Molecular Swiss Army Knives: Tardigrade CAHS Proteins Mediate Desiccation Tolerance Through Multiple Mechanisms.

Cherie S. Hesgrove[T1], Kenny H. Nguyen[T1], Sourav Biswas[T1], Charles A. Childs[T1], Shraddha KC[T1], Bryan X. Medina[V1], Vladimir Alvarado [V1], Shahar Sukenik[S1,2], Feng Yu[S2], Marco Malferrari[G1], Francesco Francia[G2], Giovanni Venturoli[G2,3], Erik W. Martin[E1], Alex S. Holehouse[A1,2], Thomas C. Boothby[T1]

T1 Molecular Biology Department, University of Wyoming, Laramie, WY, USA

V1 Chemical Engineering Department, University of Wyoming, Laramie, WY, USA

G1 Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna, via Selmi 2, I-40126 Bologna, Italy

G2 Laboratorio di Biochimica e Biofisica Molecolare, Dipartimento di Farmacia e Biotecnologie, FaBiT, Università di Bologna, via Irnerio 42, I-40126 Bologna, Italy.

G3 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

S1 Department of Chemistry and Biochemistry, University of California, Merced, CA, USA

S2 Quantitative Systems Biology Program, University of California, Merced, CA, USA

A1 Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO, USA

A2 Center for Science and Engineering of Living Systems (CSELS), Washington University in St. Louis, St. Louis, MO, USA

E1 Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN, USA

Abstract

Tardigrades, also known as water bears, make up a phylum of small but extremely robust 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 gelation of CAHS D is driven by intermolecular beta-beta interactions. We find that gelation of CAHS D promotes the slowing of diffusion, and coordination of residual water. Slowed diffusion and increased water coordination correlate with the ability of CAHS D to provide robust stabilization of an enzyme, lactate dehydrogenase, which otherwise unfolds when dried. Conversely, slowed diffusion and water coordination do not promote the prevention of protein aggregation during drying. 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] Anhydrobiosis can be a normal stage of life, as in the case of plant seeds[2,3], or it can be a response to an environmental loss of water. [4–7] In many organisms capable of anhydrobiosis, non-reducing sugars such as sucrose in plants[8] or trehalose in many organisms [9–13] are accumulated intracellularly during the process of desiccation. The acquisition of simple polysaccharides was long thought to be an essential feature in desiccation tolerance. However, many animals with robust anhydrobiotic responses, such as tardigrades and rotifers, appear to either not accumulate or accumulate very low levels of these sugars in response to drying.[5,14,15] Intriguingly, these animals utilize diverse array of intrinsically disordered proteins to provide adaptive protection against desiccation.[5,16–21]

The tardigrade *H. exemplaris* (formerly *H. dujardini*) employs a suite of intrinsically disordered proteins termed Tardigrade Disordered Proteins (TDPs) in their desiccation survival strategy.[5,19–21] Homologs of the *H. exemplaris* TDP families have also been found in several other eutardigrades species[5], but were not detected in transcriptomes of the heterotardigrade *Echiniscoides cf. sigismundi*, suggesting that these proteins are unique to desiccation-tolerant lineages.[5,22]

The TDP family that is most enriched during desiccation in *H. exemplaris* is the Cytoplasmic Abundant Heat Soluble (CAHS) family.[5,19] Previous work has established that the model CAHS protein CAHS D is both required for *H. exemplaris* anhydrobiosis and provides desiccation protection when heterologously expressed in yeast and bacteria.[5] *In vitro,* CAHS D can protect lactate dehydrogenase (LDH) from denaturation when the enzyme is subjected to desiccation and rehydration. [5,21] Despite its ability to function heterologously *in vivo* and provide protection *in vitro*, the molecular basis for how CAHS proteins confer desiccation tolerance remains unknown.

One general mechanism proposed to explain desiccation tolerance is the vitrification hypothesis[23]. This hypothesis hinges on slowed molecular motion reducing the frequency and speed of damaging processes. In the early stages of drying this is proposed to occur through high viscosity in the system, slowing diffusion and molecular motion. After a material is dried this is achieved through a vitrified solid. Once in the solid state, molecular motion is slowed so dramatically that biological processes are essentially stopped, preventing further degradation of the system. Indeed, while it has been proposed that tardigrades and their CAHS proteins undergo vitrification upon desiccation, vitrification is not mutually exclusive with other potential mechanisms of desiccation tolerance (discussed later).[5,19,21,24,25]

Here we set out to understand the molecular mechanism that underlies how CAHS D protects client proteins from desiccation-induced protein damage. We discovered that CAHS D can undergo a sol-gel transition in a concentration and temperature-reversible manner. To understand the role gelation

plays on desiccation tolerance we combined rational sequence design with a series of complementary techniques to probe the structural and biophysical properties of CAHS D. Unexpectedly, we found that the mechanisms underlying protein stabilization do not correlate with the inhibition of protein aggregation, revealing that protection against the two major forms of protein dysfunction are mechanistically distinct. We were also surprised to find that CAHS D's interactions with water, as measured through TD-NMR, 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 provide avenues for pursuing applications such as the engineering of stress tolerant crops and the stabilization of temperature sensitive biologics, such as vaccines, in a dry state.

Results

Tardigrade CAHS D undergoes a phase transition to form a gel. CAHS D is a 227-residue highly charged protein that is predicted to be fully disordered (Fig. 2b). In the course of its expression and purification, it was observed that CAHS D undergoes a sol-gel phase transition, transitioning from a liquid into a solid gel-like state (Fig. 1a). CAHS D gelation is concentration dependent, with solutions of the protein below ~10 g/L remaining diffuse, while solutions between 10 g/L − 15 g/L showing increasing viscosity, and solutions above ~15 g/L forming robust gels (Fig. 1a). Moreover, re-solvation from the gel state is possible upon the addition of buffer, as long as the final concentration after buffer addition is below the gel point (Fig 1c). Taken together, these observations suggest gelation is driven by non-covalent physical crosslinks, as opposed to chemical crosslinks which would yield an irreversible gel [26].

Gelation is also thermo-reversible; heating a gel formed at room temperature to 50 ℃ melts the gel, and this heated solution re-solidifies and gels upon cooling (Fig. 1a). The observed temperature dependence implies a dominating enthalpic contribution to the molecular basis for gel formation, as opposed to one driven by the hydrophobic effect.

High-resolution imaging revealed that CAHS D gels form reticular (web-like) networks (Fig. 1c), with larger pores between the fine meshwork of the gel (Fig. S1a). This topology is characteristic of other proteinaceous gels such as those formed by gelatin (Fig. 1c). These web-like gel structures are distinct from the crystalline solids formed by non-gelling proteins, such as lysozyme (Fig. 1c). The molecular architecture is reminiscent of that formed by synthetic polymers such as borax-crosslinked poly(vinyl alcohol), implicating relatively sparse physical crosslinks that underlie the network connectivity[27]. This in turn implies that specific regions of CAHS D may drive intermolecular interactions.

To gain further insight into the nanoscopic details of CAHS D gelation we turned to small-angle X-ray scattering (SAXS). SAXS is a solution scattering technique that provides information on length-scales ranging from Angstroms to hundreds of nanometers. [28] SAXS was performed on a concentration gradient of CAHS D gels (Fig.1d, S1a). We observed the concentration-dependent emergence of a scattering peak, indicative of a repeating structure approximately 9.5 nm in size

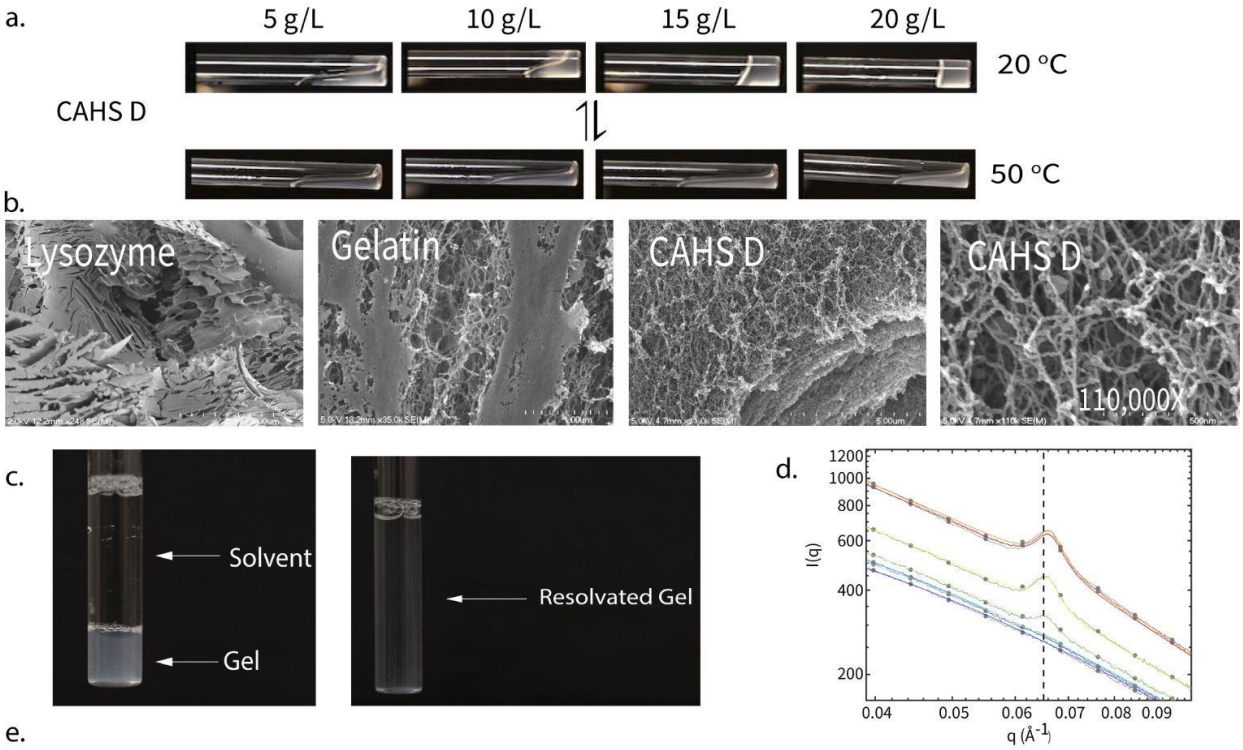

Fig. 1: CAHS D forms a reversible gel with hydration-dependent physical properties and a reticular meshwork of fibers. a) Reversible, temperature and concentration dependent gelation of CAHS D. CAHS D at 20˚C (top panel) shows gelation beginning at 15g/L, while at 50 ˚C gelation is not observed in any concentration. Once cooled, gel formation recurs at 15 g/L. b) SEM images of lysozyme, gelatin, and CAHS D at 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 done with proteins at 50 g/L. c) Dilution of 20g/L CAHS D gel results in resolvation. d) Concentration gradient gel SAXS structure probing revealed an emergent structure of approximately 9.5 nm. e) Δ Diffusion coefficients for lysozyme, CAHS D, and gelatin. Gelled proteins show more dramatic slowing of ΔDC than non-gelling lysozyme, in a concentration dependent manner. f) Kinetics of amide HDX in CAHS glassy matrices as studied by an isopiestic approach. Left panel: Extensively dehydrated CAHS glasses, equilibrated at RH=6%, were exposed to a D_2O atmosphere at RH=11% (black symbols) or RH=75% (blue symbols). Subsequently both glasses were incubated in the presence of pure D_2O (RH=100%). See supplemental text for details. The dashed line represents the value of the amide II' band area (normalized to the area of the amide I' band) measured in a fully deuterated CAHS glass, prepared by dissolving the lyophilized CAHS protein in D_2O at a temperature of about 50°C. Right panel is an expansion of the left panel, showing the initial phase of the HDX kinetics.

(Fig. 1d, S1a). This feature size corresponds well with the ~10 nm fibers observed in our SEM imaging (Fig. 1b). SAXS also allowed us to determine the mesh-size of the fibers, which shrunk in a concentration-dependent manner from ~26 Å at low CAHS D concentration, to ~20 Å at higher concentrations, consistent with greater structural integrity at higher concentrations (Fig. S1a, bottom panel).

While gelation of proteins is not rare *per se*, it is usually achieved through heating, acidic or basic treatment, or some other perturbation to a protein.[29–31] On the other hand, protein gelation in a biologically relevant context [32,33] is less commonly found in proteomes, and in these proteins gelation is often functional.[32,33] This led us to wonder if gelation of CAHS D could account for some of the protective capacity of this protein during extreme desiccation.

CAHS D gelation slows diffusion in a hydrated state and immobilizes biological material in a vitrified state. To further assess the properties of CAHS D gels, the concentration-dependent diffusion of water in CAHS D gels was assessed using time-domain nuclear magnetic resonance spectroscopy (TD-NMR). TD-NMR is a technique used commonly in food science and hydrogel research to determine the diffusion of water and oils in liquid, gel, and even solid states.[34,35] 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 of the buffer used in each experiment (water alone or 20 mM Tris), in order to emphasize the differences between proteins. This is calculated as (Buffer Diffusion − Protein Diffusion) and is referred to as Δ Diffusion Coefficient (ΔDC). TD-NMR shows that CAHS D dramatically slowed the diffusion of water, similar to the strongly gelling protein, gelatin, as a function of its concentration (Fig. 1e). In contrast, lysozyme, a non-gelling protein, slowed diffusion much less than the gelling proteins, gelatin and CAHS D (Fig. 1e).

We used hydrogen-deuterium exchange (HDX) assays on vitrified CAHS D gels (vitrified solids), assaying exchange in both fully dried and semi-hydrated samples, to determine the degree of immobilization that occurs within the dried matrix of the vitrified gel. HDX is measured by dissolving CAHS D protein in D_2O, then vacuum-drying this to create a vitrified solid. The vitrified solid is then incubated in environments with differing

relative humidity to determine water exchange with the environment (Fig. 1f; Fig. S1g). The degree to which exchange occurs is an indicator of how strongly the desiccated gel matrix is bound to the water molecules that reside within it. A solid matrix with a weak hold on its internal water will exchange more with the surrounding air than a matrix with strong internal water interactions.

HDX was measured for vitrified solids kept in low (11%) and high (75%) relative humidity (RH), as well as at 100% relative humidity. Hydrogen exchange rates were strongly correlated with humidity (Fig. 1f), demonstrating that greater hydrogen protection occurs when the matrix is fully dried. This is indicative of hydration-dependent immobilization and condensation of the CAHS D gel during vitrification. We also studied the kinetics of biological material embedded in a vitrified CAHS D gel, and found the client was also immobilized in a hydration-dependent fashion; clients within gels that were kept at high RH showed more mobility than those stored at low RH. For further detail on this, see supplemental figure 1 and supplemental text. These data indicate that the strength of the matrix formed by a dried CAHS D gel, and its ability to immobilize molecules within it (water or a client), are dependent on the hydration state of the solid.

Simulations & structural measurements support a dumbbell-like conformational ensemble for CAHS D. CAHS proteins are highly disordered,[5,19] so standard structural studies to understand the structural biases of monomeric protein are not possible.[5,19] Thus, 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, while transient beta sheets are observed in the N- and C-terminal regions (Fig.2a, S2a,b). To validate our simulations, we performed SAXS of a monomeric CAHS D. The radius of gyration (R_g) – a measure of global protein dimensions – 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 extended nature of the LR is likely due to a high density of well-mixed negatively and positively charged residues (Fig. 2b, S2c). Consistent with simulation and SAXS results, bioinformatic analysis predicts that the LR can form transient alpha-helices, leading to an amphipathic helical structure. Similar to other helical desiccation protectants, the transient helices that form in the linker are predicted to have a hydrophobic face and a charged face[19] (Fig 2c). The N- and C-termini are predicted to form beta-sheeted interactions (Fig. 2a,g). Circular dichroism (CD) spectroscopy confirms the largely disordered nature of full-length CAHS, with some propensity for alpha helical and beta-sheet formation (Fig. 2d). CD spectroscopy performed on the LR, N-, and C-terminal regions in isolation

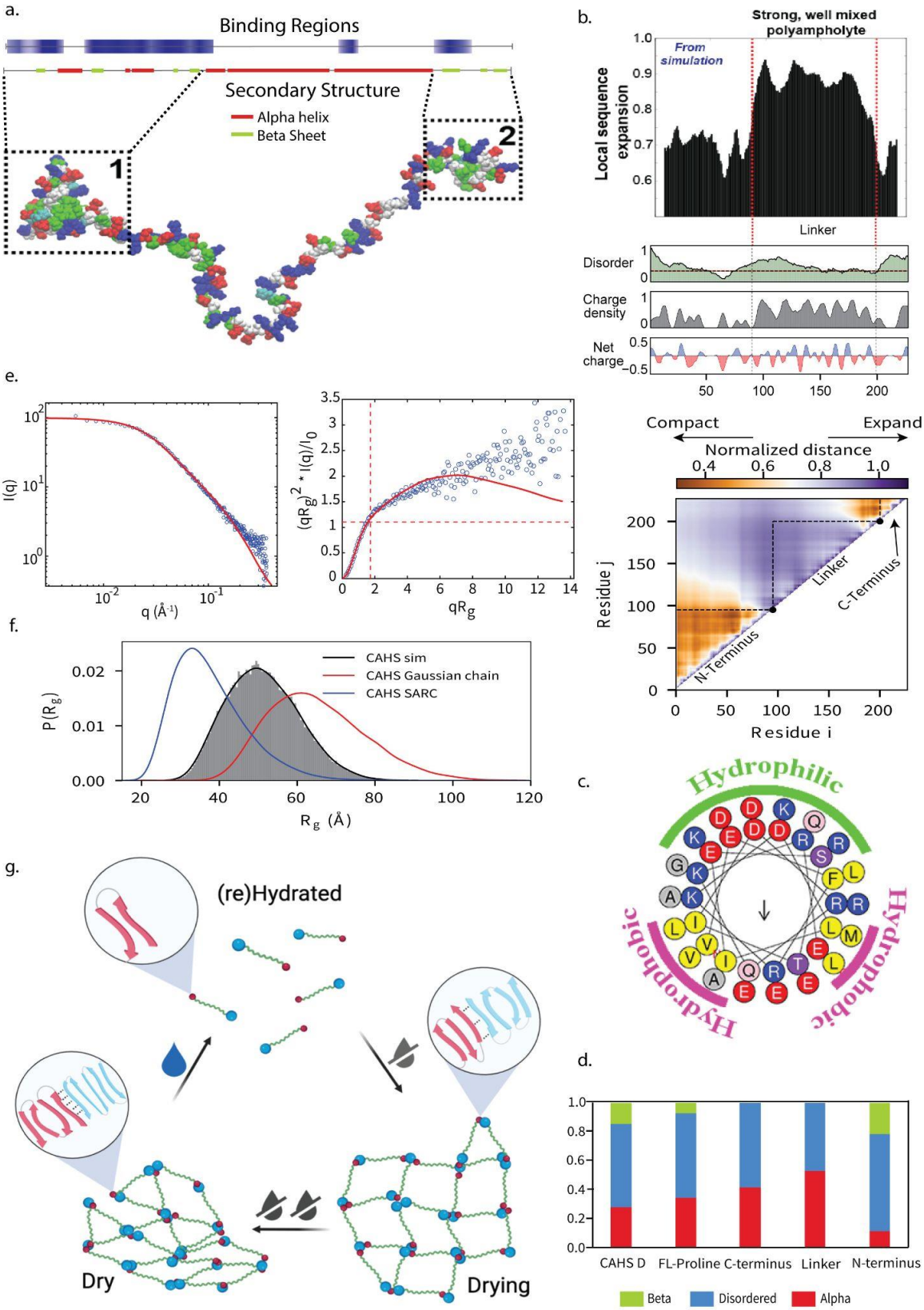

Fig 2. **CAHS D gelation is likely driven by polymerization akin to the stickers and spacers model.** a) Bioinformatic predictions of secondary structure characteristics and representative conformational global ensemble model of wildtype CAHS D. The representative ensemble structure shows the extended nature of the central linker, as well as the two terminal sticky ends. b) Top panel illustrates the relationship between sequence expansion, disorder, charge density, and net charge of CAHS D. Heat map (middle) shows normalized distance between residues as predicted in simulations. Blue indicates expansion while orange indicates compaction. The local expansion of the sequence is related to the distribution of charged residues shown in the top panel. Higher charge density correlates with a more expanded structure, as seen in the modeling. c) Helical wheel plot (HeliQuest, plothelix) demonstrating the distribution of charged faces in the 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 alpha helix, disorder, and beta sheet secondary structure propensity determined by CD spectroscopy for CAHS D, FL-Proline, C-terminus, Linker and N-terminus. Highest beta sheet content is found in the N-terminus. FL-Proline shows reduced beta-sheet content relative to wildtype, indicating that proline disruptions in the C-terminus reduced beta-sheet propensity. 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), all of equal linear size. g) Proposed mechanism of gelation for WT CAHS D. As water is removed from the system, monomers polymerize through beta-beta interactions in the termini. These interactions are strengthened as drying progresses. Upon rehydration, CAHS D gel can easily depolymerize.

confirmed substantial helical content in the LR (~50%), and beta-sheet content in N-terminal helical region (~20%) (Fig 2d). Contrary to our expectations, we observed no residual structure by CD in the C-terminal region. We sought to test if the lack of beta sheet character seen in CD of the C-terminus in isolation could be due to the loss of local context provided by the full-length protein. We therefore generated a construct with three prolines, potent breakers of beta-sheet content [36,37], inserted into regions of the C-terminus bioinformatically predicted to form beta-sheets (Fig S3a). CD on this construct (FL-Pro) resulted in a full length protein with ~50% reduced beta-sheet content (Fig. 2d; Fig. S3c). Overall these data support the predictions that CAHS D exists in a dumbbell-like ensemble consisting of 'sticky' beta-sheet containing termini held apart by an extended alpha-helical linker. [36,37]

CAHS gels form through inter-protein interactions of beta-sheeted termini.

A protein with the properties described above would be an ideal candidate for gelation via the 'stickers and spacers' model of polymerization[38]. In this model, discrete sites along a protein are designated as stickers (regions that contribute attractive intermolecular interactions) or spacers (regions that do not). In the context of CAHS D, the N- and C-terminal regions can be considered stickers, while the LR is considered a spacer. Moreover, the spacer has a large effective solvation volume, suppressing phase separation and promoting a sol-gel transition, in line with prior work.[39] Based on our biophysical characterization of the monomeric protein, we hypothesize that gel formation of CAHS D occurs through inter-protein beta-beta interactions mediated between termini (Fig. 2g). Moreover, we predict that intra-protein terminal interactions, which would suppress self-assembly through valence capping, are reduced due to the extended nature of the LR (Fig. 2c)[40].

If this model of gelation is correct, then disrupting the dumbbell-like ensemble of CAHS D should lead to a loss of polymerization. To test whether gelation of CAHS D relies on its dumbbell-like ensemble

we generated a range of CAHS D variants (Fig. 3a). Consistent with our hypothesis, all variants that are predicted to disrupt the dumbbell-like ensemble resulted in a loss of gelling potential (N, LR, C, NL1, CL1; Fig. 3a). Unexpectedly, variants that replaced one terminal region for another (NLN and CLC; Fig. 3a) did not form strong gels at the concentrations tested. The diffusion of water for all variants was tested from 1-20 mg/mL, and linear regressions were determined for the full concentration range of each variant. When compared at a uniform molar concentration of 1mM, the ΔDC linear trends for all variants except NLN were significantly different from that of CAHS D (Fig. 3d). Despite the lack of observed gelation in NLN and CLC, these variants still reduced diffusion of water dramatically. In fact, the ΔDC of NLN was not statistically different from CAHS D, yet NLN did not gel (Fig. 3a,d). This result implies heterotypic interactions between N- and C-termini may be crucial for strong gel formation, implicating molecular recognition and specificity encoded by the termini.

Interestingly, the 2X Linker variant, which maintained the wildtype heterotypic termini, but doubled the length of the linker gelled rapidly at concentrations below the gelling point of the wildtype protein (5 g/L and 15 g/L, respectively). We believe the length of the linker may tune the gel point by determining the monomeric molecular volume, and consequently setting the overlap concentration, a key determinant of the gel point.

Finally, we tested the gelation of our full-length C-terminal proline mutant (FL-Proline; described above; Fig. S3). As mentioned above, prolines are potent breakers of beta-sheet secondary structure and introduction of these three prolines results in an ~50% loss of beta-sheet content, as seen by CD spectroscopy. Consistent with our hypothesis that beta-beta interactions mediate gelation, our FL-proline variant did not gel at any concentrations tested (Fig. 3a).

Fourier Transform Infrared (FTIR) spectroscopy is a powerful tool to study secondary structure compositions in a liquid, gel, and even dried state. We employed FTIR to determine the impact of hydration levels on a hydrated CAHS D gel, and desiccated CAHS D glasses. We examined desiccated CAHS D glasses stored different ambient RH, and hydrated gels that were never dried (Fig. 3b; Fig. S1a-c). The amide I' band (measured at 1600-1800 cm^{-1}) can be decomposed into Gaussian sub-bands, which are ascribed to different secondary structure contributions. This revealed that beta-sheet interactions (green and blue bars) were maximal at 95% humidity (Fig. 3b). In hydrated gels, gels at 100% RH, and low RH, beta-beta interactions accounted for less of the total amide band area. For fully hydrated gels, lower beta-beta contacts could indicate that water can reduce the strength of induced folding between the transient beta sheets of the termini. The observation that a similar loss in beta contacts occurs at very low RH raises the possibility that a level of optimal hydration exists for stabilizing beta-beta contacts within the matrix.

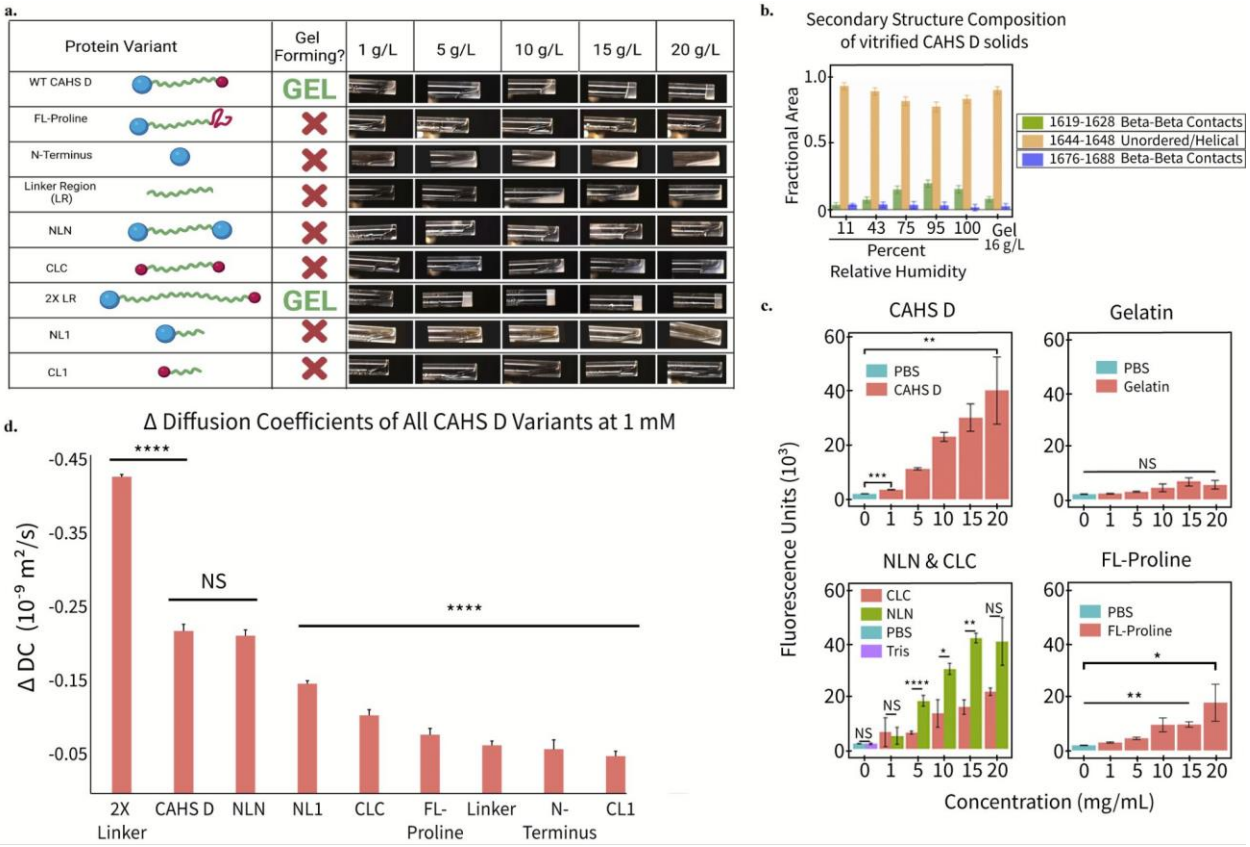

Fig. 3: Gelation of CAHS D relies on its dumbbell-like ensemble, and intermolecular beta-beta interactions. a) Schematic of the variant structure, the gel propensity of each listed concentration, and photos of the proteins in solution at indicated concentrations. Only CAHS D and the 2X Linker variant showed gelation. Future depictions of these constructs will 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. 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 beta-sheet structures. c) ThT fluorescence as a function of concentration for CAHS D, gelatin, NLN, CLC and FL-Proline. CAHS D shows concentration dependent increase in beta sheet structure. Gelatin fluorescence is no higher than background, indicating no beta-sheet structure. This emphasizes the relationship between beta-sheet stabilization and gelation in CAHS D. Once gelation is achieved, beta-sheets become more stable, reinforcing the stickers and spacers model of polymerization. NLN and CLC Tht fluorescence indicates that homotypic N-terminal interactions result in significantly higher beta sheet content than homotypic C-terminal interactions. Low ThT fluorescence in the FL-Proline construct is consistent with the disrupted nature of the C-terminus in this construct, which would reduce the potential for beta-beta interaction between molecules. Error bars represent standard deviation. Significance was determined using a Welch's t-test. d) ΔDiffusion Coefficients for all variants at 1 mM, calculated from linear fits for each concentration range (1-20 mg/mL). Gelling variants 2X Linker and CAHS D have the most slowed diffusion of all variants tested. 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 full concentration range of each variant, compared to that of CAHS D.

All experiments presented used a minimum of 3 replicates. Asterisks represent significance relative to the wildtype level of protection. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, NS is not significant.

This may relate to the need for higher stability while the matrix is undergoing drying, and when being rehydrated.

To further investigate the role in beta-beta interactions mediating polymerization of CAHS D, we assayed CAHS D solutions at ranging concentrations with thioflavin T (ThT), which is a commonly used fluorescent indicator of amyloid fibrils [41,42] and can report on beta-beta interactions. We observed increases in ThT fluorescence intensity as a function of CAHS D concentration, with the most dramatic increase taking place between 5 g/L and 10 g/L, suggesting nanoscale polymerization of CAHS D monomers prior to macroscale gelation (Fig. 3c). This is in contrast to ThT labeling of gelatin, which did not show concentration dependent changes in Tht fluorescence and was not significantly different from control samples (Fig. 3c). ThT labeling performed on NLN and CLC variants showed a concentration dependent increase of beta-beta interactions for both constructs, similar to CAHS D (Fig 3c). FL-Proline also showed beta-beta interactions, which were, as expected, at a lesser extent than variants with two folded termini. ThT fluorescence observed in the variants suggest that nanoscale polymerization events are occurring between monomers, even with a lack of observed sol-gel transition. The finding that ThT fluorescence is significantly higher in NLN than in CLC samples is consistent with bioinformatic predictions (Fig. 2a) and CD experiments (Fig. 2d), showing that the N-terminus contains a higher proportion of residues that can form beta-sheets.

Holistically, our CD, FTIR, and ThT labeling data show that CAHS D gelation is mediated by inter-protein beta-beta contacts formed through the interaction of the termini, consistent with the 'stickers and spacers' model of polymerization. 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. This is important for ensuring that inter-protein interactions, and not intra-protein interactions, dominate.

Polymerization of CAHS D correlates with better protection against protein unfolding, but not aggregation, during desiccation.

With a molecular understanding of how to induce or prevent gelation of CAHS D, we set out to determine how CAHS D gelation influences robust protection of biological material during desiccation. Extreme drying can impart a number of stresses on biological systems,[19,21,43] with protein dysfunctions being a major set of common perturbations.[19,21,43] The two prevalent, and non-mutually exclusive, forms of protein dysfunction during desiccation are protein unfolding and protein aggregation.[19] To assess protection against protein unfolding during drying, an enzymatic lactate dehydrogenase (LDH) activity assay is commonly employed.[5,20,44] LDH is a 140 kDa tetrameric enzyme that unfolds and becomes irreversibly non-functional during desiccation.[5,20,45] Furthermore, no detectable levels of aggregation of LDH have been associated with its non-functional, desiccated, state.[44,46] In contrast to LDH, the enzyme citrate synthase (CS) is known to undergo aggregation, and become non-functional as a result, after successive rounds of dehydration and rehydration.[46] We used our gelling and non-gelling variants of CAHS D (Fig. 3a) along with LDH and CS assays to determine the extent to which

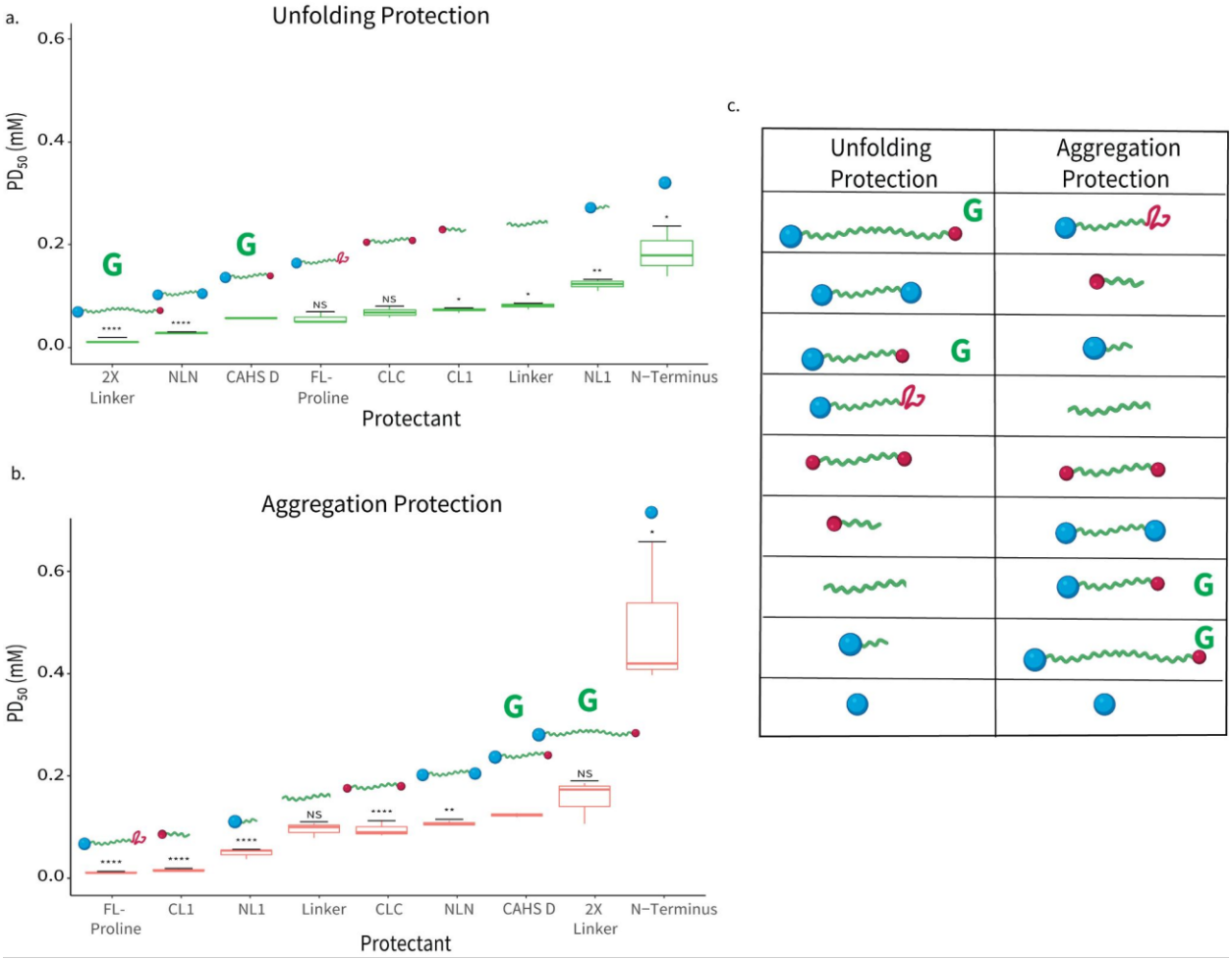

Fig. 4: Gelation promotes protection against protein unfolding, but not aggregation, during desiccation. a) PD50 values (shown left to right, in order from best to worst) of each variant tested in the unfolding assay. Lower PD50s correspond with better protective capability. b) PD50 values (shown left to right, in order from best to worst) of each variant tested in the aggregation assay. 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. All activity assays used 3 replicates. Significance was determined using a Welch's t-test, of variants protection relative to that of wildtype. Asterisks represent significance relative to the wildtype level of protection.*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, NS is not significant.

gelation is required for robust protection against protein unfolding and aggregation, respectively.

Sigmoidal curves display the percent of LDH or CS enzymatic activity that was protected as a function of concentration of each individual CAHS D variant (Fig. S4). Using these sigmoidal curves, a protective dose 50% (PD50) value was calculated for each variant (Fig. 4a,b). Lower PD50 values indicate that less protectant is needed to preserve 50% of an enzyme's activity, and thus lower values on our PD50 plots indicate better protective capacity (Fig. 4a,b).

The protection data from our LDH assay are largely consistent with the idea that gelation promotes protection against protein unfolding. In this assay, wild type CAHS D

and its variants that gel, or may form nanopolymers (Fig 3c), out-performed all other variants (Fig. 4a). Molecule for molecule, the variant that prevented LDH unfolding best was the 2x Linker, which was more protective *and* gels at a lower concentration than the wildtype CAHS D protein (Fig. 4a). Consistent with the idea that gelation and polymerization contribute to prevent protein unfolding, variants that do not gel performed the least well. Variants with little-to-no linker content (NL1 and N) performed worst at protecting from unfolding. These results suggest that the extended LR may be the protective unit of CAHS D in regards to unfolding, but that gelation, which requires terminal regions, also helps to promote stabilization of LDH during drying.

In contrast to our LDH unfolding assay, variants of CAHS D tested in our CS aggregation assay showed markedly different protective capacities. Surprisingly, gel formation seemed antagonistic to the ability to protect against aggregation, and instead the presence of un-gelled/ unpolymerized LR emerged as the primary determinant of CS aggregation (Fig. 4b). To our surprise the 2x linker variant — which is most prone to gelation, prevented LDH unfolding the best, and contains the most LR residues — performed worst out of all variants that contain a portion of the LR (Fig. 4b).

Given that both assays highlight the importance of the LR for protection, we draw from these results that gel formation sequesters CAHS D, impeding its ability to prevent aggregation when aggregate dimensions exceed the mesh-size of the gel. In contrast, the smaller LDH protein can be embedded within the gel, stabilizing it. Taken together, our LDH and CS assays show that the sequence features of CAHS D that optimally provide protection against unfolding are distinct from those that prevent aggregation (Fig. 4c), which suggests that the mechanisms underlying the prevention of these unique forms of protein dysfunction are distinct.

The mechanism most often attributed to the prevention of desiccation-induced protein aggregation is the molecular shielding hypothesis.[46] This hypothesis posits that protein aggregation during desiccation could be prevented by protectants that act as disordered spacers, either in solution or through direct interaction with client proteins (or both). These disordered spacers would impede molecular interactions between aggregation prone molecules. Shielding proteins also perform better when they have some means of interacting with their client proteins. Results from our CS aggregation assays are consistent with this mechanism; long, non-gel forming variants that include protein interaction regions are best at preventing aggregation, while conversely variants with the lower capacity for aggregation protection have the highest propensity for gel formation.

Water coordination is a good predictor of protection from protein unfolding, but not aggregation.

Various hypotheses surrounding the molecular basis for biological protection during desiccation have emerged over the last four decades. Among them, the vitrification hypothesis states that a protectant molecule functions by slowing the molecular motion in a system as water is lost, leading to the eventual formation of a vitrified solid (*i.e.,* a glass). Although tardigrades and CAHS proteins form vitrified solids when dried, [5,24,25] the effect of water loss and gelation of a CAHS D system on diffusion has not been assessed.[5,24,25] To examine

this, TD-NMR experiments were performed for CAHS D variants over a range of concentrations to measure the diffusion of water in the system (Fig. 5a). CAHS D variants that showed the greatest ability to slow the diffusion of water (2X Linker, NLN, and WT CAHS D) also prevented protein unfolding best in our LDH assay (Fig. 5d). Conversely, those same variants under-performed in our CS aggregation assay (Fig. 5d). Aside from the top-performing variants, diffusion is not a good predictor of unfolding protection, and does not correlate at all with protection from aggregation.

Since slowed diffusion could not fully account for the protective trends we observed in either our unfolding or aggregation protection assays, we looked to other hypotheses regarding how a protectant may mitigate desiccation damage. The water replacement, water entrapment, and preferential exclusion are three major hypotheses addressing how a protectant may compensate for the loss of a stabilizing hydrogen bond network (HBN) experienced by a client protein during desiccation stress. These propose, respectively, that a protectant either replace the HBN themselves[47], preserve a layer of water between a client and protectant to maintain the HBN[48], or exclude water from the protectant to crowd the client and maintain their HBN[49,50]. It should be noted that *a priori* these hypotheses regarding HBN stabilization are not mutually exclusive with the vitrification or molecular shielding hypotheses discussed earlier. For example, a protectant might help slow molecular motion through increasing intracellular viscosity *and* help coordinate residual water molecules to hydrate proteins, and form a vitrified solid when fully desiccated.

In an attempt to understand the protection trends we see in our unfolding and aggregation assays, we used TD-NMR to measure T_2 relaxation of CAHS D and its variants (Fig. 5b) to gain insight into how these protectants interact with water. Beyond diffusion, TD-NMR can also yield information regarding the coordination of water molecules in a system.[51–55,56] Water coordination occurs in phenomena like hydration shells, where intermolecular forces between a protein and its environment cause an organization of water molecules around the protein. More dramatic water coordination is generally termed structuralization, which is characterized by strongly organized water, in which the freedom of motion of a water molecule is restricted, molecular motion is reduced, and exchange of individual water molecules of the hydration layer are reduced or slowed.[57–62] T_2 (also known as spin-spin or transverse) relaxation (see supplemental text for further details) reflects the rate at which transversely magnetized nuclei decay to equilibrium, or lose coherence with one another, during the measurement period (increased relaxation kinetics; Fig. S5a). T_2 is represented as a distribution of the decay rate over time, with shorter (faster) relaxation times corresponding with strongly coordinated or structured water molecules, and longer (slower) relaxation indicating less coordinated water found in bulk liquids (Fig. S5b).

Among our CAHS D variants, we observed a remarkable diversity of water coordinating properties (Fig. 5b, S5c). The 2X Linker variant, which performed best in protecting

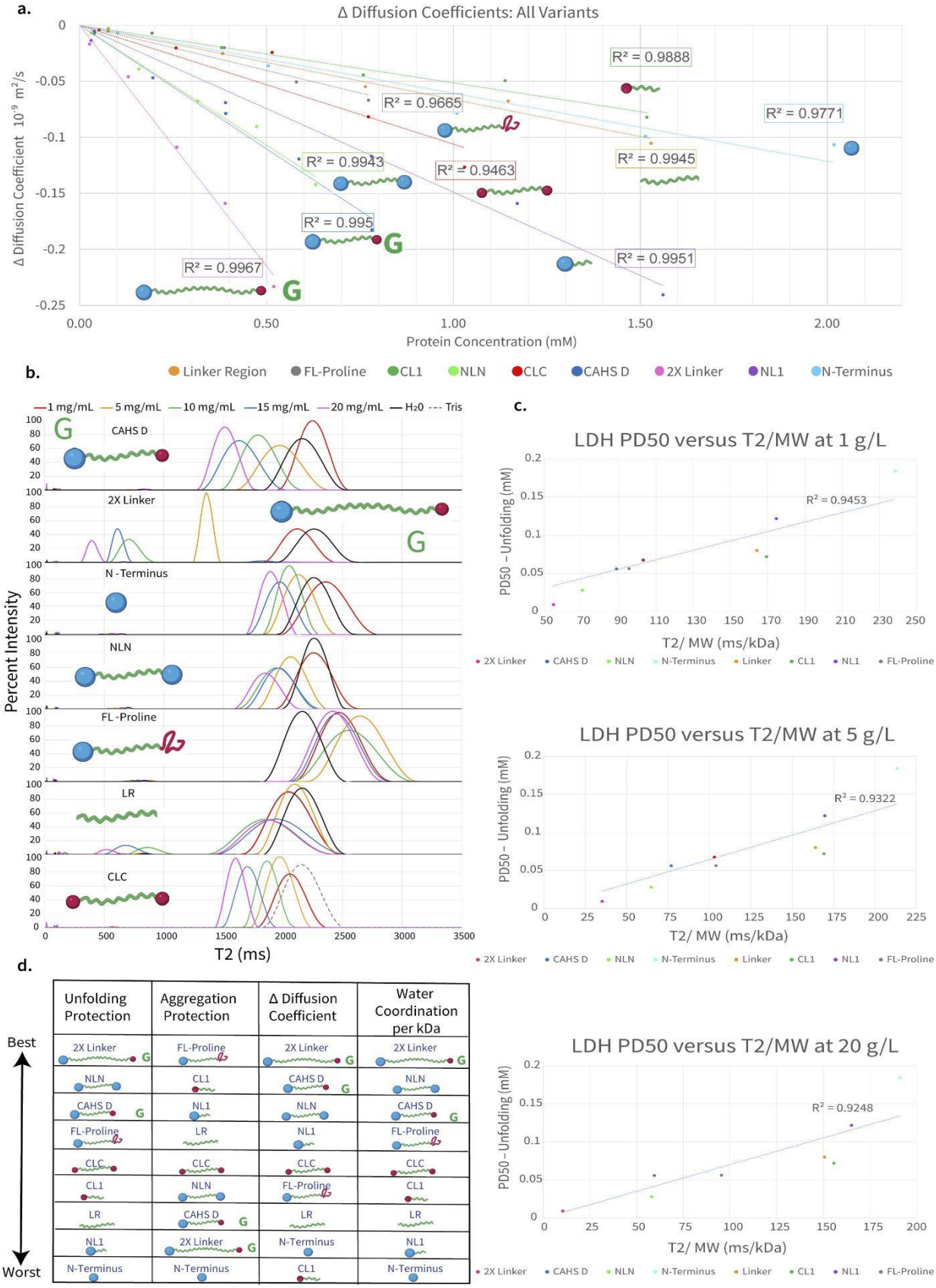

Fig. 5: Slowed diffusion and coordination of water underlie CAHS D's ability to promote protection against protein unfolding. a) ΔDC for all variants. b) T_2 distributions of variants WT CAHS D, 2X Linker, N-terminus, NLN, LR, and CLC. (All variants not shown here can be found in Fig. S5c). 2X linker, CAHS D, and NLN show the most shifted T_2 peak midpoints as a function of concentration. N-terminus also shows a concentration dependent peak shifting, however the degree to which T_2 is shifted at high concentration is not as dramatic as for the variants just described. c) Unfolding PD50 as a function of water coordination per kilodalton for the 1 g/L, 5 g/L and 20 g/L samples. 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 (d). d) Ranked table of Unfolding and Aggregation PD50, ΔDC at 1mM, and water coordination per kilodalton. 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 water coordination/kilodalton ranking shows a striking relationship. Water coordination per kilodalton is a good predictor of unfolding protection. However, no such relationship is seen for aggregation protection.

LDH from unfolding and very poorly at preventing CS from aggregating, has the most dramatically shifted T_2 distribution at 20 g/L, indicating that water molecules are highly structured and caged within a dense matrix (Fig. 5b). CAHS D also showed concentration dependent water coordination, albeit to a lesser degree that the 2X Linker variant (Fig. 5b). Concentration-dependent water coordination was demonstrated to some degree by the variants CLC, NLN, and, surprisingly, N terminus (Fig. 5b). Interestingly, our FL-Proline variant had T_2 values that shifted to the right with increasing concentration (Fig.5b), suggesting that disrupting of the C-terminus may lead to mild water exclusion, possibly through the exposure of hydrophobic residues in its disrupted C-terminus, which comprise 27.78% of that region.

To assess the relationship between water coordination and protection, we plotted the PD50 for our LDH and CS assays as a function of T_2 peak midpoints normalized to the variant's MW (T_2/MW) at 1, 5, and 20 g/L for all variants (Fig. 5c). 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. Normalizing the T_2 peak midpoint by molecular weight results in a strong correlation with prevention of protein unfolding, but shows no correlation with prevention of aggregation (Fig. 5d; Fig. S5d). We were surprised to find that the ability of a CAHS D variant to prevent protein unfolding during desiccation is directly predicted by the amount of water it can coordinate per kilodalton.

Discussion

In this work we set out to understand the mechanism of CAHS D desiccation protection. We discovered that CAHS D forms thermo-reversible physical gels, an observation explained by a model in which attractive interactions encoded by the N- and C-terminal subregions facilitate intermolecular polymerization. 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. Variants that do not gel show the best capacity to prevent protein aggregation, while, surprisingly, those that gel protect poorly against aggregation. Conversely, gelling variants promote protection against protein unfolding, while those variants that do not gel provide the least protection from denaturation during drying. Taken together, we propose that CAHS functions as a Swiss-army knife, offering multiple protective capabilities

though distinct modes of homotypic and heterotypic interactions.

How does CAHS D form a gel?

Our results provide direct molecular insight into the structural and non-structural determinants of gel formation. To our surprise, the majority of the CAHS D variants tested were unable to form a gel, implying gelation is not a generic property of CAHS D-like proteins, and instead supporting a model in which regions of the heterotypic termini allow for a degree of specificity which may help tune the gelation of the system.

A possible mechanism for gelation is one in which coiled-coils form via the transiently helical linker-region, analogous to centrosomal assemblies [63,64]. Given the linker-only variant fails to gel (Fig. 3A), this appears an unlikely mode of assembly. An alternative model is one in which - like gelatin - non-specific hydrophobic helical interactions drive assembly [65,66]. However, differences in network structure, ThT sensitivity, and T_2 relaxation rates for water all implicate a CAHS-based mode of assembly distinct from gelatin.

Bioinformatic predictions, all-atom simulations, SAXS, FTIR and CD spectroscopy, and ThT fluorescence assay all support a model in which the CAHS termini transiently flicker into and out of beta-sheeted conformations. Moreover, this beta sheet content increases upon gelation. In conjunction with the dumbbell-like conformational ensemble from all-atom simulations, we interpret this data in terms of a stickers and spacers model, whereby termini can be considered stickers while the linker is a single, well-solvated spacer.[38] The presence of a well-solvated spacer matches predictions that propose spacer solvation state can tune the balance between phase separation with gelation vs. sol-gel transitions without phase separation. As such, we propose CAHS D has evolved to undergo sol-gel transitions while simultaneously suppressing liquid-liquid phase separation.

Since only WT CAHS D and the 2x linker variant showed discernible gelation under the concentration ranges we tested, we conclude that strong gelation is driven by heterotypic interaction between the N- and C-termini. Surprisingly, variants in which the N-terminal domain is replaced by the C-terminal domain (CLC) or vice versa (NLN) failed to undergo gelation under the conditions tested. This observation hints at sequence-specific molecular interactions that may prevent unwanted CAHS D interaction with other cellular clients. Finally, the extended state of the highly charged linker supports a model in which it impedes intramolecular interaction. We note that given the CAHS D linker is among the most well-mixed high charge sequences in the tardigrade proteome, this may represent a specific evolutionarily selected trait.

Taken together, our data support a model in which CAHS D forms a novel molecular topology that encodes a mode of intramolecular interactions tuned to extended, networked gels.

Is gelation important for desiccation tolerance?

Our CAHS D variants, in conjunction with unfolding and aggregation assays, provide disparate results regarding the importance of gelation in terms of providing desiccation protection to proteins. Variants that do not gel show the best capacity to prevent protein aggregation, while those that gel protect poorly against aggregation. Conversely, gelling variants promote protection against

protein unfolding, while those variants that do not gel provide the least protection from denaturation during drying.

Protein aggregation and unfolding are not the only perturbations induced by desiccation in biological systems. With this in mind, CAHS gelation may serve to prevent other types of damage. For example, gelation may help to maintain spatial organization of the cytoplasm and/or provide spacing between membranes of organelles that would otherwise collapse and fuse with one another during drying. It is also possible that other TDPs may work in concert with CAHS D, to modulate its activity in a tardigrade experiencing drying, or other environmental stresses.

What promotes protection against aggregation?

Our data shows that protection from aggregation is likely to be mechanistically distinct from protection from unfolding. No trends observed for unfolding matched trends seen in aggregation assays. In fact, ΔDC at 1 mM protectant (Fig 3d) was almost perfectly anti-correlated with aggregation protection, seeming to indicate that slower diffusion inhibits aggregation protection. This would seem to rule out the classic vitrification hypothesis from participating in CAHS D-mediated aggregation protection.

By contrast, the molecular shielding hypothesis describes protectants as entropic springs, which physically separate clients from one another. If we envision a molecular shield as primarily functioning through aggregation protection, rather than maintaining folding, then a mechanism involving slow diffusion may not be advantageous. In order to function well, a shielding protein may need only to take up space. If we consider this in relation to gelation, we might initially expect that the gel structure would serve this purpose well. However, gel formation would also inhibit the dynamic movement of protectants around clients, which would limit a protectant's ability to isolate clients from one another. Additionally, it is possible that aggregation prone proteins could become trapped together within the space of the gel fiber, with no mobile protectant molecules available to keep the clients separate. This is consistent with our observations that gelling protectants perform poorly at preventing aggregation. We believe this may be related to the size of the mesh in which clients are embedded, as well as the size of the client itself. LDH is a large tetramer of approximately 140 kDa (R_g = 3.1 nm) , while CS is smaller, at roughly 85 kDa (R_g = 2.3 nm). It is possible that the gel meshwork formed by wild type CAHS D and its gelling variants accommodate too many CS molecules, allowing for buildup of aggregates over several cycles of drying/rehydration.

What promotes protection against protein unfolding during desiccation?

Perhaps surprisingly, protection from unfolding is not limited to variants that can gel. In all variants, water coordination per kilodalton was most correlated with unfolding protection. For example, the NLN variant did not form gels in the concentration range we tested, and its ΔDC was nearly identical to WT CAHS D, yet it had a significantly lower PD50 for unfolding protection. FL-Proline and CLC also closely matched in water coordination per kDa, and their protective capability as gauged by LDH PD50 were within 0.01 mM.

Taken together, this data 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. Based on our work, we propose that water entrapment, wherein the protectant traps a layer of hydrating water between itself and the client protein,[48,50] is the major mechanism driving protection from unfolding in CAHS D. This may occur within the larger framework of either vitrification or molecular shielding.

Different protective mechanisms for different phases of drying.

Surviving desiccation is not about surviving a singular stress, but rather surviving a continuum of inter-related stresses that can be broken down into 'drying,' 'dry,' and 'rehydrating' phases.

During drying, the need to prevent aggregation may be more pressing than the need to maintain hydration around client proteins, in order to mitigate a cascade of aggregation, where sensitive proteins that may destabilize early lead to increasing damage throughout the cell by aggregating with functional or partially destabilized proteins. The formation of protein assemblies from cellular aggregates has been implicated in many examples of proteostatic dysregulation.[67] Monomeric CAHS D could be performing important shielding functionalities, sequestering denatured proteins from the rest of the cell's proteins. This would help preserve functional proteins during the critical early stages of drying, when a cell is adapting to water loss, by preventing the formation of large, damaging protein aggregates within the cell.

As further water is lost and protectant concentration increases, CAHS D function transitions from a primarily shielding mechanism to primarily hydrating and slowing mechanism of desiccation protection, through gelation. Small, micro-scale polymers of CAHS D would form, slowing down molecular motions. As polymers grow and combine into a network, clients become locked into the meshwork of the gel. This polymer network could provide a bridge to the vitrified solid or 'dry' state; it has slowed motion more than a liquid, but not as much as is expected in a vitrified solid, and it serves to coordinate any remaining water to the surface of client proteins. This mitigates the water loss experienced by the clients, allowing them to retain their native fold for as long as possible. By the time dehydration overwhelms the protectant's capacity to hold water, a vitrified solid has formed. At this point, even though the client's native fold may no longer be stabilized by a hydrogen bond network, unfolding *still* does not occur because of the degree to which molecular motion has been reduced.

The final phase of surviving desiccation is to rehydrate and return to active biological functions. The role of CAHS proteins during rehydration, if any, has yet to be elucidated, but the ability of these proteins to quickly

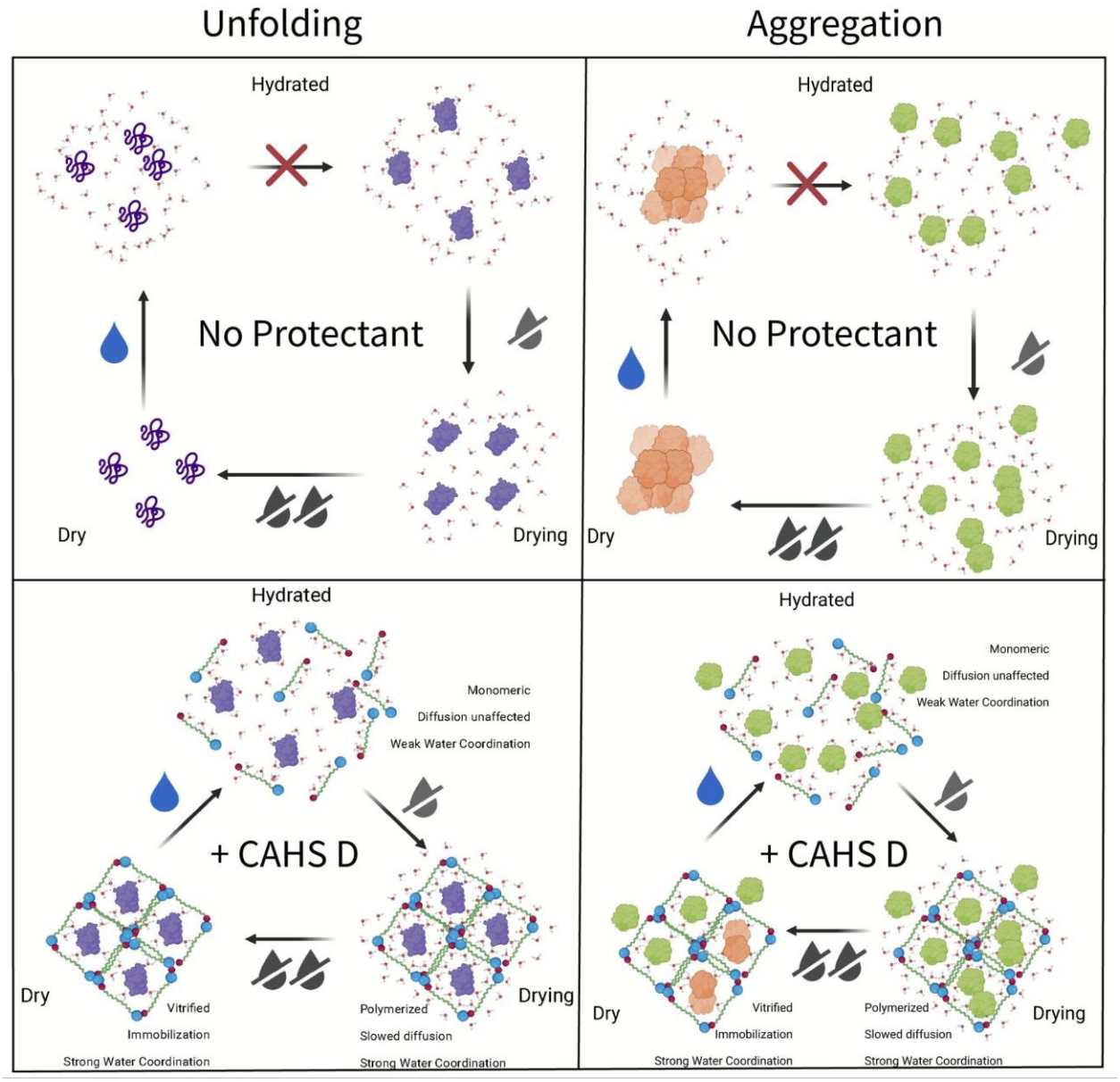

Fig. 6: Proposed mechanism for CAHS D mediated protection against desiccation induced protein aggregation and unfolding. The left hand panels show what occurs in the unfolding assay, while the right hand panels show the aggregation assay Unfolded proteins are displayed as squiggly lines, while aggregated proteins are displayed as orange clumps. Top panels of each display the system without protectants, while the lower panel of each shows the system with wildtype CAHS D. As water is lost, polymerization begins in samples with CAHS D. Water is coordinated along the linker and to some degree the N-terminus of CAHS D. In the unfolding assay, only a single LDH molecule can fit within the pores formed during polymerization, while multiple CS molecules can be present in the same pore. This is due to the relative sizes of the client proteins in each assay. As samples become fully dry and enter a vitrified solid state, some water remains coordinated by the linker, but most water has been lost. Upon rehydration, the matrix readily depolymerizes and releases molecules from the gel structure.

resolvate from a dried state is likely an essential step in efficiently returning the rehydrating system to activity. The mechanisms underlying this rapid re-solvation are intriguing, as many amyloids seen in pathologies are not easily resolved. Our FTIR data, showing that beta-sheet structure in vitrified CAHS D glass is optimal at low levels of hydration, and decreases as hydration of the glass increases, indicates that quick resolvation of CAHS D could be due to a hydration dependent disruption of beta-beta interactions. Though the mechanism underlying CAHS D re-solvation remains to be determined, the ability of this protein to break-down its glassy matrix and release the protected, functional biological material embedded within it, is likely essential for robust anhydrobiosis in tardigrades. Schematic representations of our model for how CAHS D and its variants behave during the drying, dry, and (re)hydrated phases of anhydrobiosis are shown in Fig. 6 and Fig. S6.

Multiple mechanisms likely combine to promote desiccation tolerance

Since Antonie van Leeuwenhoek first found that he could reanimate ‘animicules’ by hydrating dry dust with sterile water, the question of how life can persist in the absence of nearly all water has been one of the enduring mysteries of organismal physiology. As reviewed above, several compelling hypotheses (molecular shielding, vitrification, water replacement, water entrapment, and preferential exclusion) have been put forward to explain the potential mechanism underlying anhydrobiosis. Often, these hypotheses are considered in isolation, but many of these hypotheses are not mutually exclusive, meaning that different mediators of desiccation tolerance, or even a single mediator, could work through multiple mechanisms to promote survival during drying. Previously it has been reported that tardigrades and their CAHS proteins vitrify[5,24,25] and consistent with this we see that during gelation of CAHS D viscosity increases and diffusion is slowed. Simultaneously, we observe that CAHS D gelation leads to strong coordination of water. Both these properties: slowed diffusion and water coordination, correlate with the ability of CAHS D to prevent protein unfolding - suggesting that this phenomenon is mediated through mechanisms related to both the vitrification and water coordination hypotheses. Furthermore the ability of the FL-Proline variant, which is locked into a monomeric state, to very effectively prevent aggregation of citrate synthase implies that WT CAHS D may work as a molecular shield when the concentration is lower than its gelation point.

Thus, our work provides empirical evidence of not only multiple mechanisms likely contributing to desiccation tolerance, but also that these mechanisms may be mediated by a single molecule.

CAHS proteins function in multiple forms of stress tolerance

Interestingly, CAHS proteins not only promote desiccation tolerance, but are also known to provide protection against other forms of abiotic stress, such as freezing. While freezing and drying share some overlap in terms of stresses induced, for example loss of liquid water, there are also clear differences. Not only are the perturbations induced by these stresses generally distinct, but the global transcriptomic responses to freezing and drying in tardigrades are highly divergent[5]. This is important because owing to their lack of intramolecular bonds and relatively high

surface area, IDP ensembles are in general highly influenced by their environment. Given that freezing and drying result in changes to the intracellular environment both in terms of abiotic (e.g., temperature) and biochemical (e.g., small molecules) there is good reason to think that the ensemble state and functional properties of CAHS proteins would differ between these conditions. Given the ability of CAHS D to act through multiple mechanisms during desiccation, it will be interesting to see what similarities and differences there are in terms of how this 'molecular Swiss army knife' promotes different forms of abiotic stress tolerance.

Acknowledgements

C.S.H., K.H.N, S.B., C.A.C., S.KC, and T.C.B. are supported by DARPA (W911NF-19-2-0019).
S.S. and F.Y. were supported by the NIH under award R35GM137926 to S.S.
Financial support from MIUR of Italy (RFO2019) is gratefully acknowledged by M.M., F.F. and G.V. G.V. thanks Lorena Rebecchi (University of Modena and Reggio Emilia, UNIMORE, Italy) for stimulating discussions and valuable advice.

Methods

Cloning
All variants and wildtype CAHS D were cloned into the pET28b expression vector using Gibson assembly methods. Primers were designed using the NEBuilder tool (nebuilder.neb.com), inserts were synthesized as gBlocks and purchased from Integrated DNA Technologies (idtdna.com).

Protein Expression
Expression constructs were transformed in BL21 (DE3) E. coli (New England Biosciences, Cat. #C2527H), 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 (Eppendorf Innova S44i) 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, Cat. P2714]). Pellets were stored at -80ºC.

Protein Purification
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 (EZFlow Syringe Filter, Cat. 388-3416-OEM). The filtered lysate was diluted 1:2 in purification buffer UA (8 M Urea [Acros Organics, CAS No. 57-13-6], 50 mM sodium acetate [Tocris CAS No. 127-09-3], pH 4). The protein was then purified using a cation exchange HiPrep SP HP 16/10 (Cytiva, Cat. 29018183) on the AKTA Pure 25 L (Cytiva, Cat. #29018224), controlled using the UNICORN 7 Workstation pure-BP-exp (Cytiva, Cat. #29128116). 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, Part No. 132724). 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 (Invitrogen Qubit 4 Fluorometer, REF Q33226), aliquoted in the quantity needed for each assay, lyophilized (Labconco FreeZone 6, Cat. 7752021) for 48 hours, then stored at -20°C until use. CLC was dialyzed in 2M 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 (Millipore-Sigma) 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, WD-4001-7) 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, 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) and

supporter coating (Cressington 108 Auto). 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). 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 (GE Lifesciences) 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 1M (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.[68] 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) 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.

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

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.

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

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.

Thioflavin T Assay

Proteins were dissolved in phosphate buffered saline, pH 7.2 (Sigma Aldrich, Cat# P4244). A solution of 400 µM Thioflavin T (Sigma Life Science, Cat# T3516) was prepared in Dimethyl Sulfoxide (Sigma Aldrich, D-8779), 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, Ref. 3596) and incubated for 15 minutes at room temperature in the dark. Fluorescence was measured using a plate reader (Tecan, Spark 10M) 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.[5] 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, Cat 10127230001) 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 (SAVANT Speed Vac Concentrator ###vacuum pump). 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 P0582], 1 mM NADH [Sigma-Aldrich 10128023001], pH 6). Enzyme kinetics were measured by NAD+ absorbance at 340 nm, on the NanodropOne (Thermo Scientific). The protective capacity was calculated as a ratio of NAD+ absorbance in desiccated samples relative to non-desiccated controls.

Citrate Synthase Protection Assay

The Citrate Synthase Kit (Sigma-Aldrich Cat. CS0720) was adapted for use in this assay. 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, Gelatin from Porcine Skin; G2500, Bloom 300) 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 Cat. C3260) 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 [SAVANT, SpeedVac SC110] 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, WD-4001-7) 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 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.) circulating a constant-temperature coolant. T_2 relaxations 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 software provided by Bruker. Each variant was measured for the full concentration range, along with a buffer only sample (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 only (water or 20 mM Tris pH 7) on the same day that T_2 data was collected.

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 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**.** Amide I' bands were decomposed into Gaussian components by using a locally developed least-squares minimization routine[69]; 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, while the areas and the widths were treated as unconstrained parameters.

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[70], 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 different relative humidity values, 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.[71] 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 dry matrix, we included the CAHS protein, the RC and the detergent belt of the RC formed by 289 LDAO molecules[71] 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.[72] 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.[73]

Proteome-wide bioinformatics

The full tardigrade proteome (tg.default.maker.proteins.fasta) was taken from https://github.com/sujaikumar/tardigrade as assembled previously.[74] The proteome file was pre-processed using protfasta (https://protfasta.readthedocs.io/), IDRs predicted with metapredict[75] and IDR kappa values calculated using localCIDER.[76,77] Metapredict identified 35,511 discrete IDRs

distributed across 39,532 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.

Dimension estimates for LDH and CP

Radii of gyration for LDH and CP were calculated from the tetramer structure in PDB:1I10 and for the monomeric CP in PDB:3ENJ.[78,79]

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.[80] The combination of ABSINTH and CAMPARI has been used to generate experimentally-validated ensembles for a wide range of IDRs [81–83].

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 [84].

References

1. Crowe, J. H. & Clegg, J. S. *Anhydrobiosis*. (Hutchinson Ross Publishing Company, 1973).
2. Dorone, Y. *et al.* A prion-like protein regulator of seed germination undergoes hydration-dependent phase separation. *Cell* (2021) doi:10.1016/j.cell.2021.06.009.
3. Jenks, M. A. & Wood, A. J. *Plant Desiccation Tolerance*. (John Wiley & Sons, 2008).
4. Boothby, T. C. Desiccation of. *Cold Spring Harb. Protoc.* **2018**, (2018).
5. Boothby, T. C. *et al.* Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation. *Mol. Cell* **65**, 975–984.e5 (2017).
6. Kondo, K., Kubo, T. & Kunieda, T. Suggested Involvement of PP1/PP2A Activity and De Novo Gene Expression in Anhydrobiotic Survival in a Tardigrade, Hypsibius dujardini, by Chemical Genetic Approach. *PLoS One* **10**, e0144803 (2015).
7. Wright, J. C. The tardigrade cuticle II. Evidence for a dehydration-dependent permeability barrier in the intracuticle. *Tissue Cell* **21**, 263–279 (1989).
8. Yathisha, N. S. *et al.* Vegetative desiccation tolerance in : biochemical and physiological responses. *Heliyon* **6**, e04948 (2020).
9. Erkut, C. *et al.* Trehalose renders the dauer larva of Caenorhabditis elegans resistant to extreme desiccation. *Curr. Biol.* **21**, 1331–1336 (2011).
10. 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).
11. Tapia, H. & Koshland, D. E. Trehalose is a versatile and long-lived chaperone for desiccation tolerance. *Curr. Biol.* **24**, 2758–2766 (2014).
12. Mitsumasu, K. *et al.* Enzymatic control of anhydrobiosis-related accumulation of trehalose in the sleeping chironomid, Polypedilum vanderplanki. *FEBS J.* **277**, 4215–4228 (2010).
13. Laskowska, E. & Kuczyńska-Wiśnik, D. New insight into the mechanisms protecting bacteria during desiccation. *Curr. Genet.* **66**, 313–318 (2020).

14. Lapinski, J. & Tunnacliffe, A. Anhydrobiosis without trehalose in bdelloid rotifers. *FEBS Lett.* **553**, 387–390 (2003).
15. 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).
16. Tripathi, R., Boschetti, C., McGee, B. & Tunnacliffe, A. Trafficking of bdelloid rotifer late embryogenesis abundant proteins. *J. Exp. Biol.* **215**, 2786–2794 (2012).
17. Tripathi, R. *Functional Characterisation of LEA Proteins from Bdelloid Rotifers*. (2012).
18. 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).
19. Hesgrove, C. & Boothby, T. C. The biology of tardigrade disordered proteins in extreme stress tolerance. *Cell Commun. Signal.* **18**, 178 (2020).
20. Piszkiewicz, S. *et al.* Protecting activity of desiccated enzymes. *Protein Sci.* **28**, 941–951 (2019).
21. Boothby, T. C. & Pielak, G. J. Intrinsically Disordered Proteins and Desiccation Tolerance: Elucidating Functional and Mechanistic Underpinnings of Anhydrobiosis. *Bioessays* **39**, (2017).
22. Kamilari, M., Jørgensen, A., Schiøtt, M. & Møbjerg, N. Comparative transcriptomics suggest unique molecular adaptations within tardigrade lineages. *BMC Genomics* **20**, 607 (2019).
23. Crowe, J. H., Carpenter, J. F. & Crowe, L. M. The role of vitrification in anhydrobiosis. *Annu. Rev. Physiol.* **60**, 73–103 (1998).
24. Boothby, T. C. Water content influences the vitrified properties of CAHS proteins. *Molecular cell* vol. 81 411–413 (2021).
25. 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).

26. 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).
27. McComb, D. W., Lengyel, J. & Barry Carter, C. Cryogenic transmission electron microscopy for materials research. *MRS Bulletin* vol. 44 924–928 (2019).
28. Martin, E. W. *et al.* A multi-step nucleation process determines the kinetics of prion-like domain phase separation. *Nat. Commun.* **12**, 4513 (2021).
29. Khanna, S., Singh, A. K., Behera, S. P. & Gupta, S. Thermoresponsive BSA hydrogels with phase tunability. *Mater. Sci. Eng. C Mater. Biol. Appl.* **119**, 111590 (2021).
30. Nikfarjam, S., Jouravleva, E. V., Anisimov, M. A. & Woehl, T. J. Effects of Protein Unfolding on Aggregation and Gelation in Lysozyme Solutions. *Biomolecules* **10**, (2020).
31. Li, S., Ye, A. & Singh, H. The effect of ultrafiltration on the acid gelation properties of protein-standardised skim milk systems. *Food Res. Int.* **146**, 110432 (2021).
32. 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).
33. Böni, L. J. *et al.* Hagfish slime exudate stabilization and its effect on slime formation and functionality. *Biol. Open* **6**, 1115–1122 (2017).
34. 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).
35. 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).
36. Williams, A. D. *et al.* Mapping abeta amyloid fibril secondary structure using scanning

proline mutagenesis. *J. Mol. Biol.* **335**, 833–842 (2004).

37. Imai, K. & Mitaku, S. Mechanisms of secondary structure breakers in soluble proteins. *Biophysics* **1**, 55–65 (2005).
38. 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).
39. 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).
40. Sanders, D. W. *et al.* Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. *Cell* **181**, 306–324.e28 (2020).
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. Boothby, T. C. Mechanisms and evolution of resistance to environmental extremes in animals. *Evodevo* **10**, 30 (2019).
44. 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).
45. Rani, R. *Lactate Dehydrogenase (LDH): Biochemistry, Function and Clinical Significance*. (2019).
46. Chakrabortee, S. *et al.* Intrinsically disordered proteins as molecular shields. *Mol. Biosyst.* **8**, 210–219 (2012).
47. Clegg, J. S., Seitz, P., Seitz, W. & Hazlewood, C. F. Cellular responses to extreme water

loss: the water-replacement hypothesis. *Cryobiology* **19**, 306–316 (1982).

48. 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).
49. Timasheff, S. N. Water as ligand: preferential binding and exclusion of denaturants in protein unfolding. *Biochemistry* **31**, 9857–9864 (1992).
50. Shimizu, S. & Smith, D. J. Preferential hydration and the exclusion of cosolvents from protein surfaces. *J. Chem. Phys.* **121**, 1148–1154 (2004).
51. 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).
52. 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).
53. Hydration properties and proton exchange in aqueous sugar solutions studied by time domain nuclear magnetic resonance. *Food Chem.* **132**, 1644–1650 (2012).
54. 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).
55. 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).
56. 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).
57. Adhikari, A., Park, W.-W. & Kwon, O.-H. Hydrogen-Bond Dynamics and Energetics of Biological Water. *Chempluschem* **85**, 2657–2665 (2020).
58. Raschke, T. M. Water structure and interactions with protein surfaces. *Curr. Opin. Struct. Biol.* **16**, 152–159 (2006).

59. 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).
60. 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).
61. 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).
62. Laage, D., Elsaesser, T. & Hynes, J. T. Water Dynamics in the Hydration Shells of Biomolecules. *Chem. Rev.* **117**, 10694–10725 (2017).
63. Woodruff, J. B. *et al.* Centrosomes. Regulated assembly of a supramolecular centrosome scaffold in vitro. *Science* **348**, 808–812 (2015).
64. Woodruff, J. B. *et al.* The Centrosome Is a Selective Condensate that Nucleates Microtubules by Concentrating Tubulin. *Cell* **169**, 1066–1077.e10 (2017).
65. Djabourov, M. Architecture of gelatin gels. *Contemp. Phys.* **29**, 273–297 (1988).
66. Djabourov, M., Lechaire, J. P. & Gaill, F. Structure and rheology of gelatin and collagen gels. *Biorheology* **30**, 191–205 (1993).
67. 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).
68. 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).
69. 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).

70. Barth, A. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta* **1767**, 1073–1101 (2007).

71. 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).

72. 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).

73. 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).

74. 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).

75. 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.

76. 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).

77. 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).

78. 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).

79. 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).

80. 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).
81. 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).
82. Martin, E. W. *et al.* Valence and patterning of aromatic residues determine the phase behavior of prion-like domains. *Science* **367**, 694–699 (2020).
83. 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).
84. McGibbon, R. T. *et al.* MDTraj: a modern, open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **109**, 1528–1532 (2015).
85. Barth, A. & Zscherp, C. What vibrations tell us about proteins. *Q. Rev. Biophys.* **35**, 369–430 (2002).
86. W. L. Butler, D. W. H. Higher derivative analysis of complex absorption spectra. *Photochem. Photobiol.* **12**, 439–450 (1970).
87. 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).
88. 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).
89. 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).

90. 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).
91. Bevington, P. R. & Bevington, R. R. *Data Reduction and Error Analysis for the Physical Sciences*. (McGraw-Hill College, 1969).
92. 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).
93. 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).
94. 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).
95. 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).
96. 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).
97. 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).
98. 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).

99. Cordone, L. *et al.* Internal dynamics and protein-matrix coupling in trehalose-coated proteins. *Biochim. Biophys. Acta* **1749**, 252–281 (2005).
100. Feher, G., Allen, J. P., Okamura, M. Y. & Rees, D. C. Structure and function of bacterial photosynthetic reaction centres. *Nature* **339**, 111–116 (1989).
101. 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).
102. 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).
103. Hailwood, A. J. & Horrobin, S. Absorption of water by polymers: analysis in terms of a simple model. *Trans. Faraday Soc.* **42**, B084 (1946).
104. 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).
105. 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).
106. 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).
107. 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).
108. Remy, A. & Gerwert, K. Coupling of light-induced electron transfer to proton uptake in photosynthesis. *Nat. Struct. Biol.* **10**, 637–644 (2003).

109. 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).

110. 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).

111. 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).

112. 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).