Title: Disordered proteins interact with the chemical environment to tune their protective function during drying

Authors: Shraddha KC[1], Kenny Nguyen[1], Vincent Nicholson[1], Annie Walgren[1], Tony Trent[1], Edith Gollub[2], Sofia Romero[2], Alex S. Holehouse[3,4], Shahar Sukenik[2,*], Thomas C. Boothby[1,*]

Affiliations

1. University of Wyoming, Department of Molecular Biology. Laramie, WY

2. Department of Chemistry and Biochemistry, University of California Merced, Merced, CA, USA

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

4. Center for Biomolecular Condensates, Washington University in St. Louis, St. Louis, MO, USA

Correspondence

*Thomas Boothby: tboothby@uwyo.edu

*****Shahar Sukenik: ssukenik@ucmerced.edu

Abstract

The conformational ensemble and function of intrinsically disordered proteins (IDPs) are sensitive to their solution environment. The inherent malleability of disordered proteins combined with the exposure of their residues accounts for this sensitivity. One context in which IDPs play important roles that is concomitant with massive changes to the intracellular environment is during desiccation (extreme drying). The ability of organisms to survive desiccation has long been linked to the accumulation of high levels of cosolutes such as trehalose or sucrose as well as the enrichment of IDPs, such as late embryogenesis abundant (LEA) proteins or cytoplasmic abundant heat soluble (CAHS) proteins. Despite knowing that IDPs play important roles and are co-enriched alongside endogenous, species-specific cosolutes during desiccation, little is known mechanistically about how IDP-cosolute interactions influence desiccation tolerance. Here, we test the notion that the protective function of desiccation-related IDPs is enhanced through conformational changes induced by endogenous cosolutes. We find that desiccation-related IDPs derived from four different organisms spanning two LEA protein families and the CAHS protein family, synergize best with endogenous cosolutes during drying to promote desiccation protection. Yet the structural parameters of protective IDPs do not correlate with synergy for either CAHS or LEA proteins. We further demonstrate that for CAHS, but not LEA proteins, synergy is related to self-assembly and the formation of a gel. Our results suggest that functional synergy between IDPs and endogenous cosolutes is a convergent desiccation protection strategy seen among different IDP families and organisms, yet, the mechanisms underlying this synergy differ between IDP families.

Keywords: Intrinsically disordered proteins, cosolutes, desiccation tolerance, synergy, CAHS proteins, LEA proteins, oligomerization, sucrose, trehalose

Introduction

Intrinsically disordered proteins (IDPs) make up about 40% of the eukaryotic proteome [1,2]. Unlike typical well-folded proteins, IDPs are characterized by a lack of defined tertiary structure, and instead exist as an ensemble of dynamic, interconverting conformations [3,4]. Despite their disordered nature, IDPs are known to play important roles in many biological processes including regulation of transcription and translation, metabolic signaling, subcellular organization, molecular chaperoning and response and adaptation to environmental cues [3,5].

Despite lacking a stable three-dimensional structure, IDPs still follow a similar paradigm by which form begets function. Different from well-folded proteins, however, an IDP's sequence determines the ensemble of conformations it adopts, and this ensemble can be important for the IDP's function(s) [2,4,6]. However, sequence is not the only determinant of the conformations present in an IDP's ensemble [7–9]. This is because IDP ensembles have relatively few intramolecular bonds and a large solvent-accessible surface area, which makes their ensembles more sensitive to the physicochemical environment than the relatively rigid structures of well-folded proteins [8–10].

The sensitivity of IDP ensembles to their solution environment and their link to IDP function poses a fundamental question: how do sequence and solution combine to tune IDP ensemble and function? To explore this question, we focus on a biological phenomenon where the intracellular environment undergoes drastic physical chemical changes: desiccation.

Organisms across every biological kingdom can survive near-complete desiccation by entering a state of reversible suspended metabolism known as anhydrobiosis (from Greek for 'life without water') [11,12]. As water effluxes from the cell during drying, the concentration of cosolutes increases by orders of magnitude, dramatically changing the physicochemistry of the cell [13,14]. In addition to the decrease in water content and concomitant increase in cosolute concentrations, the composition of the intracellular environment changes massively because of a regulated metabolomic response to drying mounted by anhydrobiotic organisms [12,15].

The acquisition of desiccation tolerance has historically been linked to the intracellular buildup of cosolutes such as trehalose, sucrose, arabinose, stachyose, and raffinose in plants and trehalose in some animals, fungi, and bacteria [15–20]. More recently, the accumulation of high levels of IDPs has also been linked to desiccation tolerance in many organisms [12,21–24]. Common examples of desiccation-related IDPs include the late embryogenesis abundant (LEA) proteins, which are the most widely studied desiccation-related IDPs due to their early identification and widespread distribution among different species and kingdoms of life [21,23–25]. LEA proteins are classified into seven different families based on the presence of conserved motif sequences [25,26]. Another family of desiccation-related IDPs are the tardigrade-specific cytosolic abundant heat soluble (CAHS) proteins [12,22,27].

Simultaneous enrichment of disordered proteins and endogenous cosolutes during desiccation promotes an ideal setting in which to study IDP-cosolute interactions [19,28–30]. In the

desiccation field, these interactions have been observed to produce a functional synergy in promoting tolerance to drying. Trehalose, a cosolute enriched alongside IDPs in many desiccation tolerant systems, has previously been shown to enhance IDP protective function *in vitro* and *in vivo* [19,28–30]. These observations prompted us to ask about the specificity of these interactions and if desiccation-related IDPs may have coevolved to work synergistically alongside their endogenously enriched cosolutes to promote desiccation protection.

To test whether desiccation-related IDP sequences have evolved to work with their intracellular chemical environment, here we use representative proteins from three families of IDPs: one CAHS protein (CAHS D) and proteins from two LEA families. Our result shows that full-length CAHS D and LEA proteins derived from four different organisms synergize better with endogenous protective cosolutes compared to protective exogenous cosolutes from other organisms.

To reveal the underpinnings of cosolute:IDP synergy, we examine the secondary and tertiary structure of protective IDPs in the presence of two disaccharides that are similar in terms of size but distinct with respect to chemistry and use across taxa. In all cases, the secondary structure (residual helicity) and tertiary structure (radius of gyration) do not change significantly in the presence of synergistic cosolutes and thus cannot explain the enhancement in function observed with synergistic cosolutes. We next assessed quaternary structure as both CAHS and LEA proteins are known to oligomerize [23,31], and for CAHS proteins, oligomerization leads to gelation [31–35]. While synergistic cosolutes did not influence LEA oligomerization, CAHS D oligomerization and gelation were enhanced in the presence of synergistic cosolutes. We further show that CAHS D's synergy can be explained through direct repulsive interactions between cosolutes and CAHS D's sidechains. However, this explanation does not hold for LEA proteins, implying that synergy in different protein families occurs through distinct mechanisms.

Our study showcases that different families of protective IDPs can have orthogonal modes of action and different functions in divergent solution environments. Beyond expanding our understanding of desiccation tolerance, these findings shed light on the sensitivity of IDP ensemble and function to the chemical composition of their environment. This is important as IDPs are ubiquitous across biology and function in key developmental processes and disease states that are concomitant with large changes in intracellular chemistry. Understanding how disordered proteins interact and evolve with the solution environment will provide insights into these biological mechanisms and phenomena.

Results

Desiccation-related IDPs are enriched in organisms alongside specific cosolutes during drying

To test whether IDP sequences have evolved to be functionally-tuned by the composition of the intracellular environment during drying, we selected six desiccation-related IDPs. These IDPs come from two LEA families [25,36] as well as the CAHS family (Table 1). We selected four

LEA_4 proteins each from a different desiccation tolerant organism. These organisms include the plant *Arabidopsis thaliana* (AtLEA3-3), the nematode *Aphelenchus avenae* (AavLEA1), the tardigrade *Hypsibius exemplaris* (HeLEA68614), and the rotifer *Adineta vaga* (AvLEA1C) (Table 1). To assess whether synergy extends across LEA families, we selected a LEA_1 protein from *A. thaliana* (AtLEA4-2) (Table 1). Finally, CAHS D was selected from the tardigrade *H. exemplaris* (Table 1). These organisms were selected not only because they all utilize LEA proteins to survive desiccation, but also because they accumulate different disaccharides to varying degrees during drying [16,17,19,37–41] (Table 1). The organisms we chose use one or both of two disaccharides - trehalose and sucrose which are similar in size, but chemically distinct (Table 1).

LEA motifs are not sufficient to mediate synergistic interaction with endogenous cosolutes during desiccation

To assess whether endogenous cosolutes induce functional changes in desiccation-related IDPs, we began by testing peptides encoding LEA motifs derived from full-length LEA_4 and LEA_1 proteins (Table 1). Family 1 LEA (LEA_1) proteins are characterized by a 20-mer repeating motif, whereas Family 4 LEA (LEA_4) proteins are characterized by the repetition of an 11-mer LEA_4 motif [25,36,42–44]. These motifs are often found in multiple linear or nonlinear repeats across the length of a LEA protein [25,36,42]. Interestingly, LEA_4 motifs have previously been suggested to be sufficient to confer desiccation protection to desiccation-sensitive proteins and membranes to a degree similar to full-length LEA proteins, both *in vitro* and *in vivo [44–47]*. With this in mind, we expected to observe synergy between LEA motif repeats and their paired endogenous cosolute(s).

We generated 11-mer LEA_4 motif peptides (At11, Aav11, He11, and Av11) and a 20-mer LEA_1 peptide (At20) and measured the ability of these motifs to protect lactate dehydrogenase (LDH), a desiccation sensitive enzyme during drying. LDH assay is used to assess the function of desiccation protectants to protect the activity of LDH which otherwise retains only approximately 2% of its pre-desiccation activity when dried and rehydrated [19,22,28,48].

The protective capacity for each LEA motif and cosolute was assessed across a range of concentrations with LDH (Fig. 1A). Most LEA_4 motifs displayed levels of protection so low that a 50% level of protection could not be reached even at concentrations exceeding 1 mM (Fig. 1A&B, Fig. S1A). Additionally, higher concentrations of the LEA_4 motifs tend to inactivate the enzyme when kept under control conditions (4 °C, see Methods), over a 16 hour incubation period during the assay (Fig. S1C). The LEA_1 20-mer motif At20, however, showed robust concentration-dependent protection of LDH, demonstrating that LEA_4 and LEA_1 motifs are functionally distinct (Fig. 1A&B, Fig. S1A). Concentration-dependent protection was also observed for our cosolutes trehalose and sucrose (Fig. 1A, Fig. S1A).

Since LEA_4 motifs often exist in tandem repeats of 11-mers within a full-length LEA protein [44,45,49], we wondered if the observed lack of protection was a result of their short length or repeat number. We synthesized 2X (At22) and 4X (At44) tandem repeats of the *A. thaliana* 11-

mer LEA_4 motif (At11). At22 and At44 show minimal potency in preserving *in vitro* LDH function during drying (Fig. 1A, Fig. S1A).

Despite the low protection displayed by our LEA_4 peptides, we opted to use them in cosolute synergy assays, as we reasoned that perhaps they would become functional when in solution with trehalose or sucrose. We picked sub-optimal concentrations of protectants (Table S1) so that under instances of synergistic protection, the additive protection of cosolute:peptide mixtures would not exceed 100%. We then performed synergy assays where sucrose or trehalose was combined with LEA motifs at molar ratios of 1:100, 1:10, 1:1, 10:1, and 100:1 (cosolute:protein). The upper limit (10:1 and 100:1) of these ratios closely align with known cosolute:protein ratios that produce synergistic protection between the tardigrade disordered protein CAHS D and trehalose [19]. Here we report on synergy by showing the individual protective ability of the cosolute and IDP on its own, the sum of these protective values (hypothetical additive effect), and the actual measured protection produced by combination of the cosolute and peptide (Fig. 1C). We quantify synergy using the following equation:

$$\% \text{ synergy} = (\% \text{ LDH protection})_{\text{IDP+cosolute}} - (\% \text{ LDH protection})_{\text{IDP}} - (\% \text{ LDH protection})_{\text{cosolute}}$$

In nearly all cases, synergy was not observed for 11-mer LEA_4 peptides with either sucrose or trehalose (Fig. 1D&E, Fig. S1B). In fact, in several cases, mixing LEA_4 peptides with sucrose or trehalose elicited antagonistic, rather than synergistic, effects (Fig. 1D&E, Fig. S1B). These results suggest that LEA_4 motifs do not robustly preserve LDH function, nor do they interact with cosolutes trehalose or sucrose in a functionally productive fashion. Similar to 11-mer motifs, the *A. thaliana* 22- and 44-mer peptides did not synergize with either trehalose or sucrose (Fig. 1D&E, Fig. S1B). Unlike LEA_4 motifs, the 20-mer LEA_1 motif, At20, was both protective and synergized with both trehalose and sucrose (Fig. 1D&E, Fig. S1B).

Taken together, these experiments demonstrate a diversity in disordered protein/motif function, where LEA_4 motifs largely are not protective to the enzyme LDH during drying, nor are they synergistic with endogenous cosolutes (Fig. 1F). Conversely, the LEA_1 motif tested is highly protective and synergizes with either sucrose or trehalose (Fig. 1F).

Desiccation-related IDPs synergize with endogenous cosolutes

While LEA proteins are identified through homology in conserved LEA motif repeats, they also contain varying quantities of non-motif sequence (Table 1). Since we observed that LEA_4 motifs generally provide relatively little protection and tend not to synergize with endogenous cosolutes in LDH assays, we wondered if full-length LEA proteins might. We also included CAHS D in our analysis as it has been previously known to synergize with trehalose and, to a lesser extent, with sucrose [19].

We began by testing the baseline protection of our proteins using the LDH assay. All full-length LEA_1 and LEA_4 proteins confer protection for LDH activity up to the pre-desiccated value (Fig. 2A, Fig. S2A). Likewise, CAHS D provided concentration-dependent protection to LDH as

previously observed [19,22,48] (Fig. 2A, Fig. S2A). We also included bovine serum albumin (BSA) in these studies as a well-studied control [19,48]. Unlike the LEA motifs, most full-length proteins protected 50% LDH at concentrations less than 1 mM (Fig. 2B).

Using data derived from the concentration range of LDH assays, we chose a sub-optimal concentration that provides 15-45% protection for each protein to perform synergy experiments with (Table S2). Our results show that nearly all full-length IDPs showed synergy with either sucrose or trehalose or both (Fig. 2C&D). Exceptions to this are AvLEA1C, which is derived from a rotifer that accumulates neither trehalose nor sucrose, and BSA, which comes from cows, which of course have no capacity for anhydrobiosis. Remarkably, in cases where LEA proteins displayed synergy, they were always more synergistic with endogenous compared to exogenous cosolutes (Fig. 2E). We also calculated synergy using two different approaches: one using equilibrium constant (K_{app}) as a metric and other using the Bliss independent model. Our results for synergy are largely invariant regardless of the method we use (Fig. S3A-F).

Taken together, these experiments demonstrate that synergistic interactions between IDPs and cosolutes extend across multiple families of desiccation-related IDPs found in a variety of organisms. Our data shows that these six IDPs synergize best with their endogenous cosolute to promote desiccation tolerance and we speculate that this may apply to other desiccation-related IDPs**.** It is also of note that while many of the LEA motifs tested at the beginning of this study did not display synergy with trehalose or sucrose, corresponding full-length proteins did. Likewise, At20 showed synergy with both trehalose and sucrose, while full-length AtLEA4-2 was only synergistic with sucrose. Thus, not only do these experiments demonstrate that full-length LEA proteins synergize with their endogenous cosolute(s) (Fig. 2E), but they suggest that this synergy, at least in part, is driven by sequence features beyond conserved motifs.

Trehalose and sucrose do not elicit local ensemble changes to desiccation-related IDPs in solution or in the dry state under the conditions tested

We next wondered what mechanism(s) drive the functional synergy observed between desiccation-protective IDPs and endogenous cosolutes. We reasoned that functional synergy might be driven by cosolute-induced changes to the IDP ensemble. To test this, we first examined the secondary structure contained in the ensemble of LEA proteins and CAHS D using circular dichroism (CD) spectroscopy.

Each full-length protein was first assessed using CD in an aqueous state by itself (Fig. 3A). All full-length LEA proteins displayed a single minimum at ~200 nm (Fig. 3A, black) indicating that LEA proteins are disordered in the aqueous state. CAHS D also displayed a minimum at ~200 nm and a slight minimum around ~220 nm, indicating that while disordered, it also has some propensity for helical structure in solution (Fig. 3A, black), in line with previous studies [31,50]. This is in contrast to BSA, which showed a high propensity for helical structure as denoted by the double minima at 222 and 210 nm (Fig. 3A, black). To see if the addition of cosolutes induces secondary structural change in the aqueous state, we obtained CD spectra of cosolute:protein mixtures at 100:1 molar ratios. Somewhat to our surprise, trehalose (Fig.

3A&C, blue) and sucrose (Fig. 3A&C, green) do not induce any significant structural changes to any of the full-length IDPs tested here (Fig. S4A).

LEA proteins gain helical conformation upon drying or in response to low water availability. Drying-induced helicity has been postulated to drive their protective function [23,30,51]. We reasoned that while our cosolutes do not induce detectable changes to LEA/CAHS secondary structure in solution, cosolutes could induce structural changes in proteins in a dry state. To test this, we examined our proteins using CD in a desiccated state [52,53]. Our results show a significant structural change for all LEA proteins and CAHS D in the dry state, indicated by a shift from disordered spectra with a minimum at ~200 nm to a helical structure with two minima at ~222 and ~210 nm (Fig. 3B). This is in contrast to BSA which started out helical and showed little change in the spectrum (Fig. 3B). These changes manifest for pure proteins without any addition of synergistic cosolutes. To quantify the influence of drying on secondary structure, we examined the changes in the ratio of CD signal at 222 and 210 nm. This ratiometric value reports on secondary structure in a concentration-independent way [54]. Using this metric, all LEAs and CAHS D display a clear increase in helical propensity upon being desiccated (Fig. 3C). On the other hand, the helical propensity of BSA remains very similar to its hydrated state, indicating that no dramatic structural change took place (Fig. 3C).

Next, we examined combinations of our proteins with trehalose or sucrose in a desiccated state. As with aqueous samples, the addition of trehalose or sucrose did not induce significant changes in the secondary structure of LEA and CAHS D proteins in the dry state (Fig. 3B&C, Fig. S4B). To assess whether there is a link between the minimal structural changes we observed and functional synergy in LDH assays, the change in the ratio of signal at 210 and 222 nm in desiccated and aqueous states was compared to synergy observed for that same mixture. Synergistic protection observed in our LDH assays did not correlate with secondary structural changes with the addition of trehalose ($p = 0.9863$ for aqueous, $p = 0.1113$ for desiccated, Fig. S4C&D) or sucrose ($p = 0.6673$ for aqueous, $p = 0.9863$ for desiccated, Fig. S4E&F).

Taken together, these results indicate that while the tested LEA proteins and CAHS D undergo a structural transition during desiccation, this phenomenon does not require, nor is it affected by, the presence of trehalose and sucrose. Furthermore, synergistic protection observed in our LDH assays is not mediated by local ensemble changes in these IDPs.

Trehalose and sucrose do not elicit changes in global ensemble dimensions for desiccation-related IDPs under the tested conditions, but promote oligomerization of CAHS D

We next wondered if the synergistic interactions observed between our IDPs and cosolutes could instead be explained by a change in global dimensions, such as expansion or compaction of the protein. To measure global dimensions, which cannot be detected using CD, we used small angle X-ray scattering (SAXS), which allows a model-free estimation of the radius of gyration (R_g) of an IDP as well as a prediction of the molecular weight (see Methods) [55]. Each protein, at a concentration of 4 mg/mL, was measured with no cosolute and with different molar

ratios of trehalose and sucrose. We reasoned that a cosolute-dependent change in the radius of gyration (R_g) could indicate changes in tertiary or quaternary structure, which may correlate with increased function or synergy. We note that for SAXS experiments cosolutes were used at concentrations between 20 and 50 mM due to technical restrictions on how much protein can accurately be assayed and a desire to maintain molar ratios used in other experiments.

We began by testing the R_g of BSA in different solution conditions. As a well-folded protein, we expected that BSA would be relatively insensitive to changes in the solution environment. The R_g values obtained via this approach match existing literature (Fig. 4A) [56]. Additionally, adding cosolutes did not modulate R_g (Fig. 4A, Fig. S5A). While our molecular weight approximations trended higher than expected for monomeric BSA, we reason that this may be due to the propensity of BSA to form small populations of low-level oligomers [57,58].

We next measured the R_g of our LEA proteins in various solution conditions. We observed that regardless of the solution environment, all of our LEA proteins have an R_g that falls within error of readings in other solution environments (Fig. 4B-F). At the concentrations used here, cosolutes do not induce significant changes in the global dimensions of LEA proteins. While the predicted molecular weight (pMW) from SAXS for these proteins were somewhat variable, we see no consistent trend between the presence of cosolutes and change in pMW for any LEA protein (Fig. 4B-F).

Finally, we tested CAHS D in similar solution conditions. We obtained an R_g value for CAHS D that lies between the values reported by other groups (see Methods) [31,33]. While this R_g was consistent in 1.6:1 disaccharide solutions, 16:1 disaccharide solutions, and 160:1 sucrose solution, we found that the 160:1 trehalose solution had a Guinier region that was sharply curved upwards, even upon the protein's first exposure to the X-ray source (Fig. 4G, Fig. S5G). This is consistent with the presence of large oligomerized species. A Bayesian approximation of the molecular weight of these samples supported this result (Fig. 4G). Most of our CAHS D samples showed a pMW that was only slightly elevated from the known value for the monomeric protein. In contrast, the 160:1 sucrose sample had a pMW about 300% higher than other samples, and the 160:1 trehalose sample had a pMW about 1000% higher (Fig. 4G).

While SAXS shows little change in Rg or pMW for LEA proteins, there is evidence that not only CAHS D but also some LEA proteins tend to oligomerize [23,59]. However, LEA oligomerization appears to be weak and transient, requiring extreme crowding and/or sensitive methods to detect [23,59,60]. We therefore wanted to use a more sensitive method to assess LEA oligomers, and how they might be affected by cosolutes.

To characterize oligomerization of LEA proteins in a more sensitive fashion, we used photo-induced crosslinking of unmodified proteins (PICUP), a zero-length crosslinking method that uses a light activatable crosslinking system and is known to capture transient oligomeric species [61,62]. PICUP has previously been used to characterize oligomeric forms in LEA proteins *in vitro* [60,63]. All of our LEA proteins showed a propensity to form oligomers, even at low concentrations (Fig. S6). However, the presence of trehalose or sucrose did not elicit changes

in oligomeric populations (Fig. S6). These results confirm that LEA proteins are able to form transient oligomers, but also demonstrate that at levels where sucrose and trehalose are synergistic with these proteins oligomerization is unaffected.

Taken together, these results show a divergence in the behavior of LEA and CAHS proteins in the presence of synergistic cosolutes. For CAHS D, the SAXS data suggests a relationship between the presence of cosolutes and increased oligomerization. However, this dataset showed no evidence of a cosolute inducible increase in molecular weight or R_g for the five LEA proteins that we tested, at least at the concentrations used in this study. This was further supported by PICUP, which despite detecting LEA oligomers did not show that they were enhanced by cosolutes. Thus, our data suggests that CAHS D oligomerization is promoted by the presence of synergistic cosolutes while LEA oligomerization is not.

Synergistic cosolutes promote gelation of CAHS D but not LEA proteins

CAHS D oligomers, as well as those of other CAHS proteins, have been reported to undergo self-assembly to form a gel network [31–35,64]. We wondered if the cosolute induced oligomerization of CAHS proteins in the presence of trehalose as seen in the SAXS experiments could be attributed to the propensity of CAHS D to form gels.

To test this, we performed differential scanning calorimetry (DSC) on CAHS D to observe the presence or absence of a gel melt. To begin, we tested CAHS D at 6 mg/mL (0.235 mM), which has previously been established to be a non-gelling concentration [31]. Consistent with this, at 0.235 mM, we find that CAHS D does not undergo a characteristic gel melt indicating a lack of gelation (Fig. 5A, black). Addition of trehalose and sucrose at increasing molar ratios (1:1, 10:1, 100:1, and 500:1) showed thermal features characteristic of endothermic phase transitions (e.g., a gel melting), indicating that the presence of these cosolutes induced gelation (Fig. 5A). Measuring the area under these melt curves allows us to calculate the enthalpy of melting (Fig. S7A). Change in enthalpy measurements for the cosolute:protein mixtures relative to the protein provides us a quantification of how gelation is affected by different amounts of cosolutes. Trehalose induced significant gelation at a 100:1 ratio, while 500:1 of sucrose was required to induce a significant gel melt (Fig. 5B). This is consistent with trehalose producing larger oligomeric species in our SAXS experiments (Fig. 4G), indicating that trehalose has a larger influence than sucrose on the gelation of CAHS D (Fig. 5B).

Furthermore, to test whether synergistic cosolutes enhance gelation, we tested CAHS D at 12 mg/mL (0.47 mM), a concentration above CAHS D's gelation threshold [31]. Addition of trehalose or sucrose at increasing molar concentrations promoted the formation of stronger gels evident by the enthalpy of melting measurements (Fig. S7B-E). These experiments demonstrate that not only do trehalose and sucrose induce sub-gelling concentrations of CAHS D to form gels, but they also enhance the strength of gels formed by higher concentrations of the protein.

Unlike CAHS proteins, gelation of LEA proteins has not been commonly observed or reported. The exception to this is AfrLEA6 (a LEA protein from the brine shrimp *A. franciscana*) that

appears to undergo phase separation, forming a hydrogel-like matrix upon desiccation [65]. To see if our LEA proteins gel, we performed DSC experiments on our LEA proteins and cosolutes. We analyzed similar molar concentrations of LEA proteins on their own and in mixtures with cosolutes at equivalent ratios (1:1, 10:1, 100:1, and 500:1). None of our LEA proteins by themselves or in mixtures with trehalose or sucrose show evidence of gelation (Fig. 5B-G). Likewise, BSA also failed to form a gel (Fig. 5H).

Taken together, these results demonstrate that trehalose and sucrose affect the oligomerization and phase state of different IDPs in distinct ways. The observation that trehalose induces more synergy as well as more gelation of CAHS D relative to sucrose leads us to speculate that gelation of CAHS D and synergistic protection of LDH may be linked.

Direct cosolute:IDP interactions drive synergy for CAHS D, but not LEA proteins

To explain the possible relationship between synergy and oligomerization-driven gelation of CAHS D, we quantified the interactions between each IDP and its cosolute environment using transfer free energies (TFEs). The TFE is a measure of the change in free energy undergone by a macromolecule when transferring from water to a concentrated solution of some osmolyte (typically 1 M) [66–68]. Using transfer free energy values derived from literature, we calculated the effect of trehalose and sucrose on the ability of CAHS D to dimerize: $\Delta\Delta G_{tr}^{M\rightarrow D}$. This is calculated by finding the free energy of CAHS D's monomeric state upon transfer to an osmolyte solution (ΔG_{tr}^{M}), doing the same for the dimeric state (ΔG_{tr}^{D}), and then taking the difference (Fig. 6A). A strong negative value for $\Delta\Delta G_{tr}^{M\rightarrow D}$ indicates that the presence of the cosolute pushes the population towards dimers. A positive value indicates that the addition of the cosolute pushes the population towards monomers.

In order to perform these calculations, we utilized AlphaFold2 and AlphaFold2 Multimer to determine plausible conformations for both CAHS D's monomeric and dimeric state. These structures have a helical linker region, which is consistent with our CD measurements (Fig. 3A&B) and is consistent with previous reports for this protein [31]. We reasoned that dimerization is indicative of gelation since CAHS D dimers were especially prevalent in crosslinking data [31], and previous research suggests that CAHS D dimer formation is a necessary step toward gelation [31]. TFEs were then calculated based on the solvent accessible surface area of different chemical groups in monomeric vs. dimeric state (see methods). Our calculations reveal that trehalose has a negative $\Delta\Delta G_{tr}^{M\rightarrow D}$ with CAHS D, meaning the dimeric state is favored in the presence of this cosolute. Sucrose's $\Delta\Delta G_{tr}^{M\rightarrow D}$ is close to 0, neither stabilizing nor destabilizing the dimer (Fig. 6B). In addition to these cosolutes, we wanted to explore the effect of a cosolute with a positive $\Delta\Delta G_{tr}^{M\rightarrow D}$. Glycine betaine, a common stabilizing cosolute, but without known roles in tardigrade desiccation tolerance, displayed a positive $\Delta\Delta G_{tr}^{M\rightarrow D}$(Fig. 6B). Together these cosolutes span a range of $\Delta\Delta G_{tr}^{M\rightarrow D}$values that are expected to increase the dimer population (trehalose), have a minimal impact on dimerization (sucrose), or increase the monomer population (glycine betaine) of CAHS D [68,69]. While data for trehalose and sucrose are in line with this analysis, we next sought to determine empirically the impact of glycine betaine on CAHS D gelation.

To test the effects of glycine betaine on CAHS D dimerization predicted by these TFE calculations, we first repeated our DSC and SAXS experiments in the presence of glycine betaine. Unlike trehalose and sucrose, below the protein's gelation threshold no increase in enthalpy of melting was observed upon the addition of glycine betaine (Fig. 6C&D). To probe whether glycine betaine inhibits CAHS D oligomerization, we conducted additional DSC experiments above the protein's gelation threshold. While trehalose and sucrose enhanced gelation of CAHS D (Fig. S7B-E), we observed a decrease in enthalpy of melting when glycine betaine was present at the 500:1 molar ratio, signifying inhibition of CAHS D gelation (Fig. 6E&F).

We then performed SAXS on CAHS D at 4 mg/mL with 1000:1 molar ratios of each cosolute. The intent of using such a high molar ratio was to mimic the high concentration of cosolutes that CAHS D would experience during desiccation [13]. A non-gelling concentration of CAHS D in 1000:1 glycine betaine yielded a scattering profile similar to the protein with no cosolute indicating a lack of gelation (Fig. 6G). Meanwhile, 1000:1 sucrose and trehalose yielded scattering profiles consistent with gelation, similar to those previously reported [31]. This is made especially evident by a peak at q = 0.06 $Å^{-1}$, which reports on the width of CAHS D's gel fibers and matches previously reported scattering profiles for gelled CAHS D [31]. Consistent with our hypothesis that the more negative $\Delta\Delta G_{tr}^{M\rightarrow D}$ for trehalose will increase dimerization and subsequent gelation, we observed an increased curvature in the Guinier region indicating increased fibrillization (Fig. S8C&D).

Finally, we tested the impact of glycine betaine on CAHS D's protective capacity. If induction of self-assembly of CAHS D is a mechanism underlying trehalose/sucrose induced synergy, then one would expect that glycine betaine's inhibition of gelation would result in no synergy, or even have an antagonistic effect. Glycine betaine on its own is not protective to LDH during drying nor does it inhibit LDH activity (Fig. S8E). We found a significant antagonistic relationship between glycine betaine and CAHS D on LDH protection (Fig. 6H). Furthermore, Pearson correlation between the $\Delta\Delta G_{tr}^{M\rightarrow D}$ of a cosolute (scaled by concentration) and its ability to induce synergy in CAHS D was statistically significant (Fig. 6I).

Using a similar AlphaFold2-based approach for LEA proteins and for BSA, one observes correlations between the $\Delta\Delta$G of the disorder-to-order transition and synergy (Fig. S8F-L). Interestingly, AlphaFold2 predictions of our LEA proteins were broadly helical, which is in contrast to our experimental characterization of these proteins in aqueous solutions. However, this is not unusual for AlphaFold2 predictions and could possibly represent a "bound" conformation for the proteins [70]. For a subset of these proteins, we see a statistically significant correlation between $\Delta\Delta$G and synergy (Fig S8F-L). However, this data is purely computational. For CAHS D, we saw our predictions recapitulated in changes in the protein structure, and for LEA proteins we do not. Thus, we conclude that cosolutes do not induce synergy in our LEA proteins through a change in folding.

Taken together, these results suggest that the driving force for gelation in CAHS D and its ability to synergize with a given cosolute is rooted in the direct interaction between CAHS D and the prevalent cosolute. However, trying to apply this model to the LEAs tested in this work failed to yield meaningful correlates. Thus, we propose that direct interactions between cosolute and LEA proteins cannot explain the synergy observed with cosolutes. Overall, our study demonstrates that while synergy between desiccation-related IDPs and endogenous cosolutes appears to be a widespread and conserved behavior, the mechanisms underlying this synergy vary between IDP families.

Discussion

In this study, we examined the interplay between IDP sequence, solution environment, ensemble, and function. To do this, we have taken advantage of the dramatic changes to the cosolute content of anhydrobiotic organisms brought on by desiccation, and compared the effects of these cosolutes between representatives from three families of desiccation-related IDPs (LEA_4, LEA_1, and CAHS proteins). For the proteins we tested empirically, we demonstrate that endogenous cosolutes enriched during desiccation enhance the protective capacity of CAHS D, full-length LEA_4, and LEA_1 proteins in an *in vitro* enzyme assay. Surprisingly, the functional changes were not accompanied by any detectable structural changes to the monomeric ensemble of these IDPs under the conditions tested. However, synergistic cosolutes did induce oligomerization and gelation of CAHS D. Finally, we show that in the case of CAHS D, but not LEAs, oligomerization, gelation, and protective synergy can be traced to direct interactions between cosolute and protective protein. Our results suggest that while functional synergy between the solution environment spans multiple IDP families, different mechanisms can underlie synergistic interactions for different proteins.

Functional synergy for the full-length proteins from different organisms mirrored the endogenous cosolute environment in that organism. In all cases, an IDP protected LDH activity more with its endogenous cosolute compared to an exogenous cosolute. These differences in synergy appear to extend across even subtle variations in cosolute use. For example, nematodes and tardigrades both accumulate trehalose during desiccation, but tardigrades accumulate orders of magnitude less [19,37–39]. Consistent with this, both tardigrade proteins used in this study synergized with trehalose at an order of magnitude lower concentration than what was required to elicit synergy with the nematode LEA protein.

It is important to note that desiccation-tolerant organisms employ multiple cosolutes to counteract the effects of desiccation. The use of a single cosolute-IDP system in our *in vitro* experiments does not accurately mirror the diverse cosolute changes in desiccating systems. For instance, *Arabidopsis* seeds enrich both trehalose and sucrose, among other cosolutes. This demands the necessity of future experiments that incorporate both or multiple cosolutes and assess their synergistic effects, thus elucidating the intricate synergy in multi-cosolute systems. Additionally, we want to point out that our results cannot necessarily be generalized to all desiccation-related IDPs. More experiments will be needed to assess the relevance of

cosolute effects to functional synergy and IDP folding in the context of desiccation and beyond. This remains an important future direction for the field.

Overall, while our study found that synergy with endogenous cosolutes is observed across two families of LEA proteins as well as CAHS proteins, we observed a major difference in structural changes induced in these protein families. While in CAHS D synergy could be traced back to the interaction between the protective protein and the cosolute, no underlying mechanism was detected for the tested LEAs. What then could be driving the synergy we observe for these proteins?

One possibility is that in most of our assays we do not consider the protein being protected. Our studies of molecular mechanisms for synergy do not consider the underlying effect of both protectant protein and cosolute on LDH directly. It is possible that the presence of both endogenous cosolute and protein create a solvation environment that becomes highly protective for e.g. rehydration. In line with this, recent studies have highlighted the ability of LEA proteins to stabilize sugar glasses in a dry state [49,71–73]. Glass formation is known to preserve labile biomolecules during desiccation, contributing to survival [22,74]. Different glasses vary significantly in their protective capacity, and studies have attempted to find structural properties that explain this difference [73,75]. Because trehalose and sucrose both form glasses when dried [76,77], it is possible that our LEA proteins are inducing a change in the glass's structural properties that leads to synergy.

Another possibility is a difference in the nature of TFE-induced oligomerization. While repulsive cosolutes drive homotypic interactions between CAHS D monomers, the same thermodynamic force can promote heterotypic interactions between LEAs and other proteins [78]. For example, trehalose may stabilize electrostatic interactions between LEA proteins and LDH during our *in vitro* synergy assays. If the protective capacity of LEA proteins is dependent on direct interactions between the protectant and the client protein, then this is a plausible explanation for synergy. However, it is currently unknown whether or not this is the case.

Another major question posed by this research is why sucrose was able to elicit synergy in CAHS D. Our computational approach predicted that sucrose should be agnostic to CAHS D's ability to form dimers, and yet we clearly see that, *in vitro*, sucrose is a moderately potent driver of gelation. We believe several factors could explain this effect. While sucrose does not drive dimerization through direct 'soft' repulsion, it may still do so by acting as a crowder eliciting an excluded volume effect [79]. Another possible manifestation of an excluded volume effect is the slight increase in the melting peak that is observed with all cosolutes used in this study at high molar ratios (Fig. S7F-I, Fig. S8M-N). Alternatively, given that CAHS D must polymerize beyond the dimeric state to form a gel, sucrose may stabilize a higher-level oligomer that was not captured in our $\Delta\Delta G_{tr}^{M\rightarrow D}$ analysis.

IDPs are known to play important regulatory functions during development and disease progression that often occurs together with changes to the chemical composition of the intracellular environment. For example, there are known links between type II diabetes and Alzhiemer's disease, and it has been shown that the intrinsically disordered neurodegenerative

peptide Aβ42 undergoes pathological oligomerization in the presence of glucose whose levels mirror those found in diabetic patients [5]. Our study showcases how different cosolute environments can have a direct effect on the function of IDPs. By understanding the rules governing desiccation related IDP-cosolute interactions, we might better understand the influence of changing chemical environments on a host of other IDPs.

Acknowledgements

Support for this project came from NSF via the IntBio research program under awards 2128069 to TCB, 2128067 to SS, and 2128068 to ASH. SK and KN were supported in part by the USDA National Institute of Food and Agriculture, Hatch project #1012152. In addition, this work was made possible in part through support from an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health (Grant # 2P20GM103432). We thank members of the Water and Life Interface Institute (WALII), supported by NSF DBI grant #2213983, for helpful discussions. We thank Dr. Greg Hura and Kathryn Burnett for their correspondence and help in performing the SAXS experiments. SAXS experiments were conducted at the Advanced Light Source (ALS), operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by DOE Office of Biological and Environmental Research.

Author Contributions

Conceptualization: **SK, KN, ASH, SS, TCB**
Data Curation: **SK, KN, VN, EG, SR**
Formal Analysis: **SK, KN, VN, EG, SR**
Funding Acquisition: **ASH, SS, TCB**
Investigation: **SK, KN, VN, AW, TT, EG, SR**
- Protein expression and purification: **SK, KN, AW**
- LDH functional experiments: **SK, KN**
- Circular dichroism spectroscopy: **EG, SR**
- Small angle X-ray scattering: **VN**
- Cross-linking: **SK**
- Differential scanning calorimetry/gelation: **SK, KN**
- TFE computation/analysis: **VN**, **TT**

Methodology: **SK, KN, VN, EG, SS, TCB**
Project Administration: **VN**
Software: **TT**
Supervision: **ASH, SS, TCB**
Visualization: **SK, KN, VN**
Writing – Original Draft: **SK, KN, VN, TCB**
Writing – Review & Editing: **SK, KN, VN, AW, TT, SR, EG, ASH, SS, TCB**

Declarations of Interest

A.S.H. is a scientific consultant with Dewpoint Therapeutics and is on the Scientific Advisory Board of Prose Foods. The work reported here was not influenced by these affiliations. All other authors declare no competing interests.

Lead Contact

Further information and request for resources and reagents should be directed to and will be fulfilled by TCB (thomas.boothby@uwyo.edu)

Data availability

All data generated or analyzed during this study are included in File S1.zip (for main text figures and tables) and File S2.zip (for supplementary figures).

Code availability statement

All custom code used in this study is included in this published article (File S3.zip).

Methods

Protein Sequences

Sequences of all peptides and proteins used in this study are available in File S4.docx.

Cloning

Inserts for full-length proteins- AavLEA1, AtLEA4-2, AvLEA1C, CAHS D, and HeLEA68614 were synthesized as codon optimized gBlocks (Integrated DNA Technologies) and cloned into the pET28b expression vector using gibson assembly (New England Biosciences). AtLEA3-3 was cloned in pET28a vector by Twist Bioscience. Clones were propagated in DH5α cells (NEB, Cat. #C2987H) and verified by Sanger sequencing (Eton Bioscience).

Protein Expression

Expression constructs were transformed into BL21 (DE3) cells (New England Biosciences, Cat. #C2527H) and plated on Luria-Bertani (LB) agar plates with 50 µg/mL kanamycin. At least 3 single colonies were chosen for each construct and tested for expression. Constructs were expressed in 1 L LB/kanamycin medium and grown at 37 °C while shaking at 180 rpm (Eppendorf Innova S44i) until an OD600 of 0.6 was reached. The culture was induced with 1 mM IPTG, and grown for the next 4 hours while shaking. AvLEA1C was grown for 1 hour following IPTG addition. Cells were harvested by centrifugation at 4000 rpm for 30 minutes at 4 °C. Cell pellets were resuspended in 5 mL of 20 mM Tris buffer, pH 7.5 supplemented with 30 µL of 1X protease inhibitor [Sigma Aldrich, Cat. P2714]). Cell pellets were stored at -80 °C until further use.

Protein Purification

Frozen pellets were thawed at room temperature, subjected to heat lysis in boiling water for 10 minutes, and cooled down for 15 minutes. These were then centrifuged at 10500 rpm at 10 °C for 30 minutes, and the supernatant was later filter-sterilized through a 0.22 µm filter to remove any insoluble particles (EZFlow Syringe Filter, Cat. 388-3416-OEM). The filtrate was diluted two times the volume with buffer UA (8 M urea [Acros Organics, CAS No. 57-13-6], 50 mM sodium acetate [Tocris CAS No. 127-09-3], pH 4). This was loaded onto a HiPrep SP HP 16/10 (Cytiva, Cat. 29018183) cation exchange column and purified on an AKTA Pure (Cytiva, Cat. #29018224), controlled using the UNICORN 7-9.1 Workstation pure-BP-exp (Cytiva, Cat. #29128116). CAHS D was eluted using a 0-40% UB (8 M urea, 50 mM sodium acetate, and 1 M NaCl, pH 4) gradient and fractionated over 15 column volumes. LEA proteins were eluted using the 0-70% UB gradient over 15 column volumes. Protein fractions were assessed using SDS-PAGE and selected fractions were dialysed in a 3.5 kDa tubing (SpectraPor 3 Dialysis Membrane, Part No. 132724) in 20 mM sodium phosphate buffer pH 7, followed by six rounds of Milli-Q water (18.2 MΩcm) at four hours interval each. Concentration of the dialyzed fractions were then quantified using Qubit 4 fluorometer (Invitrogen, REF Q33226), flash frozen, then lyophilized (Labconco FreeZone 6, Cat. 7752021) for 48 hours, and stored at -20 °C until further use.

LEA Motif Sequence Identification

LEA_4 and LEA_1 sequence motifs were identified in full-length LEA proteins using RADAR (https://www.ebi.ac.uk/Tools/pfa/radar/). In cases where RADAR was unable to identify repetitive motifs (e.g., in cases where a full-length LEA protein had only one or two motif repeats), manual selection and alignment of motifs was performed.

Lactate Dehydrogenase (LDH) Protection Assay

LDH assay was adopted from previous studies [19,22,28,48]. Protectants were resuspended at a final concentration range 20 mg/mL to 0.01 mg/mL in 25 mM Tris HCl pH 7. Rabbit muscle L-lactate dehydrogenase (LDH), sourced from Sigma (Sigma-Aldrich, Cat 10127230001), was added to each solution at a concentration of 0.1 mg/mL. Half of this sample was dried in a vacuum desiccator (SAVANT Speed Vac Concentrator) for 16 hours, while the other half was refrigerated at 4 °C for the same duration. Water was added to both desiccated and non-desiccated samples to a final volume of 250 µL each. 10 µL sample was mixed with 980 µL phosphate pyruvate buffer (100 mM sodium phosphate, 2 mM sodium pyruvate; pH 6.00) supplemented with 10 µL of 10 mM NADH (Sigma-Aldrich NADH; disodium salt, grade II) in a quartz cuvette. LDH activity was measured as the kinetics of the decrease in NADH absorption at 340 nm for a minute in NanodropOne (Thermo Scientific). Percent protection was calculated as a ratio of NADH absorbance for the desiccated samples normalized to non-desiccated controls. Each sample was performed in triplicate.

Lactate Dehydrogenase (LDH) Synergy Assay

The protection data for individual protein or motif was used to select a suboptimal protective concentration. Trehalose or sucrose was mixed in equal parts with proteins at 2X concentration in 100 µL resuspension buffer (25 mM Tris HCl pH 7) at respective molar ratios. LDH assay was performed for the mixtures as described previously. For each mixture, LDH protection was assessed individually and as a mixture. The sum of the protection conferred by individual

protein and cosolute was determined, which would refer to the expected additive protection. Synergy was determined by statistical comparison of this expected additive protection with the experimental protection.

Synergy calculation using equilibrium constant (K_{app})

Synergy was calculated as an apparent equilibrium constant between functional and non-functional LDH. The equilibrium constant for the IDP - $K_{app(IDP)}$, cosolute - $K_{app(cosolute)}$ and the mixture - $K_{app(mixture)}$ were determined by the formula:

$$K_{app} = [\text{Functional LDH}] / [\text{Non Functional LDH}]$$

The fraction for expected additive effect was calculated as:

$$\text{Fraction}_{expected} = [K_{app(IDP)}+K_{app(cosolute)}] / [1+K_{app(IDP)}+K_{app(cosolute)}]$$

The expected Kapp was then determined as:

$$K_{app(expected)} = \text{Fraction}_{expected} / [1-\text{Fraction}_{expected}]$$

Synergy was calculated as:

$$\text{Synergy} = K_{app(mixture)} - K_{app(IDP)} - K_{app(cosolute)}$$

Synergy calculation using Bliss independent model

LDH assay was used to generate the percent protection for IDP - Protection_{IDP}, cosolute - $\text{Protection}_{cosolute}$ and the mixture - $\text{Protection}_{mixture}$. The expected additive protection for the IDP:cosolute mixture was calculated as:

$$\text{Protection}_{expected} = \text{Protection}_{IDP} + \text{Protection}_{cosolute} - (\text{Protection}_{IDP} * \text{Protection}_{cosolute})$$

Synergy was then calculated as:

$$\text{Synergy} = (\text{Protection}_{mixture} - \text{Protection}_{expected}) / \text{Protection}_{expected}$$

Circular Dichroism (CD) Spectroscopy

CD spectroscopy was adopted from Bremer et al. [52]. Lyophilized proteins were resuspended in 25 mM NaPi pH 7 to a concentration of 200 µM. The resuspended protein was then mixed in equal parts with the NaPi buffer, 20 mM trehalose, or 20 mM sucrose in separate samples to a 100:1 molar ratio, with a final protein concentration of 100 µM and cosolute concentration of 10 mM. Protein concentration was confirmed with either a UV-vis (Thermo Scientific, GENESYS 50 UV-visible spectrophotometer) or a Qubit (Life Technologies, Qubit 3.0 Fluorometer). 20 µL aliquots of the samples were deposited on a 0.05 mm quartz cuvette and measured in a Circular Dichroism (CD) spectrometer (JASCO, J-1500 model). New 20 µL aliquots were then deposited on one half of a 0.05 mm quartz cuvette and spread across part of the cuvette with the tip of a

pipette, to a surface area of about 1 cm^2. The samples were then desiccated in a vacuum chamber with drierite for 1 hour to create a dry film, and another CD measurement was taken immediately after the vacuum was stopped. Each measurement was performed in triplicate.

Small-Angle X-ray Scattering (SAXS) – Sample Preparation

Lyophilized protein was resuspended at high concentration in a buffer containing 20 mM tris HCl (pH = 7.0) and the correct amount of cosolute to reach the desired molar ratio. Protein samples were then quantified with the Qubit Protein Assay from ThermoFisher Scientific (catalog# Q33212). The proteins were then diluted into 8 mg/mL and 4 mg/mL stocks using the same cosolute solution. Due to the necessity for each sample to have an identical buffer blank, the concentration of cosolute in the 8 mg/mL sample had to be the same as in the 4 mg/mL sample, meaning the molar ratio would be doubled in the 4 mg/mL sample. For each sample, a small aliquot of buffer was saved and stored at 4 °C for use as a blank. Samples and buffer blanks were filtered using 0.22 µm syringe filters and loaded into an Axygen 96-well polypropylene PCR Microplate (Corning product# PCR-96-FS-C), which was then sealed with an AxyMat Sealing Mat (product# AM-96-PCR-RD) and wrapped in parafilm. Plates were shipped to Lawrence Berkeley National Labs in a styrofoam cooler filled with cold packs. All SAXS measurements were performed by the SIBYLS group at the Lawrence Berkeley National Laboratory HT-SAXS beamline (12.3.1) [80,81]. For technical restrictions, proteins were measured at 20:1 and 200:1 molar ratios of trehalose and sucrose instead of 10:1 and 100:1. The exception to this was CAHS D, which was tested with a wider range of molar ratios (1.6:1, 16:1, and 160:1).

We note that the radius of gyration that we calculated for CAHS D is approximately 5 angstroms greater than previously reported by our group [31]. We believe that this difference can be attributed to minor differences in our approach. While we used Size Exclusion Chromatography (SEC)-coupled SAXS and a relatively dilute CAHS D sample in our previous study, the number of SAXS experiments in this study necessitated a higher-throughput approach that omitted the SEC step. We therefore believe that our control samples contained a higher fraction of transient oligomeric species, which inflated the radius of gyration without significantly curving the Guinier region. Given that this fact was consistent between all of our CAHS D samples, the inter-environmental comparisons are still valid.

Small-Angle X-ray Scattering (SAXS) – Guinier and pMW Analysis

Notable aggregation, likely induced by exposure to X-rays, was present in some samples, especially in solutions that contained cosolutes. This was controlled for by excluding scattering data from samples that had already been exposed to large amounts of X-ray radiation and were thus statistically different from the initial readings. Despite some readings having been excluded, a Guinier analysis was able to be conducted for each protein:cosolute combination. Buffer subtractions and Guinier analysis were performed using BioXTAS RAW v. 2.1.4 [82,83]. A $qMaxR_g$ of 1.1 was used to establish linear fits in the Guinier region [55,84]. Samples with Guinier regions that could not be fit were excluded from the study. Molecular weight approximations were performed using the method described in Hajizadeh et al. 2018, which is

programmed directly into BioXTAS RAW [85]. The 8 mg/mL samples tended to be far more aggregation-prone than the 4 mg/mL samples, so only the 4 mg/mL data is reported here.

Photo-Induced Cross-Linking of Unmodified Proteins (PICUP)

PICUP crosslinking was performed as previously described [60,63]. Briefly, lyophilized protein, Ru(II)bpy32+, and ammonium persulfate were resuspended in 20 mM Tris pH 7.5. Each reaction mixture constituted protein at the desired concentration with 1.25 mM Tris (2,2'-bipyridyl)dichlororuthenium(II)hexahydrate (Sigma, CAS No. 50525-27-4, and 2.5 mM ammonium persulfate (Sigma, CAS No. 7727-54-0) to a final volume of 10 µL. For mixtures, cosolutes and proteins were mixed at a 100:1 molar ratio at 2X molar concentration. Photoreaction was triggered by flashing 72 W light through a 2.5 cm water filter for 10 seconds in a dark room. The reaction was immediately quenched by adding 10 µL of 2X Laemmli buffer containing 4 % SDS and 10 % β-mercaptoethanol. The reaction mixture was heated at 95 °C for 5 minutes. 8.5 µL of each sample was run in denaturing SDS-PAGE gels and stained with coomassie blue to visualize the oligomeric states.

Differential Scanning Calorimetry (DSC) Measurements

Samples were prepared in Eppendorf tubes at the desired molar ratios with cosolutes. Protein mixtures were resuspended and incubated at 55 °C for 5 minutes to ensure proper solubility. 25 µL of the sample was hermetically sealed into a previously massed pair of DSC aluminum hermetic pan and hermetic lid (Catalog 900793.901 and 901684.901, respectively, TA instruments). The sample mass was determined after the sample was sealed within the pan and lid. The sealed samples were then run on a TA DSC2500 instrument. The DSC method for heating experiments is as follows:

Samples were equilibrated at 20 °C, heated to 60 °C at a 5 °C per minute ramp, and then cooled to 20 °C at a 5 °C per minute ramp. Samples were held for a 10 minute isothermal hold at 20 °C and heated to 60 °C at a 5 °C per minute ramp. Trios software (TRIOS version #5.0.0.44608, TA Instruments) was used to analyze enthalpy for samples showing the melt curves. The changes in enthalpy for the mixtures were calculated relative to the protein.

AlphaFold2 Structural Modelling

Protein structure predictions (both monomeric and multimeric) were performed using Google's AlphaFold2 Colab notebook. The setting "relax_use_gpu" was checked to increase the speed of individual predictions. The number of recycles was left at 3 [86]. Representative images of each structure are provided (Fig. S9). All analysis was done in triplicate to account for variation in AlphaFold's predictions. Representative images of each structure are provided (Fig. S9). All pdb files can be found in supplementary data.

Transfer Free Energy (TFE) Calculations

TFE values of each cosolute for each amino acid were pulled from existing literature [68,69,87]. We used experimentally derived TFE values for the transfer of a chemical group - amino acid side chains or backbone - into 1 M solutions of trehalose, sucrose, or glycine betaine [67–69]. We then calculated the TFE for each conformation using the formula:

$$\Delta G_{tr} = \sum_{N} \sum_{i=1} \alpha_i g_N$$

Here, ΔG_{tr} is the TFE of a protein conformation from water to 1 M cosolute solution, *N* is the chemical group, *i* is a numerical index for all instances of the chemical group, α is the surface area of the specific instance of the chemical group in square angstrom, and *g* is the experimental value of the transfer free energy for that chemical group per square angstrom of exposed surface area [67]. By doing this for two conformations of a protein, one can find the change in free energy of conformational change that can be attributed to the presence of an osmolyte.

Importantly, two "end-state" protein conformations are required to calculate a $\Delta\Delta G_{tr}$ value, which is typically done for well-folded proteins [67]. However, we believe the proteins used in this study constitute an important exception. For CAHS, several different groups have noted the propensity of CAHS proteins for oligomerization, and recent research has identified the dimer as a particularly stable CAHS D conformer [31,33,35]. This notion is supported by the crosslinking data, in which the CAHS D dimer is especially prominent [31]. We thus use the following equation for the effects of cosolutes on CAHS dimerization:

$$\Delta\Delta G_{tr}^{M\rightarrow D} = \Delta G_{tr}^{D} - 2 \times \Delta G_{tr}^{M}$$

Where ΔG_{tr}^{M} and ΔG_{tr}^{D} is the TFE for the monomeric and dimeric state, respectively. The monomer and dimer conformations were obtained from AlphaFold2 Multimer and AlphaFold2 predictions, respectively (see AlphaFold2 method). Surface area for residues was calculated using SOURSOP, a python package for protein structure analysis [88].

To calculate the $\Delta\Delta G_{tr}^{folding}$ for our LEA proteins and for BSA, we compared an AlphaFold2 prediction of a monomeric protein with a theoretical conformation in which all residues are 100% exposed. This is because (a) LEAs showed no tendency to form a gel, and (b) the CD spectra switched between a primarily disordered conformation to a helical conformation upon desiccation (Fig. 3). For all LEAs, we calculated:

$$\Delta\Delta G_{tr}^{folding} = \Delta G_{tr}^{final} - \Delta G_{tr}^{initial}$$

Here, $\Delta G_{tr}^{initial}$ represents the free energy of a completely denatured protein chain where the maximum theoretical accessibility is achieved (RASA = 1). $\Delta G_{tr}^{initial}$ represents the free energy of the protein's "native" conformation (as determined by an AlphaFold2 prediction). AlphaFold2 predictions of our LEA proteins were broadly helical (Fig. S9), and thus were used to represent the disorder-to-helix transition commonly observed in LEA proteins.

Data Analysis and Visualisation

LDH protection was fitted into a sigmoidal curve by fitting a 5PL regression analysis using GraphPad Prism v9.5.1 from which the resulting PD50 values were derived. Other plots were

plotted using R-Studio. Annotation for statistical significance include: $p>0.05$: NS, $p= 0.01-0.05$: *, $p= 0.001-0.01$: **, $p<0.001$: ***

Figures and figure legends

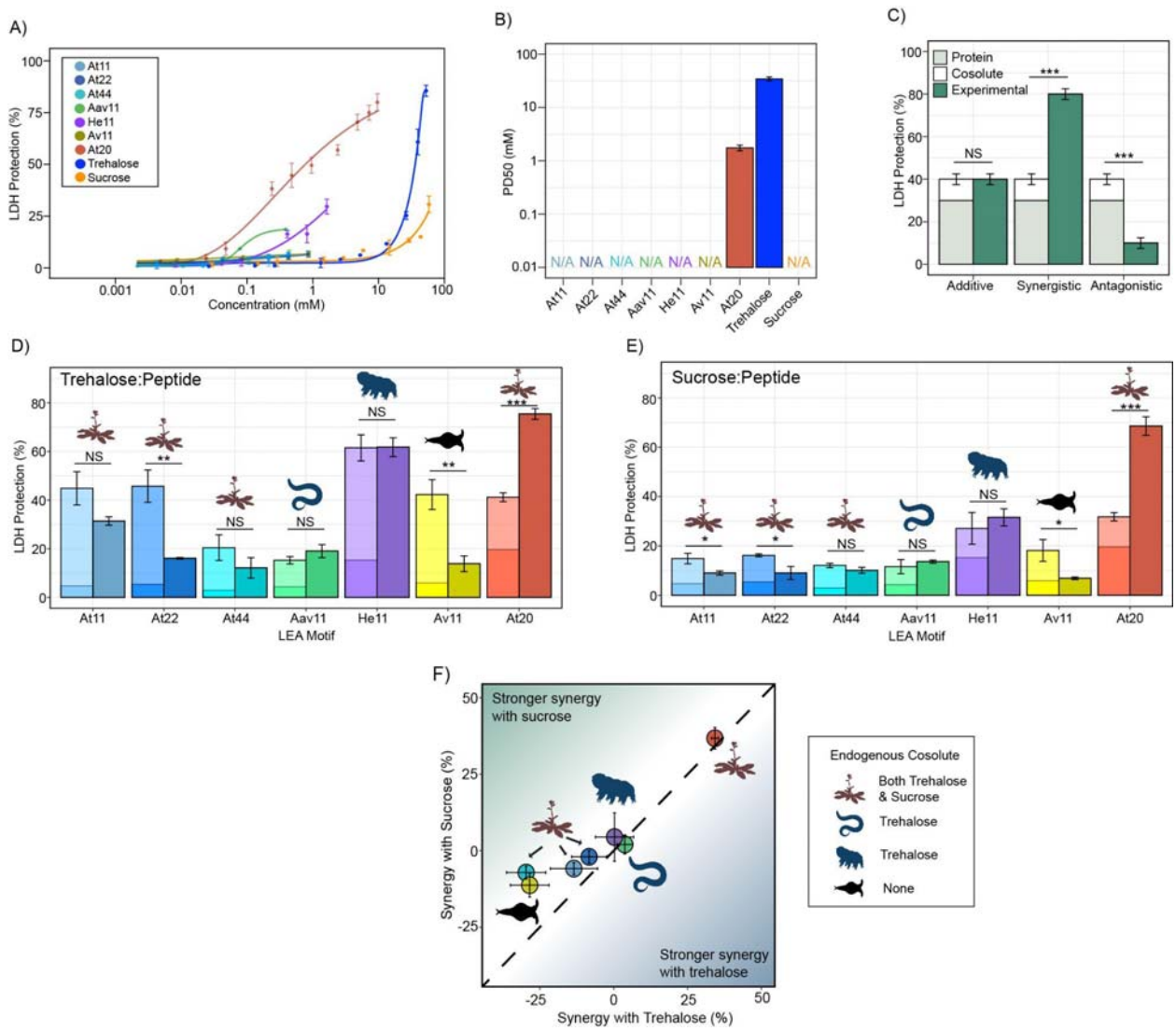


Figure 1: LEA motifs are not synergistic with endogenous cosolutes. A) Sigmoidal plot representing percent of LDH stabilization by LEA motifs and cosolutes as a function of the molar concentration. n = 3, error bars = standard deviation. **B)** Protective dose 50 (PD50) for additives obtained by sigmoidal fitting of data in Fig. 1A. n = 3, N/A represents instances where 50% protection was not achieved. **C)** Example plot showing possible outcomes (additive, synergistic, or antagonistic effect) from cosolute:IDP mixtures. "Cosolute" and "Protein" represent the percent LDH protection by cosolute and protein respectively. Experimental represents the experimental protection resulting from cosolute and protein mixtures. **D)** Synergy plots for trehalose:LEA peptide at 100:1 molar ratio. n = 3, Welch's t-test was used for statistical comparison, error bars = standard deviation. **E)** Synergy plots for sucrose:LEA peptide at 100:1 molar ratio. **F)** Plot representing synergy with trehalose vs. sucrose from Fig. 1D and Fig. 1E. Dotted line represents instances where there is equal synergy with both trehalose and sucrose.

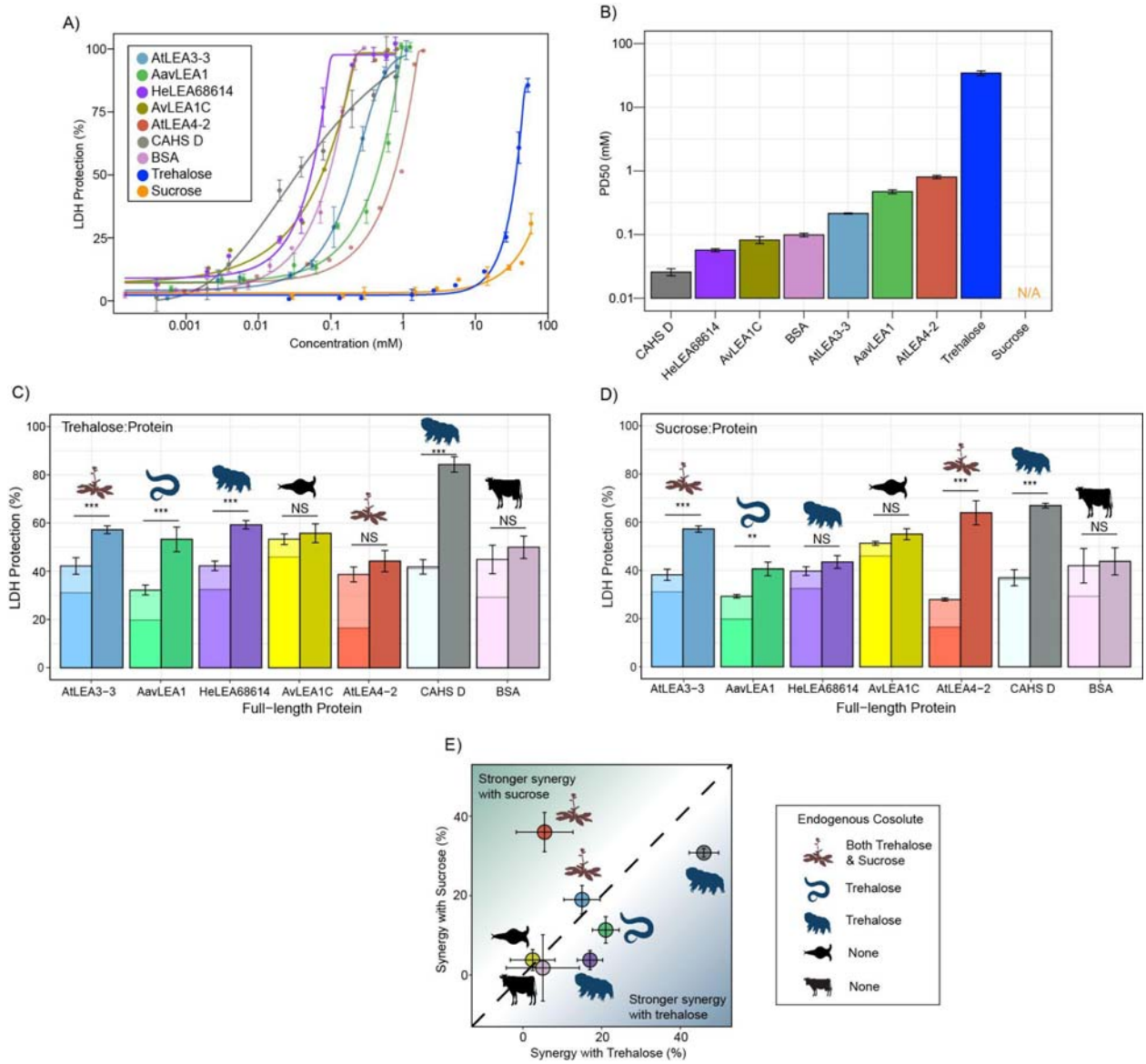


Figure 2: Full-length desiccation-related IDPs act synergistically with cosolutes. A) Concentration dependence of LDH protection by full-length proteins and cosolutes used in this study **B)** PD50 for additives obtained by sigmoidal fitting of concentration-dependent LDH protection from Fig. 2A. n = 3, N/A represents instances where 50% protection was not achieved. **C)** Synergy plots for trehalose:protein at 100:1 molar ratio. n = 3, Welch's t-test was used for statistical comparison, error bars = standard deviation. **D)** Synergy plots for sucrose: protein at 100:1 molar ratio **E)** Plot representing synergy with trehalose vs. sucrose from Fig. 2C and Fig. 2D.

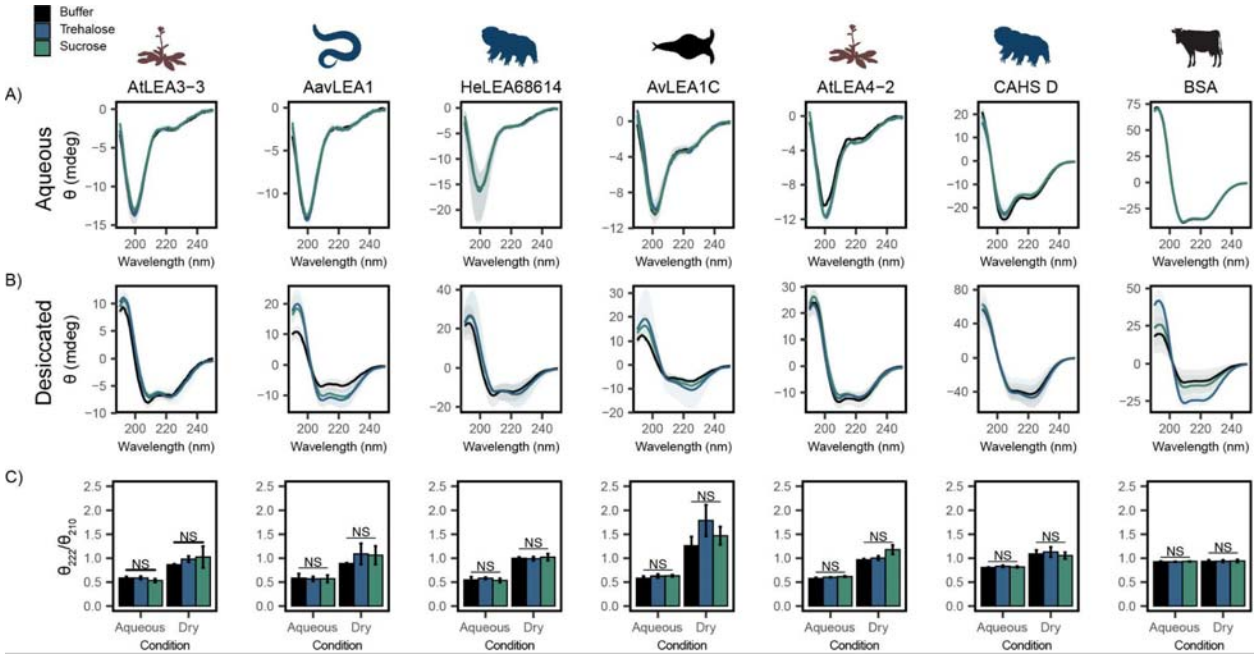


Figure 3: Functional synergy is not mediated by secondary structural changes. Buffer represents the CD analysis for the proteins without cosolutes. Trehalose and Sucrose represent the CD analysis for trehalose:protein mixture and sucrose:protein mixture at 100:1 molar ratio respectively **A, B)** CD spectra for protein and cosolute:protein mixtures at 100:1 molar ratio under aqueous **(A)** and **(B)** desiccated conditions. Each plot represents the average of three replicates, with the shaded region representing the standard deviation of the average **C)** Changes in the ratio of CD signal at 222 and 210 nm for individual protein and cosolute:protein mixtures under aqueous and desiccated conditions. n = 3, error bars = standard deviation, statistical analysis was done by Welch's t-test. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.

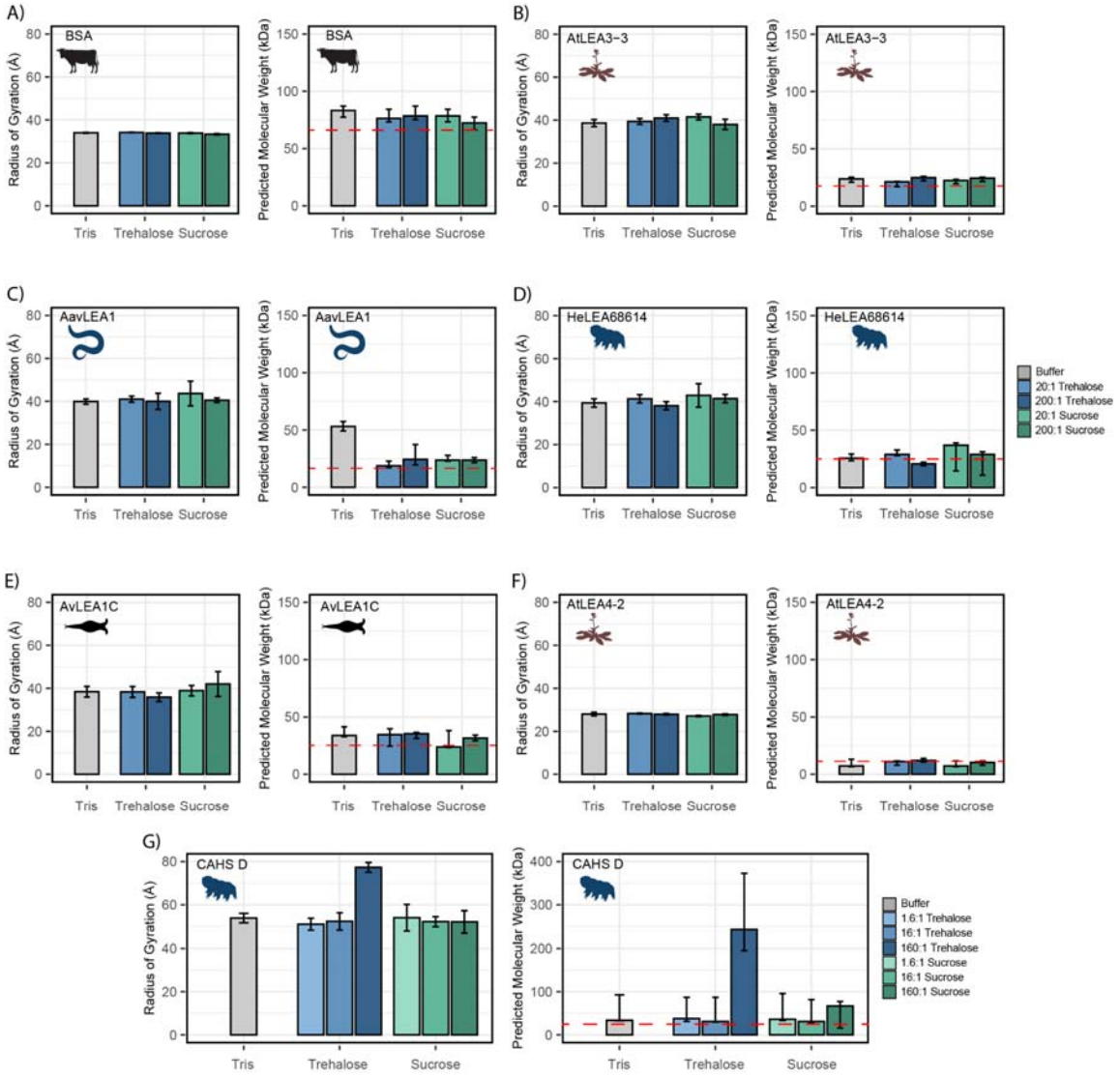


Figure 4: Cosolutes increase the global dimensions of CAHS D, but not LEA or BSA. Analyzed data from SAXS experiments of proteins at 4 mg/mL in 20 mM tris HCl pH 7. Proteins tested include **A)** BSA (0.0578 mM), **B)** AtLEA3-3 (0.22 mM), **C)** AavLEA1 (0.249 mM), **D)** HeLEA68614 (0.156 mM), **E)** AvLEA1C (0.163 mM), **F)** AtLEA4-2 (0.38 mM), and **G)** CAHS D (0.156 mM). The left plot shows the radius of gyration of the protein in the presence of no cosolute (gray), increasing molar ratios of trehalose (blue shades), and increasing molar ratios of sucrose (green shades). Error bars represent uncertainty in the measurement, provided by BioXTAS RAW. The right plot shows molecular weight values derived from Guinier analysis (see Methods). The red dashed line indicates the monomeric protein's molecular weight. Color scheme is the same as in the left figure. Error bars represent >90% confidence interval, which is directly obtained from the analysis. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.

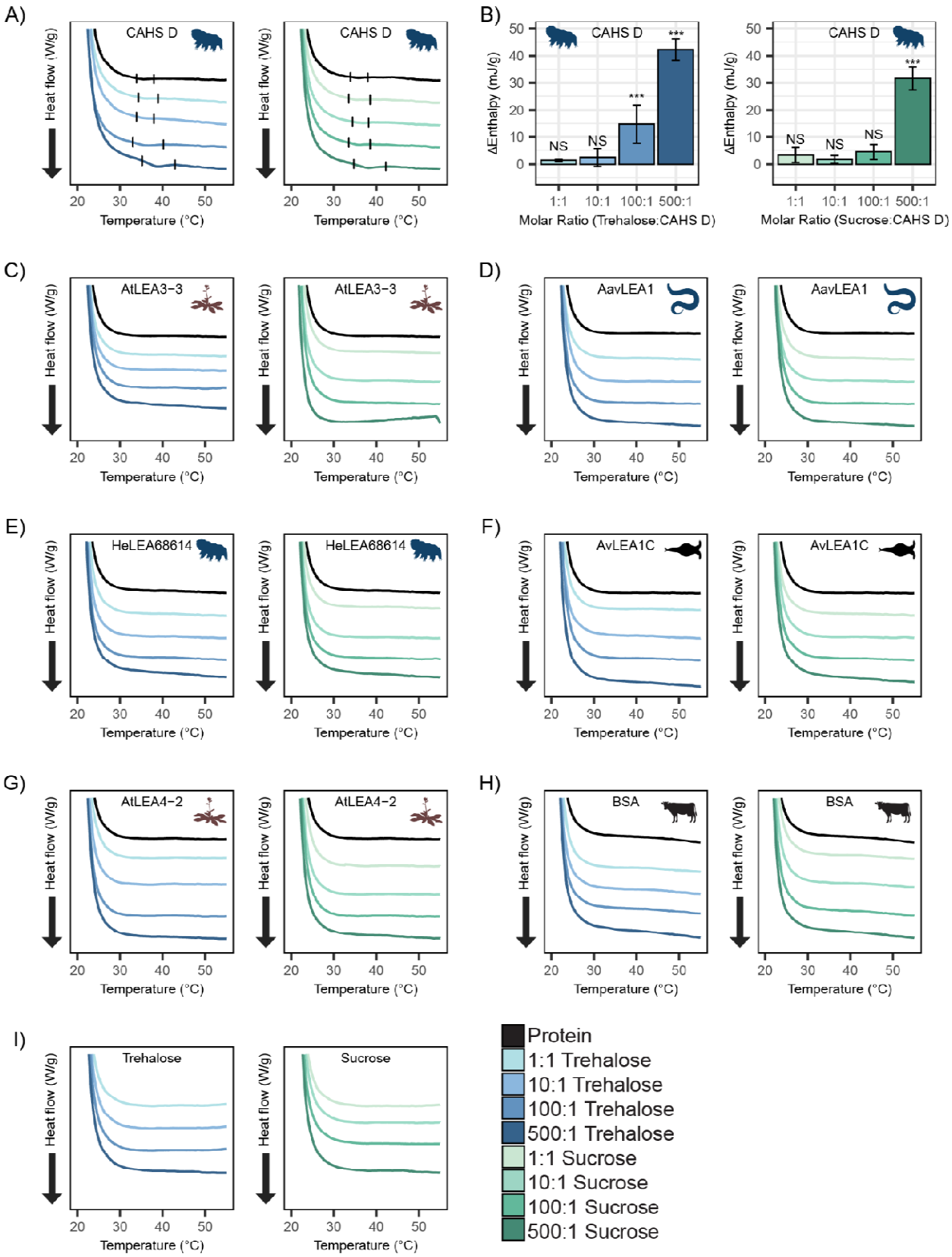

A)
CAHS D
Heat flow (W/g)
Temperature (°C)
B)
CAHS D
ΔEnthalpy (mJ/g)
NS
***
1:1
10:1
100:1
500:1
Molar Ratio (Trehalose:CAHS D)
Molar Ratio (Sucrose:CAHS D)
C)
AtLEA3-3
D)
AavLEA1
E)
HeLEA68614
F)
AvLEA1C
G)
AtLEA4-2
H)
BSA
I)
Trehalose
Sucrose
Protein
1:1 Trehalose
10:1 Trehalose
100:1 Trehalose
500:1 Trehalose
1:1 Sucrose
10:1 Sucrose
100:1 Sucrose
500:1 Sucrose

Figure 5: DSC thermograms show cosolutes promote gelation of CAHS D but not LEA proteins. A) DSC thermogram of 0.235 mM CAHS D with trehalose (left, blue lines) and sucrose (right, green lines) at increasing cosolute:protein molar ratios **B)** Change in enthalpy measurements for cosolute:CAHS D mixtures relative to CAHS D. Enthalpy measurements were done by taking the area of gel melt peaks represented by black dashes in (A) (see Methods). **C-H)** DSC thermogram of LEA proteins and BSA at 0.235 mM and with the addition of trehalose and sucrose at molar ratios of 1:1, 10:1, 100:1, and 500:1 (cosolute:protein). **I)** DSC thermogram for trehalose and sucrose in the absence of proteins at respective concentrations for different molar ratios. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.

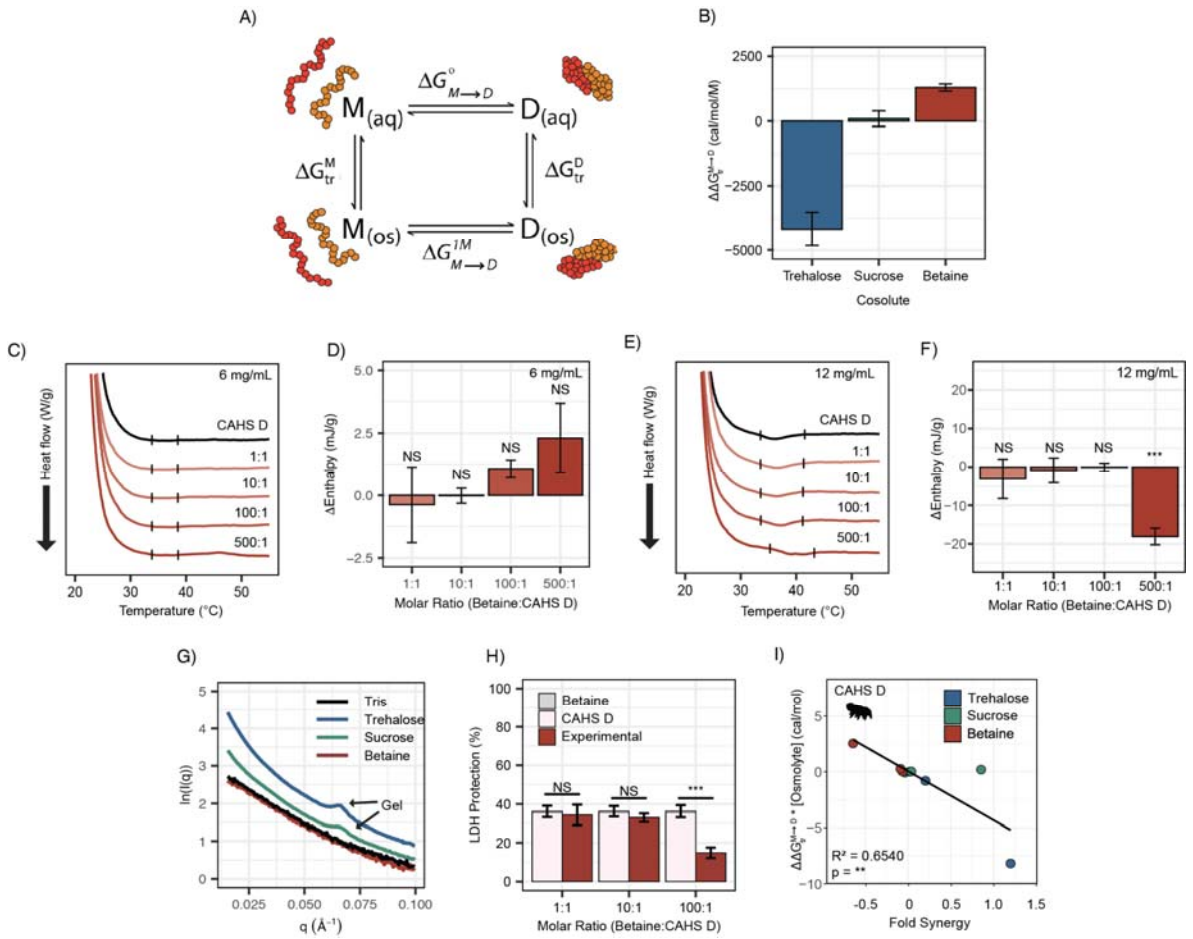
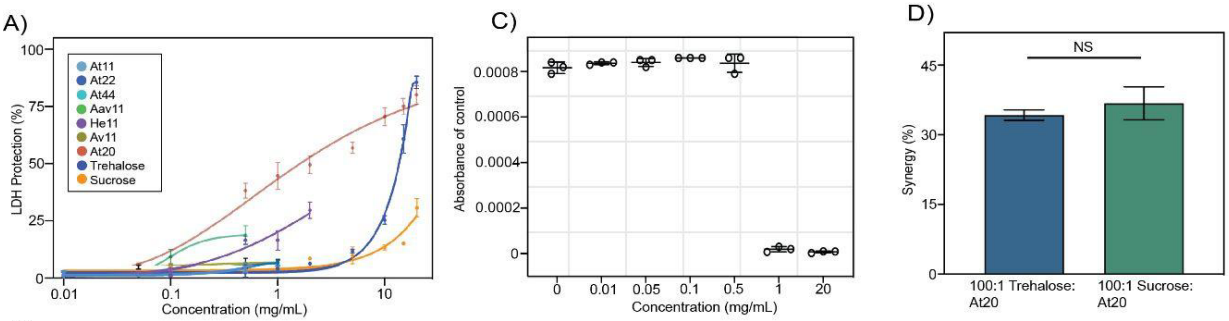


Figure 6: Transfer free energy of CAHS D highlights the difference between synergistic and non-synergistic cosolutes. A) Tanford's transfer model, depicting the effect of cosolutes on the dimerization of two proteins. "M" represents the protein's monomeric state, "D" represents the dimeric state, "aq" represents an aqueous solution, and "os" represents a solution containing some osmolyte. **B)** The difference in between two CAHS D monomers and a CAHS D dimer. All structures are predicted by AlphaFold2 [86,89]. **C)** DSC thermograms of CAHS D at 6 mg/mL (0.235 mM) in varying molar ratios of glycine betaine **D)** Change in enthalpy measurements for betaine:CAHS D mixtures relative to CAHS D at 6 mg/mL. Enthalpy measurements were done by taking the area of gel melt peaks represented by black dashes in Fig. 6C. **E)** Same as (C) but with CAHS D at 12 mg/mL (0.47 mM) **F)** Change in enthalpy measurements for betaine:CAHS D mixtures relative to CAHS D at 12 mg/mL. **G)** SAXS data depicting scattering profiles of 4 mg/mL CAHS D in tris (black), 1000:1 trehalose (blue), 1000:1 sucrose (green), and 1000:1 glycine betaine (red). **H)** LDH synergy assay for glycine betaine and CAHS D at various molar ratios. **I)** A correlation of of a given cosolute with CAHS D and its synergy in the LDH assay at different molar ratios. p-value is given from a Pearson correlation.

Organism	Enriched disaccharide during desiccation	Protein Family	LEA Motif	Representative protein	Length (amino acid)	Motif (%)	Metapredict Disorder
Arabidopsis thaliana	Sucrose (high) / Trehalose (high) Angeles-Núñez and Tiessen 2010, Fait et al. 2006	LEA_4	AKSKADETLES	AtLEA3-3	169	19.5%	
Aphelenchus avenae	Trehalose (high) Higa and Womersley 1993, Browne et al. 2004	LEA_4	LRDTAAEKLHQ	AavLEA1	143	30.06%	
Hypsibius exemplaris	Trehalose (low) Hengherr et al. 2008, Nguyen et al. 2022	LEA_4	LKDKAGSAWNQ	HeLEA68614	238	26.85%	
Adineta vaga	X Tunacliffe and Lapinski 2003, Lapinski and Tunacliffe 2003	LEA_4	KDKAAEALDAI	AvLEA1C	231	25.54%	
Arabidopsis thaliana	Sucrose (high) / Trehalose (high) Angeles-Núñez and Tiessen 2010, Fait et al. 2006	LEA_1	AKEKNIGGA KAQGHAEKTM	AtLEA4-2	97	73.19%	
Hypsibius exemplaris	Trehalose (low) Hengherr et al. 2008, Nguyen et al. 2022	CAHS	X	CAHS D	227	-	

Table 1. Summary of organisms, disaccharides, and proteins used in this study. Table displaying the organismal source of representative LEA_4, LEA_1 and CAHS proteins used in this study. In addition, the table displays endogenous cosolute reported in the literature to be co-enriched alongside LEA and CAHS proteins during desiccation in the given organism. The consensus sequence of 11-mer LEA_4 or 20-mer LEA_1 motifs as well as the length of the full-length proteins and predicted disorder using Metapredict [90] are shown. Shaded areas in the disorder plot correspond to the motif coordinates in the full-length LEA proteins. We note that the reason many of these profiles contain large folded regions is because the amphipathic LEA and CAHS proteins are predicted to form helices, which metapredict infers and incorrectly highlights these regions as ‘folded’ when really they are disordered in isolation.

Supplemental Figures



B)

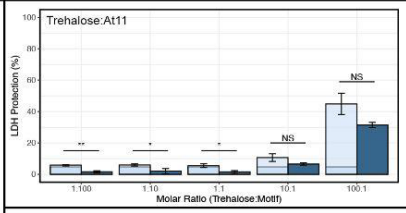
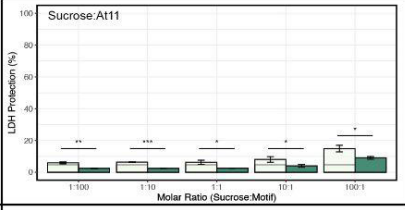
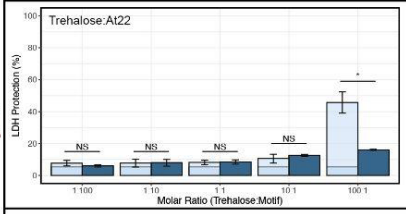
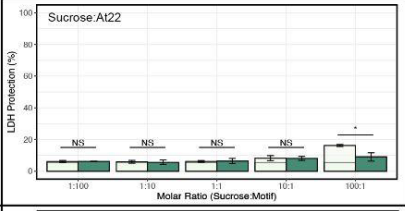
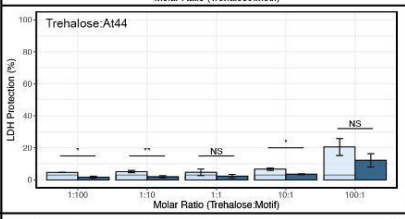
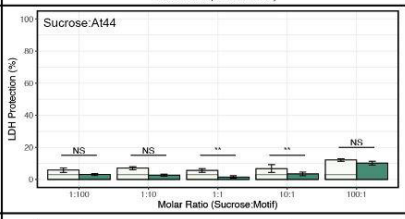
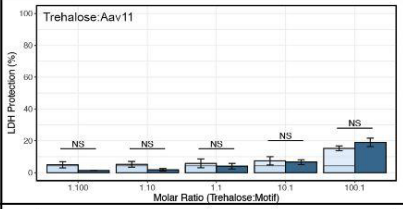
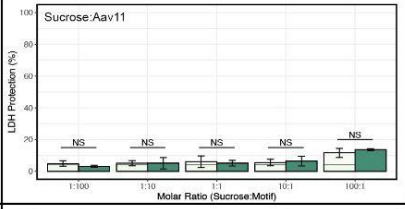
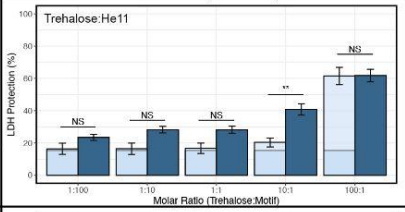
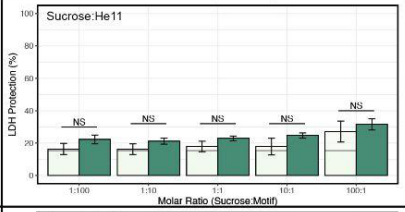
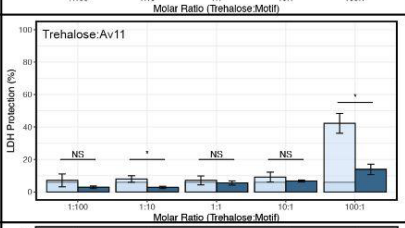
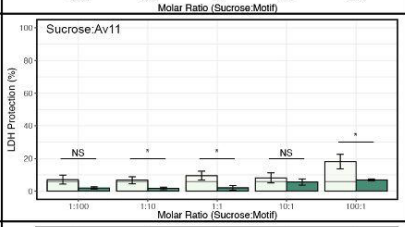
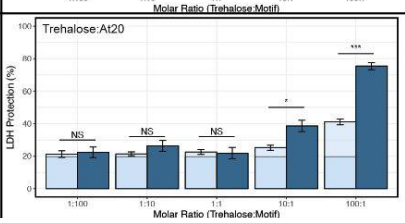
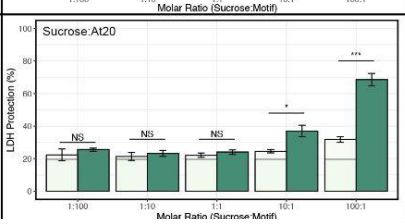
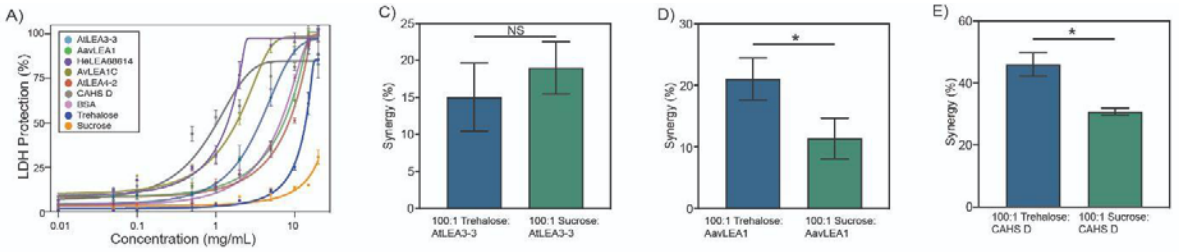
Organism	Trehalose:Motif	Sucrose:Motif	Synergy with Trehalose	Synergy with Sucrose
			✗	✗
			✗	✗
			✗	✗
			✗	✗
			✓ 19.4±6.1	✗
			✗	✗
			✓ 13.3±1.9, 34.2±1.1	✓ 12.4±2.5, 36.7±3.5

Figure S1. A) Sigmoidal plot representing concentration dependence of LDH protection by LEA motifs and cosolutes as a function of the weight/volume concentration **B)** Synergy plots for cosolutes with LEA motifs at 1:100, 1:10, 1:1, 10:1, and 100:1 molar ratios. n = 3, Welch's t-test was used for statistical comparison, error bars = standard deviation. Numbers under the checkmarks represent the mean percentage synergy and standard deviation **C)** Change in NADH absorbance values for the refrigerated controls of Aav11 at different concentrations after 16 hours **D)** Percent synergy comparison for At20 with trehalose and sucrose at 100:1 ratio. Welch's t-test was used for statistical comparison. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.



B)

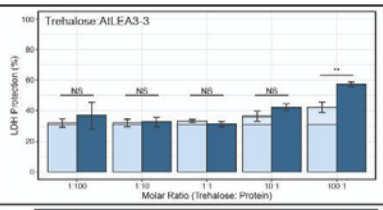
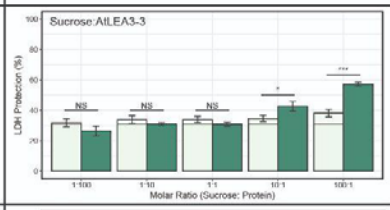
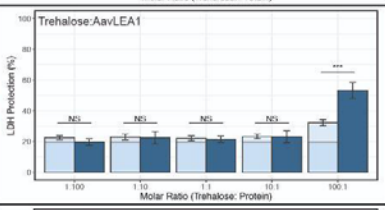
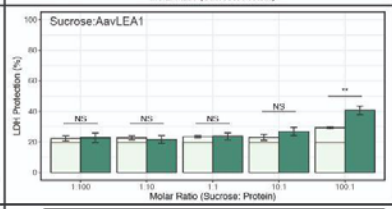
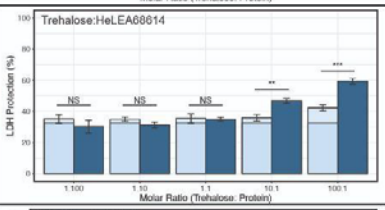
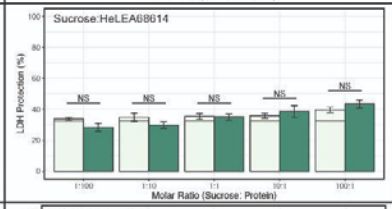
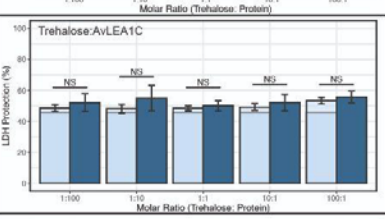
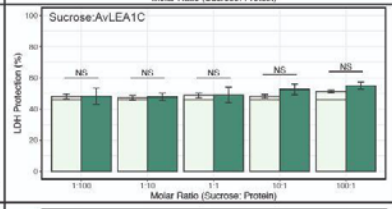
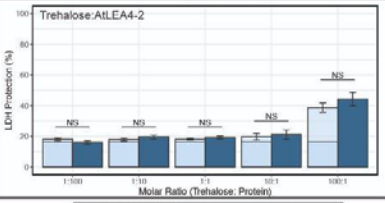
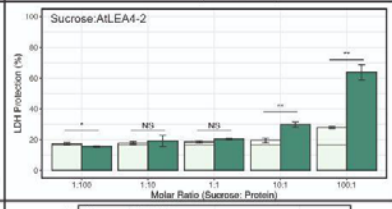
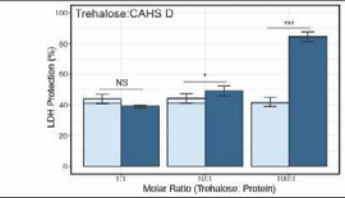
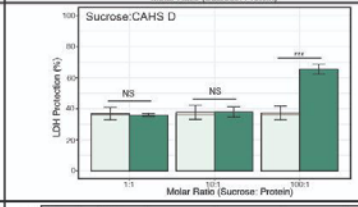
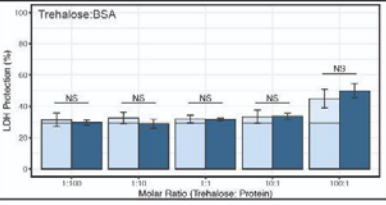
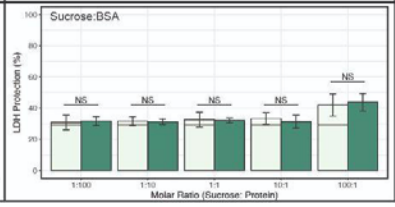
Organism	Trehalose:Protein	Sucrose:Protein	Synergy with	
			Trehalose	Sucrose
			✓ 15.3±4.6	✓ 7.9±3.6, 18.3±3.5
			✓ 21.0±3.4	✓ 11.3±3.3
			✓ 10.1±3.6, 17.0±3.2	✗
			✗	✗
			✗	✓ 10.2±0.2, 35.9±4.9
			✓ 8.3±4.3, 45.9±3.7	✓ 30.7±1.1
			✗	✗

Figure S2. A) Sigmoidal plot representing concentration dependence of LDH protection by full-length proteins and cosolutes as a function of the weight/volume concentration. **B)** Synergy plots for cosolute with full-length proteins at respective molar ratios. n = 3, Welch's t-test was used for statistical comparison, error bars = standard deviation. Numbers under the checkmarks represent the mean percentage synergy and standard deviation **C-E)** Percent synergy comparison for **C)** AtLEA3-3 **D)** AavLEA1 **E)** CAHS D with trehalose and sucrose at 100:1 ratio. n = 3, error bars = standard deviation. Welch's t-test was used for statistical comparison. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.

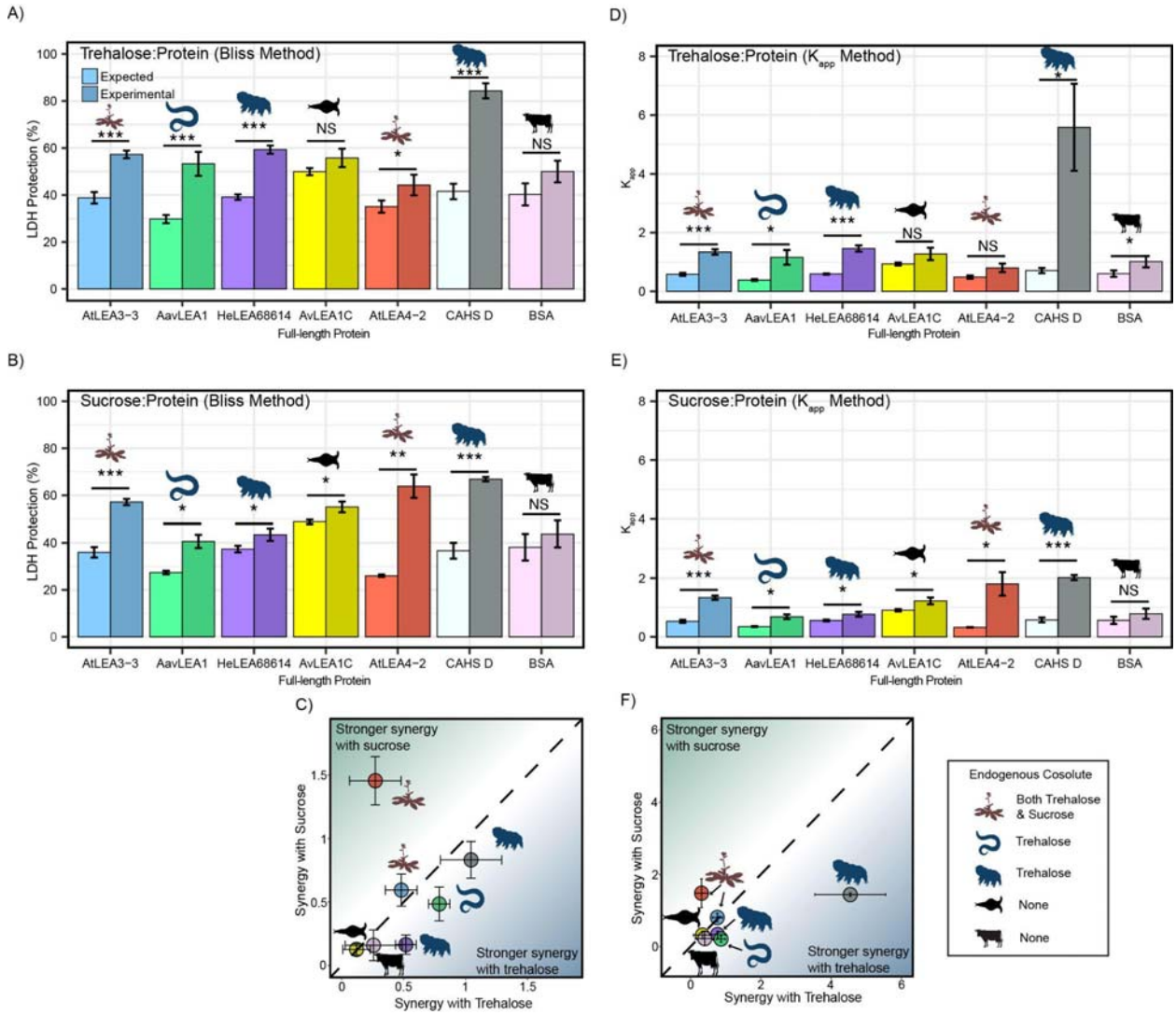


Figure S3. Synergy calculation using alternative methods. Synergy plots for **A)** trehalose:protein and **B)** sucrose:protein calculated using the Bliss method. **C)** Plot represents synergy with trehalose vs. sucrose using the Bliss method. **D-E)** Synergy plots for trehalose:protein and sucrose:protein respectively calculated using the K_{app} method. **F)** Plot represents synergy with trehalose vs. sucrose using the K_{app} method.

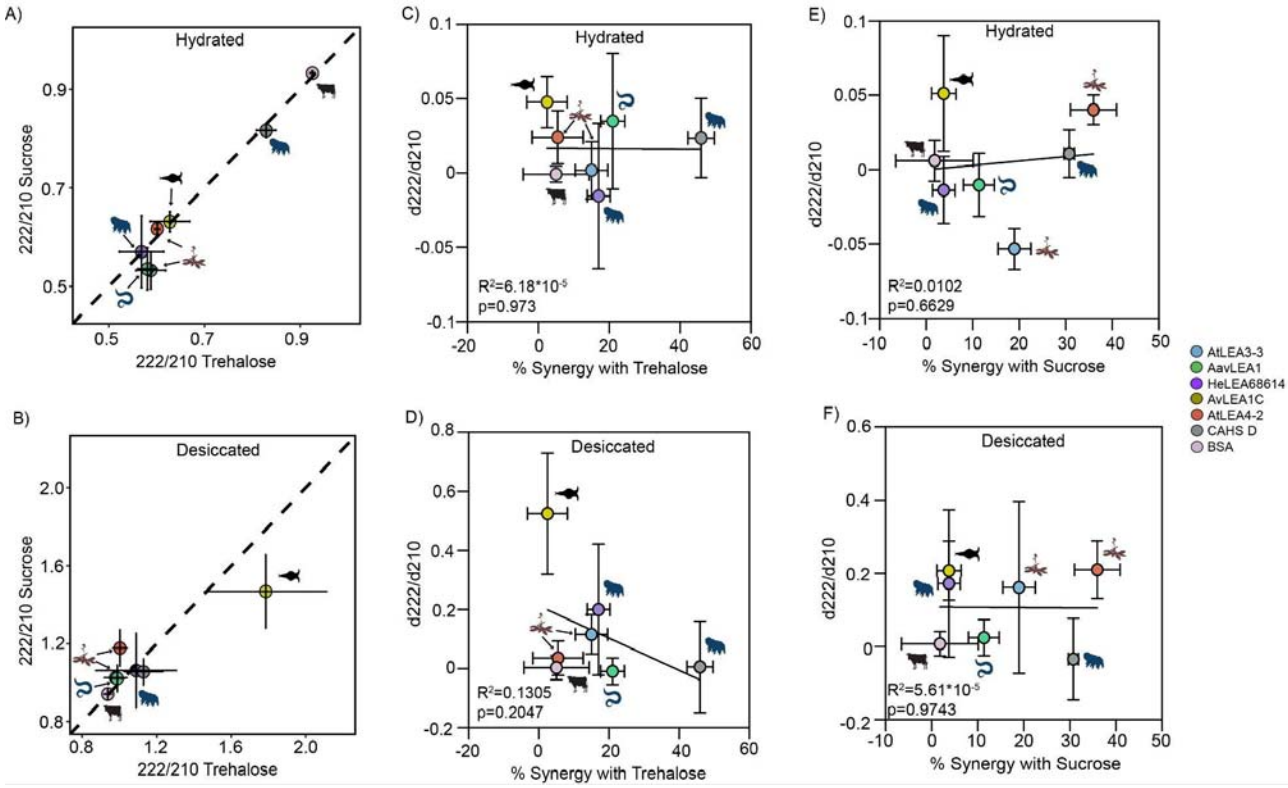


Figure S4. A) Plot comparing secondary structural changes in proteins with trehalose vs. sucrose at 100:1 molar ratio under hydrated conditions. Dotted line represents cases where there is equal structural shift in both trehalose and sucrose mixture. **B)** Plot comparing secondary structural alterations in proteins with trehalose vs. sucrose at 100:1 molar ratio under desiccated conditions. Dotted line represents cases where there is equal structural shift in both trehalose and sucrose mixture. **C)** Correlation plots for synergy vs. secondary structural shift in presence of trehalose in the hydrated state. **D)** Correlation plots for synergy vs. secondary structural shift in presence of trehalose in the desiccated state. **E)** Correlation plots for synergy vs. secondary structural shift in presence of sucrose in the hydrated state. **F)** Correlation plots for synergy vs. secondary structural shift in presence of sucrose in the desiccated state. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.

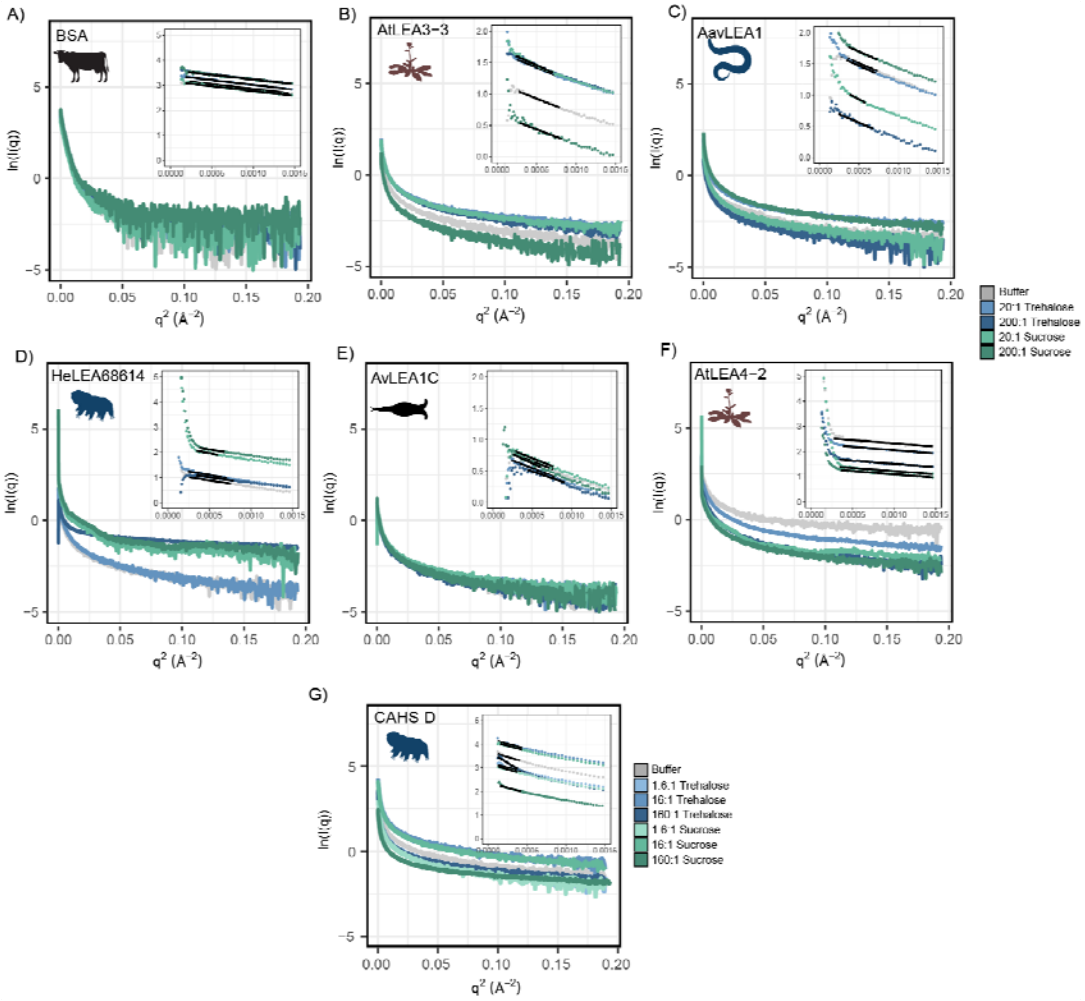


Figure S5. Raw SAXS data from the experiments shown in Fig. 4 for 4 mg/mL of **A)** BSA (0.0578 mM), **B)** AtLEA3-3 (0.22 mM), **C)** AavLEA1 (0.249 mM), **D)** HeLEA68614 (0.156 mM), **E)** AvLEA1C (0.163 mM), **F)** AtLEA4-2 (0.38 mM), and **G)** CAHS D (0.156 mM). Each plot displays raw SAXS data on a Guinier scale, with a zoomed-in portion in the top right displaying the Guinier region. The Guinier fit is displayed as a black trendline. In each plot, the protein's scattering profile in several solution environments is shown, including protein in the presence of no cosolute (black), a 20:1 molar ratio of trehalose (light blue), a 200:1 molar ratio of trehalose (dark blue), a 20:1 molar ratio of sucrose (light green), a 200:1 molar ratio of sucrose (dark green). CAHS D data contains slightly different molar ratios of 1.6:1 trehalose (very light blue), 16:1 trehalose (light blue), 160:1 trehalose (dark blue), 1.6:1 sucrose (very light green), 16:1 sucrose (light green), 160:1 sucrose (dark green). Unlike other samples, a Guinier analysis with a $qMaxR_g$ of 1.1 could not be obtained. A best attempt to establish the R_g of this sample ($qMaxR_g$ ~ 1.35) revealed a significant increase in the R_g. Given the clear curvature of the Guinier region, this data was consistent with the presence of large oligomeric species. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.

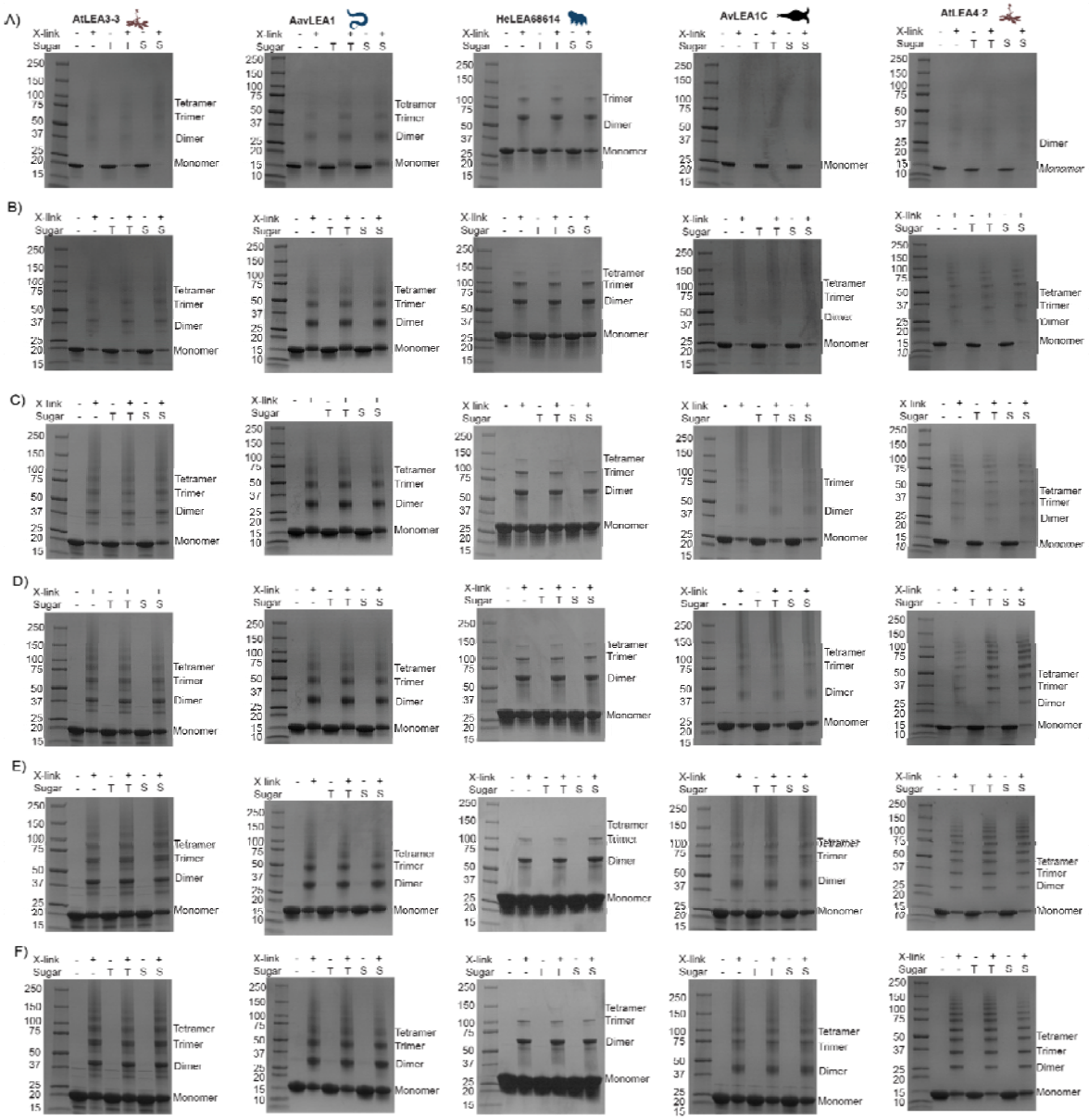


Figure S6. Crosslinking gels for full-length LEA proteins and in mixtures with trehalose or sucrose. A range of protein concentrations were used to cover the concentrations used in the synergy assay. Protein concentrations used were **A)** 25 μM, **B)** 50 μM, **C)** 75 μM, **D)** 100 μM, **E)** 150 μM, and **F)** 200 μM. Each gel contains protein ladder (Precision Plus Protein™ Dual Xtra Standards, Bio Rad, Catalog #161-0377) in the first lane, protein at desired concentration in the second lane, crosslinked protein in the third lane, 100:1 trehalose:protein in the fourth lane, crosslinked 100:1 trehalose:protein in the fifth lane, 100:1 sucrose:protein in the sixth lane, and crosslinked 100:1 sucrose:protein in the last lane. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.

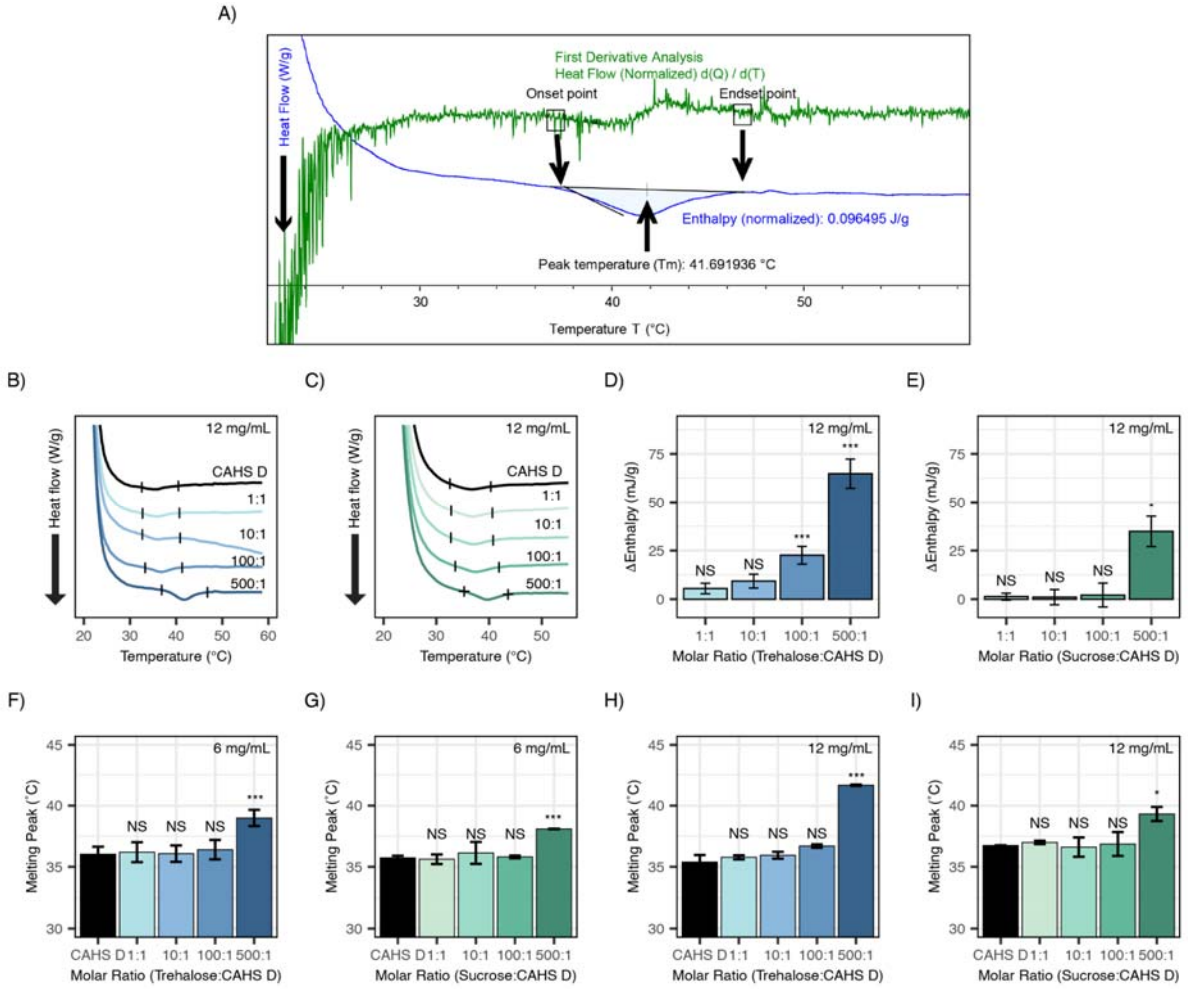


Figure S7. A) Sample enthalpy analysis using TRIOS software. First derivative analysis was used to aid in selecting the onset and endset in the thermogram representative of protein gelation. This is done by identifying observable spikes that correspond to where the curve begins to dip (onset) or level out (endset). The area between the approximate onset and endset is selected and TRIOS software is used to calculate defined onset and endset values, enthalpy (J/g), and peak temperature (Tm, ˚C) by analyzing the area above this feature within the identified range. **B)** DSC thermograms of CAHS D at 12 mg/mL (0.47 mM) in varying molar ratios of trehalose (0:1, 1:1, 10:1, 100:1, and 500:1). **C)** DSC thermograms of CAHS D at 12 mg/mL (0.47 mM) in varying molar ratios of sucrose (0:1, 1:1, 10:1, 100:1, and 500:1). **D)** Change in enthalpy measurements for trehalose:CAHS D mixtures relative to CAHS D at 12 mg/mL. Enthalpy measurements were done by taking the area of gel melt peaks represented by black dashes in Fig. S7B **E)** Change in enthalpy measurements for sucrose:CAHS D mixtures relative to CAHS D at 12 mg/mL. Enthalpy measurements were done by taking the area of gel melt peaks represented by black dashes in Fig. S7C. **(F-I)** Melting temperature for the gel melts as calculated from the DSC thermograms for CAHS D and cosolute:CAHS D mixtures at respective ratios.

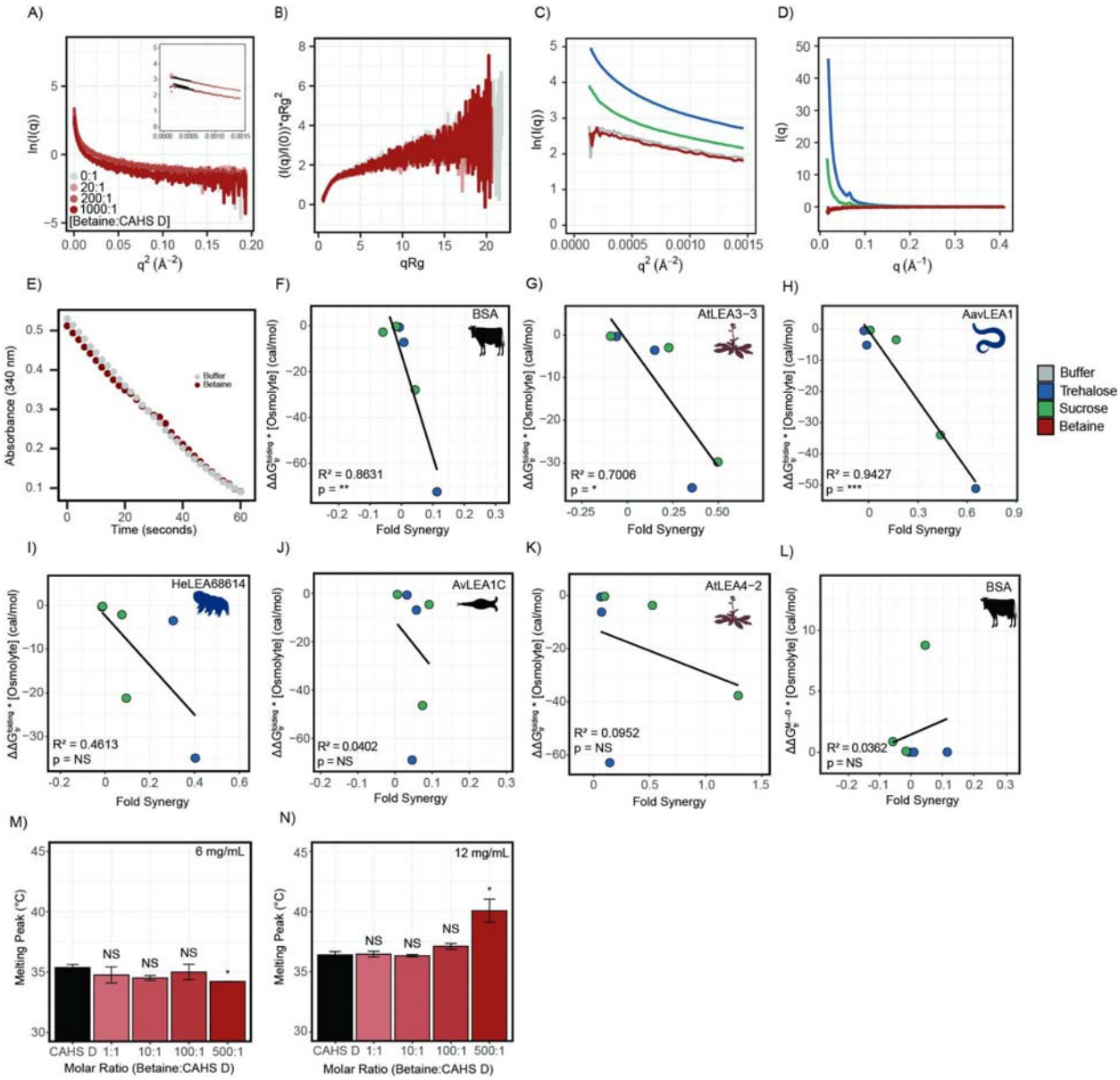


Figure S8. Raw SAXS data from the experiments shown in Fig. 6 for 4 mg/mL of CAHS D in various solution conditions. **A)** Scattering curves for CAHS D in tris (gray), 10:1 glycine betaine (light red), 100:1 glycine betaine (medium red), 1000:1 glycine betaine (dark red). Guinier regions are displayed in the top left with Guinier fits drawn in black. **B)** Dimensionless Kratky analysis of the data from A. **C)** Zoomed in Guinier regions of CAHS D in tris (gray), 1000:1 glycine betaine (red), 1000:1 trehalose (blue), and 1000:1 sucrose (green). **D)** Scattering profiles for CAHS D in glycine betaine, trehalose, and sucrose normalized to the protein's scattering profile in tris. **E)** Positive controls of LDH in Tris buffer (20 mM, pH 7.0) and highest concentration of glycine betaine used in this study. Absorbance readings at 340 nm measure NADH reporter as a function of time. **F-K)** Scatterplots of $\Delta\Delta G_{tr}^{folding}$ and synergy for BSA and LEA proteins. All fits are done using linear regression. All p-values derived from Pearson

correlation. **L)** Scatterplot of $\Delta\Delta G_{tr}^{M \rightarrow \boldsymbol{D}}$ and synergy for BSA. **M-N)** Melting temperature for the gel melts as calculated from the DSC thermograms for CAHS D and Betaine:CAHS D mixtures at respective ratios. Organismal icons that are colored blue indicate the species uses trehalose as an endogenous cosolute, brown indicates the species uses both trehalose and sucrose, black indicates the species uses neither trehalose nor sucrose as an endogenous cosolute.

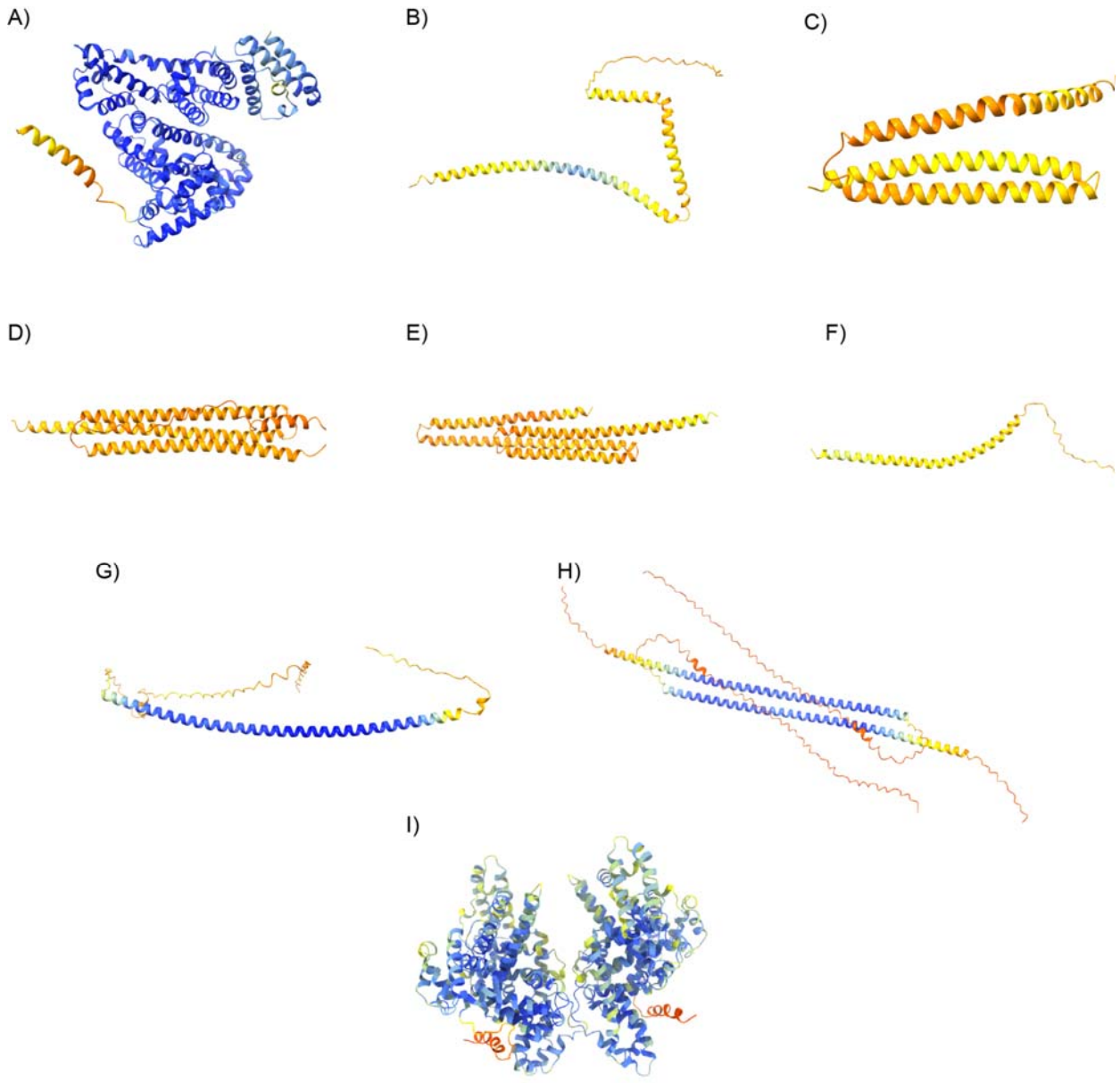


Figure S9. Representative images of the pdb structures obtained from our AlphaFold2 predictions. All structures are colored by their pLDDT on a scale from 0 (red) to 100 (blue). **A**) Monomeric BSA. **B**) Monomeric AtLEA3-3. **C**) Monomeric AavLEA1. **D**) Monomeric HeLEA68614. **E**) Monomeric AvLEA1C. **F**) Monomeric AtLEA4-2. **G**) Monomeric CAHS D. **H**) Dimeric CAHS D. **I)** Dimeric BSA. See supplementary data for pdb files of each prediction & their exact pLDDT. We note that the structural predictions from AlphaFold2 contain largely ordered structures. We believe this is due to the propensity of these proteins to form helices in the context of drying or when interacting with a client. This has been shown in cases where an IDR contains residual helicity or is folded upon binding [70].

Organism	Motif	Concentration (mM)
	At11	0.297
	At22	0.277
	At44	0.161
	Aav11	0.0562
	He11	0.287
	Av11	0.262
	At20	0.264

Table S1. Protein concentrations used in synergy assays for LEA motifs in this study.

Organism	Protein	Concentration (mM)
	AtLEA3-3	0.1215
	AavLEA1	0.112
	HeLEA68614	0.0392
	AvLEA1C	0.068
	AtLEA4-2	0.217
	CAHS D	0.0078
	BSA	0.054

Table S2. Protein concentrations used in synergy assays for full-length proteins in this study.

References

1. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol. 2004;337: 635–645.

2. Holehouse AS, Kragelund BB. The molecular basis for cellular function of intrinsically disordered protein regions. Nat Rev Mol Cell Biol. 2023. doi:10.1038/s41580-023-00673-0

3. van der Lee R, Buljan M, Lang B, Weatheritt RJ, Daughdrill GW, Dunker AK, et al. Classification of intrinsically disordered regions and proteins. Chem Rev. 2014;114: 6589–6631.

4. Mao AH, Lyle N, Pappu RV. Describing sequence-ensemble relationships for intrinsically disordered proteins. Biochem J. 2013;449: 307–318.

5. Kedia N, Almisry M, Bieschke J. Glucose directs amyloid-beta into membrane-active oligomers. Phys Chem Chem Phys. 2017;19: 18036–18046.

6. Das RK, Ruff KM, Pappu RV. Relating sequence encoded information to form and function of intrinsically disordered proteins. Curr Opin Struct Biol. 2015;32: 102–112.

7. Soranno A, Koenig I, Borgia MB, Hofmann H, Zosel F, Nettels D, et al. Single-molecule spectroscopy reveals polymer effects of disordered proteins in crowded environments. Proc Natl Acad Sci U S A. 2014;111: 4874–4879.

8. Moses D, Yu F, Ginell GM, Shamoon NM, Koenig PS, Holehouse AS, et al. Revealing the Hidden Sensitivity of Intrinsically Disordered Proteins to their Chemical Environment. J Phys Chem Lett. 2020;11: 10131–10136.

9. Moses D, Ginell GM, Holehouse AS, Sukenik S. Intrinsically disordered regions are poised to act as sensors of cellular chemistry. Trends Biochem Sci. 2023;48: 1019–1034.

10. Theillet F-X, Binolfi A, Frembgen-Kesner T, Hingorani K, Sarkar M, Kyne C, et al. Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). Chem Rev. 2014;114: 6661–6714.

11. Crowe JH. Anhydrobiosis: an unsolved problem. Plant, cell & environment. 2014. pp. 1491–1493.

12. Hesgrove C, Boothby TC. The biology of tardigrade disordered proteins in extreme stress tolerance. Cell Commun Signal. 2020;18: 178.

13. Romero-Perez PS, Dorone Y, Flores E, Sukenik S, Boeynaems S. When Phased without Water: Biophysics of Cellular Desiccation, from Biomolecules to Condensates. Chem Rev. 2023;123: 9010–9035.

14. Oliver MJ, Farrant JM, Hilhorst HWM, Mundree S, Williams B, Bewley JD. Desiccation Tolerance: Avoiding Cellular Damage During Drying and Rehydration. Annu Rev Plant Biol. 2020;71: 435–460.

15. Koster KL, Leopold AC. Sugars and desiccation tolerance in seeds. Plant Physiol. 1988;88:

829–832.

16. Angeles-Núñez JG, Tiessen A. Arabidopsis sucrose synthase 2 and 3 modulate metabolic homeostasis and direct carbon towards starch synthesis in developing seeds. Planta. 2010;232: 701–718.

17. Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie AR, et al. Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. Plant Physiol. 2006;142: 839–854.

18. Ryabova A, Cornette R, Cherkasov A, Watanabe M, Okuda T, Shagimardanova E, et al. Combined metabolome and transcriptome analysis reveals key components of complete desiccation tolerance in an anhydrobiotic insect. Proc Natl Acad Sci U S A. 2020;117: 19209–19220.

19. Nguyen K, Kc S, Gonzalez T, Tapia H, Boothby TC. Trehalose and tardigrade CAHS proteins work synergistically to promote desiccation tolerance. Commun Biol. 2022;5: 1046.

20. Tapia H, Koshland DE. Trehalose is a versatile and long-lived chaperone for desiccation tolerance. Curr Biol. 2014;24: 2758–2766.

21. Hincha DK, Thalhammer A. LEA proteins: IDPs with versatile functions in cellular dehydration tolerance. Biochem Soc Trans. 2012;40: 1000–1003.

22. Boothby TC, Tapia H, Brozena AH, Piszkiewicz S, Smith AE, Giovannini I, et al. Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation. Mol Cell. 2017;65: 975–984.e5.

23. Hernández-Sánchez IE, Maruri-López I, Martinez-Martinez C, Janis B, Jiménez-Bremont JF, Covarrubias AA, et al. LEAfing through literature: late embryogenesis abundant proteins coming of age—achievements and perspectives. J Exp Bot. 2022;73: 6525–6546.

24. Hand SC, Menze MA, Toner M, Boswell L, Moore D. LEA proteins during water stress: not just for plants anymore. Annu Rev Physiol. 2011;73: 115–134.

25. Hundertmark M, Hincha DK. LEA (late embryogenesis abundant) proteins and their encoding genes in Arabidopsis thaliana. BMC Genomics. 2008;9: 118.

26. Artur MAS, Zhao T, Ligterink W, Schranz E, Hilhorst HWM. Dissecting the Genomic Diversification of Late Embryogenesis Abundant (LEA) Protein Gene Families in Plants. Genome Biol Evol. 2019;11: 459–471.

27. Yamaguchi A, Tanaka S, Yamaguchi S, Kuwahara H, Takamura C, Imajoh-Ohmi S, et al. Two novel heat-soluble protein families abundantly expressed in an anhydrobiotic tardigrade. PLoS One. 2012;7: e44209.

28. Goyal K, Walton LJ, Tunnacliffe A. LEA proteins prevent protein aggregation due to water stress. Biochem J. 2005;388: 151–157.

29. Kim SX, Çamdere G, Hu X, Koshland D, Tapia H. Synergy between the small intrinsically disordered protein Hsp12 and trehalose sustain viability after severe desiccation. Elife. 2018;7. doi:10.7554/eLife.38337

30. LeBlanc BM, Hand SC. Target enzymes are stabilized by AfrLEA6 and a gain of α-helix coincides with protection by a group 3 LEA protein during incremental drying. Biochim Biophys Acta: Proteins Proteomics. 2021;1869: 140642.

31. Sanchez-Martinez S, Nguyen K, Biswas S, Nicholson V, Romanyuk AV, Ramirez J, et al. Labile assembly of a tardigrade protein induces biostasis. bioRxiv. 2023. p. 2023.06.30.547219. doi:10.1101/2023.06.30.547219

32. Eicher JE, Brom JA, Wang S, Sheiko SS, Atkin JM, Pielak GJ. Secondary structure and stability of a gel-forming tardigrade desiccation-tolerance protein. Protein Sci. 2022;31: e4495.

33. Malki A, Teulon J-M, Camacho-Zarco AR, Chen S-WW, Adamski W, Maurin D, et al. Intrinsically Disordered Tardigrade Proteins Self-Assemble into Fibrous Gels in Response to Environmental Stress. Angew Chem Int Ed Engl. 2022;61: e202109961.

34. Yagi-Utsumi M, Aoki K, Watanabe H, Song C, Nishimura S, Satoh T, et al. Desiccation-induced fibrous condensation of CAHS protein from an anhydrobiotic tardigrade. Sci Rep. 2021;11: 21328.

35. Tanaka A, Nakano T, Watanabe K, Masuda K, Honda G, Kamata S, et al. Stress-dependent cell stiffening by tardigrade tolerance proteins that reversibly form a filamentous network and gel. PLoS Biol. 2022;20: e3001780.

36. Battaglia M, Olvera-Carrillo Y, Garciarrubio A, Campos F, Covarrubias AA. The enigmatic LEA proteins and other hydrophilins. Plant Physiol. 2008;148: 6–24.

37. Higa LM, Womersley CZ. New insights into the anhydrobiotic phenomenon: The effects of trehalose content and differential rates of evaporative water loss on the survival of *Aphelenchus avenae*. J Exp Zool. 1993;267: 120–129.

38. Browne JA, Dolan KM, Tyson T, Goyal K, Tunnacliffe A, Burnell AM. Dehydration-specific induction of hydrophilic protein genes in the anhydrobiotic nematode Aphelenchus avenae. Eukaryot Cell. 2004;3: 966–975.

39. Hengherr S, Heyer AG, Köhler H-R, Schill RO. Trehalose and anhydrobiosis in tardigrades--evidence for divergence in responses to dehydration. FEBS J. 2008;275: 281–288.

40. Tunnacliffe A, Lapinski J. Resurrecting Van Leeuwenhoek's rotifers: a reappraisal of the role of disaccharides in anhydrobiosis. Philos Trans R Soc Lond B Biol Sci. 2003;358: 1755–1771.

41. Lapinski J, Tunnacliffe A. Anhydrobiosis without trehalose in bdelloid rotifers. FEBS Lett. 2003;553: 387–390.

42. Dure L 3rd. A repeating 11-mer amino acid motif and plant desiccation. Plant J. 1993;3: 363–369.

43. Knox-Brown P, Rindfleisch T, Günther A, Balow K, Bremer A, Walther D, et al. Similar Yet Different-Structural and Functional Diversity among LEA_4 Proteins. Int J Mol Sci. 2020;21. doi:10.3390/ijms21082794

44. Furuki T, Sakurai M. Group 3 LEA protein model peptides protect enzymes against

desiccation stress. Biochim Biophys Acta. 2016;1864: 1237–1243.

45. Furuki T, Shimizu T, Chakrabortee S, Yamakawa K, Hatanaka R, Takahashi T, et al. Effects of Group 3 LEA protein model peptides on desiccation-induced protein aggregation. Biochim Biophys Acta. 2012;1824: 891–897.

46. Furuki T, Sakurai M. Group 3 LEA protein model peptides protect liposomes during desiccation. Biochim Biophys Acta. 2014;1838: 2757–2766.

47. Hibshman JD, Goldstein B. LEA motifs promote desiccation tolerance in vivo. BMC Biol. 2021;19: 263.

48. Piszkiewicz S, Gunn KH, Warmuth O, Propst A, Mehta A, Nguyen KH, et al. Protecting activity of desiccated enzymes. Protein Sci. 2019;28: 941–951.

49. Shimizu T, Kanamori Y, Furuki T, Kikawada T, Okuda T, Takahashi T, et al. Desiccation-induced structuralization and glass formation of group 3 late embryogenesis abundant protein model peptides. Biochemistry. 2010;49: 1093–1104.

50. Wang S, Eicher J, Pielak GJ. Trifluoroethanol and the behavior of a tardigrade desiccation-tolerance protein. Protein Sci. 2023;32: e4716.

51. Cuevas-Velazquez CL, Saab-Rincón G, Reyes JL, Covarrubias AA. The Unstructured N-terminal Region of Arabidopsis Group 4 Late Embryogenesis Abundant (LEA) Proteins Is Required for Folding and for Chaperone-like Activity under Water Deficit. J Biol Chem. 2016;291: 10893–10903.

52. Bremer A, Wolff M, Thalhammer A, Hincha DK. Folding of intrinsically disordered plant LEA proteins is driven by glycerol-induced crowding and the presence of membranes. FEBS J. 2017;284: 919–936.

53. Navarro-Retamal C, Bremer A, Alzate-Morales J, Caballero J, Hincha DK, González W, et al. Molecular dynamics simulations and CD spectroscopy reveal hydration-induced unfolding of the intrinsically disordered LEA proteins COR15A and COR15B from Arabidopsis thaliana. Phys Chem Chem Phys. 2016;18: 25806–25816.

54. Greenfield NJ. Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc. 2006;1: 2876–2890.

55. Kachala M, Valentini E, Svergun DI. Application of SAXS for the Structural Characterization of IDPs. Adv Exp Med Biol. 2015;870: 261–289.

56. Graewert MA, Franke D, Jeffries CM, Blanchet CE, Ruskule D, Kuhle K, et al. Automated pipeline for purification, biophysical and x-ray analysis of biomacromolecular solutions. Sci Rep. 2015;5: 10734.

57. Babcock JJ, Brancaleon L. Bovine serum albumin oligomers in the E- and B-forms at low protein concentration and ionic strength. Int J Biol Macromol. 2013;53: 42–53.

58. Levi V, González Flecha FL. Reversible fast-dimerization of bovine serum albumin detected by fluorescence resonance energy transfer. Biochim Biophys Acta. 2002;1599: 141–148.

59. Shou K, Bremer A, Rindfleisch T, Knox-Brown P, Hirai M, Rekas A, et al. Conformational

selection of the intrinsically disordered plant stress protein COR15A in response to solution osmolarity – an X-ray and light scattering study. Phys Chem Chem Phys. 2019;21: 18727–18740.

60. Rivera-Najera LY, Saab-Rincón G, Battaglia M, Amero C, Pulido NO, García-Hernández E, et al. A group 6 late embryogenesis abundant protein from common bean is a disordered protein with extended helical structure and oligomer-forming properties. J Biol Chem. 2014;289: 31995–32009.

61. Fancy DA, Kodadek T. Chemistry for the analysis of protein-protein interactions: rapid and efficient cross-linking triggered by long wavelength light. Proc Natl Acad Sci U S A. 1999;96: 6020–6024.

62. Preston GW, Wilson AJ. Photo-induced covalent cross-linking for the analysis of biomolecular interactions. Chem Soc Rev. 2013;42: 3289–3301.

63. Liu G, Liu K, Gao Y, Zheng Y. Involvement of C-Terminal Histidines in Soybean PM1 Protein Oligomerization and Cu2+ Binding. Plant Cell Physiol. 2017;58: 1018–1029.

64. Eicher J, Hutcheson BO, Pielak GJ. Properties of a tardigrade desiccation-tolerance protein aerogel. Biophys J. 2023;122: 2500–2505.

65. Belott C, Janis B, Menze MA. Liquid-liquid phase separation promotes animal desiccation tolerance. Proc Natl Acad Sci U S A. 2020;117: 27676–27684.

66. Tanford C. Protein denaturation. C. Theoretical models for the mechanism of denaturation. Adv Protein Chem. 1970;24: 1–95.

67. Auton M, Bolen DW. Application of the transfer model to understand how naturally occurring osmolytes affect protein stability. Methods Enzymol. 2007;428: 397–418.

68. Auton M, Bolen DW. Predicting the energetics of osmolyte-induced protein folding/unfolding. Proc Natl Acad Sci U S A. 2005;102: 15065–15068.

69. Hong J, Gierasch LM, Liu Z. Its preferential interactions with biopolymers account for diverse observed effects of trehalose. Biophys J. 2015;109: 144–153.

70. Alderson TR, Pritišanac I, Kolarić Đ, Moses AM, Forman-Kay JD. Systematic identification of conditionally folded intrinsically disordered regions by AlphaFold2. Proc Natl Acad Sci U S A. 2023;120: e2304302120.

71. Wolkers WF, McCready S, Brandt WF, Lindsey GG, Hoekstra FA. Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses in vitro. Biochim Biophys Acta. 2001;1544: 196–206.

72. Shih M-D, Hsieh T-Y, Jian W-T, Wu M-T, Yang S-J, Hoekstra FA, et al. Functional studies of soybean (Glycine max L.) seed LEA proteins GmPM6, GmPM11, and GmPM30 by CD and FTIR spectroscopy. Plant Sci. 2012;196: 152–159.

73. Hincha DK, Zuther E, Popova AV. Stabilization of Dry Sucrose Glasses by Four LEA_4 Proteins from. Biomolecules. 2021;11. doi:10.3390/biom11050615

74. Sakurai M, Furuki T, Akao K-I, Tanaka D, Nakahara Y, Kikawada T, et al. Vitrification is

essential for anhydrobiosis in an African chironomid, Polypedilum vanderplanki. Proc Natl Acad Sci U S A. 2008;105: 5093–5098.

75. Water content, transition temperature and fragility influence protection and anhydrobiotic capacity. BBA Advances. 2024; 100115.

76. Crowe JH, Carpenter JF, Crowe LM. The role of vitrification in anhydrobiosis. Annu Rev Physiol. 1998;60: 73–103.

77. Simperler A, Kornherr A, Chopra R, Bonnet PA, Jones W, Motherwell WDS, et al. Glass transition temperature of glucose, sucrose, and trehalose: an experimental and in silico study. J Phys Chem B. 2006;110: 19678–19684.

78. Riback JA, Zhu L, Ferrolino MC, Tolbert M, Mitrea DM, Sanders DW, et al. Composition-dependent thermodynamics of intracellular phase separation. Nature. 2020;581: 209–214.

79. Sarkar M, Li C, Pielak GJ. Soft interactions and crowding. Biophys Rev. 2013;5: 187–194.

80. Dyer KN, Hammel M, Rambo RP, Tsutakawa SE, Rodic I, Classen S, et al. High-throughput SAXS for the characterization of biomolecules in solution: a practical approach. Methods Mol Biol. 2014;1091: 245–258.

81. Trame C, MacDowell AA, Celestre RS, Padmore HA, Cambie D, Domning EE, et al. SIBYLS—a SAXS and Protein Crystallography Beamline at the ALS. AIP Conference Proceedings. American Institute of Physics; 2004. pp. 502–505.

82. Nielsen SS, Toft KN, Snakenborg D, Jeppesen MG, Jacobsen JK, Vestergaard B, et al. BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. J Appl Crystallogr. 2009;42: 959–964.

83. Hopkins JB, Gillilan RE, Skou S. BioXTAS RAW: a free open-source program for small-angle X-ray scattering data reduction. Foundations of Crystallography. 2018;74: a219.

84. Rieloff E, Skepö M. Determining Rg of IDPs from SAXS Data. Methods Mol Biol. 2020;2141: 271–283.

85. Hajizadeh NR, Franke D, Jeffries CM, Svergun DI. Consensus Bayesian assessment of protein molecular mass from solution X-ray scattering data. Sci Rep. 2018;8: 7204.

86. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596: 583–589.

87. Auton M, Bolen DW. Additive transfer free energies of the peptide backbone unit that are independent of the model compound and the choice of concentration scale. Biochemistry. 2004;43: 1329–1342.

88. Lalmansingh JM, Keeley AT, Ruff KM, Pappu RV, Holehouse AS. SOURSOP: A Python package for the analysis of simulations of intrinsically disordered proteins. bioRxiv. 2023. doi:10.1101/2023.02.16.528879

89. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Tunyasuvunakool K, et al. AlphaFold 2. Fourteenth Critical Assessment of Techniques for Protein Structure Prediction; DeepMind: London, UK. 2020. Available:

https://predictioncenter.org/casp14/doc/presentations/2020_12_01_TS_predictor_AlphaFold2.pdf

90. Emenecker RJ, Griffith D, Holehouse AS. Metapredict: a fast, accurate, and easy-to-use predictor of consensus disorder and structure. Biophys J. 2021;120: 4312–4319.