Full title:

Tardigrades' cytoplasmic abundant heat soluble proteins serve as membrane protectors during dehydration.

Short title:

CAHS protects cell membranes during dehydration

Claire K. Zhang[1]*

[1]Brentwood High School, Brentwood, Tennessee, United States of America

* Corresponding author

E-mail: clairezhang168@gmail.com

Abstract

Tardigrades are microscopic organisms with extraordinary tolerance to environmental stresses such as desiccation and thus offer unique solutions to bio-preservation, anti-aging, and interstellar travel. Recent studies revealed a collection of cytoplasmic and secretory abundant soluble proteins (CAHSs and SAHSs respectively) contributing to tardigrades' extreme resilience. Using computational tools, I examined 39 CAHSs and 28 SAHSs from three representative tardigrade species. Both protein families possess a conserved central region and two highly variable terminal regions. Phylogenetic analysis suggests that CAHSs and SAHSs have distinct sequences despite functional similarity. Modeling by AlphaFold shows that CAHSs' central region forms α-helix whereas SAHSs' folds into β-barrel. Since dehydration leads to the increase of intracellular protein concentration, I used AlphaFold to simulate CAHS oligomerization and find that they preferably dimerize via their central helix motifs. Examination of CAHS dimers reveals a strong inter-helix interaction and an alternating pattern of hydrophobicity and hydrophilicity along the helix bundle. The latter is a common feature found in lipid-interacting proteins. Empirical tests using mammalian cells expressing the representative RvCAHS3 show that CAHSs concentrated on intracellular membranes instead of in proteinaceous condensates during dehydration. Moreover, RvCAHS3 significantly improves cell survival as measured by the stimulation-evoked release of Ca^{2+} from membrane-bound internal stores like endoplasmic reticulum. Taken together, these results suggest that CAHS dimers form network covering cell membranes and thus protect the integrity as well as function of membrane-bound organelles during anhydrobiosis. This finding implicates new strategies to preserve biomolecules, cells, and tissues under challenging conditions or for easy transportation.

Introduction

Tardigrades, commonly known as water bears or moss piglets, are microscopic invertebrates comprising the distinct phylum Tardigrada [1]. They have a segmented body with four pairs of legs and live in marine

or freshwater environments as well as semi-terrestrial habitats [1]. Tardigrades are renowned for withstanding extreme environments including desiccation, extreme temperatures, and cosmic radiation [2]. Most notably, they can remain lifeless in a vacuum for decades until rehydration [3]. Thus, tardigrades are exceptional models for studying anhydrobiosis (i.e., an organism loses almost all its water and enters a state of reversible ametabolism) [4]. For most organisms, dehydration causes hyperosmosis and damages cellular structures, leading to cell deformation and eventually death. Tardigrades, on the other hand, have various protective measures for such challenges.

Recent studies have found that two protein families, cytosolic and secretory abundant heat soluble proteins (CAHSs and SAHSs), serve as intracellular and extracellular protectants in tardigrades, respectively [5]. Early research had concluded that both were intrinsically disordered and vitrified upon dehydration, hypothetically sequestering biomolecules, organelles, and cellular apparatuses [5]. Lately, computational as well as empirical studies have indicated that those proteins have a defined tertiary structure, at least partially [6-9]. Moreover, the latest investigations have shown that CAHSs are protective for protein complexes and organelles [6] whereas SAHSs shelter extracellular biomolecules and structures [10]. To better appraise their protective mechanisms, especially in case of desiccation, I have employed *in silico* and *in vitro* tests with a focus on the oligomerization of CAHSs owing to dehydration. As more and more tardigrades have undergone genomic sequencing [11-12], more and more CAHS and SAHS gene sequences have been deposited in GenBank and become available to the general public. More importantly, sequence-based structure modeling fueled by the latest development in artificial intelligence has become highly reliable (e.g., AlphaFold 3.0 achieved over 97% accuracy in predicting protein complex) [13]. In conjunction with computer simulation, empirical tests in model systems like cultured cells yield insights bridging proteins' functionality with their structure.

Results

From among ~1,500 species of tardigrades, I selected three representative ones: *Ramazzottius varieornatus* (*Rv*) is best known for its extremotolerance [14]; *Hypsibius exemplaries* (*He*) is the most studied for evolutionary biology and astrobiology [15]; and *Paramacrobiotus metropolitanus* (*Pm*) is a popular genetic model [16]. Most of all, their genomes have been sequenced and almost all extremotolerance-related genes have been identified and deposited in public-accessible databases like GenBank [17]. After exhaustive searching, I acquired all DNA and protein sequences of CAHSs and SAHSs in those three species (S1 and S2 Tables).

First, the protein sequences of 39 CAHSs and 28 SAHSs were aligned using COBALT with default settings. A distinct consensus region was observed in each group (Fig 1). In CAHSs, the conserved region of approximately 130 amino acid residues was flanked by N- and C-terminal regions with highly variable lengths and amino acid compositions (Fig 1A). Notably, the most conserved amino acid residues (shown in red in Fig 1A) are hydrophilic (i.e., charged or polar) and distributed evenly across the consensus region (S1A Fig). In SAHSs, the conserved region is made of approximately 100 amino acid residues and is closer to the C-terminals (Fig 1B). Different from that of CAHSs, it is less consistent and can be segmented into three subregions of shorter and variable sequences, about 20 ~ 40-residues long. The most conserved amino acid residues in the conserved regions of SAHSs are either hydrophilic or hydrophobic (S1B Fig). The phylogenetic tree plots suggest that CAHSs are more conserved across different species because there are more species than gene differences between the neighboring CAHSs (S2A Fig). Notably, PmCAHS89226-like was found to be phylogenetically distant from all other 38 CAHS, consistent with its sequence alignment, indicating an incorrect categorization. As for SAHSs, the closest genes are always from the same species (S2B Fig). Even combined together, CAHSs and SAHSs form two separate branches (S2C Fig). In summary, the sequence analyses suggests that CAHSs and SAHSs are two very different families of proteins despite their shared names and functional similarity. Furthermore, their conserved regions have very different amino acid compositions, implicating differences in protein structure and function.

101 **Figure 1. Protein sequence alignment identifies highly conserved central motifs and highly variable**
102 **terminal regions in CAHSs and SAHSs.** Alignments of CAHSs (A) and SAHSs (B) were generated by
103 COBALT. Red represents highly conserved amino acid residues, blue for less conserved ones, and gray for
104 highly variable ones. The conserved central motif for CAHSs is about 120 amino acids long, and SAHSs'
105 is about 100 amino acids long. RC, Rv CAHS; HC, He CAHS; PC, Pm CASH. They are all numbered in
106 the same order as that of S1 Fig and S1 and S2 Tables.
107
108 To model CAHSs and SAHSs, the AI-based AlphaFold was used because its latest version (3.0) offers
109 unprecedented accuracy and reliability, especially in predicting protein complex including oligomers [13].
110 Consistent with the sequence alignment, all of CAHSs' long conserved regions formed a single α-helix (Fig
111 2A and S3A Fig) except for PmCAHS89226-like, again suggesting that it was miscategorized. So, it is
112 excluded from all analyses thereafter. In case of SAHSs, their conserved regions form several consecutive
113 β-sheets with α-helixes or coils in between (Fig S3B). Unlike CAHSs, there is no outlier in SAHSs, again
114 consistent with the phylogenetic analysis result. As for the highly variable N- and C-terminal regions in
115 CAHSs and SAHSs, they are generally deemed to be disordered by AlphaFold with low predicted local
116 distance difference test score (plDDT) (Fig 2A and S3 Fig).
117
118 **Figure 2**. **CAHS's central motif forms long helix.** (A) the 3D structures of CAHS monomers predicted
119 by AlphaFold. Blue indicates high confidence for the highly ordered central helix whereas yellow/orange
120 indicates low confidence for the disordered terminals. (*, PmCAHS89226-like). (B) the hydrophobic
121 moment plot of RvCAHS3's central motif shows hydrophobic residues are arranged in a periodic manner.
122 (C) two central motifs of RvCAHS3 form an anti-parallel dimer via electrostatic interactions (highlighted
123 in green) and π-π stacking (highlighted in yellow). Most hydrophobic residues face outside.
124

125 Using RvCAHS3 as an example, the hydrophobic moments along the conserved central regions were
126 calculated, which clearly exhibits periodic peaks (Fig 2B). This is consistent with the evenly distributed
127 hydrophobic amino acid residues shown in S1A Fig. Dehydration effectively concentrates biomolecules
128 and thus promotes CAHS/SAHS oligomerization inside and outside of cells. So, AlphaFold 3.0 was
129 deployed to model the formation of CAHS and SAHS oligomers. For example, when a second helix of
130 RvCAHS3 conserved regions was introduced, AlphaFold yielded a dimer of two helixes in an anti-parallel
131 fashion (Fig 2C). A close examination of inter-peptide interactions revealed strong electrostatic interactions
132 and π-π stacking along the interface of the two helixes (highlighted in Fig 2C), indicating a high stability
133 of such dimer. Moreover, the hydrophobic (red) and hydrophilic (blue) segments of both helixes were well
134 aligned in the dimer (Fig 2C), very much reminiscent to helix bundles in the lipid-binding apolipoproteins
135 [9, 18].
136
137 Like RvCAHS3 (Fig 3A1), all CAHS dimers and trimers were formed in an anti-parallel fashion with
138 moderate increase or decrease of prediction confidence (i.e., predicted template modeling score, pTM) in
139 comparison to monomers (Fig 3A2). For the central helix motif alone, the overall confidence scores were
140 the highest for dimers but dropped sharply for trimers (Fig 3B). Due to the low prediction scores of the
141 disordered regions obscuring interactions between the central helix motifs, the consensus regions of CAHSs
142 were used for the subsequent modeling of CAHS oligomerization thereafter. Fig 3C illustrates a clear trend
143 of decrease in pTM as oligomerization progresses. Due to the strong binding in the dimers and the high
144 confidence in dimer prediction, I hypothesize that desiccation promotes CAHSs to dimerize and the dimers
145 connect to each other to form a protective mesh on cell membranes. I further speculate that the covering of
146 such a CAHS oligomer network can prevent the merge or collapse of intracellular membranes and
147 consequently prevent the breakdown of membrane-bound organelles, a requisite for surviving desiccation.
148
149 **Figure 3**. **A helical dimer is the most likely form during CAHS oligomerization.** (A) predicted structures
150 of full-length RvCAHS3 in mono-, di-, and trimer forms (A1) and the prediction confidence (measured as

151 pTM) for the mono-, di-, and trimers of all 39 full-length CAHSs (A2), which lacks an overall trend of
152 changes. (B) predicted structures of RvCAHS3 central motif in mono-, di-, and trimer forms (B1) and the
153 prediction confidence (measured as pTM scores) of the central motifs of all 39 CAHSs (B2), which shows
154 an overall increase for dimers and a drastic decrease for timers. (C) the box and whisker plot combined with
155 point-line plot shows the trend of prediction confidence (measured as ipTM) during the oligomerization
156 (from dimers to decamers) of the central motifs of six representative CAHSs. There is a progressive
157 decrease from dimers to decamers.
158
159 In order to empirically test that hypothesis, 3T3 cells (i.e., immortalized mouse embryonic fibroblast cells)
160 growing on Matrigel-coated glass coverslips were transfected with a mammalian-expressing plasmid
161 encoding RvCAHS3, which is tagged with green fluorescent protein (i.e., CAHS3-AcGFP1) for detection
162 by fluorescence microscopes [19]. About 1 day after the transfection, more than 80% cells expressed
163 CAHS3-AcGFP1 (estimated by AcGFP1 fluorescence). In order to simulate dehydration, cells growing on
164 the coverslips were air-dried in a laminar flow cabinet at room temperature (~25°C) for various periods of
165 time (i.e., 0, 1, 2, 5, 10, and 20 minutes). Previous studies suggested that CAHSs underwent gel-transition
166 or liquid-liquid phase separation (LLPS) upon dehydration-like treatments [9, 19]. To visualize LLPS, the
167 same cells were co-transfected with a DsRed-expressing plasmid. It is well documented that DsRed inclines
168 to aggregate and form LLPS-like protein condensates [20]. To visualize membrane-bound organelles, the
169 transfected cell cultures were incubated with FM4-64, a far-red fluorescent dye that can reversibly insert
170 into lipid bilayers and label intracellular membranes after being endocytosed. After loading, the FM4-64
171 remaining on the cell surface membrane was readily washed off by a 5-minute perfusion with dye-free
172 Tyrode's solution. Fig 4A and supplementary movies exemplify such triple-labeled 3T3 cells (blue, cell
173 membrane; green, CASH3; and red, protein condensates). After 5-min air drying in a laminar flow cabinet,
174 the majority of AcGFP1 fluorescence was found to be colocalized with that of FM4-64 but not DsRed (Fig
175 4A). More importantly, there is a significant correlation between AcGFP1 and FM4-64 but not DsRed (Fig

176 4B&C), meaning more membranes came with more CAHS3. This result suggests that dehydration drove
177 most CAHS3 onto intracellular membranes instead of LLPS-like protein condensates.
178
179 **Figure 4**. **During dehydration, CAHS is mostly associated with cell membranes rather than**
180 **proteinaceous condensates.** (A) sample confocal fluorescence images of DsRed and CAHS3-AcGFP1-
181 expressing cells pre-loaded with FM4-64 and air-dried at 25°C for 10 minutes. Scale bar, 20μm. (B) a
182 scatter plot of FM4-64 and CAHS3-AcGFP1 fluorescence intensities within the puncta defined by FM4-64
183 (i.e., membrane-bound organelles). The two values are significantly correlated ($p < 0.0001$), i.e., CAHS3-
184 AcGFP1 colocalized with the membranes during dehydration. (C) a scatter plot of DsRed and CAHS3-
185 AcGFP1 fluorescence intensities within the puncta defined by DsRed (i.e., proteinaceous condensates). The
186 two values are not correlated ($p = 0.3114$), i.e., CAHS3-AcGFP1 did not co-condense with DsRed during
187 dehydration.
188
189 Next, Ca^{2+}-imaging was used to test if CAHS3 protects membrane-bound organelles and made cells more
190 resilient to dehydration. For that, transfected 3T3 cells and controls were pre-loaded with a cell membrane-
191 permeable red fluorescent Ca^{2+}-indicator (i.e., X-Rhod-1AM) [21] before they were air dried. Immediately
192 after drying, those cultures were rehydrated and continuously perfused with the normal Tyrode's solution
193 (in mM: NaCl, 140; KCl, 2; CaCl2, 2; MgCl2, 2; HEPES, 10; D-Glucose, 10mM. pH7.35; 305 Osm/L).
194 During imaging, the application of ATP stimulated the release of Ca^{2+} from internal stores like endoplasmic
195 reticulum (ER, a major membrane-bound organelles), allowing the testing of cell survival and particularly
196 the functional integrity of membrane-bound organelles. Fig 5A shows that longer air-drying caused less
197 cells responsive to ATP stimulation. Notably, there were significantly more cells responding to ATP
198 stimulation in the CAHS3-expressing group than the control. Statistical significance was reached at 2, 5,
199 10, and 20 minutes (Fig 5B), demonstrating that CAHS3 when present in mammalian cells is indeed
200 protective against dehydration.
201

202 **Figure 5**. **CAHS3 increased cell survival during dehydration.** (A) sample fluorescence images of control
203 and CAHS3-expressing 3T3 cells still responding to 50μM ATP stimulation (visualized by X-Rhod-1AM,
204 a red Ca^{2+} indicator) after 1, 5, and 20-min air-drying at 25°C. Scale bar, 100μm. (B) Percentage of
205 responding cells (mean ± SEM) after air-drying at 25°C. The CAHS3-expressing group had more
206 responding cells than the control group since 5-min drying (all p <0.05)
207

208 # Discussion

209 The fascinating ability of tardigrades to sustain and survive extreme environments such as the vacuum of
210 space ignites great interest in using them as model organisms to study biological mechanisms for bio-
211 preservation, anti-aging, and space travel [1-3]. The physiological basis for tardigrades' extremotolerance
212 is anhydrobiosis, for which intrinsically disordered proteins, namely CAHSs and SAHSs, are known to be
213 essential. As more and more CAHSs and SAHSs have been discovered in different species of tardigrades
214 [14-17], it becomes clear that both of them are indispensable for the preservation of intracellular and
215 extracellular structures and functions during anhydrobiosis [5, 6, 10]. Previously, these unstructured
216 proteins were believed to work as absorbents for intra- and extracellular biomolecules [5]. However, it is
217 puzzling how a single mechanism can deal with very different needs by intracellular and extracellular
218 apparatuses. Furthermore, significant differences in protein sequences and subcellular localizations
219 between CAHSs and SAHSs speak against the notion of a shared mechanism between the two.
220

221 To better understand CAHSs and SAHSs, I started with structural analysis. As proteins' functions are largely
222 determined by their peptide sequences, I collected all CAHSs and SAHSs sequences in three representative
223 tardigrade species from GenBank. The sequence alignments unveiled highly conserved regions in both
224 CAHSs and SAHSs, which are significantly different from each other (Fig 1), which is confirmed by their
225 separation in the phylogenetic trees (Fig S2C). Next, AlphaFold 3.0 consistently predicted a single α-helix
226 for the consensus region of CAHSs (Figs 2, 3, and S3A&C) and a mix of β-sheets and short α-helixes for

227 that of SAHSs (Fig S3B&D), which suggests that neither CAHSs nor SAHSs are completely disordered.
228 Given their structural difference, CAHSs and SAHSs very likely act differently for cell protection.
229
230 Due to their unique and highly stable helical dimers (Fig 3), I focused on CAHSs. Intriguingly, the highly
231 conserved hydrophilic amino acid residues and repeated hydrophobic moments (Fig 2B&C) result in the
232 periodic hydrophobicity and hydrophilicity across the helical bundle, a characteristic structure found in
233 lipid-binding proteins like ApoE [9, 18]. This indicates that CAHS dimers favorably interact with lipid
234 membranes. Furthermore, the connection of CAHS dimers via their unstructured terminal regions likely
235 promotes the formation of CAHS-dimer networks covering cell membranes. This prediction is supported
236 by the experimental observation that RvCAHS3 mostly co-localizes with intracellular membrane label (i.e.,
237 FM4-64) upon dehydration (Fig 4). Hence, I propose that dehydration promotes the formation of a web of
238 CAHS dimers, which attaches to intracellular membranes and creates a barrier to prevent the collapsing or
239 merging of intracellular membranes. By doing so, CAHSs can help membrane-bound organelles to retain
240 their integrity when cells undergo desiccation. Again, this idea is supported by the observation that CAHS3-
241 expressing 3T3 cells exhibited significantly better survival rates than the control after prolonged
242 dehydration (>2 minutes) (Fig 5).
243
244 Due to the constraints of AlphaFold, the oligomer modeling could not account for changes in biomolecule
245 mixing, ion concentration, or other extracellular and intracellular changes during dehydration. Additionally,
246 there were unaccounted errors due to the limitations in AI algorithms and training datasets used by
247 AlphaFold. Nevertheless, the fact that AlphaFold consistently predicts the helical central motifs and helix
248 dimers for most CAHSs reassures the structural prediction. The cell-based assays so far only investigated
249 RvCAHS3 in the cytoplasm. Thus, it is worthwhile to expand such empirical study to other CAHSs from
250 different species of tardigrades or bearing structural difference from RvCAHS3. It is also interesting to
251 investigate if such a mechanism by RvCAHS3 can protect cell surface membrane, which can be achieved
252 by adding a secretory signaling sequence to RvCAHS3, relocating it to extracellular spaces. In addition,

253 alternative challenges like hyperosmotic stress or different types of cells like the more fragile neurons can
254 be used to explore the protection capacity of CAHSs and SAHSs. Future research on their protective
255 mechanisms should be extended to whole animals using model organisms such as *C. elegans*, which is
256 certainly more informative for translational applications. Last but not least, such hybrid studies combing
257 computational and empirical analyses can be applied to intrinsically disordered proteins native to
258 mammalian cells (e.g., late embryogenesis abundant proteins) to investigate and improve their protective
259 effects for clinical use.
260

261 Materials and Methods

262 All nucleotide and protein sequences used for this study were obtained from GenBank and UniProt using
263 keyword searching (i.e., abundant heat soluble protein, CAHS, or SAHS) and filtered by selected tardigrade
264 species. All protein sequence alignments were performed using constraint-based multiple alignment tools
265 (i.e., COBALT) [22] available from the National Center for Biotechnology Information. The default
266 alignment parameters were used. To generate phylogenetic trees based on the sequence alignments, I used
267 the ETE3 toolkit with default settings available from GenomeNet (www.genome.jp). All structural models
268 of CAHSs and SAHSs were generated using AlphaFold 3.0 [13]. The default settings were used to ensure
269 a fairness to all proteins. Resulting structures were downloaded and visualized using UCSF Chimera
270 program [23].
271

272 All chemical reagents were acquired from Thermo Fisher Scientific unless specified. 3T3 cells were gifted
273 from Dr. Henriette van Praag. All DNA plasmids were acquired from Addgene. DNA extraction and
274 purification were completed using MaxiPrep kit from Zymo Research. DNA transfection to 3T3 cells were
275 done using Lipofectamine. Confocal fluorescence imaging was carried out using Nikon A1R confocal
276 system, and Ca^{2+}-imaging were conducted with a Nikon Ti-E microscope controlled by μManager [24].

Image analyses were executed using FIJI [25]. Statistical analyses and plots were done using Excel and/or Prism.

Acknowledgment

I would like to thank Dr. Cristina Fenollar Ferrer in the Stiles-Nicholson Brain Institute at Florida Atlantic University for mentoring me and supporting my research.

References

1. McInnes SJ, Jørgensen A, Michalczyk Ł. 20 years of Zootaxa: Tardigrada (Ecdysozoa: Panarthropoda). Zootaxa2021.
2. Goldstein B. Tardigrades. Nature Methods. 2022;19(8):904–5.
3. Arakawa K. Examples of Extreme Survival: Tardigrade Genomics and Molecular Anhydrobiology. Annual Review of Animal Biosciences2022.
4. Keilin D. The problem of anabiosis or latent life: history and current concept. Proceedings of the Royal Society of London Series B, Biological sciences. 1959;150(939).
5. Boothby TC, Tapia H, Brozena AH, Piszkiewicz S, Smith AE, Giovannini I, et al. Tardigrades Use Intrinsically Disordered Proteins to Survive Desiccation. Molecular Cell. 2017;65(6):975–84.e5.
6. Bino T, Goto Y, Maryu G, Arakawa K, Aoki K. Possible roles of CAHS proteins from Tardigrade in osmotic stress tolerance in mammalian cells. Cell Structure and Function. 2024;49(2):123–33.
7. Fukuda Y, Inoue T. Crystal structure of secretory abundant heat soluble protein 4 from one of the toughest “water bears” micro-animals Ramazzottius Varieornatus. Protein Science. 2018;27(5).
8. Fukuda Y, Miura Y, Mizohata E, Inoue T. Structural insights into a secretory abundant heat-soluble protein from an anhydrobiotic tardigrade, Ramazzottius varieornatus. FEBS Letters. 2017;591(16).
9. Veling MT, Nguyen DT, Thadani NN, Oster ME, Rollins NJ, Brock KP, et al. Natural and Designed Proteins Inspired by Extremotolerant Organisms Can Form Condensates and Attenuate Apoptosis in Human Cells. ACS Synth Biol. 2022;11(3):1292–302.
10. Lim S, Reilly CB, Barghouti Z, Marelli B, Way JC, Silver PA. Tardigrade secretory proteins protect biological structures from desiccation. Communications Biology. 2024;7(1).
11. Li L, Ge Z, Liu S, Zheng K, Li Y, Chen K, et al. Multi-omics landscape and molecular basis of radiation tolerance in a tardigrade. Report No.: 011163281.1.
12. Erözden AA, Tavsanli N, Çalışkan M. Advances in bioinformatic approaches to tardigrade phylogeny. Computational Biology and Chemistry: Elsevier Ltd; 2024.

309 13. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of
310 biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493–500.
311 14. Neves RC, Hvidepil LKB, Sorensen-Hygum TL, Stuart RM, Mobjerg N. Thermotolerance
312 experiments on active and desiccated states of Ramazzottius varieornatus emphasize that tardigrades are
313 sensitive to high temperatures. Sci Rep. 2020;10(1):94.
314 15. Goldstein B. The Emergence of the Tardigrade Hypsibius exemplaris as a Model System. Cold
315 Spring Harb Protoc. 2018;2018(11).
316 16. Sugiura K, Matsumoto M, Kunieda T. Description of a model tardigrade Paramacrobiotus
317 metropolitanus sp. nov. (Eutardigrada) from Japan with a summary of its life history, reproduction and
318 genomics. Zootaxa. 2022;5134(1):92–112.
319 17. Fleming JF, Pisani D, Arakawa K. The Evolution of Temperature and Desiccation-Related Protein
320 Families in Tardigrada Reveals a Complex Acquisition of Extremotolerance. Genome Biol Evol.
321 2024;16(1).
322 18. Narayanaswami V, Kiss RS, Weers PM. The helix bundle: a reversible lipid binding motif. Comp
323 Biochem Physiol A Mol Integr Physiol. 2010;155(2):123–33.
324 19. Tanaka A, Nakano T, Watanabe K, Masuda K, Honda G, Kamata S, et al. Stress-dependent cell
325 stiffening by tardigrade tolerance proteins that reversibly form a filamentous network and gel. PLoS Biol.
326 2022;20(9):e3001780.
327 20. Bracha D, Walls MT, Brangwynne CP. Probing and engineering liquid-phase organelles. Nat
328 Biotechnol. 2019;37(12):1435–45.
329 21. Zhou X, Belavek KJ, Miller EW. Origins of Ca(2+) Imaging with Fluorescent Indicators.
330 Biochemistry. 2021;60(46):3547–54.
331 22. Papadopoulos JS, Agarwala R. COBALT: Constraint-based alignment tool for multiple protein
332 sequences. Bioinformatics. 2007;23(9):1073–9.

333 23. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera
334 - A visualization system for exploratory research and analysis. Journal of Computational Chemistry.
335 2004;25(13):1605–12.
336 24. Edelstein AD, Tsuchida MA, Amodaj N, Pinkard H, Vale RD, Stuurman N. Advanced methods of
337 microscope control using muManager software. J Biol Methods. 2014;1(2).
338 25. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: An open-
339 source platform for biological-image analysis. Nature Methods2012. p. 676–82.

340

341

342 **Supporting information**

343 **S1 Fig**. **Characterization of amino acid residues within the consensus sequences of CAHSs (A) and**
344 **SAHSs (B).** Red, hydrophobic; blue, hydrophilic; gray, neutral.

345 **S2 Fig**. **Phylogenetic trees of CAHSs (A), SAHSs (B), and both (C), generated by ETE3.**

346 **S3 Fig**. **Structural prediction of 39 CAHSs and 28 SAHSs in monomer (A and B).** The prediction
347 confidence is measured as the predicted local distance difference test (plDDT). Dark blue indicates high
348 confidence (i.e., plDDT > 90), blue to cyan indicates good confidence (i.e., 90 > plDDT > 70), yellow
349 indicates low confidence (i.e., 70 > pTM > 50), and orange indicates very low confidence (i.e., pTM < 50).

350 **S1 Table. All CAHSs used in this study.**

351 **S2 Table. All SAHSs used in this study.**

352 **S1 Files. Ca^{2+}-imaging of Controls after 1, 5, and 20-minute airdry (3 AVI files).**

353 **S2 Files. Ca^{2+}-imaging of CAHS3-expressing cultures after 1, 5, and 20-minute airdry (3 AVI files).**

354 **S3 Files. 3D view of reconstructed confocal images showing CAHS3-AcGFP1 (green), DsRed (red),**
355 **and FM4-64 (blue) in 3T3 cells after 5-minute airdry (3 AVI files).**

Tardigrades' cytoplasmic abundant heat soluble proteins serve as membrane protectors during dehydration.

Claire K. Zhang[1]

Figures and Legends

[1]C. K. Zhang is with the Brentwood High School, Brentwood, TN 37027 USA (corresponding author to e-mail: clairezhang168@gmail.com).

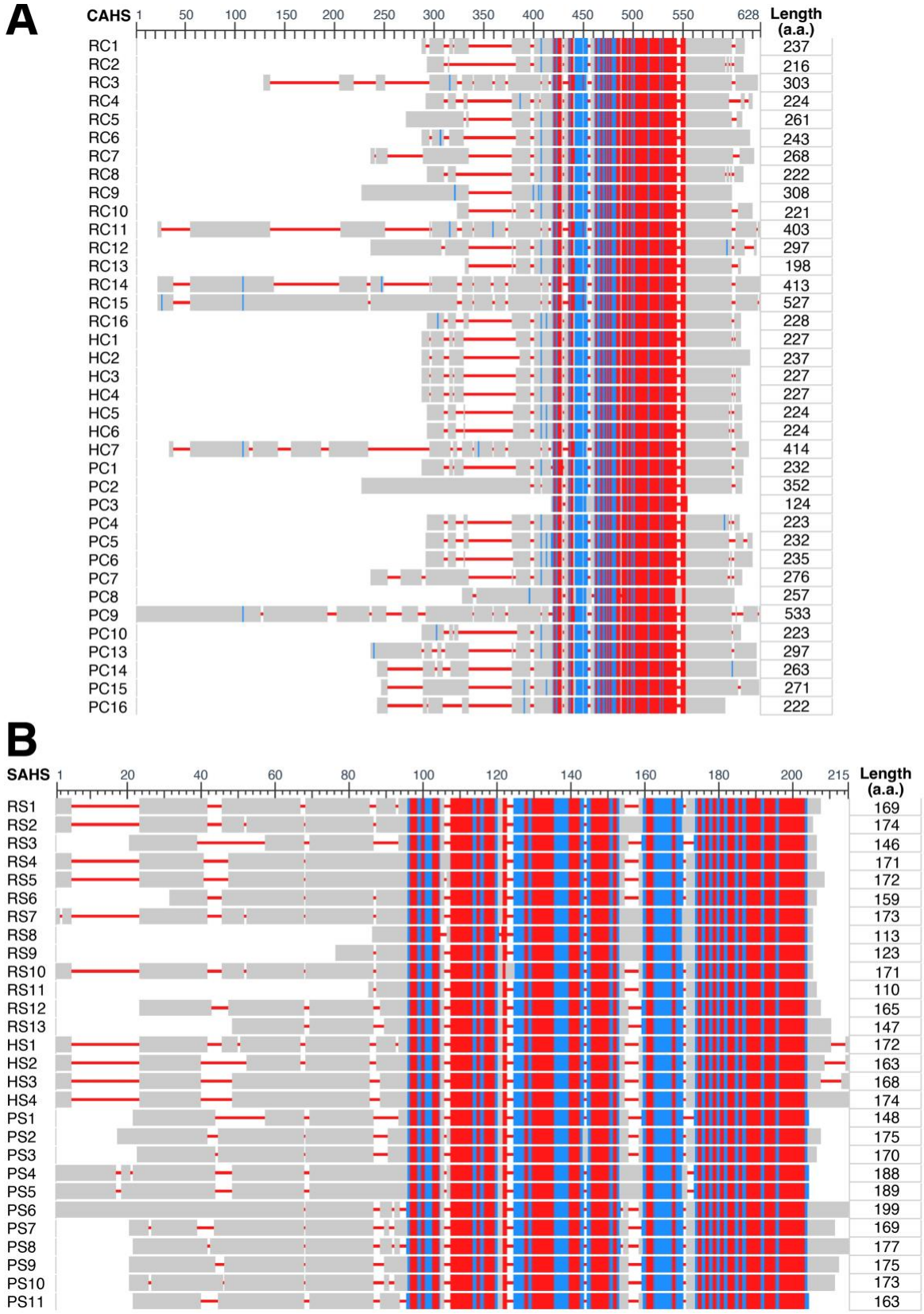

Fig 1. Protein sequence alignment identifies highly conserved central motifs and highly variable terminal regions in CAHSs and SAHSs.
Alignments of CAHSs (A) and SAHSs (B) were generated by COBALT. Red represents highly conserved amino acid residues, blue for less conserved ones, and gray for highly variable ones. The conserved central motif for CAHSs is about 120 amino acids long, and SAHSs' is about 100 amino acids long. RC, Rv CAHS; HC, He CAHS; PC, Pm CAHS. They are all numbered in the same order as that of Figure S1 and Table S1 & S2.

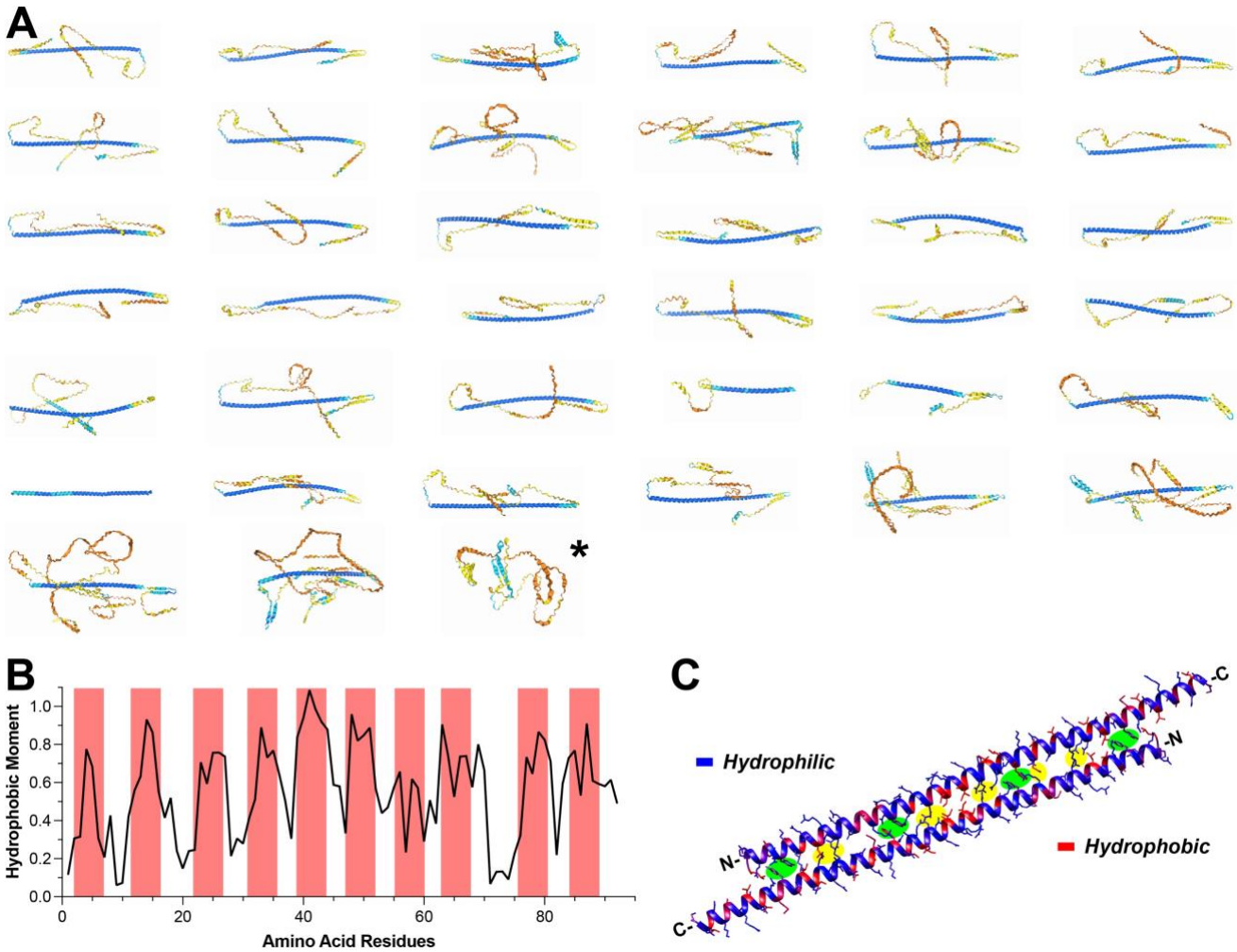

Fig 2. CAHS's central motif forms long helix.
(A) the 3D structures of CAHS monomers predicted by AlphaFold. Blue indicates high confidence for the highly ordered central helix whereas yellow/orange indicates low confidence for the disordered terminals. (*, PmCAHS89226-like). (B) the hydrophobic moment plot of RvCAHS3's central motif shows hydrophobic residues are arranged in a periodic manner. (C) two central motifs of RvCAHS3 form an anti-parallel dimer via electrostatic interactions (highlighted in green) and π-π stacking (highlighted in yellow). Most hydrophobic residues face outside.

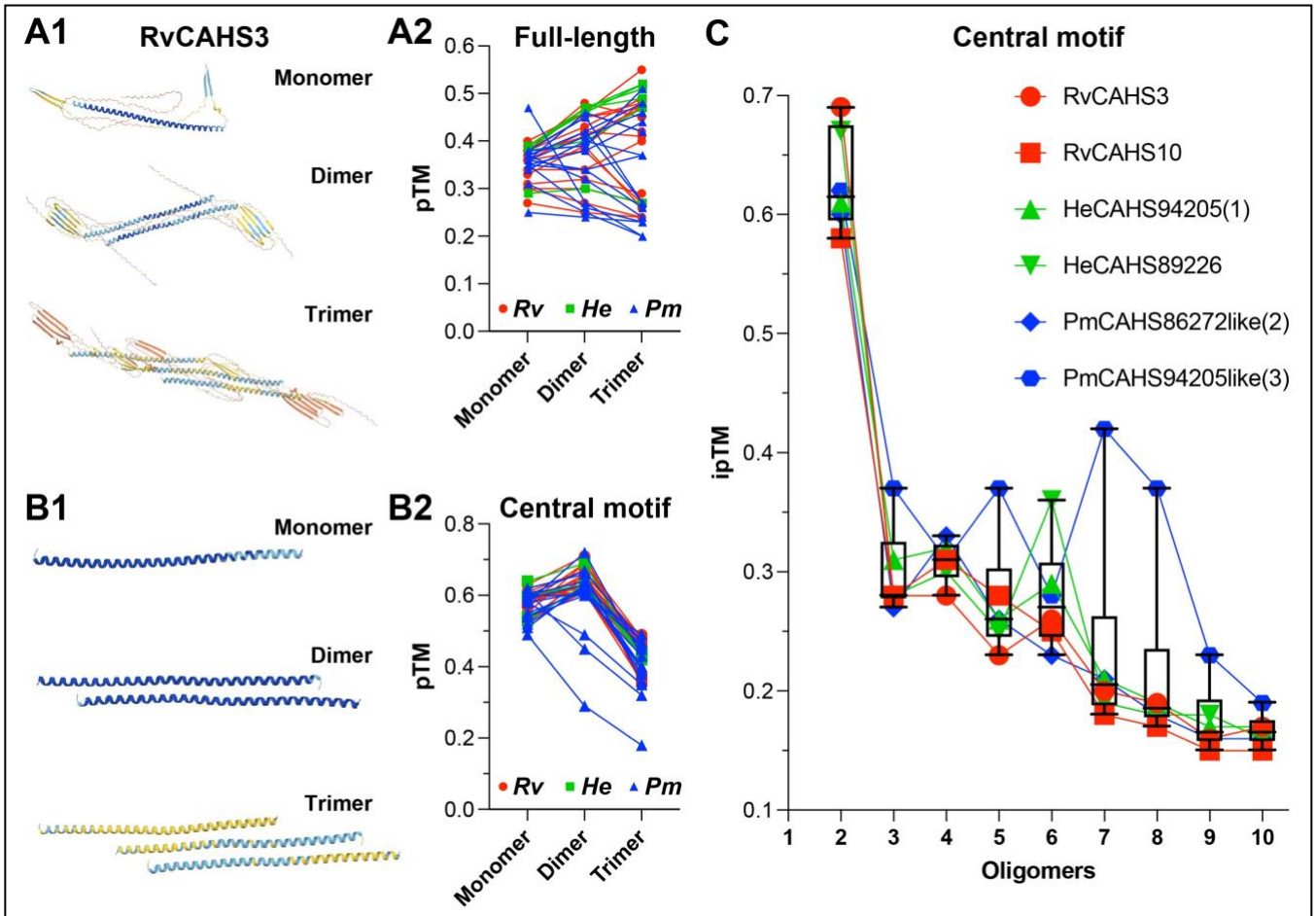

Fig 3. A helical dimer is the most likely form during CAHS oligomerization.

(A) predicted structures of full-length RvCAHS3 in mono-, di-, and trimer forms (A1) and the prediction confidence (measured as pTM) for the mono-, di-, and trimers of all 39 full-length CAHSs (A2), which lacks an overall trend of changes. (B) predicted structures of RvCAHS3 central motif in mono-, di-, and trimer forms (B1) and the prediction confidence (measured as pTM scores) of the central motifs of all 39 CAHSs (B2), which shows an overall increase for dimers and a drastic decrease for timers. (C) the box and whisker plot combined with point-line plot shows the trend of prediction confidence (measured as ipTM) during the oligomerization (from dimers to decamers) of the central motifs of six representative CAHSs. There is a progressive decrease from dimers to decamers.

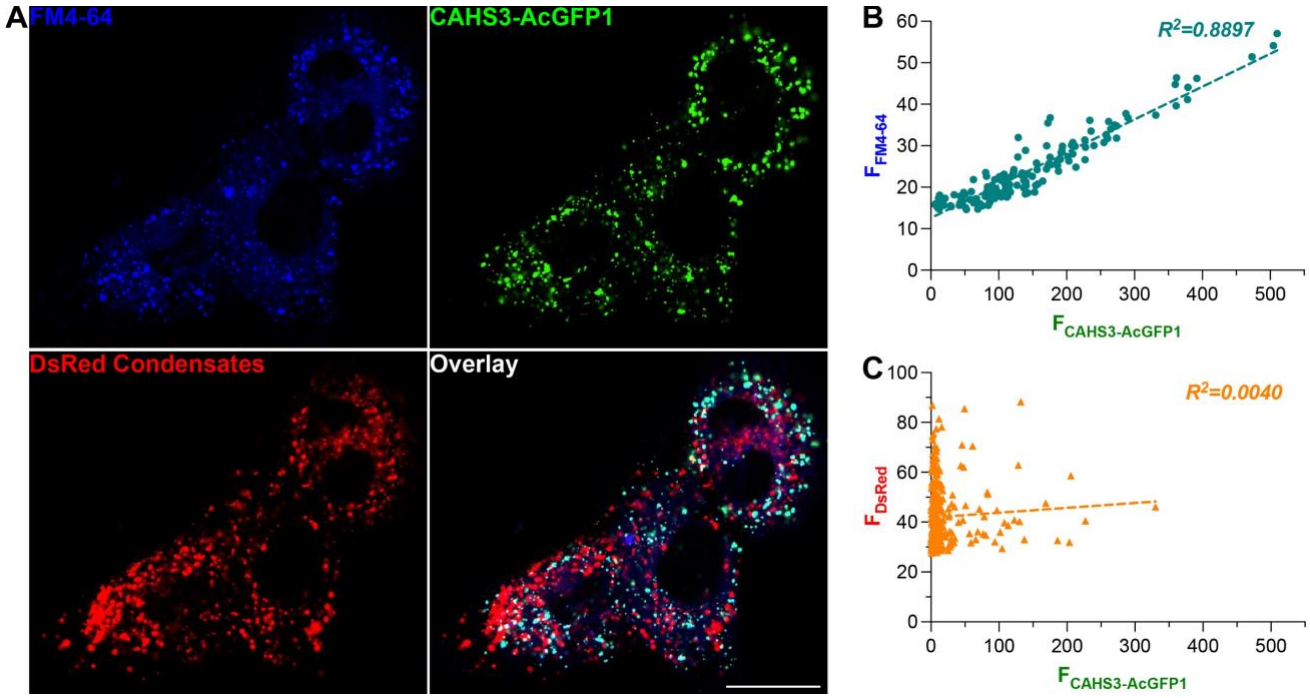

Fig 4. During dehydration, CAHS is mostly associated with cell membranes rather than proteinaceous condensates.
(A) sample confocal fluorescence images of DsRed and CAHS3-AcGFP1-expressing cells pre-loaded with FM4-64 and air-dried at 25°C for 10 minutes. Scale bar, 20μm. (B) a scatter plot of FM4-64 and CAHS3-AcGFP1 fluorescence intensities within the puncta defined by FM4-64 (i.e., membrane-bound organelles). The two values are significantly correlated ($p < 0.0001$), i.e., CAHS3-AcGFP1 colocalized with the membranes during dehydration. (C) a scatter plot of DsRed and CAHS3-AcGFP1 fluorescence intensities within the puncta defined by DsRed (i.e., proteinaceous condensates). The two values are not correlated ($p = 0.3114$), i.e., CAHS3-AcGFP1 did not co-condense with DsRed during dehydration.

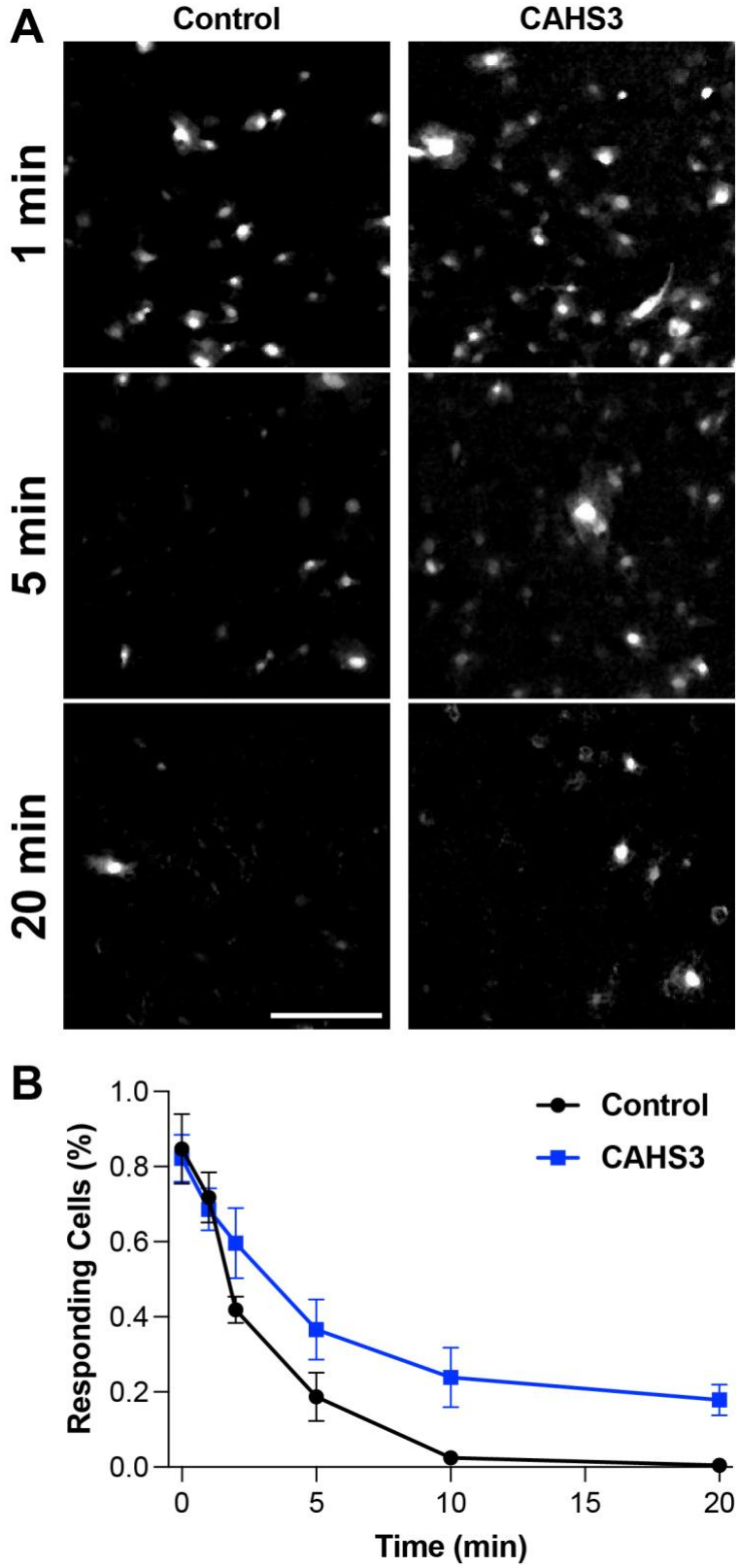

Fig 5. CAHS3 increased cell survival during dehydration.
(A) sample fluorescence images of control and CAHS3-expressing 3T3 cells still responding to 50μM ATP stimulation (visualized by X-Rhod-1AM, a red Ca^{2+} indicator) after 1, 5, and 20-min air-drying at 25°C. Scale bar, 100μm. (B) Percentage of responding cells (mean ± SEM) after air-drying at 25°C. The CAHS3-expressing group had more responding cells than the control group since 5-min drying (all p <0.05)

Supporting Information

Full title:

Tardigrades' cytoplasmic abundant heat soluble proteins serve as membrane protectors during dehydration.

Short title:

CAHS protects cell membranes during dehydration

Claire K. Zhang[1]*

[1]Brentwood High School, Brentwood, Tennessee, United States of America

* Corresponding author

E-mail: clairezhang168@gmail.com

[1]C. K. Zhang is with the Brentwood High School, Brentwood, TN 37027 USA (corresponding author to e-mail: clairezhang168@gmail.com).

I. Supplementary Figures

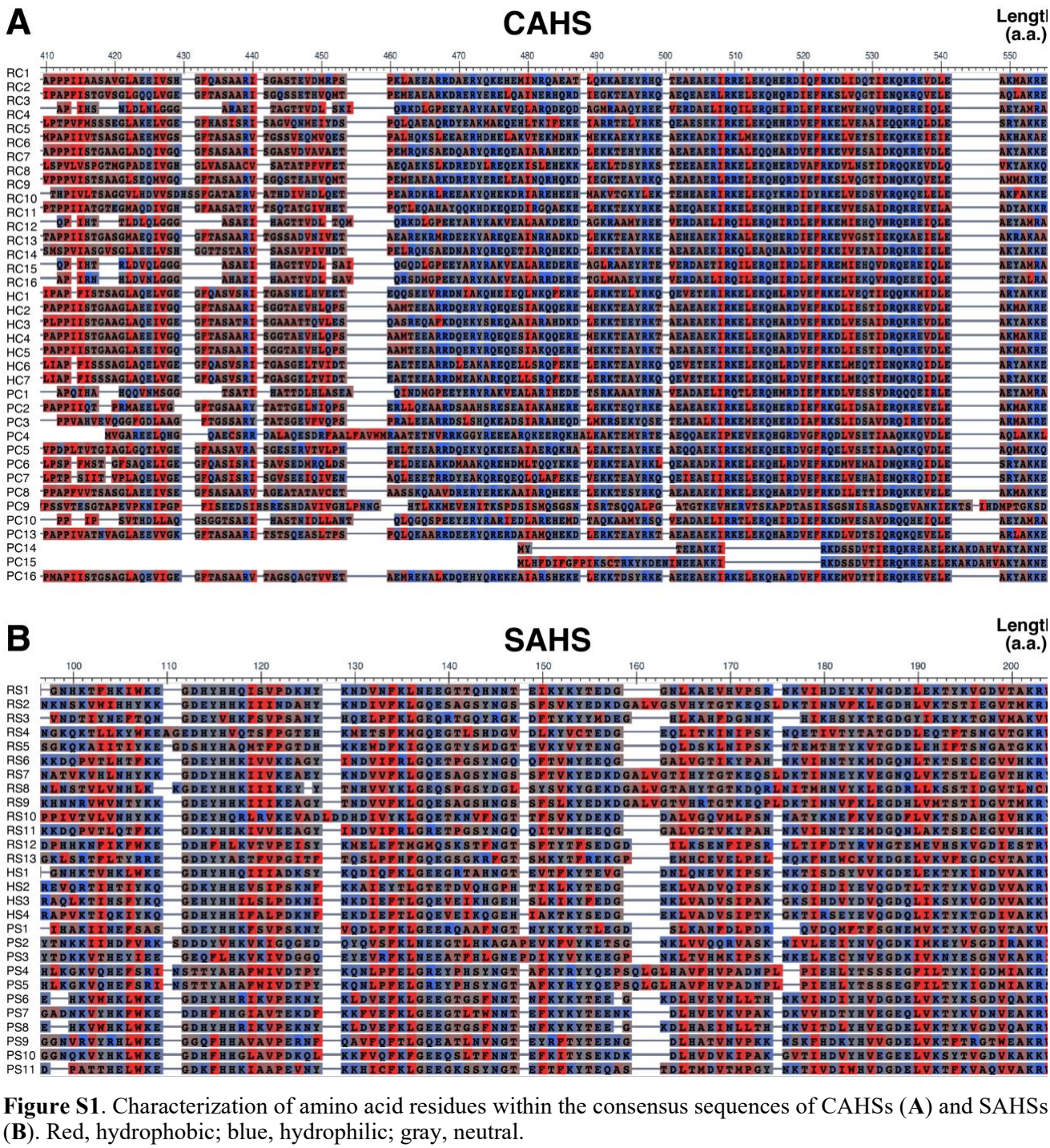

Figure S1. Characterization of amino acid residues within the consensus sequences of CAHSs (**A**) and SAHSs (**B**). Red, hydrophobic; blue, hydrophilic; gray, neutral.

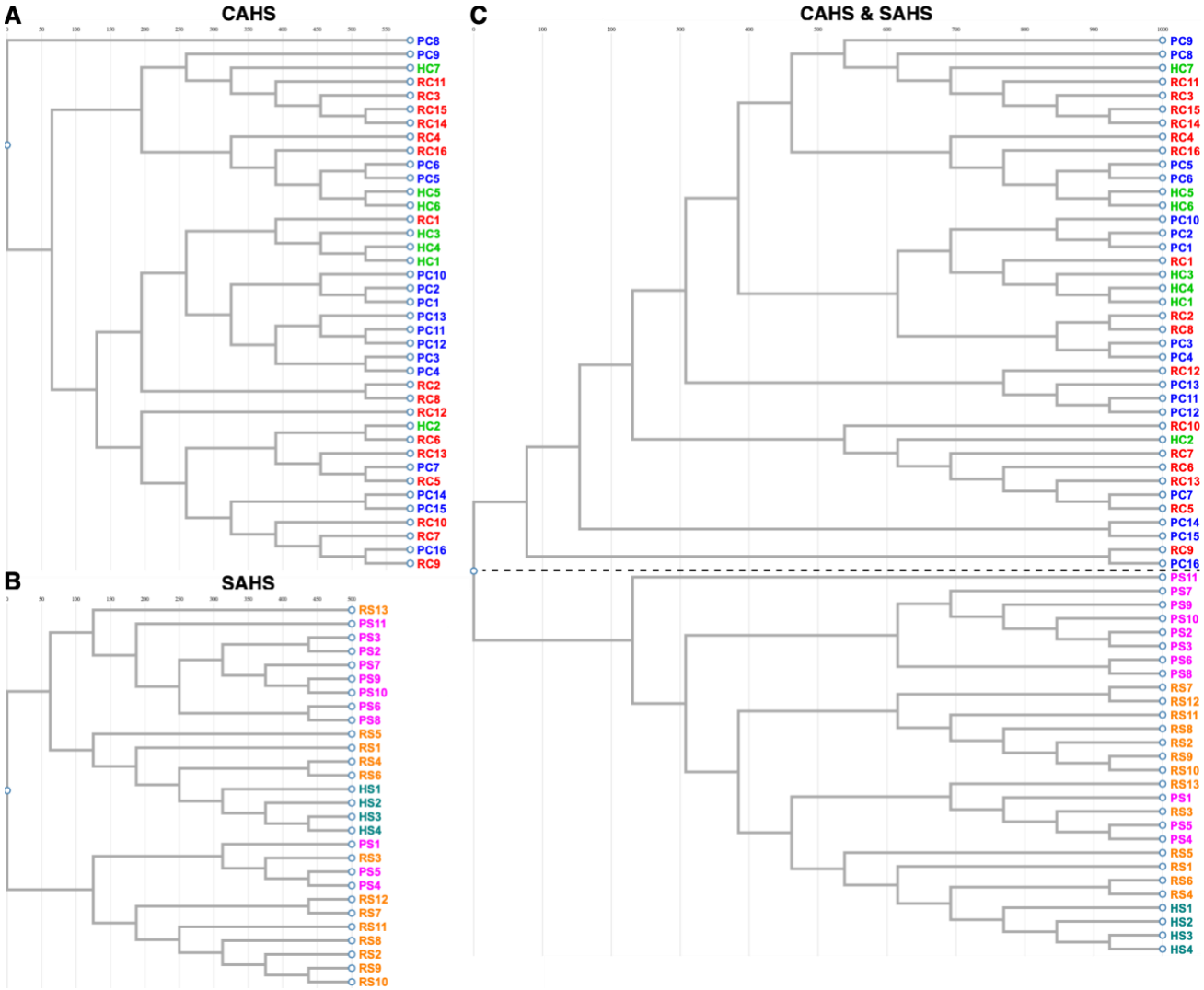

Figure S2. Phylogenetic trees of CAHSs (**A**), SAHSs (**B**), and both (**C**), generated by ETE3.

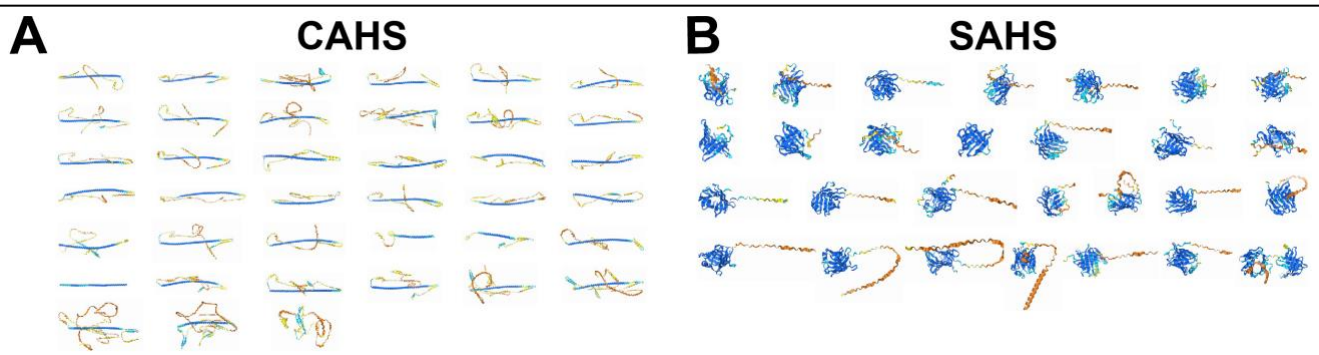

Figure S3. Structural prediction of 39 CAHSs and 28 SAHSs in monomer (**A** and **B**). The prediction confidence is measured as the predicted local distance difference test (plDDT). Dark blue indicates high confidence (i.e., plDDT > 90), blue to cyan indicates good confidence (i.e., 90 > plDDT > 70), yellow indicates low confidence (i.e., 70 > pTM > 50), and orange indicates very low confidence (i.e., pTM < 50).

II. Supplementary Tables

Table S1: All CAHSs used in this study

Name	GenBank ID	Abbreviation
Rv CAHS1	402783758	RC1
Rv CAHS2	402783760	RC2
Rv CAHS3	402783762	RC3
Rv CAHS4	1101393619	RC4
Rv CAHS5	1101393709	RC5
Rv CAHS6	1101393893	RC6
Rv CAHS7	1101396435	RC7
Rv CAHS8	1101389352	RC8
Rv CAHS9	1101387114	RC9
Rv CAHS10	1101380788	RC10
Rv CAHS11	1101379708	RC11
Rv CAHS12	1101379978	RC12
Rv CAHS13	1101378429	RC13
Rv CAHS14	1101375151	RC14
Rv CAHS15	1101375152	RC15
Rv CAHS16	1101372861	RC16
He CAHS 94063	1209747827	HC1
He CAHS 86272	1209747819	HC2
He CAHS 94205 (1)	1209747816	HC3
He CAHS 94205 (2)	1209747825	HC4
He CAHS 77611	1209747824	HC5
He CAHS 77580	1209747822	HC6
He CAHS 89226	1209747829	HC7
Pm CAHS 94205-like (1)	2493494936	PC1
Pm CAHS 94205-like (2)	2493480804	PC2
Pm CAHS 2-like	2493480067	PC3
Pm CAHS 94205-like (3)	2493479718	PC4
Pm CAHS 107838-like (1)	2493469471	PC5
Pm CAHS 107838-like (2)	2493468852	PC6
Pm CAHS 86272-like	2493445383	PC7
Pm CAHS 89226-like	2493438951	PC8
Pm CAHS 89226-like (2)	2493438943	PC9
Pm CAHS 94205-like (4)	2493428999	PC10
Pm CAHS 86272-like (X2)	2493423542	PC11
Pm CAHS 86272-like (X1)	2493423539	PC12
Pm CAHS 86272-like (2)	2493413587	PC13
Pm CAHS 86272-like (3)	2493398612	PC14
Pm CAHS 86272-like (4)	2493398609	PC15
Pm CAHS 86272-like (5)	2493397557	PC16

Table S2: All SAHSs used in this study

Name	GenBank ID	Abbreviation
Rv SAHS1	402783754	RS1
Rv SAHS2	402783756	RS2
Rv SAHS3	1101394461	RS3
Rv SAHS4	1101395371	RS4
Rv SAHS5	1101395434	RS12
Rv SAHS6	1101395435	RS5
Rv SAHS7	1101395438	RS6
Rv SAHS8	1101395445	RS7
Rv SAHS9	1101395447	RS8
Rv SAHS10	1101395452	RS9
Rv SAHS11	1101395453	RS10
Rv SAHS12	1101395666	RS11
Rv SAHS13	1101389123	RS13
He SAHS 63681	1210274786	HS1
He SAHS 53582	1210274784	HS2
He SAHS 33020	1210274782	HS3
He SAHS 68234	1210274787	HS4
Pm SAHS 1-like (1)	2493482411	PS1
Pm SAHS 1-like (2)	2493456970	PS2
Pm SAHS 1-like (3)	2493420974	PS3
Pm SAHS 1-like X2	2493398470	PS4
Pm SAHS 1-like X1	2493398467	PS5
Pm SAHS 1-like (4)	2493395038	PS6
Pm SAHS 1-like (5)	2493392926	PS7
Pm SAHS 1-like (6)	2493392853	PS8
Pm SAHS 1-like (7)	2493395084	PS9
Pm SAHS 646811-like (1)	2493394257	PS10
Pm SAHS 646811-like (2)	2493392930	PS11

III. Supplementary Movies

1. Ca^{2+}-imaging of Controls after 1, 5, and 20-minute airdry (3 AVI files)
2. Ca^{2+}-imaging of CAHS3-expressing cultures after 1, 5, and 20-minute airdry (3 AVI files)
3. 3D view of reconstructed confocal images showing CAHS3-AcGFP1 (green), DsRed (red), and FM4-64 (blue) in 3T3 cells after 5-minute airdry (3 AVI files)

Tardigrades' cytoplasmic abundant heat soluble proteins serve as membrane protectors during dehydration.

Claire K. Zhang[1]

Figures and Legends

[1]C. K. Zhang is with the Brentwood High School, Brentwood, TN 37027 USA (corresponding author to e-mail: clairezhang168@gmail.com).

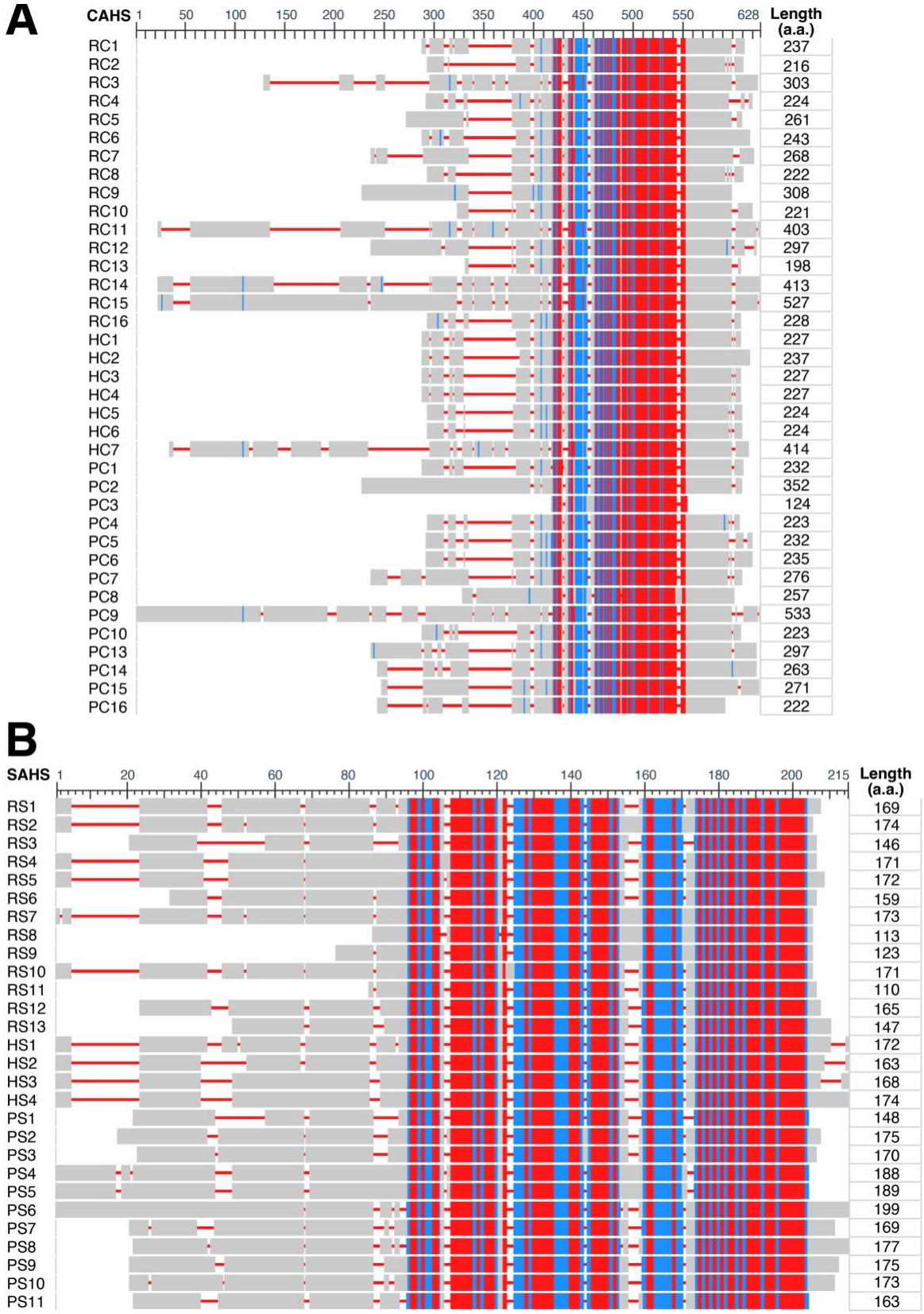

Fig 1. Protein sequence alignment identifies highly conserved central motifs and highly variable terminal regions in CAHSs and SAHSs.
Alignments of CAHSs (A) and SAHSs (B) were generated by COBALT. Red represents highly conserved amino acid residues, blue for less conserved ones, and gray for highly variable ones. The conserved central motif for CAHSs is about 120 amino acids long, and SAHSs' is about 100 amino acids long. RC, Rv CAHS; HC, He CAHS; PC, Pm CAHS. They are all numbered in the same order as that of Figure S1 and Table S1 & S2.

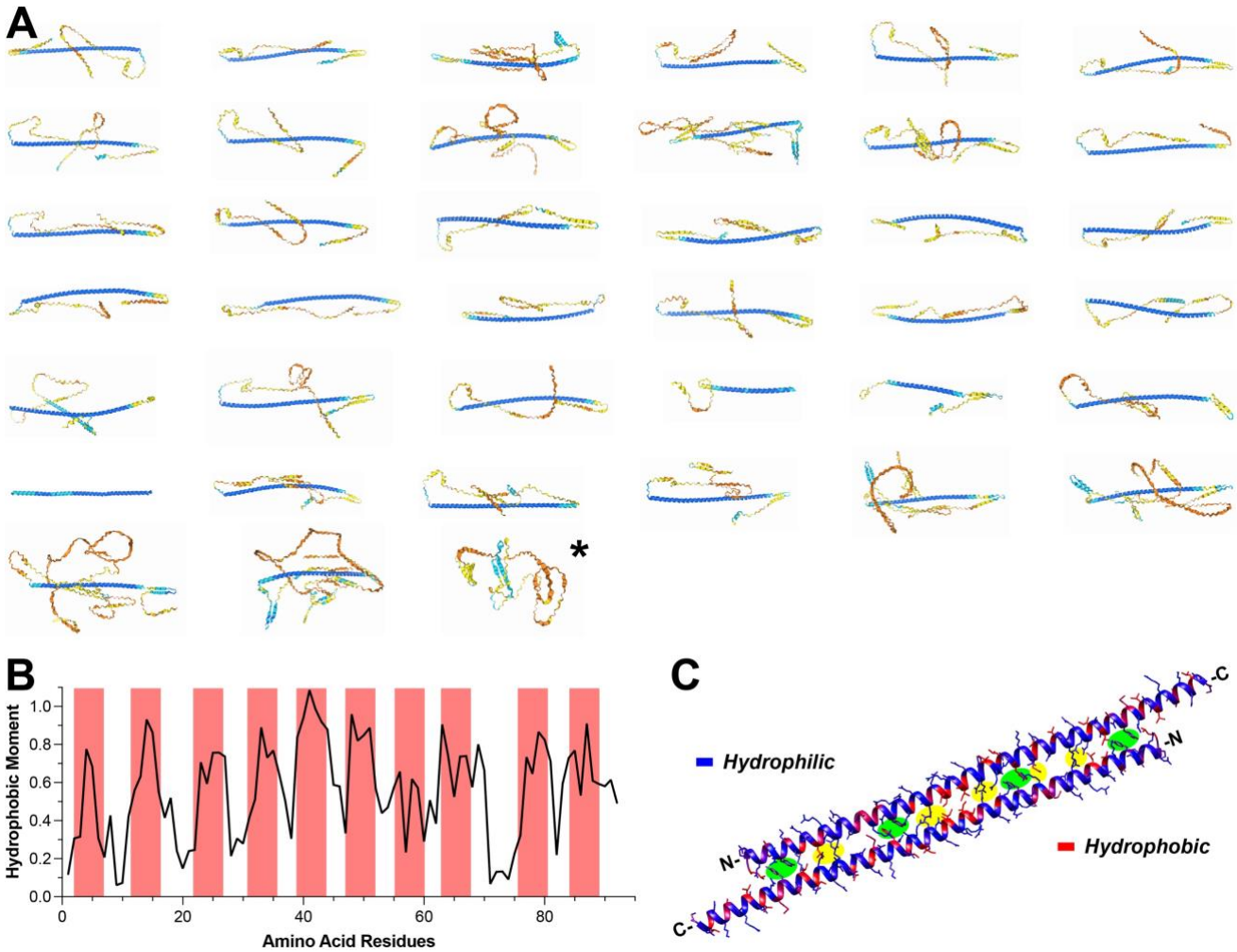

Fig 2. CAHS's central motif forms long helix.
(A) the 3D structures of CAHS monomers predicted by AlphaFold. Blue indicates high confidence for the highly ordered central helix whereas yellow/orange indicates low confidence for the disordered terminals. (*, PmCAHS89226-like). (B) the hydrophobic moment plot of RvCAHS3's central motif shows hydrophobic residues are arranged in a periodic manner. (C) two central motifs of RvCAHS3 form an anti-parallel dimer via electrostatic interactions (highlighted in green) and π-π stacking (highlighted in yellow). Most hydrophobic residues face outside.

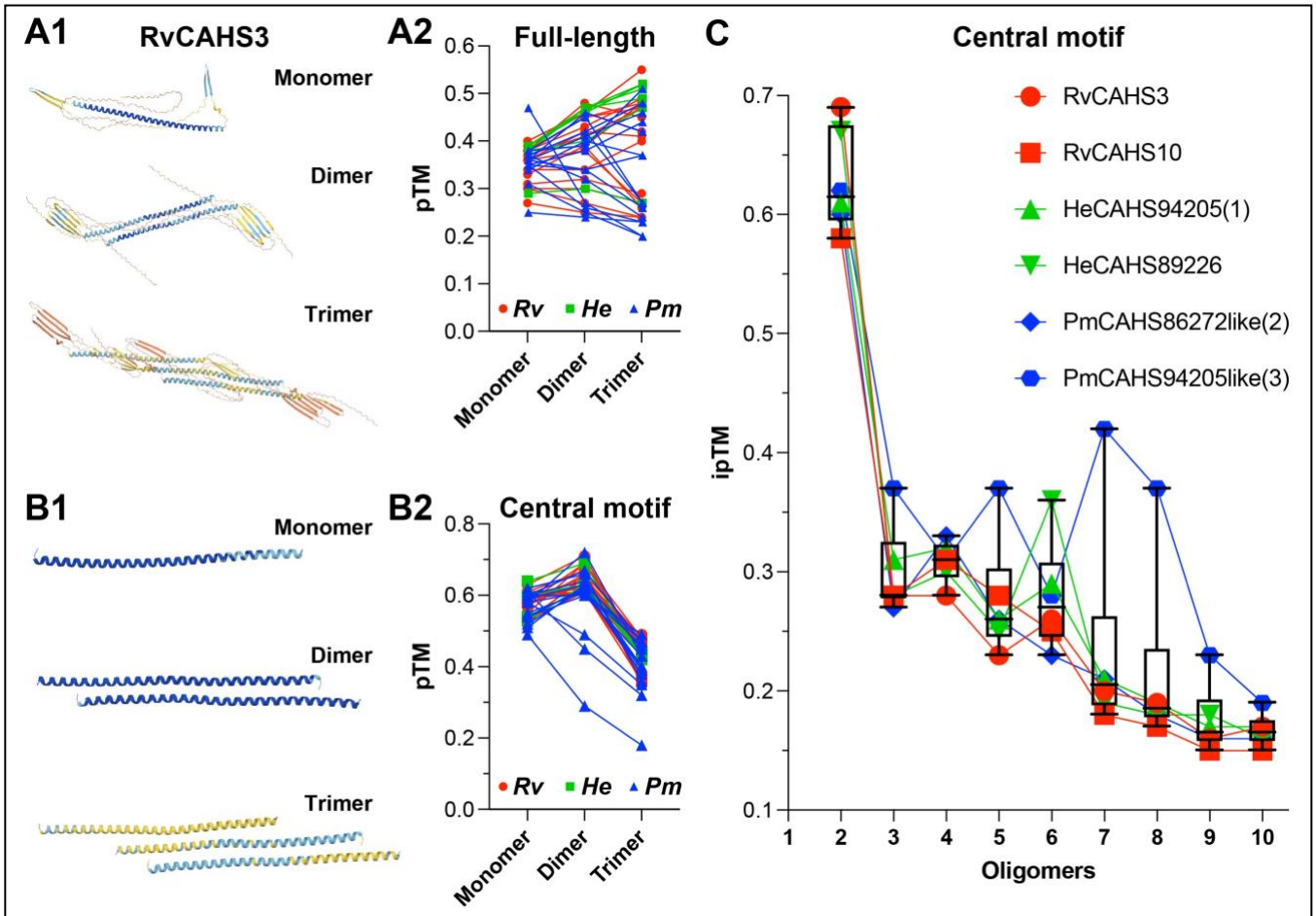

Fig 3. A helical dimer is the most likely form during CAHS oligomerization.

(A) predicted structures of full-length RvCAHS3 in mono-, di-, and trimer forms (A1) and the prediction confidence (measured as pTM) for the mono-, di-, and trimers of all 39 full-length CAHSs (A2), which lacks an overall trend of changes. (B) predicted structures of RvCAHS3 central motif in mono-, di-, and trimer forms (B1) and the prediction confidence (measured as pTM scores) of the central motifs of all 39 CAHSs (B2), which shows an overall increase for dimers and a drastic decrease for timers. (C) the box and whisker plot combined with point-line plot shows the trend of prediction confidence (measured as ipTM) during the oligomerization (from dimers to decamers) of the central motifs of six representative CAHSs. There is a progressive decrease from dimers to decamers.

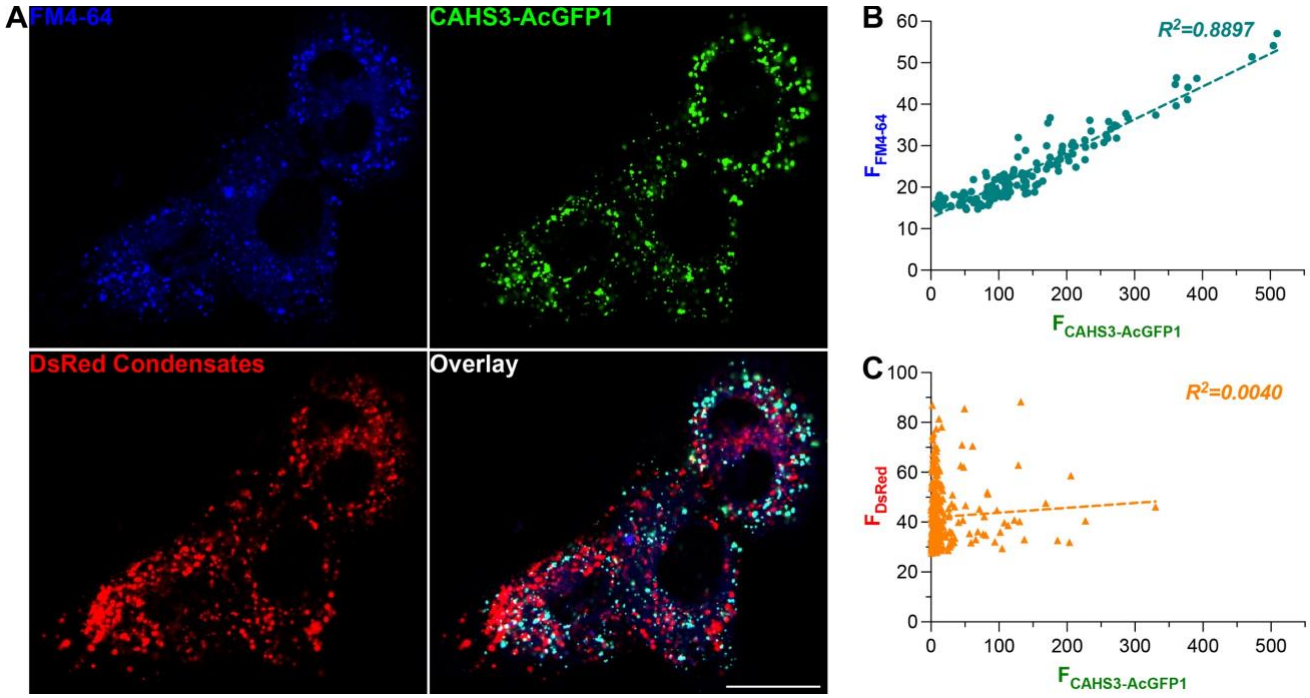

Fig 4. During dehydration, CAHS is mostly associated with cell membranes rather than proteinaceous condensates.
(A) sample confocal fluorescence images of DsRed and CAHS3-AcGFP1-expressing cells pre-loaded with FM4-64 and air-dried at 25°C for 10 minutes. Scale bar, 20μm. (B) a scatter plot of FM4-64 and CAHS3-AcGFP1 fluorescence intensities within the puncta defined by FM4-64 (i.e., membrane-bound organelles). The two values are significantly correlated ($p < 0.0001$), i.e., CAHS3-AcGFP1 colocalized with the membranes during dehydration. (C) a scatter plot of DsRed and CAHS3-AcGFP1 fluorescence intensities within the puncta defined by DsRed (i.e., proteinaceous condensates). The two values are not correlated ($p = 0.3114$), i.e., CAHS3-AcGFP1 did not co-condense with DsRed during dehydration.

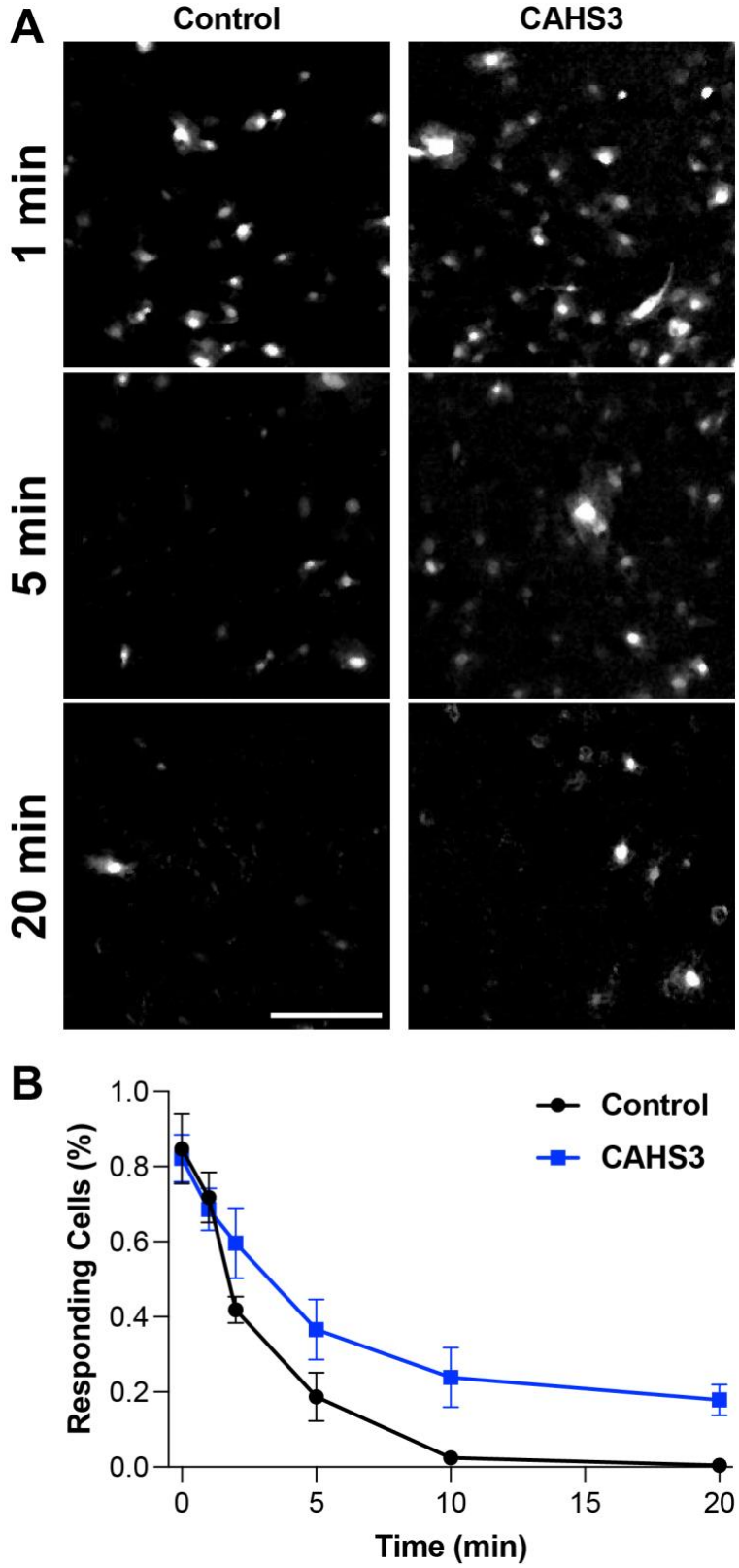

Fig 5. CAHS3 increased cell survival during dehydration.
(A) sample fluorescence images of control and CAHS3-expressing 3T3 cells still responding to 50μM ATP stimulation (visualized by X-Rhod-1AM, a red Ca^{2+} indicator) after 1, 5, and 20-min air-drying at 25°C. Scale bar, 100μm. (B) Percentage of responding cells (mean ± SEM) after air-drying at 25°C. The CAHS3-expressing group had more responding cells than the control group since 5-min drying (all p <0.05)