrated gel state the unstructured region again dominates (~90% of the FA). This suggests that the content of interprotein β-sheet structures is finely tuned in the glass by its hydration state, being favored in the strongly hydrated glassy matrix, while in the extensively dried glass and in the fully hydrated gel the unordered structures dominate.

Amide hydrogen/deuterium exchange kinetics in a CAHS glassy matrix at different hydration levels

The kinetics of amide (peptide NH) proton exchange can in principle provide information on the flexibility of proteins. The retardation or inhibition of hydrogen/deuterium exchange can in fact be brought back to the involvement of the amide hydrogen in H bonding and/or to its limited

accessibility to the solvent. It has been shown that a hydrogen bonded amide proton is protected from exchange even when it is exposed to solvent, slowing down exchange rates by more than six orders of magnitude.[7] On the other hand solvent exclusion can retard the exchange by a similar extent. Specific amide groups can be accessible to the solvent only in conformational substates that are rarely populated in the time average structure. Hydrogen exchange can therefore be extremely sensitive to the protein conformational dynamics. Starting from this rationale we studied the extent and the kinetics of hydrogen/deuterium exchange in CAHS glassy matrices by using an isopiestic approach, at different hydration levels.

The amide H/D exchange can be followed by using the amide II band of proteins, a mode which is essentially the combination of the NH in-plane bending and the CN stretching vibration. Following H/D exchange the N–D bending no longer couples with the CN stretching vibration converting the mode to a largely CN stretching vibration around 1450 cm^{-1}.[7] The wavenumber of this mode (the so-called amide II' band) is approximately 100 cm^{-1} downshifted compared to the amide II mode detectable in H_2O (around 1550 cm^{-1}). Thus the two bands are clearly separated in the spectrum.

In order to examine the extent and kinetics of the amide H/D exchange we prepared two CAHS glassy matrices as described in the previous subsection , except that the lyophilized CAHS protein was initially solubilized in H_2O rather than D_2O. The glasses were extensively dried by equilibration within gas-tight sample holders at a relative humidity RH=6% in the presence of $NaOH{\cdot}H_2O$. The attainment of a minimum, steady hydration level was checked by monitoring during 3 days the (v_2+v_3) combination band of water around 5150 cm^{-1},[11] using the amide I band as an internal standard. Subsequently, in one of the sample holders, the saturated $NaOH{\cdot}H_2O$ was replaced by a saturated solution of LiCl in D_2O, starting the H/D exchange at a relative humidity RH=11%.

A series of FTIR spectra was recorded in sequence, at selected time intervals from the time of replacement. For each spectrum 100 interferograms were averaged to allow a sufficiently rapid acquisition (3 minutes) particularly at the beginning of the H/D exchange. In a parallel experiment, within the second sample holder, the saturated $NaOH{\cdot}H_2O$ was replaced by a saturated solution of NaCl in D_2O, in order to follow the H/D exchange at a relative humidity RH=75%.

Fig. S3e shows a series of spectra in the amide region (1350 – 1750 cm^{-1}) acquired in the latter experiment at selected time intervals from the time (t=0) when H/D exchange started, following incubation in the presence of the saturated NaCl solution in D_2O. We have previously shown,[21,22] by using a film formed upon dehydration by a model membrane protein (the bacterial photosynthetic reaction center complex), that such an isopiestic approach for isotopic exchange is quite effective and rapid (hour time scale).

As expected for a progressive amide H/D exchange, an intense amide II' band appears on the minute time scale, centered around 1450 cm^{-1}. The progressive appearance and increase of the amide II' band is accompanied by a progressive absorbance decrease in the spectral region of the amide II band at 1550 cm^{-1}. Such a decrease interferes however with the appearance of a band peaking around 1575 cm^{-1}, which is clearly detected in the fully deuterated glass (see Fig. S3c), attributed to the contribution of the amino acid residue side chains. This band has been assigned to an overlap of $v_{as}(COO^-)$ from Asp and Glu and of Arg $v_s(CN_3H_5^+)$.[7,16] We evaluated the extent of amide H/D exchange from the area of the amide II' band at 1450 cm^{-1} after

subtraction of a straight baseline drawn between the minima on either side of the band.[23] The contribution of the small background band present at t=0 was also subtracted, and the resulting area was normalized to the area of the amide I/I' band.

The dependence of the obtained amide II' areas upon the deuteration time is shown in Fig. 1f, which collects also the results obtained by using the same approach in the other CAHS glass, in which the amide H/D exchange has been studied during equilibration with a saturated LiCl solution in D_2O at a relative humidity RH=11%. At both relative humidities, the kinetics appear biphasic, including a major phase (in the minute and tens of minutes time scale, at RH=75% and RH=11% respectively) (see Fig. 1f) and a much slower kinetic component (in the time scale of days).

As indicated in Fig. 1f , when an essentially steady amide deuteration level was reached, in both samples the saturated salt solutions in D_2O (LiCl and NaCl, respectively) were replaced by pure D_2O (RH=100%) and the additional increase in the area of the amide II' band was monitored. Under this condition, a full amide deuteration is attained, as shown by comparison with the normalized amide II' area measured in a CAHS glass, prepared by solubilizing the lyophilized protein in D_2O at about 50°C (dashed line in Fig. 1f). At variance, when the CAHS glass is rehydrated in the presence of D_2O vapor at RH=11%, hydrogen/deuterium exchange occurred only partially, in about 37% of the total amide population, even after 20 days of incubation. This percentage of NH deuteration increases to about 83% when H/D exchange takes place in a D_2O atmosphere at RH=75%.

As mentioned above, an incomplete hydrogen/deuterium exchange could be due to the involvement of the amide proton in H bonding. However, analysis of the amide I' band (see Fig. 3b) indicates that in the CAHS glass equilibrated at RH=11%, the relative content of interprotein β-sheet secondary structures, in which the amide acts as a H bond donor, accounts for <10% of the total secondary structure, while under the same conditions, >60% of the amide groups does not undergo H/D exchange. This suggests that the limited exchange observed in the CAHS glass at RH=11% is likely due to the solvent exclusion from a substantial fraction of the CAHS amide groups in the matrix. Such an exclusion is expected to result not so much from the limited amount of residual water in the glass, but rather from the tight packing of the CAHS protein within the matrix, and by the strong inhibition of conformational fluctuations of the protein which blocks the sampling of conformational substates characterized by a potential accessibility of peptide NH groups to the solvent. This view is consistent with the observation that, when the CAHS matrix is incubated in the presence of D_2O at RH=76%, the fraction of deuterated amide groups is more than doubled. The larger hydration of the matrix under these conditions (see the stretching D_2O band in Fig. S3c) will act as a plasticizer of the protein, activating CAHS conformational fluctuations, thus increasing the accessibility to the solvent of peptide NH groups. In conclusion, the observed hydrogen/deuterium exchange behavior strongly suggests that dehydration of the CAHS matrix leads to a tightly packed, although disordered, architecture, in which conformational fluctuations of the protein are drastically inhibited.

Immobilization of biological materials in dried CAHS D gels

Using differential scanning calorimetry it has been found that CAHS protein gels vitrify when dried.[24] A number of experimental evidences obtained *in vivo* and *in vitro* indicated that the vitrified state of CAHS proteins correlates with their protective capabilities.[24] On this basis, a

study of the gelation and vitrification properties of CAHS proteins, has led to suggest that non-desiccation tolerant proteins can be encapsulated within the pores of CAHS gels, which, upon dehydration, vitrify and, by locking the internal dynamics of the embedded proteins, prevent their unfolding and thermal denaturation. The protective action of CAHS glasses would be therefore brought back to their capability of reducing the internal degrees of freedom of the embedded proteins, as shown for glassy matrices formed upon dehydration by the disaccharide trehalose.[25] In order to test such a suggestion we have used as a model protein the bacterial photosynthetic reaction center, which has been incorporated into amorphous matrices formed by CAHS proteins upon dehydration. We aimed at investigating the capability of CAHS glassy matrices to inhibit the conformational dynamics of the reaction center associated with photoinduced electron transfer processes (see below).

Inhibition of the conformational dynamics of bacterial reaction centers as probed by the kinetics of charge recombination

The photosynthetic reaction center (RC) from the purple bacterium *Rhodobacter* (*Rb.*) *sphaeroides* represents an ideal model system to probe matrix effects on conformational protein dynamics. This membrane-spanning pigment-protein complex catalyzes the primary photochemical events of bacterial photosynthesis. Within the RC, following absorption of a photon, the primary electron donor P, a bacteriochlorophyll dimer situated near the periplasmic side of the protein, delivers an electron (via an interposed bacteriopheophytin molecule) to the primary quinone acceptor, Q_A, located 25 Å away from P and closer to the cytoplasmic side of the RC. This electron transfer process, occurring in about 200 ps, generates the primary charge separated state, $P^+Q_A^-$. In the absence of the secondary quinone acceptor bound at the Q_B site (or in the presence of inhibitors which block electron transfer from Q_A^- to Q_B), the electron on Q_A^- recombines with the hole on P^+ by direct electron tunneling.[26]

The kinetics of $P^+Q_A^-$ recombination after a short (nanosecond) flash of light provides an endogenous probe of the RC conformational dynamics. In fact, in room temperature solutions, following light-induced $P^+Q_A^-$ charge separation, the RC protein undergoes a dielectric relaxation from a dark-adapted to a light-adapted conformation, which stabilizes thermodynamically the $P^+Q_A^-$ state (lifetime for charge recombination of about 100 ms). When this conformational relaxation is inhibited (by freezing the RC to cryogenic temperature in the dark,[27] or by incorporating it into dehydrated trehalose glasses[28]) the recombination kinetics are accelerated (lifetime of about 20 ms) and become strongly distributed, mirroring the immobilization of the protein over a large ensemble of conformational substates. Under these conditions, we describe therefore the charge recombination kinetics by using a continuous distribution *p(k)* of rate constants *k*:[28]

$$N(t) = \int_0^\infty p(k) exp(-kt)\, dk = \left(1 + \frac{\sigma^2}{<k>} t\right)^{-<k>^2/\sigma^2} \qquad (1)$$

where *N(t)* is the survival probability of the $P^+Q_A^-$ state at time *t* after the photo-activating light pulse, <k> is the average rate constant, and σ is the width of the rate distribution. The latter two parameters provide quantitative information on the conformational dynamics of the RC on the time scale of milliseconds. An increase in either or both parameters (<k>, σ) reflects a retardation

of the RC relaxation from the dark- to the light-adapted conformation (<k>), or of the fluctuation between conformational substates (σ).[13,28]

Methods

A volume of 78 µL of RCs purified from *Rb. sphaeroides* R26, suspended at a 76 µM concentration in 10 mM TRIS buffer, 0.025% LDAO, pH 8.0, was mixed with 64 µL of a 16 g/L CAHS protein solution in water and 8 µL of 200 mM o-phenanthroline in ethanol. This inhibitor blocks Q_A^- to Q_B electron transfer, thus allowing the recombination kinetics of the $P^+Q_A^-$ state to be monitored. The lyophilized CAHS protein dissolved in water, and heated at about 50°C for a few minutes, was allowed to cool at room temperature and, before gelifying, it was rapidly mixed with the RC suspension. The obtained mixture (volume 150 µL) was immediately layered on a 50 mm diameter CaF_2 optical window and dried under N_2 flow for 5 minutes. The sample was subsequently inserted into the gas-tight holder and equilibrated at a relative humidity r=11% in the presence of LiCl. The glassy matrix is characterized by a (CAHS/RC) molar ratio of approximately 6.6, corresponding to a mass ratio of about 1.7. Such a mass ratio was chosen to allow a comparison with results obtained previously in glassy trehalose matrices incorporating the RC. A (trehalose/RC) mass ratio of 1.7 corresponds to a ratio of 500 trehalose molecules per RC complex, which has been shown to strongly inhibit the RC conformational dynamics.[13]

The kinetics of $P^+Q_A^-$ recombination was measured by time-resolved spectrophotometry , by recording the absorbance change at 422 nm after photoexcitation by a laser pulse (7 ns width) at 532 nm , essentially as described.[29] We have previously shown that in RC embedded into trehalose glasses $P^+Q_A^-$ recombination kinetics accelerates progressively upon progressive dehydration of the matrix below a threshold value of residual water content, demonstrating a progressive inhibition of the RC conformational relaxation for the dark- to the light-adapted conformation. In order to test the capability of the CAHS glassy matrix to similarly inhibit RC dynamics, we have measured $P^+Q_A^-$ recombination kinetics in CAHS-RC glasses equilibrated at different values of relative humidity by using the isopiestic method described above . At each relative humidity the residual water content of the CAHS-RC matrices was determined from the area of the ($\nu_2+\nu_3$) combination band of water at 5155 cm^{-1}, using the absorption band of the RC at 802 as an internal standard (Malferrari et al., 2011; Malferrari et al., 2015).[11,13] The water content, determined as (H_2O/RC) molar ratio, has been subsequently expressed as the ratio (mass water/mass dry matrix) as described in Methods. Since we do not know how the residual water is distributed between the different components which form the matrix, this evaluation of the residual water content allows a physically meaningful comparison of the overall average hydration of the matrix between samples differing in their chemical composition.

Fig. S1d, upper panel, shows the hydration isotherm obtained by using this approach in the CAHS-RC glassy matrices described above. For the sake of comparison the ratio (mass water/mass dry matrix) has been similarly determined in RC-trehalose matrices, characterized by a (trehalose/RC) mass ratio of 1.7, i.e. equal to the (CAHS/RC) mass ratio of the CAHS-RC glasses.

The isotherms indicate that the overall propensity for water adsorption is very similar in the disaccharide-RC and in the CAHS-RC protein matrix when the relative humidity is varied over a large range. Sorption data of the CAHS-RC matrix well fit the Hailwood and Horrobin equation.[30]

$$h = h_0 \left(\frac{K_1 r}{1+K_1 r} + \frac{K_2 r}{1-K_2 r} \right) \qquad (2)$$

where h represents the equilibrium water content of the matrix, h_0 and K_1 are constants, proportional to the number and activity of the hydration sites, respectively, and K_2 is related to the water activity of the water condensing at the surface of the absorbing components. Best fitting to eq.2 yields h_0=5.59 $\cdot 10^{-2}$ g water/g dry matrix, K_1=3.44 for the strong binding sites, and K_2=1.04 for the weak condensing sites.

Results

Fig. S1c shows the kinetics of $P^+Q_A^-$ recombination after a laser pulse (fired at t=0) recorded in RCs incorporated into glassy matrices formed by the CAHS protein, at different hydration levels obtained by equilibration at different values of relative humidity.

Incorporation of the RC into an extensively dehydrated (RH=11%) CAHS glassy matrix leads to a substantial acceleration of the decay kinetics (<k> = 35.5 s^{-1}) and to a dramatic broadening of the rate constant distribution (σ = 23.8 s^{-1}) as compared to solution (<k>=8.8 s^{-1}; σ=3.5 s^{-1}). Upon progressive rehydration of the CAHS-RC matrix both these effects revert progressively. At increasing values of relative humidity, between RH=23%-53%, the decay kinetics become progressively slower and lose the strongly distributed, non-exponential character displayed under the most dehydrated conditions. This behavior clearly indicates that the hydration state of the CAHS matrix modulates the conformational dynamics of the embedded RC, resulting, upon extensive dehydration, in a severe RC immobilization on the time scale of charge recombination.

The results of a systematic analysis of $P^+Q_A^-$ recombination is presented in Fig. S1d (middle and lower panels) which shows the dependence of the average rate constant, <k>, (Fig. S1d middle panel) and of the rate constant distribution width, σ, (Fig. S1d, lower panel) upon the residual water content of the CAHS-RC amorphous matrices. For the sake of comparison Fig. S1d also includes the results of a similar analysis performed in amorphous RC films formed upon dehydration in the absence of the CAHS protein, and data obtained in trehalose-RC glassy matrices characterized by (trehalose/RC) molar ratios equal to 500 and 5000. In the trehalose matrices we find similar dependencies of <k> and σ on hydration, indicating the same efficacy of trehalose glasses in inhibiting the RC conformational dynamics over this range of (trehalose/RC) molar ratios, in agreement with previous results.[31] Also in the absence of any excipient, i.e. in RC films, dehydration below a threshold of about 0.1 g H_2O per g of dry matrix leads to a drastic retardation of the RC relaxation from the dark- to the light-adapted conformation, as well as of the interconversion between conformational substates. Such a behavior, already observed in dehydrated films formed by RCs purified from *Rb. sphaeroides* strain 2.4.1,[11,31] has been interpreted as due to the slaving of the RC relaxation and conformational fluctuations to the solvent fluctuations of the hydration shell which surrounds the protein complex, which is progressively broken down and depleted upon dehydration of the RC film (for a detailed discussion see Malferrari et al., 2011 and Malferrari et al., 2015). Interestingly, a strong inhibition of the RC conformational dynamics, characterized by <k> > 20 s^{-1} and σ > 15 s^{-1}, is found in RC films only at very low contents of residual water, systematically lower than those leading to comparable inhibition in trehalose-RC glassy matrices. The larger conformational freedom

observed in the absence of trehalose can be brought back to the difference in the strength and extension of the hydrogen bond networks which characterize the amorphous RC film and the trehalose-RC matrices, leading to a more severe conformational locking of the RC dynamics in trehalose-RC glasses. This conclusion[13] is in line with results of molecular dynamic simulations and inelastic neutron scattering measurements of lysozyme-trehalose glasses at low hydration.[32]

The kinetic analysis of $P^+Q_A^-$ recombination in RC embedded into CAHS glassy matrices indicates that both <k> and σ increase progressively upon dehydration, similarly to that observed in trehalose-RC glasses. Quite interestingly, however, as shown in Fig. S1d, the dependencies of <k> and σ upon the hydration level appear to be systematically shifted in the CAHS-RC glassy matrices at higher contents of residual water, as compared to those found in trehalose-RC glasses. We infer that on the time scale of $P^+Q_A^-$ recombination (hundreds of milliseconds) CAHS glasses are better, at any given content of residual water, at inhibiting the RC conformational dynamics than trehalose matrices.

It is worth noting that the structural basis of the RC relaxation from the dark- to the light-adapted conformation has not yet been unequivocally fixed, and a number of different conformational events, not mutually exclusive, has been proposed to assist the stabilization of the primary $P^+Q_A^-$ charge separated state. These include: (a) small scale protein structural rearrangements on the quinone acceptor side of the RC,[33] (b) dielectric relaxation of water molecules weakly bound near the Q_A binding site,[34,35] (c) protonation of aspartate or glutamate residues on the cytoplasmic RC side,[36] and (d) rearrangements involving groups of the primary donor P, as well as amino acid residues and bound water molecules in the proximity of P.[37,38] All these events involve localized structural rearrangements, occurring in the sub-millisecond time scale. Therefore the kinetics of $P^+Q_A^-$ recombination that we use to probe RC relaxation and interconversion provides information on short-time, small-scale conformational dynamics. The control exerted on protein (RC) dynamics by the solvent dynamics can be rationalized within the framework of a “unified model of protein dynamics”, developed by Frauenfelder and coworkers[39] which summarizes decades of theoretical and experimental studies. The model envisages two classes of solvent-controlled protein dynamical processes: class I, following dielectric alpha relaxation, coupled to thermal fluctuations of bulk water, typically involving large scale, protein domain conformational rearrangement, and class II, obeying dielectric βα relaxations driven by hydration shell dynamics. Based on the local small-scale character of the above described RC conformational events, and on a number of experimental results concerning matrix effects,[11,13,31,40] the RC dynamics have been ascribed to class II processes, slaved to the dynamics of the protein hydration layer.

In view of the above considerations, the observation that CAHS glassy matrices are able to block fast, small scale RC conformational dynamics at hydration levels higher than trehalose glasses suggests that water of the hydration layer, shared by the RC-detergent complex and by the RC coating CAHS proteins, is more dynamically constrained on the fast time scale than residual water within trehalose-RC matrices. In the case of the latter system, a number of experimental observations and of molecular dynamics simulations [25,29,31] are consistent with the notion that protein conformational immobilization by trehalose glassy matrices is largely due to the formation of an extended HBN involving sugar molecules, residual water and surface protein residues anchored to the matrix via the thin H-bonded hydration layer of the protein (water entrapment/anchorage hypothesis). Upon extensive reduction of the water content, a few direct

protein-trehalose H bond interactions are likely to contribute to anchor the protein to the solid matrix.[32] It can be speculated that, similarly to what proposed for trehalose-protein matrices, in the case of RC-CAHS glasses the RC-CAHS interactions are mainly mediated by H bonded water molecules belonging to the hydration layers of the two proteins. The similar water adsorption properties of the trehalose-RC and CAHS-RC matrices (see Fig. S1d upper panel) are compatible with this possibility. Moreover, the intrinsically disordered nature of the CAHS proteins, which leads to gelification in the hydrated system, is expected to result within the dehydrated glass in a large number of water-mediated or direct RC-CAHS contacts, favoring a severe conformational immobilization of the RC even in the presence of significant contents of residual water. This might be relevant to explain why a substantial block of RC conformational dynamics is observed in the CAHS-RC matrix at higher hydration as compared to trehalose-RC matrix. The occurrence of tight protein-protein interactions and the strong inhibition of protein conformational fluctuation within a CAHS glassy matrix is also suggested by the impaired hydrogen/deuterium exchange found in CAHS glasses even in the presence of significant amounts of residual water (Fig. 1e).

References

1. Emsley, J. W. & Feeney, J. Magnetic Resonance, Historical Perspective. *Encyclopedia of Spectroscopy and Spectrometry* 1232–1240 (1999) doi:10.1006/rwsp.2000.0165.
2. Ectors, D., Goetz-Neunhoeffer, F., Hergeth, W.-D., Dietrich, U. & Neubauer, J. In situ 1 H-TD-NMR: Quantification and microstructure development during the early hydration of alite and OPC. *Cement and Concrete Research* vol. 79 366–372 (2016).
3. Zipp, A., James, T. L., Kuntz, I. D. & Shohet, S. B. Water proton magnetic resonance studies of normal and sickle erythrocytes Temperature and volume dependence. *Biochimica et Biophysica Acta (BBA) - General Subjects* vol. 428 291–303 (1976).
4. Valori, A., McDonald, P. J. & Scrivener, K. L. The morphology of C–S–H: Lessons from 1H nuclear magnetic resonance relaxometry. *Cement and Concrete Research* vol. 49 65–81 (2013).
5. Atta-ur-Rahman, Atta-ur-Rahman & Choudhary, M. I. Spin-Echo and Polarization Transfer. *Solving Problems with NMR Spectroscopy* 91–147 (1996) doi:10.1016/b978-012066320-0/50004-5.
6. Antalek, B. Accounting for spin relaxation in quantitative pulse gradient spin echo NMR mixture analysis. *J. Am. Chem. Soc.* **128**, 8402–8403 (2006).
7. Barth, A. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta* **1767**, 1073–1101 (2007).
8. Barth, A. & Zscherp, C. What vibrations tell us about proteins. *Q. Rev. Biophys.* **35**, 369–430 (2002).
9. W. L. Butler, D. W. H. Higher derivative analysis of complex absorption spectra. *Photochem. Photobiol.* **12**, 439–450 (1970).
10. Maréchal, Y. The molecular structure of liquid water delivered by absorption spectroscopy in the whole IR region completed with thermodynamics data. *J. Mol. Struct.* **1004**, 146–155

(2011).

11. Malferrari, M., Francia, F. & Venturoli, G. Coupling between electron transfer and protein-solvent dynamics: FTIR and laser-flash spectroscopy studies in photosynthetic reaction center films at different hydration levels. *J. Phys. Chem. B* **115**, 14732–14750 (2011).
12. Greenspan, L. Humidity Fixed Points of Binary Saturated Aqueous Solutions. *Journal of Research of the National Bureau of Standards—A Physics and Chemistry* **81A**, 89–96 (1977).
13. Malferrari, M., Francia, F. & Venturoli, G. Retardation of Protein Dynamics by Trehalose in Dehydrated Systems of Photosynthetic Reaction Centers. Insights from Electron Transfer and Thermal Denaturation Kinetics. *J. Phys. Chem. B* **119**, 13600–13618 (2015).
14. Bevington, P. R. & Bevington, R. R. *Data Reduction and Error Analysis for the Physical Sciences*. (McGraw-Hill College, 1969).
15. Francia, F., Malferrari, M., Sacquin-Mora, S. & Venturoli, G. Charge recombination kinetics and protein dynamics in wild type and carotenoid-less bacterial reaction centers: studies in trehalose glasses. *J. Phys. Chem. B* **113**, 10389–10398 (2009).
16. de Johgn Jean-Marie Ruysschaert, E. G. H. H. J. Relevance of Protein Thin Films Prepared for Attenuated Total Reflection Fourier Transform Infrared Spectroscopy: Significance of the pH. *Appl. Spectrosc.* **50**, 1519–1527 (1996).
17. Haris, P. I., Coke, M. & Chapman, D. Fourier transform infrared spectroscopic investigation of rhodopsin structure and its comparison with bacteriorhodopsin. *Biochim. Biophys. Acta* **995**, 160–167 (1989).
18. Clark, A. H., Saunderson, D. H. & Suggett, A. Infrared and laser-Raman spectroscopic studies of thermally-induced globular protein gels. *Int. J. Pept. Protein Res.* **17**, 353–364 (1981).
19. Wolkers, W. F., McCready, S., Brandt, W. F., Lindsey, G. G. & Hoekstra, F. A. Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses in vitro. *Biochim.*

Biophys. Acta **1544**, 196–206 (2001).

20. Koubaa, S., Bremer, A., Hincha, D. K. & Brini, F. Structural properties and enzyme stabilization function of the intrinsically disordered LEA_4 protein TdLEA3 from wheat. *Sci. Rep.* **9**, 3720 (2019).
21. M. Malferrari, G. Venturoli, F. Francia, A. Mezzetti. A new method for D2O/H2O exchange in infrared spectroscopy of proteins. *Spectroscopy* **12**, 337–342 (2012).
22. M. Malferrari, F. Francia, A. Mezzetti, G. Venturoli. Kinetic effects in dehydration, rehydration, and isotopic exchange of bacterial photosynthetic reaction centers. *Biomedical Spectroscopy and Imaging* **5**, 185–196 (2016).
23. Goormaghtigh, E., Vigneron, L., Scarborough, G. A. & Ruysschaert, J. M. Tertiary conformational changes of the Neurospora crassa plasma membrane H(+)-ATPase monitored by hydrogen/deuterium exchange kinetics. A Fourier transformed infrared spectroscopy approach. *J. Biol. Chem.* **269**, 27409–27413 (1994).
24. Boothby, T. C. *et al.* Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation. *Mol. Cell* **65**, 975–984.e5 (2017).
25. Cordone, L. *et al.* Internal dynamics and protein-matrix coupling in trehalose-coated proteins. *Biochim. Biophys. Acta* **1749**, 252–281 (2005).
26. Feher, G., Allen, J. P., Okamura, M. Y. & Rees, D. C. Structure and function of bacterial photosynthetic reaction centres. *Nature* **339**, 111–116 (1989).
27. McMahon, B. H., Müller, J. D., Wraight, C. A. & Nienhaus, G. U. Electron transfer and protein dynamics in the photosynthetic reaction center. *Biophys. J.* **74**, 2567–2587 (1998).
28. Palazzo, G., Mallardi, A., Hochkoeppler, A., Cordone, L. & Venturoli, G. Electron transfer kinetics in photosynthetic reaction centers embedded in trehalose glasses: trapping of conformational substates at room temperature. *Biophys. J.* **82**, 558–568 (2002).
29. Francia, F. *et al.* Protein-matrix coupling/uncoupling in ‘dry’ systems of photosynthetic reaction center embedded in trehalose/sucrose: the origin of trehalose peculiarity. *J. Am.*

Chem. Soc. **130**, 10240–10246 (2008).

30. Hailwood, A. J. & Horrobin, S. Absorption of water by polymers: analysis in terms of a simple model. *Trans. Faraday Soc.* **42**, B084 (1946).
31. Malferrari, M., Savitsky, A., Lubitz, W., Möbius, K. & Venturoli, G. Protein Immobilization Capabilities of Sucrose and Trehalose Glasses: The Effect of Protein/Sugar Concentration Unraveled by High-Field EPR. *J. Phys. Chem. Lett.* **7**, 4871–4877 (2016).
32. Lerbret, A. *et al.* How strongly does trehalose interact with lysozyme in the solid state? Insights from molecular dynamics simulation and inelastic neutron scattering. *J. Phys. Chem. B* **116**, 11103–11116 (2012).
33. Nabedryk, E. *et al.* A protein conformational change associated with the photoreduction of the primary and secondary quinones in the bacterial reaction center. *FEBS Lett.* **266**, 59–62 (1990).
34. Iwata, T., Paddock, M. L., Okamura, M. Y. & Kandori, H. Identification of FTIR bands due to internal water molecules around the quinone binding sites in the reaction center from Rhodobacter sphaeroides. *Biochemistry* **48**, 1220–1229 (2009).
35. Malferrari, M., Mezzetti, A., Francia, F. & Venturoli, G. Effects of dehydration on light-induced conformational changes in bacterial photosynthetic reaction centers probed by optical and differential FTIR spectroscopy. *Biochim. Biophys. Acta* **1827**, 328–339 (2013).
36. Remy, A. & Gerwert, K. Coupling of light-induced electron transfer to proton uptake in photosynthesis. *Nat. Struct. Biol.* **10**, 637–644 (2003).
37. Deshmukh, S. S., Williams, J. C., Allen, J. P. & Kálmán, L. Light-induced conformational changes in photosynthetic reaction centers: dielectric relaxation in the vicinity of the dimer. *Biochemistry* **50**, 340–348 (2011).
38. Deshmukh, S. S., Williams, J. C., Allen, J. P. & Kálmán, L. Light-Induced Conformational Changes in Photosynthetic Reaction Centers: Redox-Regulated Proton Pathway near the Dimer. *Biochemistry* vol. 50 3321–3331 (2011).

39. Fenimore, P. W. *et al.* Coupling Of Solvent And Protein Dynamics: Mossbauer And Incoherent Neutron Scattering From Dielectric Relaxation Data. *Biophysical Journal* vol. 96 323a (2009).

40. Savitsky, A., Malferrari, M., Francia, F., Venturoli, G. & Möbius, K. Bacterial photosynthetic reaction centers in trehalose glasses: coupling between protein conformational dynamics and electron-transfer kinetics as studied by laser-flash and high-field EPR spectroscopies. *J. Phys. Chem. B* **114**, 12729–12743 (2010).

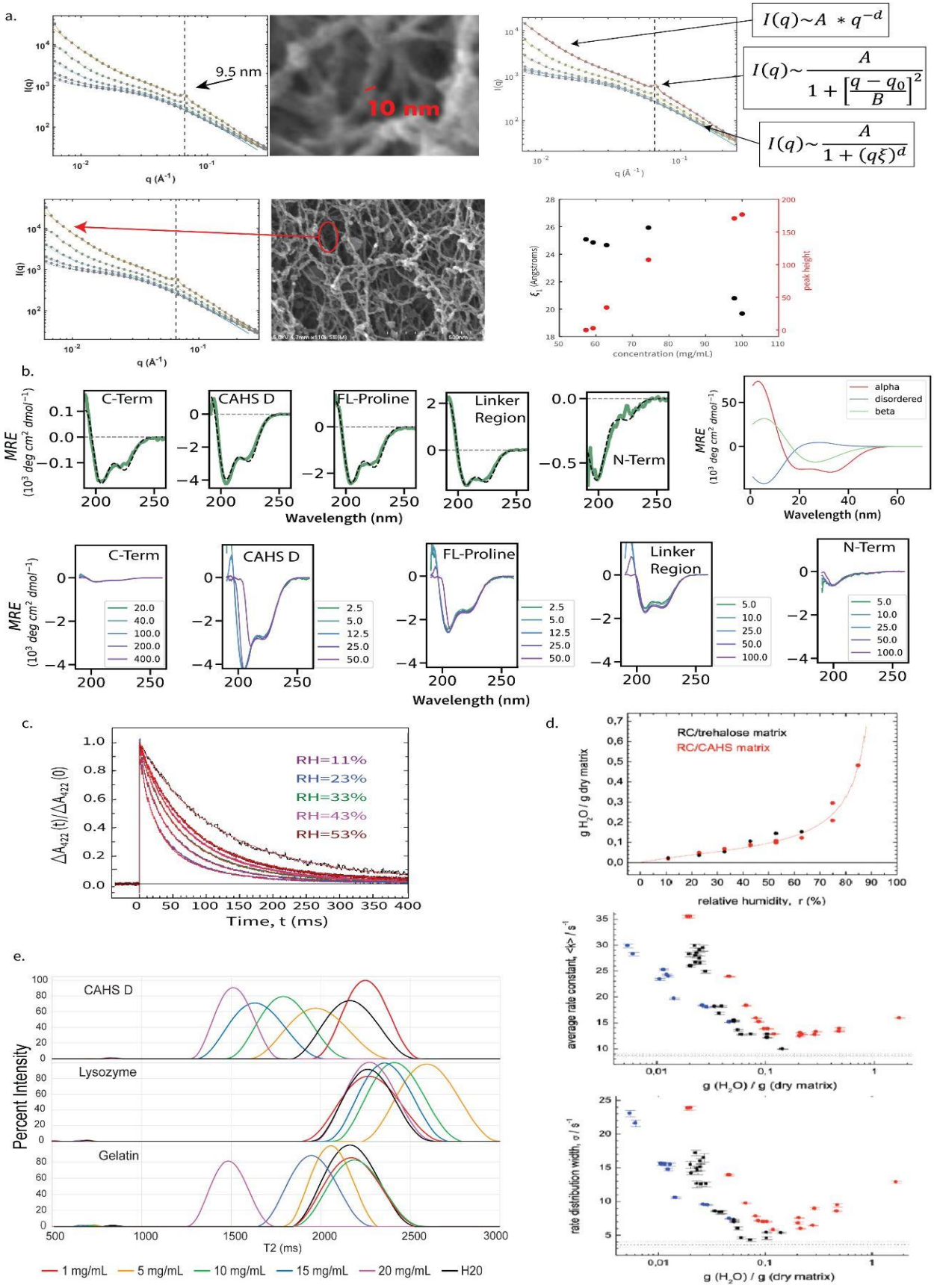

a.
9.5 nm
10 nm
I(q)
q (Å⁻¹)
$I(q) \sim A * q^{-d}$
$I(q) \sim \frac{A}{1+\left[\frac{q-q_0}{B}\right]^2}$
$I(q) \sim \frac{A}{1+(q\xi)^d}$
ξ_1 (Angstroms)
peak height
concentration (mg/mL)
b.
C-Term
CAHS D
FL-Proline
Linker Region
N-Term
MRE (10^3 deg cm^2 dmol^{-1})
Wavelength (nm)
alpha
disordered
beta
20.0
40.0
100.0
200.0
400.0
2.5
5.0
12.5
25.0
50.0
10.0
100.0
c.
RH=11%
RH=23%
RH=33%
RH=43%
RH=53%
$\Delta A_{422}(t)/\Delta A_{422}(0)$
Time, t (ms)
d.
RC/trehalose matrix
RC/CAHS matrix
g H_2O / g dry matrix
relative humidity, r (%)
average rate constant, <k> / s⁻¹
g (H_2O) / g (dry matrix)
rate distribution width, σ / s⁻¹
e.
CAHS D
Lysozyme
Gelatin
Percent Intensity
T2 (ms)
1 mg/mL
5 mg/mL
10 mg/mL
15 mg/mL
20 mg/mL
H20

Fig. S1: Properties of hydrated and desiccated CAHS D gels. a) Concentration gradient gel SAXS structure probing revealed an emergent structure of approximately 9.5 nm. Areas indicated in the SAXS plot correspond with structures shown in SEM images, with the 9.5nm feature corresponding to approximate fiber sizes, while larger features correspond with the void spaces between gel fibers. Top right panel shows the derivation of sizes for each region of the SAXS plot. Lower right panel shows a projected relationship of the size of fibers with increasing protein concentration. b) CD spectra for C-terminus, CAHS D, Fl-Proline, Linker Region, and N-terminus. Top panels show the spectra for each variant listed at 20 or 25 μM (in green), overlayed with the resulting fit of the base spectra (dashed black line). Rightmost top panel shows the base spectra used to fit the experimental spectra (Perczel 1992). Lower panels show the concentration dependence of the spectra for each variant, at a range of concentrations. No structural changes are observed in the range tested. Premature rise at low wavelengths occuring at high concentrations in the CAHS D, FL-Proline, and Linker Region spectra are a result of reduced signal reaching the detector. Concentrations in μM are shown in the legend. c) Hydration dependent immobilization of biological material with the CAHS D glassy matrix. Charge recombination kinetics after a laser pulse (fired at t=0) are shown for RC incorporated into CAHS D glasses equilibrated at different relative humidity. Continuous curves represent best fit to eq.1. Reaction centers in solution (black line) and embedded into the CAHS D glass held 11% RH (purple line) have decay kinetics characterized by an average rate constant $\langle k \rangle = 8.8\ s^{-1}$ and $\langle k \rangle = 35.5\ s^{-1}$ respectively, and a rate constant distribution width of $\sigma = 3.6\ s^{-1}$ and $\sigma = 23.8\ s^{-1}$, respectively. With increasing relative humidity (23%-53% RH) the RC becomes more mobile. d) (Top) Hydration isotherms at 298 K determined in CAHS-RC (red circles) and trehalose-RC (black circles) amorphous matrices. In the glasses the CAHS/RC and trehalose/RC mass ratios were equal to 1.7. Continuous curve representing a best fit to eq. 2 of the water sorption data obtained in the CAHS-RC matrix (see supplemental text for details). Middle & lower panels show dependence of recombination kinetics on the residual water content in CAHS-RC matrices (red circles), in trehalose-RC glasses characterized by trehalose/RC molar ratios of 5000 and 500 (black squares and circles, respectively), and in a dehydrated RC film in the absence of any excipient (blue circles). Middle panel shows the average rate constant, $\langle k \rangle$, bottom panel shows the rate constant distribution width, σ. Vertical bars represent confidence intervals within two standard deviations. The dashed lines give the value of $\langle k \rangle$ and σ obtained in solution. Dotted lines indicate the confidence intervals within two standard deviations. e) T_2 relaxation of lysozyme, gelatin, and CAHS D. Concentration-dependent increase in relaxation kinetics is observed for CAHS D indicating strong water coordination, while gelatin shows a shift in T_2 relaxation only after gelation occurs (20g/L). Lysozyme, a globular non-gelling protein, does not show strong water coordination.

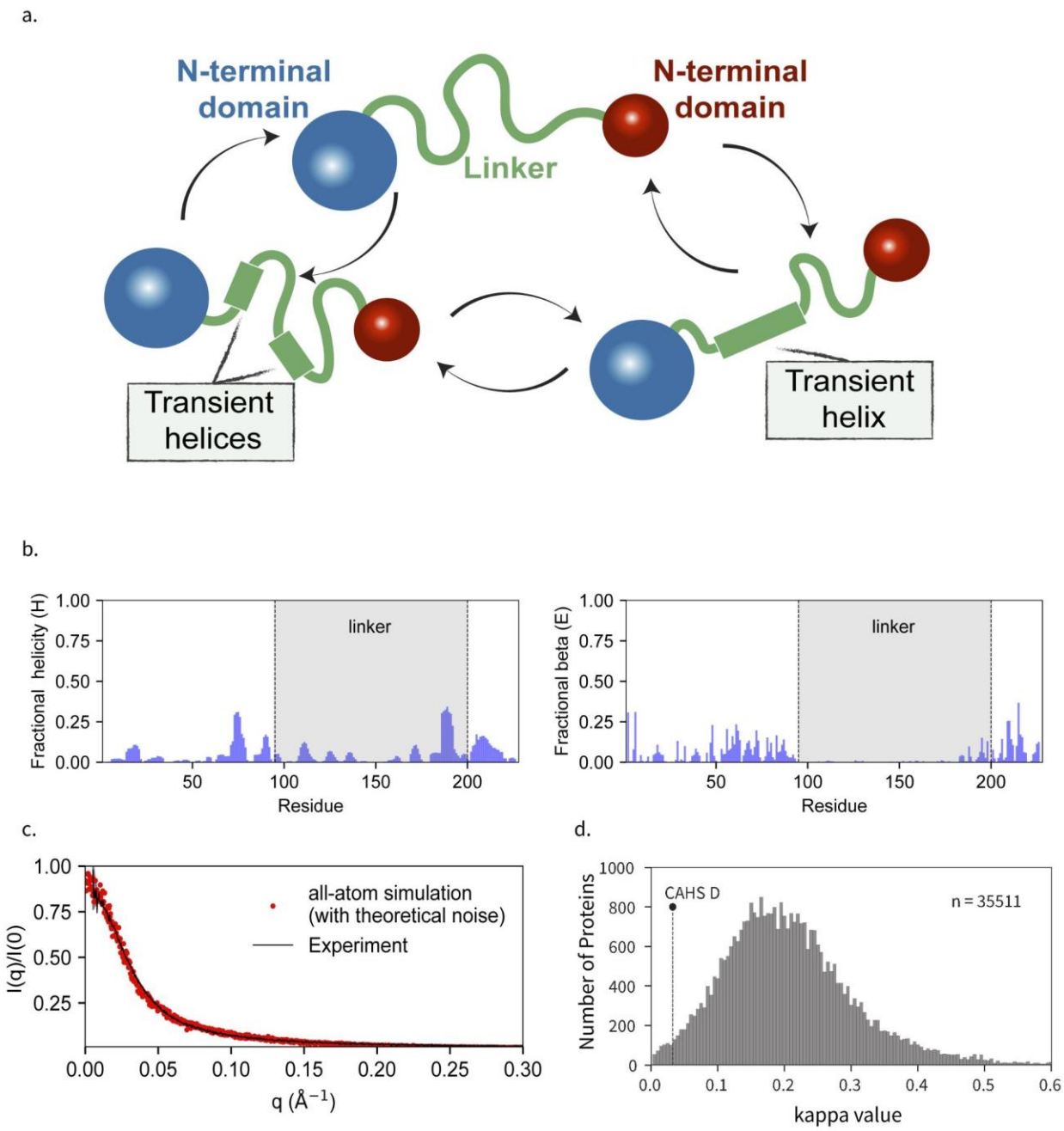


Fig S2. Transient structural properties of CAHS D. a) Model of interconversion of disordered linker helix, gaining and losing transient α helices. b) Calculated fractional α-helix (left) and β-sheet (right) content of the wildtype CAHS D as determined from simulations. c) Curve-fitting of radius of gyration determined from simulations (red points) overlaid with SAXS experimental data (black line). d) Calculated kappa value (amino acid charge distribution) of all proteins in the tardigrade proteome. Kappa provides a metric to quantify the degree of charge patterning, such that sequences with a lower kappa value have a more even distribution of charged residues. As shown, CAHS D has some of the most evenly distributed charges of all proteins in the proteome.

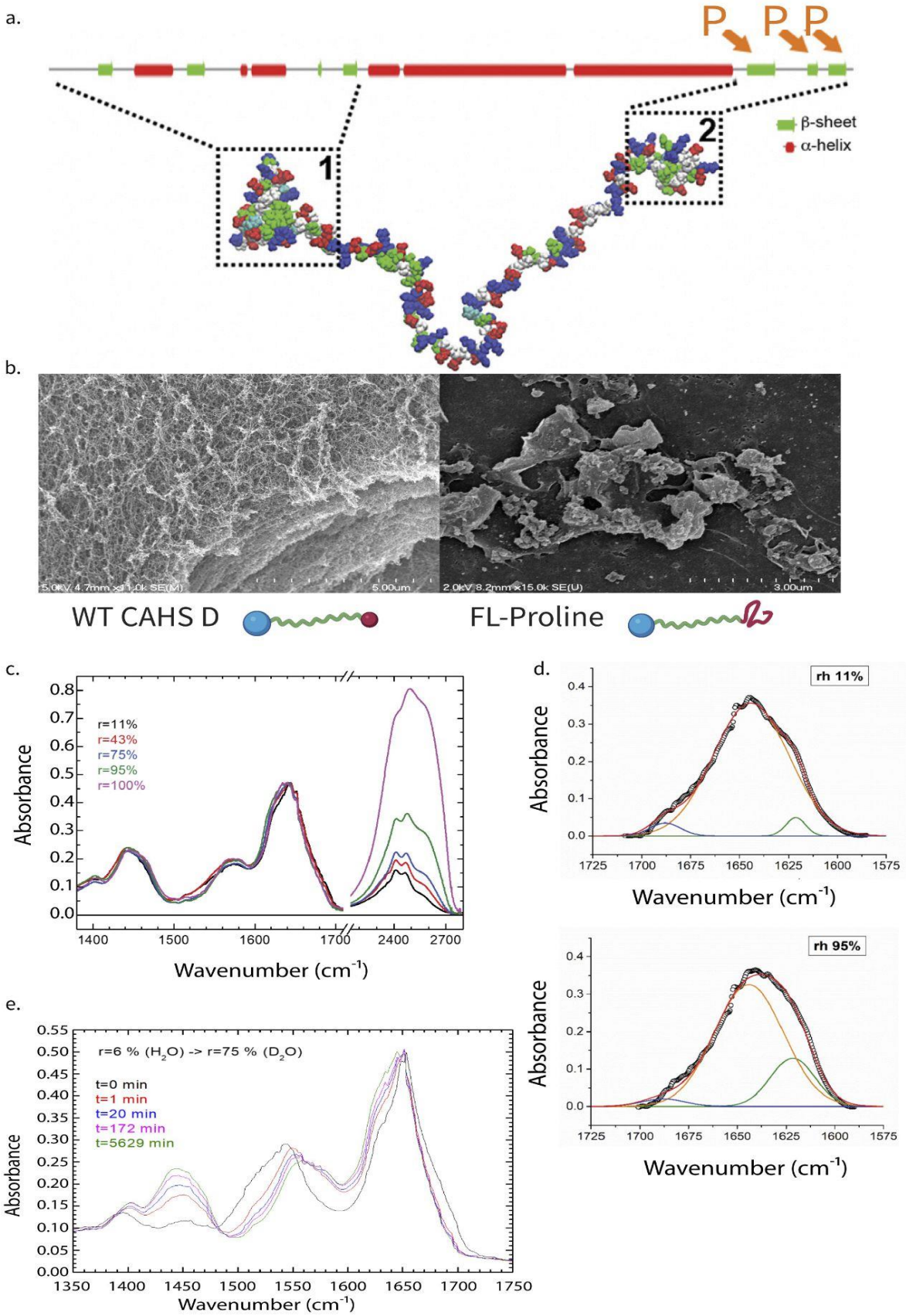

a.
P
P
P
β-sheet
α-helix
1
2
b.
5.00um
3.00um
WT CAHS D
FL-Proline
c.
r=11%
r=43%
r=75%
r=95%
r=100%
Absorbance
Wavenumber (cm-1)
d.
rh 11%
rh 95%
Absorbance
Wavenumber (cm-1)
e.
r=6 % (H2O) -> r=75 % (D2O)
t=0 min
t=1 min
t=20 min
t=172 min
t=5629 min
Absorbance
Wavenumber (cm-1)

Fig S3. a) Bioinformatic predictions of the secondary structure of WT CAHS D, with arrows indicating where prolines were inserted into the sequence to disrupt secondary structures. b) SEM images of CAHS D and FL-Proline, showing that FL-Proline does not form the reticular meshwork gel structure found in CAHS D. c) Hydration-dependent FTIR spectra for CAHS D glasses stored at different relative humidities. Spectra for all hydration levels are overlaid. d) Deconvolution of FTIR spectra shown in (c) for the 1725-1575 cm^{-1} range. Top panel shows the deconvolution for 11% RH, and the bottom panel shows the same for 95% RH. Individual Gaussian components are shown in blue, orange and green; the red curve represents the corresponding best fit e) Evolution of the infrared spectrum between 1350 and 1750 cm^{-1} of an extensively dehydrated CAHS glass as a function of the deuteration time. The CAHS glassy matrix, previously equilibrated at RH=6% in the presence of $NaOH{\cdot}H_2O$, was exposed at t=0 to the atmosphere of a saturated NaCl solution in D_2O, at RH=75%. Spectra are normalized to the amplitude of the amide I/I' band.

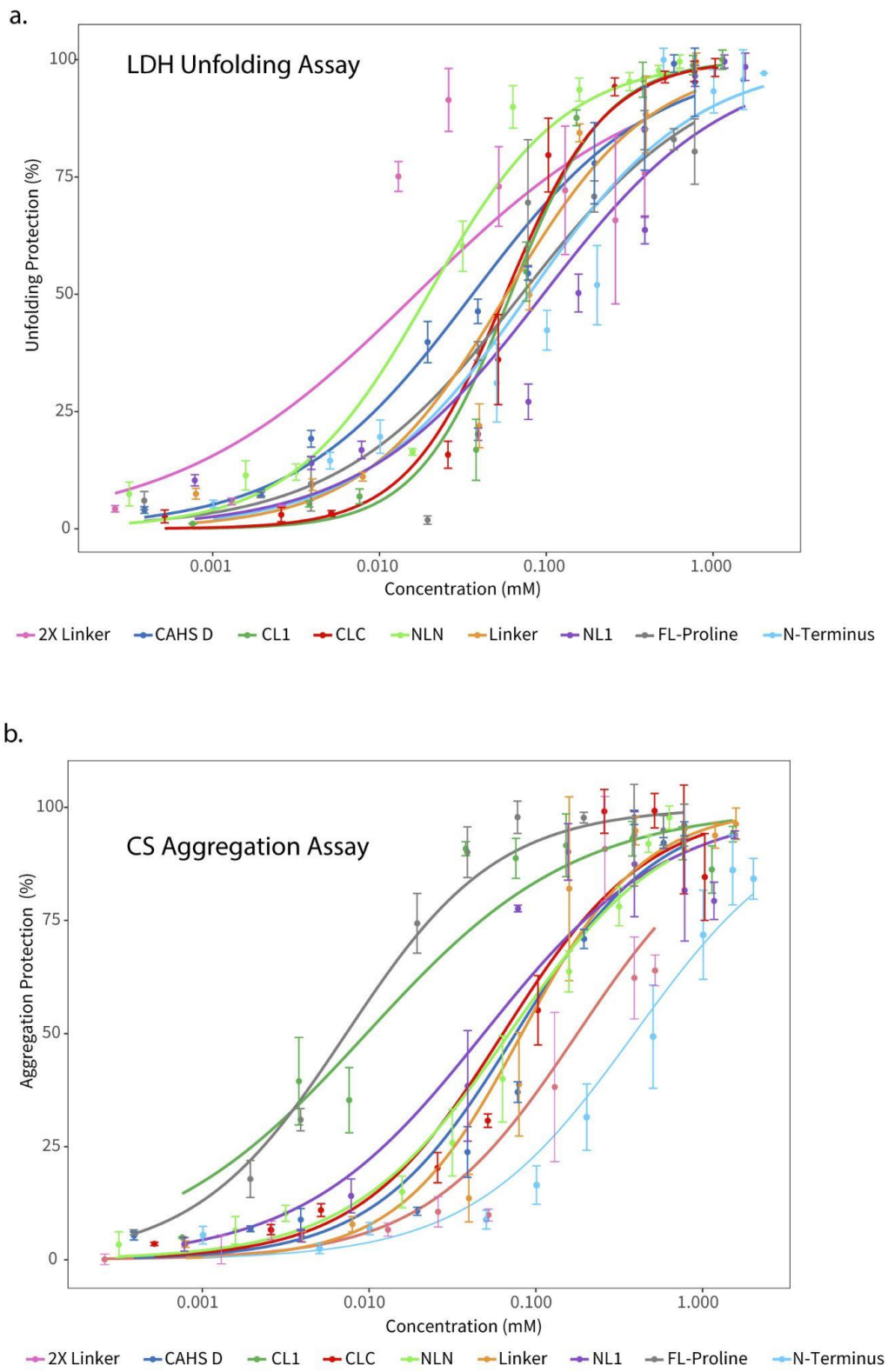


Figure S4. Sigmoidal curves representing the full range of protection for each variant using the LDH unfolding assay. b) Sigmoidal curves representing the full range of protection for each variant using the CS aggregation assay.

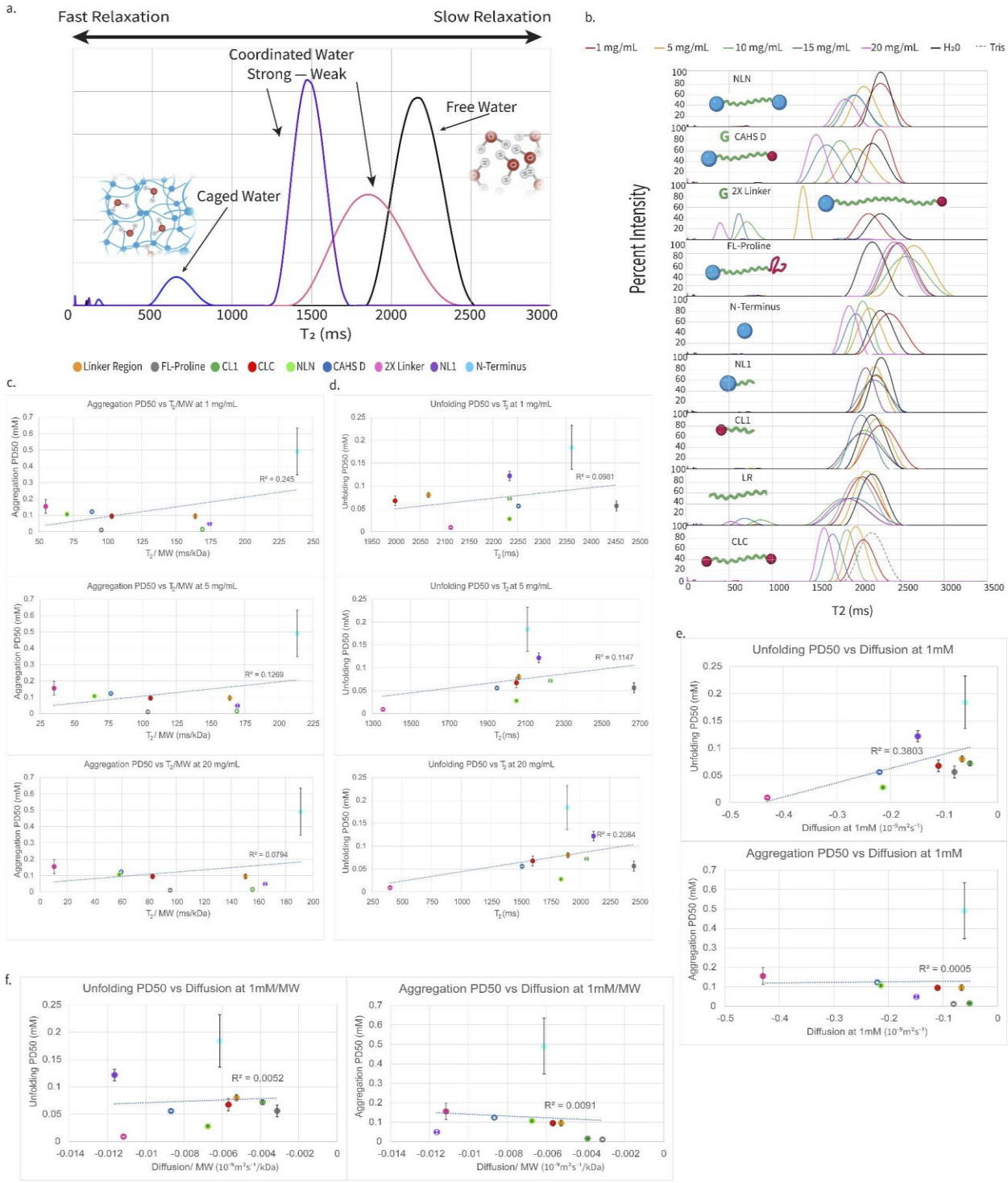
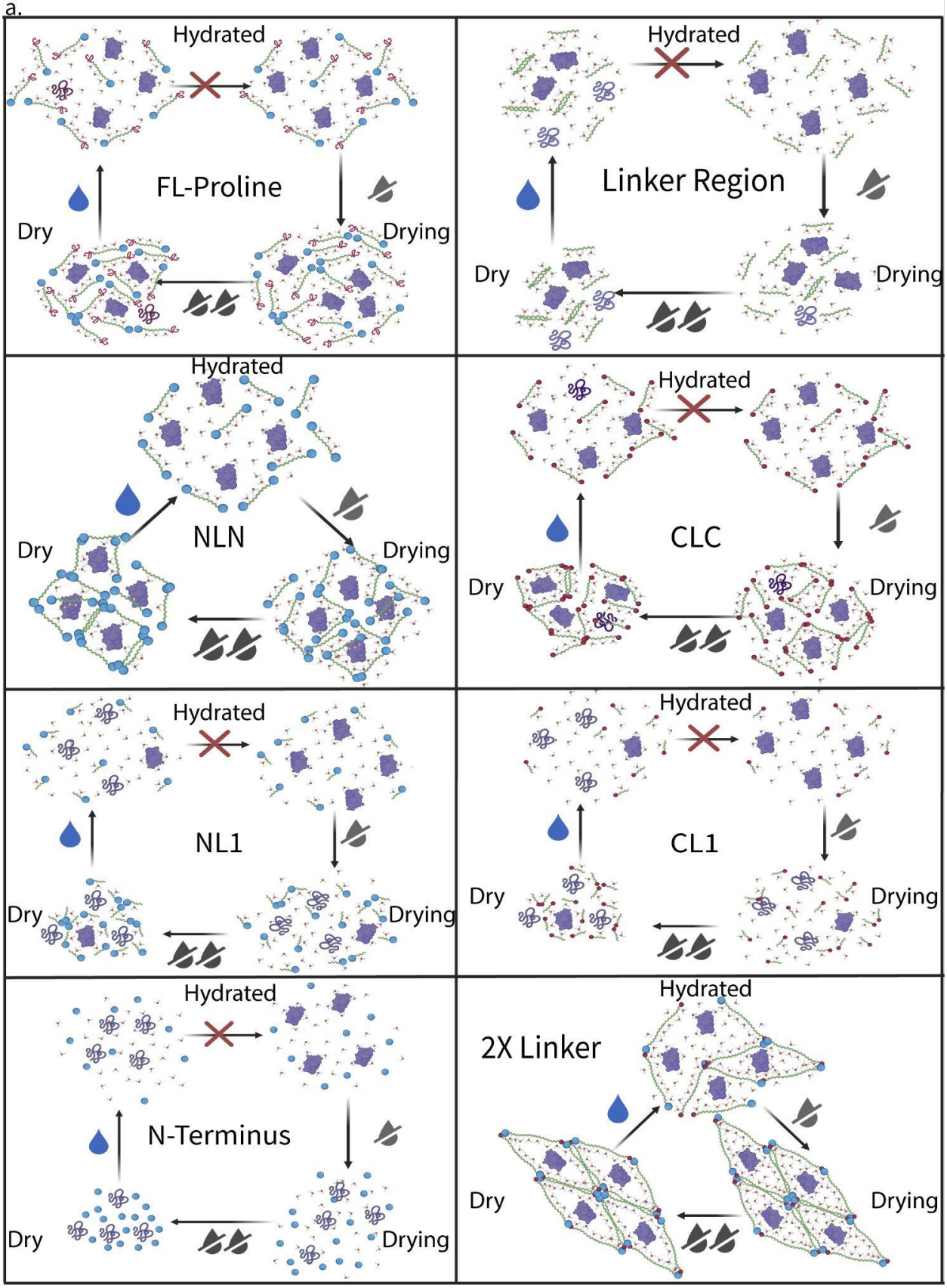


Fig S5. a) 1H nuclei transverse spin magnetization Free Induction Decays (FID) are subjected to an inverse Laplace transformation, which provides a distribution of spin-spin decay times referred to as T_2 relaxation. The leftward shift of T_2 peak midpoints indicates how water is ordered in the sample being measured. Bulk liquids (free water) have slower T_2 relaxation than samples which display water coordination. This figure shows example T_2 distributions, with representative illustrations of the water behavior in each sample. b) T_2 relaxation distributions for all variants. c)

Aggregation PD50 plotted against T2/MW for all variants at 1 g/L (top), 5 g/L (middle) and 20 g/L (bottom). No correlations are observed. d) Unfolding PD50 plotted against T2 peak midpoint for all variants at 1 g/L (top), 5 g/L (middle) and 20 g/L (bottom). Only very weak correlations are seen between non-normalized T2 and unfolding PD50.. e) Diffusion at 1mM plotted against PD50 for unfolding (top) and aggregation (bottom). f) Diffusion at 1mM normalized to MW, plotted against PD50 for unfolding (left) and aggregation (right). No strong correlations are found for diffusion and PD50. For parts (c)-(f) error bars represent standard deviation for PD50. Bars smaller than the size of the points are shown as white bars within points.

a.


Hydrated
FL-Proline
Dry
Drying
Hydrated
Linker Region
Dry
Drying
Hydrated
NLN
Dry
Drying
Hydrated
CLC
Dry
Drying
Hydrated
NL1
Dry
Drying
Hydrated
CL1
Dry
Drying
Hydrated
N-Terminus
Dry
Drying
Hydrated
2X Linker
Dry
Drying

b.

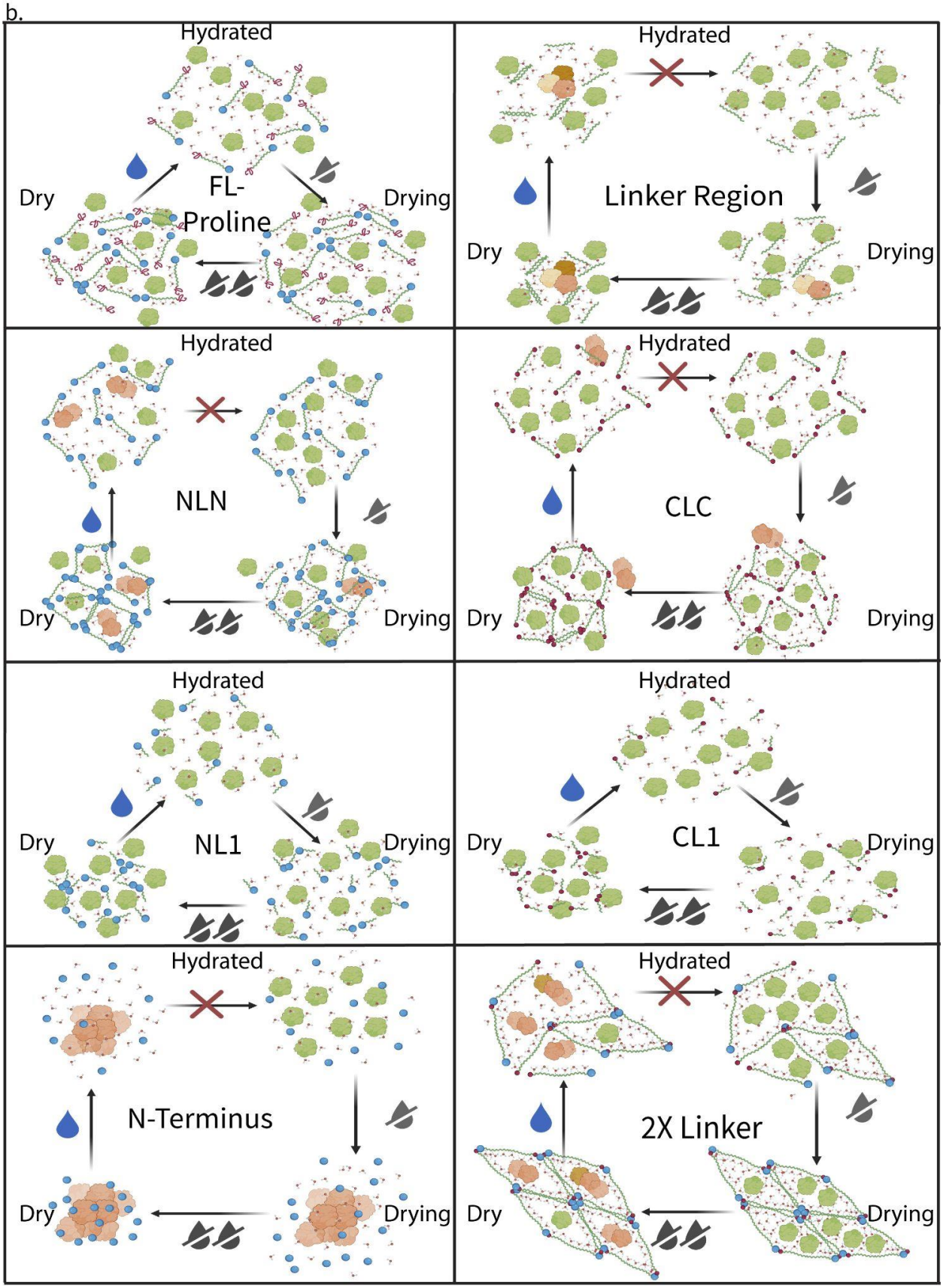

Hydrated
Dry
FL-
Proline
Drying
Hydrated
Linker Region
Dry
Drying
Hydrated
NLN
Dry
Drying
Hydrated
CLC
Dry
Drying
Hydrated
Dry
NL1
Drying
Hydrated
Dry
CL1
Drying
Hydrated
N-Terminus
Dry
Drying
Hydrated
2X Linker
Dry
Drying

Fig S6. Working models for CAHS D variant behavior and prevention of protein unfolding (a) and aggregation (b) in the (re)hydrated, drying, and dry state. Schematics representation interpreted modes and degree of protection and gelation for each of our variants is presented.