View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by University of Louisville

University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Faculty Scholarship

8-24-2018

Role of Intrinsic Disorder in Animal Desiccation Tolerance

Brett Janis
University of Louisville

Clinton Belott
University of Louisville

Michael Menze
University of Louisville, michael.menze@louisville.edu

Follow this and additional works at: https://ir.library.louisville.edu/faculty

Part of the Animal Experimentation and Research Commons, and the Biology Commons

ThinkIR Citation

Janis, Brett; Belott, Clinton; and Menze, Michael, "Role of Intrinsic Disorder in Animal Desiccation Tolerance" (2018). *Faculty Scholarship*. 367.
https://ir.library.louisville.edu/faculty/367

This Article is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. For more information, please contact thinkir@louisville.edu.

Role of Intrinsic Disorder in Animal Desiccation Tolerance

Brett Janis[1#], Clinton Belott[1#], and Michael A. Menze[1*]

[1]Department of Biology, University of Louisville, Louisville, KY, 40292 USA

Emails BJ: Brett.Janis@louisville.edu
CB: Clinton.Belott@louisville.edu
MAM: Michael.Menze@louisville.edu

[#]Both authors contributed equally to this work.
[*]Corresponding Author:
Phone: 502-852-8962; Fax 502-852-0725; email: michael.menze@louisville.edu

Abbrevations

ARID, Anhydrobiosis Related Intrinsically Disordered Protein
ARIds, Anhydrobiosis Related Gene Island Clusters
CIDP, C-flavored Intrinsically Disordered Protein
FCR, Frequency of Charged Residue
IDP, Intrinsically Disordered Protein
LEA, Late Embryogenesis Abundant
MoRFs, Molecular Recognition Features
SIDP, S-flavored Intrinsically Disordered Protein
TDP, Tardigrade Intrinsically Disordered Protein
VIDP, V-flavored Intrinsically Disordered Protein

Word count: 8576

Key words: Lyopreservation, Membranless organelles, Water stress, Cryptobiosis, Extremophiles

Abstract

This review compares the molecular strategies employed by anhydrobiotic invertebrates to survive extreme water stress. Intrinsically disordered proteins (IDPs) play a central role in desiccation tolerance in all species investigated. Various hypotheses about the functions of anhydrobiosis-related intrinsically disordered (ARID) proteins, including late embryogenesis abundant (LEA) and tardigrade-specific intrinsically disordered proteins, were evaluated by broad sequence characterization. A surprisingly wide range in sequence characteristics including hydropathy and the frequency and distribution of charges was discovered. Interestingly, two clusters of similar proteins were found that potentially correlate with distinct functions. This may indicate two broad groups of ARID proteins, composed of one group that folds into functional conformations during desiccation and a second group that potentially displays functions in the hydrated state. A broad range of physiochemical properties suggest that folding may be induced by factors such as hydration level, molecular crowding, and interactions with binding partners. This plasticity may be required to fine tune the ARID-proteome response at different hydration levels during desiccation. Furthermore, the sequence properties of some LEA proteins share qualities with IDPs known to undergo liquid-liquid phase separations during environmental challenges.

1. Introduction

Unfavorable conditions for life can arise from changes in abiotic factors, including the availability of water. Some animals have evolved remarkable strategies to withstand virtually complete water loss for prolonged periods of time, despite the cellular damage associated with desiccation, such as membrane destabilization, protein and nucleic acid denaturation, oxidative stress, and metabolic dysregulation [1-5]. This transient state of life has been an enigma since 1702 when Van Leeuwenhoek first noted anhydrobiosis in rotifers or 'wheel animals' [6, 7]. Since then, desiccation tolerance has been confirmed to occur in several other animal phyla including Arthropoda, Tardigrada, and Nematoda [8, 9]. Remarkably, many desiccation-tolerant species can survive in an anhydrobiotic state for years, or even decades, with limited impacts on viability [10]. Understanding the constrains that govern desiccation tolerance has obvious biotechnological applications, particularly in crop-drought resistance and stabilization of clinically relevant cells and tissues at ambient temperatures [11]. To translate insights from anhydrobiotic animals into clinical applications, it is imperative to compare and contrast the molecular principles among these organisms to distinguish between fundamental and unique strategies. The goal of this review is to briefly introduce currently known anhydrobiotic strategies in animals and then to focus on the properties of intrinsically disordered proteins (IDPs) found in animals in preparation or response to desiccation stress.

We will use the term 'anhydrobiosis-related intrinsically disordered (ARID) proteins' as an umbrella term for late embryogenesis abundant (LEA) proteins, tardigrade-specific intrinsically disordered proteins (TDPs), and other intrinsically disordered proteins with confirmed or strongly suggested roles in animal desiccation tolerance. However, this grouping should not imply that all ARID proteins are functionally similar but distinguishes them from more ordered proteins, such as enzymes involved in ROS detoxification, small HSPs, and aquaporins. The term ARID is similar

to the term ‘anhydrobiosis-related gene island clusters (ARIds)’ presented by Gusev and colleagues to describe the grouping of anhydrobiosis-related genes in *P. vanderplanki*, where all genes in a given cluster were upregulated during desiccation [12]. Indeed, it would be highly instructive to know whether similar patterns of ARId regulation are observed in other anhydrobiotic animals. We anticipate that proteins employed in anhydrobiotic organisms from other kingdom such as plants, eubacteria, archaebacteria, and fungi display physicochemical similarities to ARID proteins found in animals. However, some divergence is to be expected and warrants a closer investigation of IDPs from theses kingdoms in the future. Furthermore, several excellent broader reviews covering anhydrobiosis are available [1, 13-16]. For a recent comparative review of desiccation tolerance between plants and animals, please see Leprince and Buitnik [1].

2. Molecular strategies in anhydrobiotic animals

The minimum molecular requirements to enter and exit anhydrobiosis remain undefined. In general, animals that rely only on the expression of protective proteins can be distinguished from animals that combine proteins with non-proteinaceous compounds. The expression of specific IDPs to survive desiccation and rehydration seems to be a common strategy in all known anhydrobiotic invertebrates (Fig. 1). Furthermore, a variety of stress-response proteins might be upregulated during or before the onset of desiccation [5].

2.1 Antioxidant enzymes, heat shock proteins (HSPs) and aquaporins

Upregulation of antioxidant systems to scavenge ROS is a common strategy among anhydrobiotes in response or preparation for desiccation [17-20]. ROS production in a dehydrating cell, due to enzyme dysfunction and unregulated metabolic activities, leads to peroxidation of membrane

lipids, carbonylation of proteins, and damage to nucleic acids [21]. Intuitively, the predominant source of ROS production during desiccation is the mitochondrion and detoxification at the organelle level might be necessary to prevent excessive cellular damage. Indirect evidence from *Polypedilum vanderplanki* supports this hypothesis, given that the most strongly upregulated gene in response to desiccation was mitochondrial thioredoxin [22]. However, it is important to note that several anhydrobiotic animals enter a state of metabolic arrest (e.g. diapause or quiescence) in preparation or response to desiccation stress and thus limit excessive ROS production from dysregulated metabolic activities [3, 23].

Small HSPs, characterized by their α-crystallin domain, are also associated with anhydrobiosis and have been found to be upregulated in several animals during water stress [22, 24-27]. However, only minor changes in the expression levels of small HSPs and their encoding mRNAs was observed in tardigrades during desiccation compared to fully hydrated controls [28, 29]. Most small HSPs, unlike their larger counterparts, do not require ATP to prevent protein aggregation or aid in refolding [30]. Since the adenylate energy charge of the cell will likely be dramatically lowered during desiccation, small HSPs might offer a substantial advantage over ATP-dependent HSPs considering the limited metabolic ATP regeneration.

Aquaporins have been suggested to play some role in animal anhydrobiosis, particularly in *P. vanderplanki* and tardigrades [31-33]. However, no current evidence suggests a role in species like *Artemia franciscana* and nematodes, and relatively low expression levels were found in resting eggs of rotifers [34]. An attractive hypothesis on the role of aquaporins includes the regulation of water loss to fine tune desiccation kinetics [14]. While aquaporins may offer applications for engineering water-loss kinetics for cell and tissue preservation, more data will be required before any firm conclusions about their role in anhydrobiotic animal can be made.

2.2 Trehalose and polyamines

The accumulation, or absence, of trehalose, a non-reducing disaccharide, in anhydrobiotic animals has been well studied [35-37]. In brief, trehalose is hypothesized to confer protection during desiccation by three distinct mechanisms: 1) by replacing water with its hydroxyl groups, 2) by undergoing vitrification at lower water contents to prevent molecular movement, and 3) by stabilizing globular proteins in solution, where the unfolded protein displays a greater preferential hydration than the folded state but water interacts more preferable with trehalose than the dissolved protein thereby reducing the thermodynamic driving force for unfolding [38, 39]. It is important to point out that these hypotheses are not mutually exclusive and each mechanism may contribute to the protective properties of trehalose especially at different hydration levels and temperatures that the animal may encounter during desiccation. Early works on anhydrobiotic animals, including *Artemia salina* [40] and *Aphelenchus avenae* [41], suggested that trehalose may be required for anhydrobiosis. However, later evidence demonstrated that this is not the case since some species of rotifers and tardigrades do not accumulate trehalose prior to desiccation [35, 42, 43].

Polyamines are another group of compounds that was more recently associated with protection during desiccation. In the anhydrobiotic dauer larvae of *Caenorhabditis elegans*, mutants unable to synthesize polyamines (particularly spermidine) were rendered extremely sensitive to desiccation [13], but the role of these compounds in other anhydrobiotic animals has not been confirmed. Furthermore, it is unclear if spermidine production by spermidine synthase (SPDS-1) has direct protective properties during desiccation or if it affects other processes that modulate desiccation sensitivity. However, it was found for the brine shrimp *A. franciscana* that polyamine concentrations further increased as the organism developed past its anhydrobiotic stage

of life [44]. This may imply a limited role of polyamines for this organism, or that the concentrations present prior to desiccation are sufficient for protection.

2.3 Late embryogenesis abundant (LEA) and tardigrade intrinsically disordered proteins (TDPs)

The proteins most readily associated with desiccation tolerance are LEA proteins and more recently TDPs [45-48]. LEA proteins were discovered by Dure et al. in the late embryogenic stage of cotton (*Gossypium hirsutum*) seeds and have more recently be linked to anhydrobiosis in animals [49, 50]. Plant LEA proteins were initially grouped based on the presence of specific sequence motifs [51]. Since then, several nomenclatures have been proposed for LEA proteins [52, 53]. For the purposes of this review, we will be following the classification scheme by Tunnacliffe and Wise [15].

Group 1 LEA proteins contain one or more repeats of a hydrophilic 20 amino acid motif, while group 2 LEA proteins, termed 'dehydrins', contain two or more specific motifs denoted as Y, S, and K. Group 3 contains the largest number of LEA proteins and are characterized by a specific 11 amino acid motif [15, 54]. While most LEA proteins in plants seem to fall into groups 1-3, other minor groups have been described. Group 4 LEA proteins lack any consensus sequence, and group 5 LEA proteins are characterized by an unusually high content of hydrophobic residues [55]. Finally, group 6 LEA proteins are characterized by the presence of at least one seed maturation protein motif and have recently been associated with the long-term stability of seeds in the desiccated state [56]. Interestingly, only group 3 LEA proteins have been identified in anhydrobiotic

animals with the exception of *Artemia*, which expresses LEA proteins from groups 1, 3, and 6 in their desiccation tolerant embryos.

The reason(s) why different groups of anhydrobiotic animals rely on different types of ARID proteins are unresolved. One hypothesis may be that the lack of trehalose or presence of other protective compounds encourages different adaptive trajectories for proteins involved in desiccation tolerance. Data presented on three *Triops* species demonstrated that the cysts undergo vitrification in absence of trehalose [57]. These data are strikingly similar to those presented for tardigrades [46]. Blasting LEA protein sequence data from *Artemia* against EST libraries for *Triops* yielded no significant results, while searches against EST libraries derived from tardigrades yielded low-identity hits (*data not shown*). This may support the hypothesis that the absence of trehalose requires proteins with a different set of physicochemical properties compared to animals that accumulate substantial levels of this sugar (e.g. <0.5% dry weight in *Triops longicaudatus, Triops cancriformis, and Triops australiensis vs.* 13-18% dry weight in *A. franciscana* and *P. vanderplanki* [57]).

3. Disorder: regulatory element or essential property for hydration-level specific function(s)?

The intrinsic disorder of LEA proteins has been hypothesized to serve a role in several functions such as 1) molecular shields that block protein aggregation [58], 2) regulating desiccation rates as hydration buffers [59], 3) binding divalent metal ions [60, 61], and 4) reinforcing trehalose glasses [62]. Recently discovered TDPs from tardigrades have been demonstrated to form protective hydrogels and glasses that reinforce structural integrity of the animal during desiccation [63]. An additional hypothesis for LEA proteins states that they undergo *in vivo* conformational transitions from randomly-coiled hydrated chains to semi-folded, activated proteins at lower hydration levels [64].

Furthermore, LEA proteins in plants have been found to confer membrane protection and freeze tolerance even in the hydrated, disordered state [65]. These properties can be mechanistically explained by binding-partner induced conformational transitions *via* specific molecular recognition features (MoRFs) [66]. Protein-protein interactions may be modulated by conformational transitions as water is reduced during freezing or drying [67]. Multiple intrinsically disordered regions found on a single protein may fold upon recognition of distinct binding partners, thus allowing one-to-many targeting or protein moonlighting [68]. Some of these regions may only fold under specific crowding and hydration conditions such that one protein may have several functions and targets during desiccation. Nevertheless, the specific role of disorder in dehydration-related proteins warrants further investigations and unresolved questions remain: 1) Is protein disorder an intrinsic mode of function or a regulatory element of functional properties? 2) Is the impact of the hydration state on protein function(s) similar among proteins? Although many of the above stated hypotheses were initially developed on LEA proteins, their relevance to other IDPs will also be discussed.

3.1 Molecular shielding

The molecular shielding hypothesis predicts that protection is conferred by entropic chains that act as steric and/or electrostatic barriers against protein aggregation during water stress [58]. The protective IDP has low target specificity which is limited to fuzzy and small MoRF-induced structural interactions with specific surfaces on target proteins. The target protein may partly unfold upon binding to the IDP but refolds by entropic energy transfer upon release from the protein [69]. Molecular shielding-like anti-aggregation has been observed in LEA proteins expressed by the sleeping chironomid *P. vanderplanki* [70]. Most of the IDP's structure exists as a

highly plastic ensemble of conformations that encompasses a large hydrodynamic radius. This hypothesis is consistent with the finding that IDPs, compared to globular proteins, in general have a greatly enlarged hydrodynamic radius [71], but it does not fully address the functional relevance of the observed increase in secondary structure of LEA proteins during desiccation.

One challenge to this hypothesis might be that a high ratio of protective proteins to targets is required for shielding. To protect each target would likely require several molecular shields to insulate them from multiple angles of interaction. Target selection might be complicated by increasing crowding during desiccation and recruitment of additional molecular shields might be hindered by already interacting IDPs. Furthermore, LEA proteins can gain as much as an additional 40% of defined secondary structure and reduce their hydrodynamic radii substantially as cellular water depletes [67]. The reduced hydrodynamic radii in turn would likely reduce the molecular shielding efficiency due to the decrease in area of steric and electrostatic repulsion. Therefore, the observed gain in secondary structure of some LEA proteins during desiccation would be counterintuitive for this model. Molecular shielding does not offer a mechanistic explanation for the membrane stabilizing effect observed for several LEA proteins. Consequently, molecular shielding by LEA proteins may only be one of several possible functions.

3.2 Hydration buffers

The intrinsic disorder of anhydrobiosis-related proteins gives rise to unique solution properties when compared to globular proteins, such as large hydration shells and conformational plasticity. These properties have been hypothesized to play a role in desiccation tolerance by modulating the solution properties and desiccation rates of anhydrobiotic organisms [59]. The large protein hydration shell may act as a water reservoir that is released upon assuming secondary

structures during desiccation. Solid state NMR studies have demonstrated enlarged hydration shells of IDPs, and also suggest that those of dehydrins from plants are particularly extensive [72]. Although loss of water upon disorder-to-order transitions is likely to occur, a functional role of this released water would likely be small. While there may be some impact on the overall drying kinetics of the animal due to water release, the relatively small fraction of water that could be released, even assuming relatively high levels of LEA expression, is likely insignificant in the context of variable drying conditions found in nature. Furthermore, the lack of free water, once desiccation is complete, is the major source of cellular protection because unregulated chemical reactions are inhibited in the vitrified state. A similar principle is noted in cryobiology, wherein the relatively mobile water found at temperatures above the glass-transition temperature of water (~136 °K) contributes to degradation in frozen samples [73].

3.2 Ion sequestration

During water removal, dissolved molecules and ions will concentrate and, eventually, precipitate from the solution depending on their physicochemical properties. Precipitation, if not carefully regulated, often produces damaging aggregation events that permanently inactivate proteins and other biomolecules. The ion-sequestration hypothesis states that, during water loss, ions are buffered by binding to LEA proteins [60, 61]. Considering the large amounts of inorganic ions relative to protein in the cell, this effect is likely limited to metal cations that are found in low concentrations. However, highly charged LEA proteins might serve as nucleation sites for salt precipitation during desiccation. Promoting precipitation of salts would only be possible for LEA proteins with relatively high hydrophilicity and frequency of charged residues (FCR).

3.4 Reinforcing sugar glasses and protein vitrification

During the past decade, group 3 LEA proteins have been associated with the vitrification of the intracellular space by reinforcing sugar-based glasses [62] and, more recently, proteins and LEA model peptides have been demonstrated to form glasses themselves [74]. Protein and sugar glasses are a non-crystalline, physical state that is characterized by a high viscosity above 10^8 Pa-s which greatly impedes molecular movement and prevents chemical reactions [75]. Sugar vitrification is a well-established mechanism in anhydrobiosis for both plants and animals, which is generally associated with membrane and protein stabilization [36]. Trehalose vitrifies at low water contents, but the capacity for protection is impacted by the glass transition temperature (Tg), which dictates the temperature and degree of hydration where the sugar will form or maintain a glassy state [73, 76]. Although the mechanism by which IDPs can reinforce and/or stabilize sugar glasses is not well understood, some LEA proteins and peptides have been shown to increase the Tg of sugar glasses [62, 74]. In the case of tardigrades, vitrification occurs rapidly during desiccation provided that the rate of drying is sufficiently slow for the organism to accumulate sufficient amounts of TDPs. Protein-based vitrification was further demonstrated by ectopic expression of TDPs in both yeast and bacteria [63].

3.4 Disorder-to-order transitions and membrane stabilization

LEA proteins have been shown to undergo conformational changes in response to various solutes and crowding effects [77] and two group 3 LEA proteins from *A. franciscana* increased their amount of α-helical structures *in vitro* when exposed to sodium dodecyl sulfate, tetrafluoroethylene, or when desiccated [67]. Proteins from *P. vanderplanki* and *A. franciscana* undergo conformational transitions when dried and protect enzymatic activity of lactate dehydrogenase during desiccation.

Additionally, both LEA proteins were shown to prevent aggregation of casein better than bovine serum albumin [67, 70]. Furthermore, two LEA proteins from *Arabidopsis thaliana* were recently shown to undergo conformational transitions when crowded by glycerol and when localized near a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylcholine liposome [78]. The conformational state was maintained in the desiccated state, and the protein inserted into the phospholipid bilayer as measured using FTIR. These findings support the hypothesis of direct membrane interactions and insertion of some LEA proteins, although the exact thermodynamics are not fully understood. This evidence also suggests desiccation induces conformational transitions, rather than MoRF-induced membrane interactions alone, because the protein needed to be folded by glycerol crowding before associating with the liposome membrane [78].

Membrane interactions of some folded LEA proteins are hypothesized to occur through bundled, amphipathic α-helices. Two hydrated group 3 LEA proteins from *A. franciscana* contain regions of high α-helical propensity that, if folded, may form amphipathic helices with stripes of positive and negative residues separating the hydrophobic face from the hydrophilic portion of the protein [79]. These proteins protect the membranes of liposomes during desiccation and rehydration, potentially via interactions between the phospholipids and the hydrophobic face of the amphipathic α-helices. Stripes of positive and negative charge distributed may not be a prerequisite for all LEA proteins to stabilize phospholipid bilayers and some LEA proteins may actually insert into the membrane during desiccation [78].

3.5. Phase separation and the formation of desiccation induced 'membraneless' organelles

With the discovery of the glass and gel propensities of TDPs, the question remains if or which ARID proteins may form fibrils, glasses, or gels during desiccation. To date, no stable super-

molecular LEA assembles have been observed. However, the question if LEA proteins form higher order oligomers has been discussed in the literature, and some LEA proteins form multimers in solution [80, 81]. A more recently described physicochemical property of some IDPs is the separation of the protein from the solvent as another liquid phase [82-84]. Protein liquid-liquid phase separations, comprised of loose associations of LEA proteins, could serve different functions including molecular shielding and hydration buffering. Protein liquid-phase separations have become a rapidly expanding topic in biology that has explained the behavior of the disordered tail of DEAD-box helicase 4 (Ddx4) and several other previously uncharacterized IDPs [85-87]. These liquid protein droplets, termed "membraneless organelles," are found under osmotic and oxidative stress in eukaryotes and are predicted for some LEA proteins in *A. francsicana* [88-90].

The observed desiccation-induced folding of LEA proteins suggests that intramolecular interactions occur as surface water is depleted. Protein interactions between similar LEA proteins or partner molecules may result in a protein droplet with specific physicochemical properties as governed by the amino acid sequence of the nucleating protein. Similar to other proteins that undergo liquid-liquid phase separations, LEA proteins are highly repetitive, have low complexity, and tend to have high overall charges and charge separation [53, 91]. Furthermore, the amyotrophic lateral sclerosis (ALS) related fused in sarcoma (FUS) protein still forms liquid protein droplets when its multivalent interaction sites are substituted with LEA motifs [92]. Unlike the molecular shielding hypothesis, where the induced folding of several LEA proteins in response to interactions with each target protein is necessary for protection, membraneless organelles can nucleate off a core material and incorporate a variety of targets [93]. The interior of the droplet should then function like a molecular shield, but each added target further expands the droplet radius and increases the odds of collision and fusion with other LEA proteins and/or target molecules. If these

droplets form early during drying, or in preparation for drying, then they may require a lower protein content for protection compared to molecular shields. Although not all ARID proteins should be expected to form anhydrobiosis associated membraneless organelles, the sequence characteristics of some LEA proteins and TDPs suggest that they are candidates for this super-molecular structure (e.g. MH351624).

4. Protein sequence analysis

4.1 Datasets

Amino acid sequences for animal LEA proteins were retrieved using "LEA" or "Late Embryogenesis Abundant Protein" as search terms in the protein and nucleotide sequence databases at NCBI. The retrieved sequences were individually cross-referenced using BLAST-P and rejected if they did not share sequence similarity within at least $E < 1*10^{-3}$ with a confirmed LEA protein (Supplemental File 1). 101 LEA protein sequences failed to be rejected and comprise the LEA protein dataset for the following analysis (Tab. 1). Published tardigrade protein sequences were used [46] and no additional proteins were found using BLAST algorithms limiting the total dataset to only 14 sequences (Tab. 2). The globular protein dataset was retrieved from RCSB PDB selecting for proteins between 10 and 80 kDa with structures verified by X-ray crystallography to avoid intrinsically disordered regions (Supplemental File 2). All sequences were analyzed in bulk using localCIDER, a freely accessible program designed by Holehouse et al. [94], and flavor predictions were made using values generated from the VL2 predictor with a window length of 21 amino acids [95]. SigmaPlot 12.5 (Systat Software Inc., San Jose, CA) was used to generate graphs.

4.2. Mean net charge and hydropathy analysis

The ratio of the mean net hydropathy and the absolute value of the mean net charge indicates the likelihood for disorder based on an arbitrary boundary established by statistical analysis of known proteins [96]. As shown in Fig. 2, the distribution of LEA proteins overlaps with the diverse range of known IDPs originally used to generate the disorder-order boundary and both cluster separately from known globular proteins as expected (Fig. 2A, B). Although TDPs have been confirmed as IDPs, they distribute evenly between the areas of the plot where most globular proteins or IDPs are located (Fig. 2A, C). Some LEA proteins are also rather hydrophobic and overlap with the distribution pattern of globular proteins (Fig. 2A, D). In case desiccation induces conformational transitions to 'activate' proteins, then more hydrophobic proteins may begin folding earlier during water stress than the more hydrophilic ones. A large range of hydropathies could indicate a temporal regulation of the LEA-proteome response during drying. Alternatively, more hydrophobic ARID proteins may separate from solution more readily into gels or liquid droplets than their hydrophilic counterparts.

4.3 Phase diagrams – charge ratios and distributions

The reduced range of amino acid expressed in IDPs results in a lower overall complexity than observed in globular proteins [97]. This low complexity increases the impact of overrepresented amino acids in a protein sequence. IDP properties are often governed by the frequency of charged residues (FCR), the frequency of order-promoting residues (e.g. alanine and phenylalanine), and the spatial distributions of both [98]. The frequency of positive and negative charges, for example, offers insights into a protein's capacity to maintain intramolecular interactions and has been used to define three compositional categories: polar tracts,

polyelectrolytes, and polyampholytes [98]. These categories were further expanded into five; weak polyampholytes and polyelectrolytes (R1), Janus sequences (R2), strong polyampholytes (R3), negatively charged polyampholytes (R4), and positively charges polyampholytes (R5) [99]. R1 proteins tends to be globular or IDPs with a globular domain. R2 proteins tend to be more plastic, with conformations that depend on salt concentration, ligand binding, or other factors. R3 proteins are highly charged but have a low mean net charge due to a balance of charges. R3 proteins with regularly distributed charges have strong self-repulsion and exhibit more coiled structures, whereas proteins with more localized charges increase their likelihood to form intramolecular interactions such as hairpins or chimeras. R4 and R5 proteins form expansive coils due to polyanionic or polycationic repulsion [98]. Employing this analysis predicts that LEA proteins have a wide range of potential behaviors, which is represented by proteins distributing across regions R1, R2, and R3 of the phase diagram (Fig. 3A). However, LEA proteins cluster into two major groups across R2 and R3 with few outliers in R1 (Fig. 3B). Conversely, all known TDPs are found in region R2, and the group of globular protein used for comparison is mainly represented in regions R1 and R2 (Fig. 3C, D). Although the tight clustering of TDPs may be due to some trend in this protein family, the small sample size likely accounts for some lack in variance (Tab.1).

The separation of LEA proteins into two clusters at the border of R1/R2 and R2/R3 may indicate differences in function. Proteins in R1/R2 are likely more environmentally regulated than those in R2/R3 but may undergo a limited amount of charge-mediated folding due to their relatively low FCR. This suggests that hydropathy and MoRF regions should have a higher impact than charge distribution on induced-fit or desiccation-induced conformational transition for LEA proteins falling at the R1/R2 border. LEA proteins in the R2/R3 cluster have a higher FCR and should be more influenced by their charge distribution, pH fluctuations, and ion concentrations.

For proteins plotting in R2, which predicts environmentally modulated structures, conformational transitions during desiccation, such as induced folding at phospholipid bilayers or in response to ligand binding, are supported. The localization of LEA proteins in R1 and R3, however, also suggest that some of these proteins may not need to gain much additional structure during desiccation and, therefore, function at high water contents before environmentally-induced folding.

4.4 Flavor categorization based on amino acid composition

Like compositional categories, flavors of disorder are an established criterion that, while mainly used in disorder algorithms, correlates broadly to predicted functions and partner-binding behaviors. The flavors described by Vucetic et al include “V”, “C”, and “S”-flavored proteins (VIDPs, CIDPs, and SIDPs, respectively) [95]. VIDPs are enriched in structure-promoting and hydrophobic amino acids residues, most notable cysteine, phenylalanine, isoleucine, and tyrosine, relative to CIDPs and SIDPs. VIDPs are in generally associated with ribosomal proteins. CIDPs are enriched in alanine, histidine, and methionine and are generally associated with DNA and RNA binding. Finally, SIDPs are relatively depleted in histidine and are generally associated with protein-binding. The proteins examined showed a clear bias towards SIDPs (Fig. 4A) and each of the four TDPs that were categorized as CIDPs are secreted and not cytoplasmic localized proteins. LEA proteins were mostly classified as SIDPs but some VIDPs and CIDPs were also discovered. Interestingly, no tardigrade VIDPs were identified and all rotifer LEA proteins were CIDPs. Given the propensity of TDPs to form protein-glasses, the SIDP flavor that suggests protein-protein interactions is not surprising. However, flavor categories cannot predict the potential membrane interactions associated with some LEA proteins.

A more in-depth analysis using composition profiling reveals that LEA proteins and TDPs both are each enriched in alanine, acidic residues, and lysine (Fig. 4 B, C). The hydrophilic nature of ARID proteins probably allows them to maintain disorder in the hydrated state despite their remarkable depletion in the major structure-breaking amino acids, glycine and proline. A possible explanation for the enrichment in alanine observed for LEA proteins and, to a lesser extent, TDPs might be the propensity to form α-helices in the desiccated state. Furthermore, relative to ordered proteins TDPs are enriched in all positively charge amino acids whereas LEA proteins are enriched in lysine but depleted in arginine (Fig. 4 D). This bias towards lysine over arginine may allow for a greater variety of post-translational modifications and lower surface-charge interactions that should increases the relative solubility of LEA proteins [100].

4.5 Differences in protein properties among animal genera

Proteins were separated into 5 groups to investigate genera-specific properties. Not surprisingly, some genera-specific clustering was observed in the charge-hydropathy plot and phase diagram (Fig. 5A, B). Rotifers express less hydrophilic LEA proteins, which may interact more readily with membranes than other animal LEA proteins. LEA proteins in animals that utilize trehalose are mainly localized in the disordered region of the plot, although *Artemia* and *Polypedilum* each express at least one rather hydrophobic protein (Tab. 1). TDPs have a lower absolute mean net charge and higher mean net hydropathy on average. LEA proteins are generally longer than TDPs, perhaps allowing for larger hydrodynamic radii even after folding (Fig. 5C). LEA proteins found in *Caenorhabditis* were on average substantially longer than observed in the other groups. However, a larger variation in the frequency of charged residues was found for *Artemia* and *Polypedilum* compared to the other groups (Fig. 5D).

An overabundance of like-charges promotes self-repulsion, whereas distinct regions of different charges may facilitate intermolecular and intramolecular electrostatic interactions. This distribution of charges can be represented by a scale from 0 to 1, denoted by the variable κ, where 0 represents a completely uniform charge distribution and 1 describes complete separation of charges [98]. LEA proteins with larger size such as observed in *Caenorhabditis* and *Polypedilum* should have increased odds of intramolecular interactions forming tertiary structures compared to shorter proteins. Both, the FCR and protein length will modify the impact of κ on the probability of electrostatic interactions. For example, longer proteins will in general have a higher probability to form favorable intramolecular interactions which in turn reduces the impact of κ. However, as the FCR increases, the impact of kappa also increases. Overall, larger proteins appear to exhibit larger FCRs and smaller kappa values which may increases electrostatic repulsion within the polypeptide chain (Fig. 4E). The small sample sizes of protein sequences available for *Adineta* and *Hypsibius/Paramacrabiotus* (Tab. 1) may contribute to the smaller range in metrics compared to proteins from the other genera. However, based on our analysis, rotifer LEA proteins and TDPs in tardigrades are both characterized by high FCRs and moderate to high kappa values. Interestingly, anhydrobiosis does not depend on trehalose in the species from either genus [46, 101].

5. Conclusions

Returning to the initial question of intrinsic disorder as a regulatory element or functional property of ARID proteins, we conclude that it may be either, depending on the specific LEA protein or TDP in question. For any given protein, the degree of desiccation-mediated regulation could be governed by hydropathy and the quantity and distribution of charged residues. Furthermore, our analyses demonstrate that the impact of the hydration state on protein function may vary

substantially among proteins. Anhydrobiotic animals may require a physiochemically diverse LEA or TDP proteome to elicit a temporal progression of responses during desiccation. A temporal distribution of responses would better accommodate variable rates of desiccation and cellular structures may require targeted protection from different sources of deterioration depending on the hydration state of the animal.

6. Acknowledgments

This work was supported by the National Science Foundation [grant number IOS-1659970/IOS-1457061].

7. Conflict of Interest

The authors declare that they have no conflict of interest.

Figure legends

Figure 1: Common and phyla-specific molecular strategies found in anhydrobiotic animals in desiccation tolerance. IDPs such as LEA proteins and TDPs are found in all phyla. The role of small HSPs in tardigrades is currently unresolved (*small HSPs may contain extended intrinsically disordered regions with functional relevance). Ferritin homologues, hemoglobin, and polyubiquitin may play a role in animal desiccation tolerance, but current data suggests that these are phyla-specific strategies. Non-proteinaceous strategies include trehalose and polyamines, such as spermidine produced by spermidine synthase (SPDS-1) in *C. elegans*. Please refer to text for more information.

Figure 2: Mean net charge versus mean net hydropathy plots. (A) Intrinsically disordered proteins (white circles) and globular proteins (black circles) separate, which allows for deducing an arbitrary border (black line) for order-disorder prediction [96]. (B) LEA proteins share distribution patterns with other IDPs. (C) TDPs distribute similarly to either side of the arbitrary border and (D) the comparison group of globular proteins mirrors the previously-established boundary.

Figure 3: Phase diagrams of anhydrobiosis-related IDPs and globular proteins. (A) Summary of phase-diagram regions for LEA proteins (black), TDPs (light gray), and globular proteins (dark gray). Individual phase diagrams for LEA proteins (B), TDPs (C), and globular proteins (D) are shown.

Figure 4: VL2 flavor and Sequence Compiler Analysis. (A) LEA proteins (black) and TDPs (gray) show different proportions in flavor of disorder. Frequencies of amino acid that compose LEA proteins relative to ordered proteins (B), TDPs relative to ordered proteins (C), and LEA proteins relative to TDPs (D). Bootstrap significance values are shown (<0.05 (*), <0.01 (**), <0.001 (***)).

Figure 5: Analysis of anhydrobiosis-related proteins by selected genera. (A) Mean net charge versus mean net hydropathy plot and phase diagram (B) for IDPs from *Adineta* (red), *Artemia* (dark green), *Caenorhabditis* (blue), *Polypedilum* (cyan), and *Hypsibius/Paramacrabiotus* (magenta). Globular comparison group in gray. Protein length (C), FCR (D), separation of charges (E), and isoelectric point, for proteins from *Adineta* (Adi.), *Artemia* (Art.), *Caenorhabditis* (Cae.), *Polypedilum* (Pol.), and *Hypsibius/Paramacrabiotus* (Tar.).

Table 1: Selected properties of LEA-related proteins in animal species. Protein sequences were deduced from full-length nucleotide sequences* as indexed by the US National Center for Biotechnology Information (NCBI).

Organism	Accession Nr.	Length (aa)	Frac. Dis. Prom.	Frac. Char. Res.	GRAVY	References
Adineta ricciae	ABU62808	376	0.75	0.35	-0.457	[102]
Adineta ricciae	ABU62809	421	0.74	0.35	-0.460	[102]
Adineta ricciae	ABU62810	376	0.75	0.35	-0.465	[102]
Adineta ricciae	ABU62811	420	0.74	0.35	-0.461	[102]
Adineta vaga	ADD91471	354	0.75	0.35	-0.627	
Adineta vaga	ADD91479	354	0.76	0.35	-0.626	
Ancylostoma ceylanicum	EPB73657	387	0.80	0.28	-0.466	
Ancylostoma duodenale	KIH53544	314	0.83	0.37	-0.887	
Ancylostoma duodenale	KIH57747	359	0.78	0.26	-0.350	
Aphelenchus avenae	Q95V77	143	0.85	0.39	-1.585	[103]
Aphelenchus avenae	AAL18843	143	0.85	0.39	-1.585	[103]
Aphelenchus avenae	ABQ23232	102	0.82	0.44	-1.376	[104]
Aphelenchus avenae	ABQ23233	85	0.84	0.52	-1.832	[104]
Artemia franciscana	ABR67402	182	0.83	0.27	1.365	
Artemia franciscana	ABX89317	182	0.84	0.27	-1.410	[105]
Artemia franciscana	ACM16586	307	0.79	0.37	-1.295	[106]
Artemia franciscana	ACX81197	97	0.74	0.27	-1.158	
Artemia franciscana	ACX81198	217	0.79	0.27	-1.257	
Artemia franciscana	ADE45145	142	0.84	0.27	-1.418	
Artemia franciscana	ADE45146	122	0.83	0.27	-1.312	
Artemia franciscana	ADE45147	62	0.82	0.27	-1.234	
Artemia franciscana	MH351624	257	0.70	0.21	-0.418	
Artemia franciscana	ACA47267	357	0.81	0.27	-1.027	[107]
Artemia franciscana	ACA47268	364	0.80	0.27	-0.884	[107]
Artemia parthenogenetica	AEM72698	85	0.78	0.27	-1.235	
Artemia parthenogenetica	AEM72699	182	0.83	0.26	-1.396	

Artemia persimilis	AEM72697	85	0.78	0.30	-1.235	
Artemia sinica	AMQ80946	182	0.83	0.30	-1.412	[108]
Artemia sinica	AOV81545	364	0.79	0.30	-0.885	[108]
Bemisia tabaci	XP_018915417	136	0.75	0.37	-0.839	
Brachionus plicatilis	ADE05593	613	0.84	0.38	-1.248	[109]
Brachionus plicatilis	ADE05594	248	0.85	0.33	-1.219	[109]
Caenorhabditis brenneri	EGT57645	935	0.86	0.39	-1.255	
Caenorhabditis brenneri	EGT57648	789	0.84	0.39	-1.020	
Caenorhabditis brenneri	EGT59057	917	0.85	0.39	-1.244	
Caenorhabditis brenneri	EGT59115	379	0.84	0.37	-0.870	
Caenorhabditis brenneri	EGT59117	724	0.85	0.40	-1.032	
Caenorhabditis briggsae	CAP25432	324	0.82	0.36	-0.833	
Caenorhabditis briggsae	CAP25462	379	0.79	0.38	-1.125	
Caenorhabditis briggsae	CAP25449	925	0.82	0.39	-1.252	
Caenorhabditis elegans	AAB69446	733	0.83	0.38	-1.126	
Caenorhabditis elegans	CCF23420	821	0.81	0.36	-1.104	
Caenorhabditis elegans	CCF23421	1166	0.82	0.36	-1.067	
Caenorhabditis elegans	CCF23422	1214	0.83	0.38	-1.066	
Caenorhabditis elegans	CCF23423	1381	0.82	0.37	-1.065	
Caenorhabditis elegans	CCF23424	1198	0.83	0.36	-1.066	
Caenorhabditis elegans	CCF23425	1349	0.82	0.37	-1.065	
Caenorhabditis elegans	CCF23426	1397	0.82	0.37	-1.066	
Caenorhabditis elegans	NP_001256160	1397	0.82	0.36	-1.065	
Caenorhabditis elegans	NP_001256161	1214	0.83	0.37	-1.066	
Caenorhabditis elegans	NP_001256162	821	0.81	0.37	-1.104	
Caenorhabditis elegans	CAB05543	733	0.83	0.36	-1.126	[110]
Caenorhabditis elegans	NP_001256163	1381	0.82	0.37	-1.065	
Caenorhabditis elegans	NP_001256164	1198	0.83	0.38	-1.066	
Caenorhabditis elegans	NP_001256165	805	0.81	0.40	-1.105	
Caenorhabditis elegans	NP_001256166	1349	0.82	0.37	-1.066	
Caenorhabditis elegans	NP_001256167	1166	0.82	0.36	-1.067	
Caenorhabditis elegans	NP_001256168	773	0.81	0.37	-1.108	

Caenorhabditis elegans	NP_001256169	1309	0.83	0.36	-1.075	
Caenorhabditis elegans	NP_001256170	1126	0.83	0.37	-1.078	
Caenorhabditis elegans	NP_001256171	733	0.83	0.40	-1.126	
Caenorhabditis elegans	NP_001256172	409	0.86	0.37	-1.226	
Caenorhabditis elegans	CAB05548	497	0.84	0.36	-1.054	
Caenorhabditis elegans	NP_001256173	556	0.82	0.37	-0.997	
Caenorhabditis elegans	NP_001256174	497	0.84	0.37	-1.054	
Caenorhabditis elegans	CAI46598	556	0.82	0.36	-0.997	
Caenorhabditis elegans	CBZ01819	1126	0.83	0.37	-1.078	[110]
Caenorhabditis elegans	CCA65580	409	0.86	0.36	-1.226	
Caenorhabditis elegans	CCA65581	1309	0.83	0.37	-1.075	[110]
Caenorhabditis elegans	CCF23418	805	0.81	0.36	-1.105	
Caenorhabditis elegans	CCF23419	773	0.81	0.36	-1.108	
Caenorhabditis remanei	EFO95235	821	0.83	0.38	-1.184	
Caenorhabditis remanei	EFO95236	843	0.83	0.38	-1.189	
Caenorhabditis remanei	EFO95291	1172	0.82	0.42	-1.369	
Caenorhabditis remanei	XP_003116339	821	0.83	0.38	-1.184	
Caenorhabditis remanei	XP_003116340	843	0.83	0.38	-1.189	
Caenorhabditis remanei	XP_003116395	1172	0.82	0.42	-1.369	
Cherax quadricarinatus	ALC79587	169	0.73	0.23	-0.057	
Dictyocaulus viviparus	KJH51853	535	0.71	0.25	-0.595	
Drosophila hydei	XP_023160045	233	0.77	0.26	-0.829	
Limulus polyphemus	XP_013783717	198	0.66	0.32	-0.196	
Oesophagostomum dentatum	KHJ93211	740	0.84	0.37	-0.921	
Oesophagostomum dentatum	KHJ93212	453	0.79	0.30	-0.510	
Polypedilum vanderplanki	BAE92617	180	0.74	0.45	-1.263	[48]
Polypedilum vanderplanki	BAE92618	484	0.58	0.27	-0.340	[48]
Polypedilum vanderplanki	BAN67644	143	0.78	0.44	-1.487	[12]
Polypedilum vanderplanki	BAN67645	709	0.76	0.44	-1.082	[12]
Polypedilum vanderplanki	BAE92616	742	0.67	0.35	-0.643	[48]
Ramazzottius varieornatus	BAQ94586	293	0.83	0.25	-0.942	[111, 112]
Ramazzottius varieornatus	A0A0E4AVP3	293	0.83	0.25	-0.942	[111, 112]

Saccoglossus kowalevskii	XP_006818499	118	0.52	0.26	-0.566	
Steinernema carpocapsae	ABQ23230	87	0.79	0.40	-1.211	
Steinernema carpocapsae	ABQ23231	95	0.76	0.41	-1.194	[104, 110]
Steinernema carpocapsae	ABQ23240	70	0.81	0.43	-1.139	
Teladorsagia circumcincta	PIO62605	594	0.81	0.34	-0.874	
Teladorsagia circumcincta	PIO73643	595	0.74	0.39	-0.905	
Teladorsagia circumcincta	PIO73975	634	0.74	0.41	-0.996	
Teladorsagia circumcincta	PIO74047	1580	0.83	0.34	-0.815	
Toxocara canis	KHN83840	600	0.82	0.27	-0.504	
Trichinella papuae	KRZ64074	66	0.74	0.17	0.170	

*ESTs were excluded from the analysis.

Table 2: Selected properties of tardigrade protein sequences retrieved from the supplemental materials of Boothby et al [46].

Organism	Accession Nr.	Length (aa)	Frac. Dis. Prom.	Frac. Char. Res.	GRAVY
Hypsibius dujardini	P0CU39.1	224	0.73	0.35	-0.314
Hypsibius dujardini	P0CU40.1	224	0.73	0.35	-0.529
Hypsibius dujardini	P0CU41.1	237	0.77	0.34	-0.635
Hypsibius dujardini	P0CU42.1	414	0.74	0.27	-0.41
Hypsibius dujardini	P0CU43.1	227	0.69	0.34	-1.175
Hypsibius dujardini	P0CU44.1	229	0.69	0.34	-1.167
Hypsibius dujardini	P0CU45.1	238	0.79	0.31	-0.984
Hypsibius dujardini	P0CU46.1	227	0.76	0.35	-0.878
Hypsibius dujardini	P0CU47.1	227	0.76	0.35	-0.825
Hypsibius dujardini	P0CU48.1	172	0.67	0.27	-0.985
Hypsibius dujardini	P0CU48.1	168	0.63	0.26	-0.985
Hypsibius dujardini	P0CU49.1	163	0.67	0.26	-1.019
P. richtersi	P0CU52.1	174	0.68	0.28	-1.074
P. richtersi	P0CU51.1	227	0.76	0.35	-0.894

References:

[1] O. Leprince, J. Buitink, *Planta* **2015**, *242*, 369.

[2] S.C. Hand, M.A. Menze, *Planta* **2015**, *242*, 379.

[3] W. Welnicz, M.A. Grohme, L. Kaczmarek, R.O. Schill, M. Frohme, *J. Insect. Physiol.* **2011**, *57*, 577.

[4] F.A. Hoekstra, E.A. Golovina, J. Buitink, *Trends Plant Sci.* **2001**, *6*, 431.

[5] J.H. Crowe, F.A. Hoekstra, L.M. Crowe, *Annu. Rev. Physiol*. **1992**, *54*, 579.

[6] A. Tunnacliffe, J. Lapinski, *Philos. T. Roy. Soc. B* **2003**, *358*, 1755.

[7] A. Van Leeuwenhoek, On certain animalcules found in the sediment in gutters of the roofs of houses. In: Hoole S, translator. The select works of Antony van Leeuwenhoek. London, 1798, 1807, Vol. 2, Letter 144. p 207–213.

[8] S.C. Hand, M.A. Menze, M. Toner, L. Boswell, D. Moore, *Annu. Rev. Physiol.* **2011**, *73*, 115.

[9] P. Alpert, *J. Exp. Biol.* **2006**, *209*, 1575.

[10] R. Guidetti, K.I. Jonsson, *J. Zool.* **2002**, *257*, 181.

[11] J.H. Crowe, *Subcell. Biochem.* **2015**, *71*, 263.

[12] O. Gusev, Y. Suetsugu, R. Cornette, T. Kawashima, M.D. Logacheva, A.S. Kondrashov, A.A. Penin, R. Hatanaka, S. Kikuta, S. Shimura, H. Kanamori, Y. Katayose, T. Matsumoto, E. Shagimardanova, D. Alexeev, V. Govorun, J. Wisecaver, A. Mikheyev, R. Koyanagi, M. Fujie, T. Nishiyama, S. Shigenobu, T.F. Shibata, V. Golygina, M. Hasebe, T. Okuda, N. Satoh, T. Kikawada, *Nat. Commun.* **2014**, *5*, 4784.

[13] C. Erkut, A. Vasilj, S. Boland, B. Habermann, A. Shevchenko, T.V. Kurzchalia, *PLoS One* **2013**, *8*, e82473.

[14] R. Cornette, T. Kikawada, *Iubmb Life* **2011**, *63*, 419.

[15] A. Tunnacliffe, M.J. Wise, *Naturwissenschaften* **2007**, *94*, 791.

[16] T.C. Boothby, G.J. Pielak, *Bioessays* **2017**, *39*.

[17] O. Gusev, Y. Nakahara, V. Vanyagina, L. Malutina, R. Cornette, T. Sakashita, N. Hamada, T. Kikawada, Y. Kobayashi, T. Okuda, *PLoS One* **2010**, *5*, e14008.

[18] A.M. Rizzo, M. Negroni, T. Altiero, G. Montorfano, P. Corsetto, P. Berselli, B. Berra, R. Guidetti, L. Rebecchi, *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2010**, *156*, 115.

[19] A. Vasilj, M. Gentzel, E. Ueberham, R. Gebhardt, A. Shevchenko, *J. Proteome Res.* **2012**, *11*, 3680.

[20] C. Boschetti, N. Pouchkina-Stantcheva, P. Hoffmann, A. Tunnacliffe, *J. Exp. Biol.* **2011**, *214*, 59.

[21] M.B. Franca, A.D. Panek, E.C. Eleutherio, *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **2007**, *146*, 621.

[22] R. Cornette, Y. Kanamori, M. Watanabe, Y. Nakahara, O. Gusev, K. Mitsumasu, K. Kadono-Okuda, M. Shimomura, K. Mita, T. Kikawada, T. Okuda, *J. Biol. Chem.* **2010**, *285*, 35889.

[23] J.E. Podrabsky, S.C. Hand, *J. Exp. Biol.* **2015**, *218*, 1897.

[24] A.M. King, T.H. MacRae, *Annu. Rev. Entomol.* **2015**, *60*, 59.

[25] Y. Sun, T.H. MacRae, *Cell Mol. Life Sci.* **2005**, *62*, 2460.

[26] A.M. King, J. Toxopeus, T.H. MacRae, *FEBS J.* **2013**, *280*, 4761.

[27] P.V. Mazin, E. Shagimardanova, O. Kozlova, A. Cherkasov, R. Sutormin, V.V. Stepanova, A. Stupnikov, M. Logacheva, A. Penin, Y. Sogame, R. Cornette, S. Tokumoto, Y. Miyata, T. Kikawada, M.S. Gelfand, O. Gusev, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E2477.

[28] E. Schokraie, A. Hotz-Wagenblatt, U. Warnken, M. Frohme, T. Dandekar, R.O. Schill, M. Schnolzer, *J. Zool. Syst. Evol. Res.* **2011**, *49*, 111.

[29] A. Reuner, S. Hengherr, B. Mali, F. Forster, D. Arndt, R. Reinhardt, T. Dandekar, M. Frohme, F. Brummer, R.O. Schill, *Cell Stress Chaperon* **2010**, *15*, 423.

[30] F. Narberhaus, *Microbiol. Mol. Biol. Rev.* **2002**, *66*, 64.

[31] T. Kikawada, A. Saito, Y. Kanamori, M. Fujita, K. Snigorska, M. Watanabe, T. Okuda, *Biochim. Biophys. Acta* **2008**, *1778*, 514.

[32] M.A. Grohme, B. Mali, W. Welnicz, S. Michel, R.O. Schill, M. Frohme, *Bioinform. Biol. Insights* **2013**, *7*, 153.

[33] B. Mali, M.A. Grohme, F. Forster, T. Dandekar, M. Schnolzer, D. Reuter, W. Welnicz, R.O. Schill, M. Frohme, *BMC Genomics* **2010**, *11*, 168.

[34] N.Y. Denekamp, Reinhardt, R., Albrecht, M. W., Drungowski, M., Kube, M., and Lubzens, E., *Hydrobiologia* **2011**, *662*, 51.

[35] J. Lapinski, A. Tunnacliffe, *FEBS Lett.* **2003**, *553*, 387.

[36] J.H. Crowe, J.F. Carpenter, L.M. Crowe, *Annu. Rev. Physiol.* **1998**, *60*, 73.

[37] J.H. Crowe, *Adv. Exp. Med. Biol.* **2007**, *594*, 143.

[38] S.N. Timasheff, in: G.N., S., C.B., O., C.L., B. (Eds.), Springer, Springer, Berlin, Heidelberg 1992, pp. 70.

[39] S.N. Timasheff, *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 67.

[40] J.S. Clegg, *Comp. Biochem. Physiol.* **1965**, *14*, 135.

[41] K. Madin, J. Crowe, *J. Exp. Biol.* **1975**, *193*, 335.

[42] J.H. Crowe, L.M. Crowe, A.E. Oliver, N. Tsvetkova, W. Wolkers, F. Tablin, *Cryobiology* **2001**, *43*, 89.

[43] M. Caprioli, A. Krabbe Katholm, G. Melone, H. Ramlov, C. Ricci, N. Santo, *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **2004**, *139*, 527.

[44] S.A. Watts, Lee K. J., and Cline G. B., *J. Exp. Biol.* **1994**, *270*, 426.

[45] T.H. MacRae, *Cell Stress Chaperon* **2016**, *21*, 9.

[46] T.C. Boothby, H. Tapia, A.H. Brozena, S. Piszkiewicz, A.E. Smith, I. Giovannini, L. Rebecchi, G.J. Pielak, D. Koshland, B. Goldstein, *Mol. Cell* **2017**, *65*, 975.

[47] K. Goyal, L.J. Walton, A. Tunnacliffe, *Biochem. J.* **2005**, *388*, 151.

[48] T. Kikawada, Y. Nakahara, Y. Kanamori, K. Iwata, M. Watanabe, B. McGee, A. Tunnacliffe, T. Okuda, *Biochem. Biophys. Res. Commun.* **2006**, *348*, 56.

[49] L. Dure, C. Chlan, *Plant Physiol.* **1981**, *68*, 180.

[50] A. Solomon, R. Salomon, I. Paperna, I. Glazer, *Parasitology* **2000**, *121 (Pt 4)*, 409.

[51] L. Dure, 3rd, M. Crouch, J. Harada, T.H. Ho, J. Mundy, R. Quatrano, T. Thomas, Z.R. Sung, *Plant Mol. Biol.* **1989**, *12*, 475.

[52] M.J. Wise, A. Tunnacliffe, *Trends Plant. Sci*. **2004**, *9*, 13.

[53] E. Jaspard, D. Macherel, G. Hunault, *PLOS one* **2012**, *7*, e36968.

[54] A.C. Cuming, in Seed Proteins,Vol. (Eds: Shewry, P. R., Casey, R.), Springer Netherlands, Dordrecht **1999**,Ch.

[55] M.-D. Shih, F.A. Hoekstra, Y.-I.C. Hsing, in *Advances in Botanical Research*,Vol. Volume 48 (Eds: Jean-Claude, K., Michel, D.), Academic Press **2008**,Ch.

[56] M. Battaglia, A. Covarrubias, Frontiers in Plant Science **2013**, *4*.

[57] S. Hengherr, A.G. Heyer, F. Brummer, R.O. Schill, *Physiol. Biochem. Zool.* **2011**, *84*, 147.

[58] S. Chakrabortee, R. Tripathi, M. Watson, G.S. Schierle, D.P. Kurniawan, C.F. Kaminski, M.J. Wise, A. Tunnacliffe, *Mol. Biosyst.* **2012**, *8*, 210.

[59] J.M. Mouillon, P. Gustafsson, P. Harryson, *Plant Physiol*. **2006**, *141*, 638.

[60] M. Hara, M. Fujinaga, T. Kuboi, *J. Exp. Bot.* **2005**, *56*, 2695.

[61] J. Svensson, E.T. Palva, B. Welin, *Prot. Exp. Pur.* **2000**, *20*, 169.

[62] W.F. Wolkers, S. McCready, W.F. Brandt, G.G. Lindsey, F.A. Hoekstra*, Biochim. Biophys. Acta* **2001**, *1544*, 196.

[63] T.C. Boothby, H. Tapia, A.H. Brozena, S. Piszkiewicz, A.E. Smith, I. Giovannini, L. Rebecchi, G.J. Pielak, D. Koshland, B. Goldstein, *Mol. Cell* **2017**, *65*, 975.

[64] D. Li, X. He, *Biomacromolecules* **2009**, *10*, 1469.

[65] K. Sasaki, N.K. Christov, S. Tsuda, R. Imai, *Plant Cell. Physiol.* **2014**, *55*, 136.

[66] Z. Dosztányi, B. Mészáros, I. Simon, *Bioinformatics* **2009**, *25*, 2745.

[67] L.C. Boswell, M.A. Menze, S.C. Hand, *Physiol. Biochem .Zool.* **2014**, *87*, 640.

[68] C.J. Jeffery, *Trends Genet*. **2003**, *19*, 415.

[69] P. Tompa, P. Csermely, *FASEB J.* **2004**, *18*, 1169.

[70] R. Hatanaka, Y. Hagiwara-Komoda, T. Furuki, Y. Kanamori, M. Fujita, R. Cornette, M. Sakurai, T. Okuda, T. Kikawada, *Insect Biochem. Mol. Biol.* **2013**, *43*, 1055.

[71] V.N. Uverskii, *Biofizika* **1998**, *43*, 416.

[72] M. Bokor, V. Csizmok, D. Kovacs, P. Banki, P. Friedrich, P. Tompa, K. Tompa, *Biophys. J.* **2005**, *88*, 2030.

[73] T. Chen, A. Fowler, M. Toner, *Cryobiology* **2000**, *40*, 277.

[74] T. Shimizu, Y. Kanamori, T. Furuki, T. Kikawada, T. Okuda, T. Takahashi, H. Mihara, M. Sakurai, *Biochemistry* **2010**, *49*, 1093.

[75] J. Buitink, O. Leprince, *C. R. Biol.* **2008**, *331*, 788.

[76] H.Z. Cummins, H. Zhang, J. Oh, J.-A. Seo, H.K. Kim, Y.-H. Hwang, Y.S. Yang, Y.S. Yu, Y. Inn, *J. Non-Cryst. Solids* **2006**, *352*, 4464.

[77] K. Goyal, L. Tisi, A. Basran, J. Browne, A. Burnell, J. Zurdo, A. Tunnacliffe, *J. Biol. Chem.* **2003**, *278*, 12977.

[78] A. Bremer, M. Wolff, A. Thalhammer, D.K. Hincha, *FEBS J.* **2017**, *284*, 919.

[79] D.S. Moore, R. Hansen, S.C. Hand*, Biochim. Biophys. Acta* **2016**, *1858*, 104.

[80] K. Nakayama, K. Okawa, T. Kakizaki, T. Honma, H. Itoh, T. Inaba, *Plant Physiol.* **2007**, *144*, 513.

[81] L.Y. Rivera-Najera, G. Saab-Rincon, M. Battaglia, C. Amero, N.O. Pulido, E. Garcia-Hernandez, R.M. Solorzano, J.L. Reyes, A.A. Covarrubias, *J. Biol. Chem.* **2014**, *289*, 31995.

[82] D.M. Mitrea, R.W. Kriwacki, *Cell Commun. Signal.* **2016**, *14*, 1.

[83] C.P. Brangwynne, C.R. Eckmann, D.S. Courson, A. Rybarska, C. Hoege, J. Gharakhani, F. Jülicher, A.A. Hyman, *Science* **2009**, *324*, 1729.

[84] H.-Y. Lee, K.-Y. Cheng, J.-C. Chao, J.-Y. Leu, *Microbial. Cell* **2016**, *3*, 109.

[85] T.J. Nott, E. Petsalaki, P. Farber, D. Jervis, E. Fussner, A. Plochowietz, T.D. Craggs, D.P. Bazett-Jones, T. Pawson, J.D. Forman-Kay, A.J. Baldwin, *Mol. Cell* **2015**, *57*, 936.

[86] S. Wegmann, B. Eftekharzadeh, K. Tepper, K.M. Zoltowska, R.E. Bennett, S. Dujardin, P.R. Laskowski, D. MacKenzie, T. Kamath, C. Commins, C. Vanderburg, A.D. Roe, Z. Fan, A.M. Molliex, A. Hernandez-Vega, D. Muller, A.A. Hyman, E. Mandelkow, J.P. Taylor, B.T. Hyman, *EMBO J.* **2018**, *37*, e98049.

[87] J.A. Riback, C.D. Katanski, J.L. Kear-Scott, E.V. Pilipenko, A.E. Rojek, T.R. Sosnick, D.A. Drummond, *Cell* **2017**, *168*, 1028.

[88] O. Bounedjah, L. Hamon, P. Savarin, B. Desforges, P.A. Curmi, D. Pastre, *J. Biol. Chem.* **2012**, *287*, 2446.

[89] C. Rabouille, S. Alberti, *Curr. Opin. Cell Biol.* **2017**, *47*, 34.

[90] B. Janis, V.N. Uversky, M.A. Menze, *J.Biomol. Struct. Dyn.* **2017**, 10.1080/07391102.2017.1387177.

[91] T.S. Harmon, A.S. Holehouse, M.K. Rosen, R.V. Pappu, *eLife* **2017**, eLife.30294.

[92] D.S.W. Protter, B.S. Rao, B. Van Treeck, Y. Lin, L. Mizoue, M.K. Rosen, R. Parker, *Cell* Reports **2018**, *22*, 1401.

[93] J.R. Wheeler, T. Matheny, S. Jain, R. Abrisch, R. Parker, *eLife* **2016**, *5*, e18413.

[94] A.S. Holehouse, J. Ahad, R.K. Das, R.V. Pappu, *Biophys. J.* **2015**, *108*, 228a.

[95] S. Vucetic, C. Brown, K. Dunker, Z. Obradovic, *Proteins* **2003**, *52*.

[96] V.N. Uversky, J.R. Gillespie, A.L. Fink, Proteins: Structure, Function, and Bioinformatics **2000**, *41*, 415.

[97] A.K. Dunker, C.J. Oldfield, J. Meng, P. Romero, J.Y. Yang, J.W. Chen, V. Vacic, Z. Obradovic, V.N. Uversky, *BMC Genomics* **2008**, *9 Suppl 2*, S1.

[98] R.K. Das, K.M. Ruff, R.V. Pappu, *Curr. Opin. Struct. Biol.* **2015**, *32*, 102.

[99] A.H. Mao, N. Lyle, R.V. Pappu, *Biochem J.* **2013**, *449*, 307.

[100] J. Warwicker, S. Charonis, R.A. Curtis, *Mol. Pharm.* **2014**, *11*, 294.

[101] S. Hengherr, A.G. Heyer, H.R. Kohler, R.O. Schill, *FEBS J.* **2008**, *275*, 281.

[102] N.N. Pouchkina-Stantcheva, B.M. McGee, C. Boschetti, D. Tolleter, S. Chakrabortee, A.V. Popova, F. Meersman, D. Macherel, D.K. Hincha, A. Tunnacliffe, *Science* **2007**, *318*, 268.

[103] J. Browne, A. Tunnacliffe, A. Burnell, *Nature* **2002**, *416*, 38.

[104] T. Tyson, W. Reardon, J.A. Browne, A.M. Burnell, *Int. J. Parasitol* **2007**, *37*, 763.

[105] M.A. Sharon, A. Kozarova, J.S. Clegg, P.O. Vacratsis, A.H. Warner, *Biochem. Cell Biol.* **2009**, *87*, 415.

[106] M.A. Menze, L. Boswell, M. Toner, S.C. Hand, *J. Biol. Chem* **2009**, *284*, 10714.

[107] S.C. Hand, D. Jones, M.A. Menze, T.L. Witt, *J. Exp. Zool. A* **2007**, *307*, 62.

[108] W. Zhao, F. Yao, M. Zhang, T. Jing, S. Zhang, L. Hou, X. Zou, *PLoS ONE* **2016**, *11*, e0162272.

[109] N.Y. Denekamp, M.A. Thorne, M.S. Clark, M. Kube, R. Reinhardt, E. Lubzens, *BMC Genomics* **2009**, *10*, 108.

[110] T.Z. Gal, I. Glazer, H. Koltai, *FEBS Lett.* **2004**, *577*, 21.

[111] S. Tanaka, J. Tanaka, Y. Miwa, D.D. Horikawa, T. Katayama, K. Arakawa, A. Toyoda, T. Kubo, T. Kunieda, *PLoS ONE* **2015**, *10*, e0118272.

[112] A. Yamaguchi, S. Tanaka, S. Yamaguchi, H. Kuwahara, C. Takamura, S. Imajoh-Ohmi, D.D. Horikawa, A. Toyoda, T. Katayama, K. Arakawa, A. Fujiyama, T. Kubo, T. Kunieda, *PLoS One* **2012**, *7*, e44209.

Fig. 1

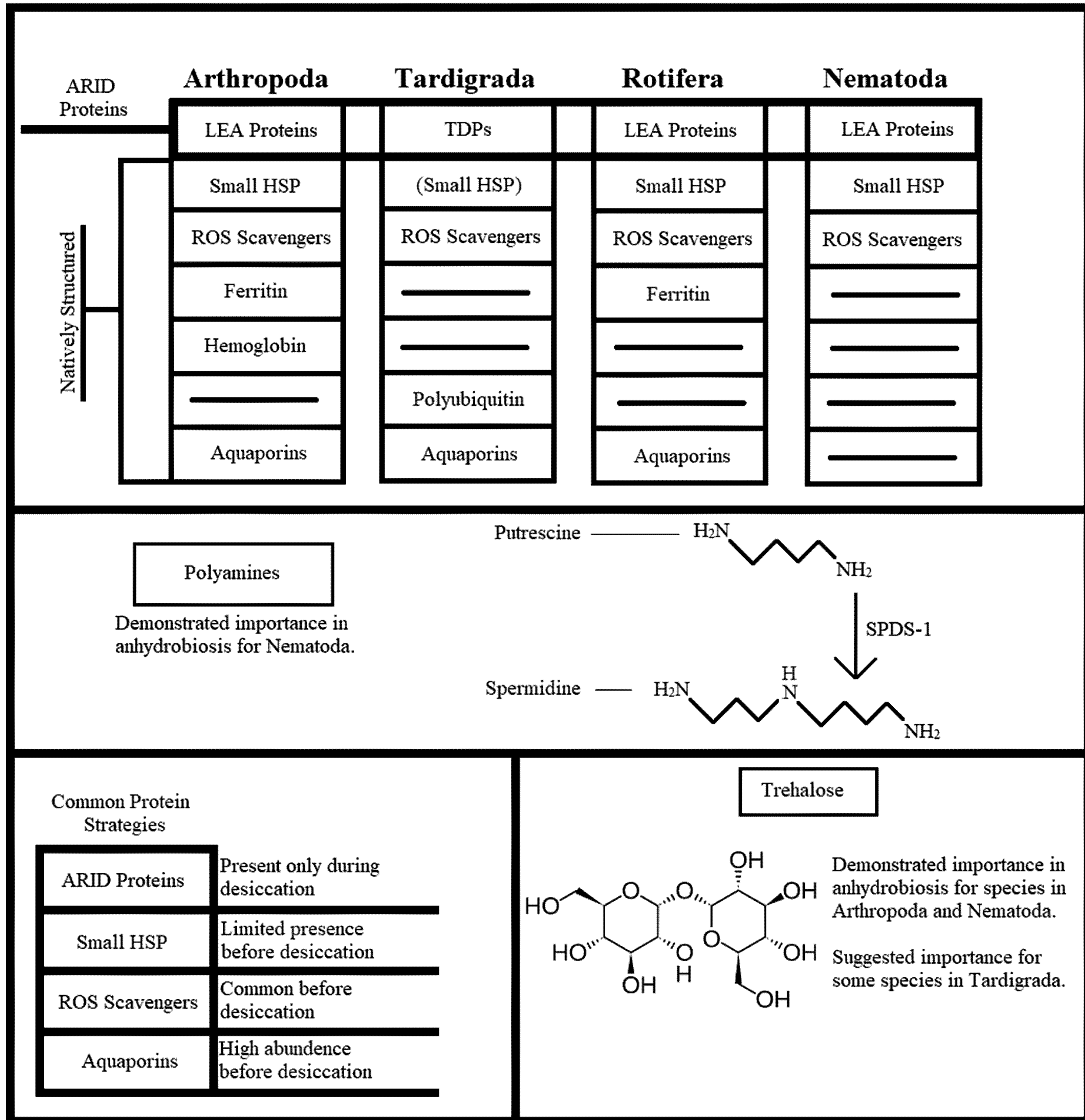

Fig. 2

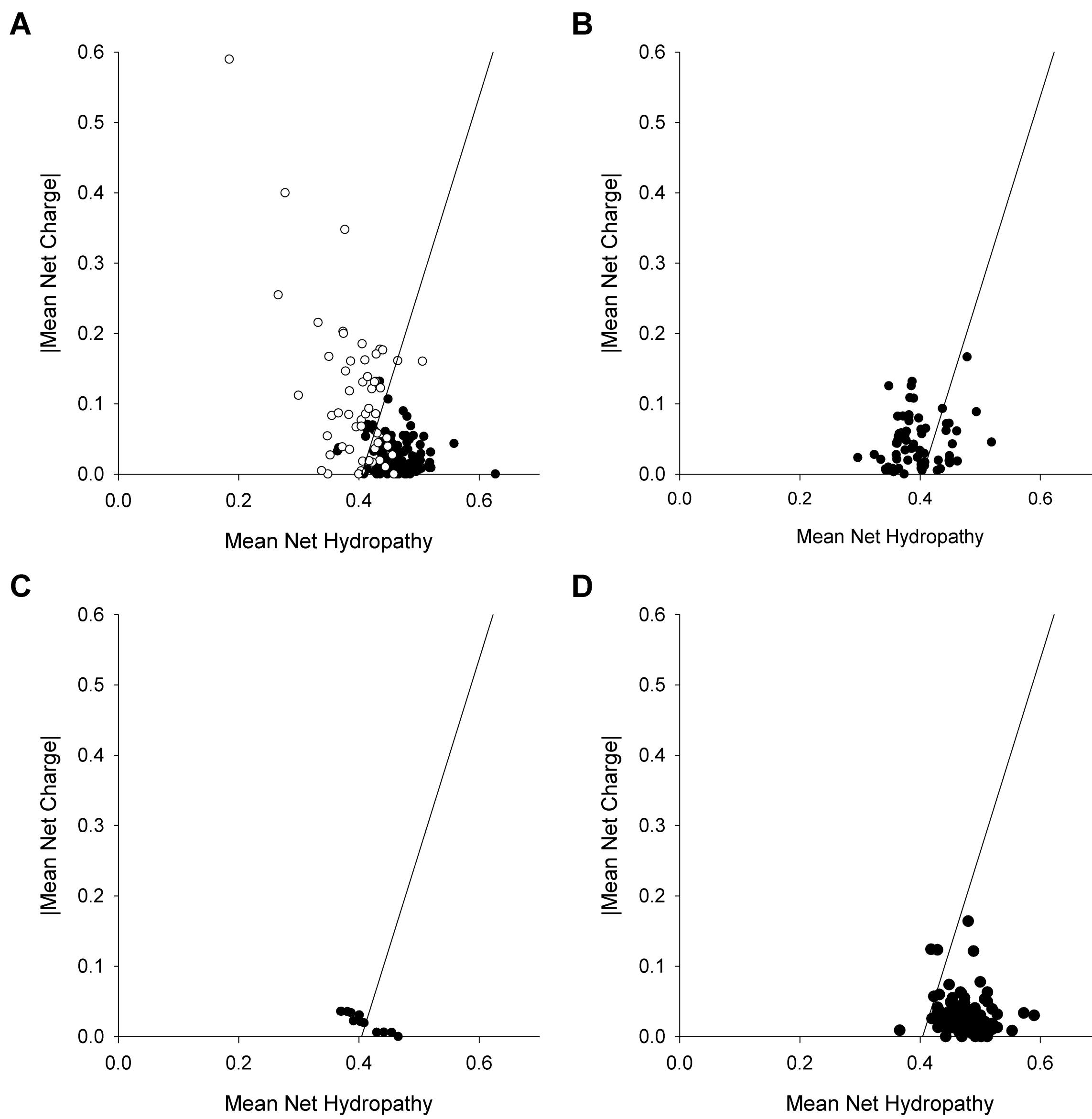

Fig. 3

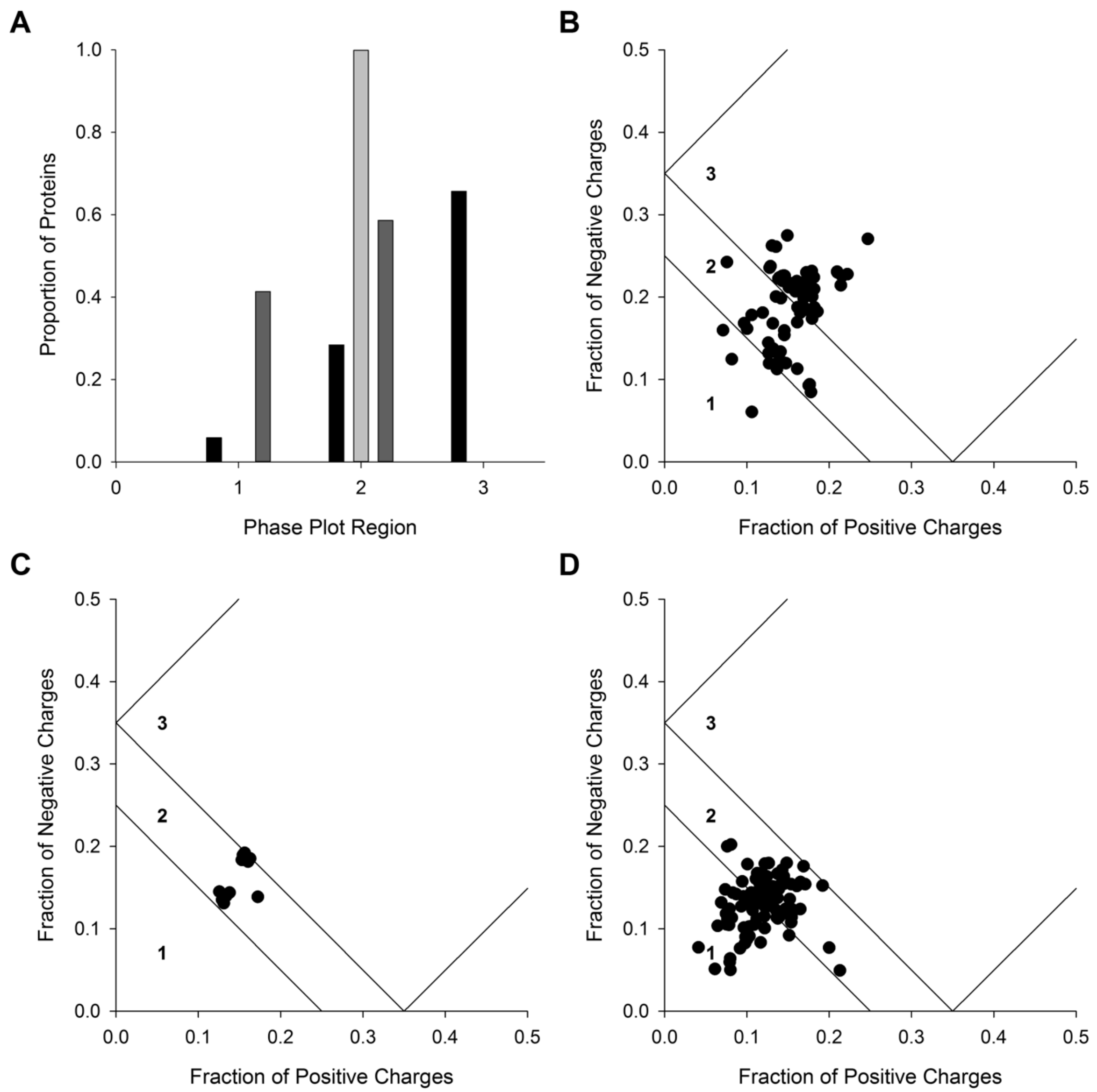

Fig. 4

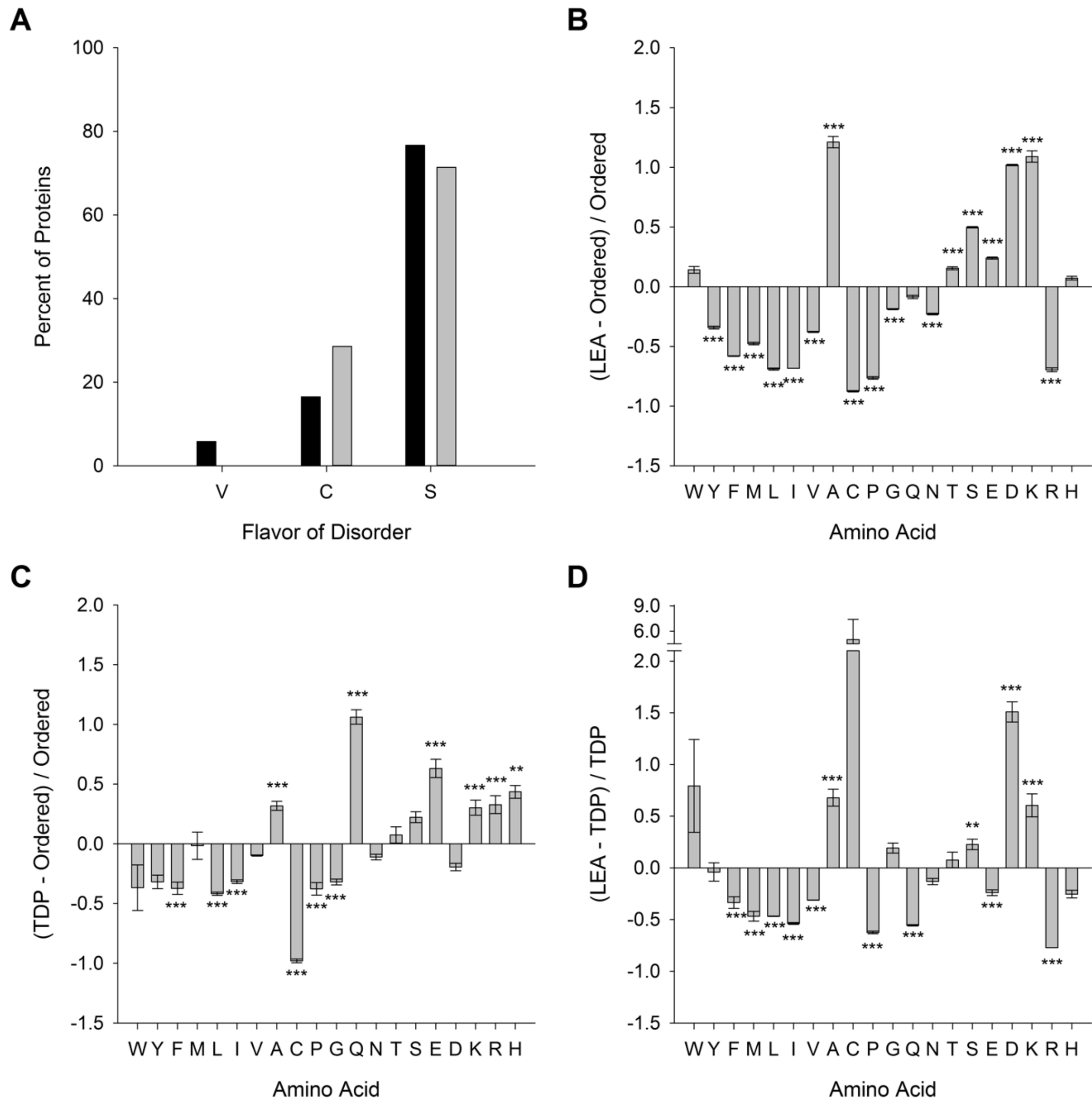

Fig. 5