Natural and designed proteins inspired by extremotolerant organisms can form condensates and attenuate apoptosis in human cells

Mike T. Veling[1,2,8], Dan T. Nguyen[1,2,8], Nicole N. Thadani[1], Michela E. Oster[1,2], Nathan J. Rollins[1,2], Kelly P. Brock[1], Neville P. Bethel[3,4], Samuel Lim[1,2], David Baker[3,4,5], Jeffrey C. Way[1,2], Debora S. Marks[1,6], Roger L. Chang[1,2,7,*], Pamela A. Silver[1,2,*]

Affiliations

1. Department of Systems Biology, Harvard Medical School, Boston, MA, 02115, USA;
2. Wyss Institute for Biologically Inspired Engineering, Boston, MA, 02115, USA;
3. Institute for Protein Design, University of Washington, Seattle, WA, 98195, USA;
4. Department of Biochemistry, University of Washington, Seattle, WA, 98195, USA;
5. Howard Hughes Medical Institute, University of Washington, Seattle, WA, 98195, USA;
6. Broad Institute of Harvard and MIT, Cambridge, MA, 02142, USA;
7. Department of Systems & Computational Biology, Albert Einstein College of Medicine, Bronx, NY, 10461
8. Co-first authors

*. Co-corresponding authors (pamela_silver@hms.harvard.edu, roger.chang@einsteinmed.edu)

Key words

Stress tolerance, Phase separation, Intrinsically disordered proteins, Apoptosis

ABSTRACT

Many organisms can survive extreme conditions and successfully recover to normal life. This extremotolerant behavior has been attributed in part to repetitive, amphipathic, and intrinsically disordered proteins that are upregulated in the protected state. Here, we assemble a library of approximately 300 naturally-occurring and designed extremotolerance-associated proteins to assess their ability to protect human cells from chemically-induced apoptosis. We show that proteins from tardigrades, nematodes, and the Chinese giant salamander are apoptosis protective. Notably, we identify a region of the human ApoE protein with similarity to extremotolerance-associated proteins that also protects against apoptosis. This region mirrors the phase separation behavior seen with such proteins, like the tardigrade protein CAHS2. Moreover, we identify a synthetic protein, DHR81, that shares this combination of elevated phase separation propensity and apoptosis protection. Finally, we demonstrate that driving protective proteins into the condensate state increases apoptosis protection, and highlight the ability for DHR81 condensates to sequester caspase-7. Taken together, this work draws a link between extremotolerance-associated proteins, condensate formation, and designing human cellular protection.

INTRODUCTION

When exposed to extreme environmental stress, extremotolerance-associated (*ExTol*) organisms enter a state of reduced cellular metabolic activity that promotes survival in harsh conditions. This state has been shown to protect these organisms from high salinity, desiccation, and even the vacuum of outer space [1]. To acclimate to these conditions,

47 *ExTol* organisms significantly upregulate the production of protective metabolites, like
48 trehalose [2], as well as tolerance-conferring proteins [3], many of which are intrinsically
49 disordered proteins (*IDPs*) with repetitive sequence and predicted amphipathic helices.
50
51 While the mechanism of protection is unknown, one model suggests that at least some
52 of these proteins mediate formation of a gel- or glass-like material state within an entire
53 cell to protect it from damage [3]. The model posits that such proteins promiscuously
54 interact with membranes and other proteins, increasing intracellular viscosity and
55 protecting vulnerable structures [4-6] (Figure 1A Top). Such interactions would also work to
56 prevent stress-induced protein unfolding and protein aggregation [7]. This vitrification
57 hypothesis is the subject of active investigation at the time of this writing, with evidence
58 both for and against its validity [8-10]. An alternative model suggests that these proteins
59 form intracellular phase-separated membraneless compartments with distinct physical
60 parameters (referred to here as *condensates*) that partition key stress-sensitive cellular
61 components [11] (Figure 1A Bottom). Within these dense protein structures, essential
62 components are sequestered and protected from damage. The models are not mutually
63 exclusive; for example, in the case of desiccation tolerance, continued dehydration could
64 lead phase-separated compartments to fuse, comprise much (if not all) of the cell, and
65 promote a glassy material state.

66

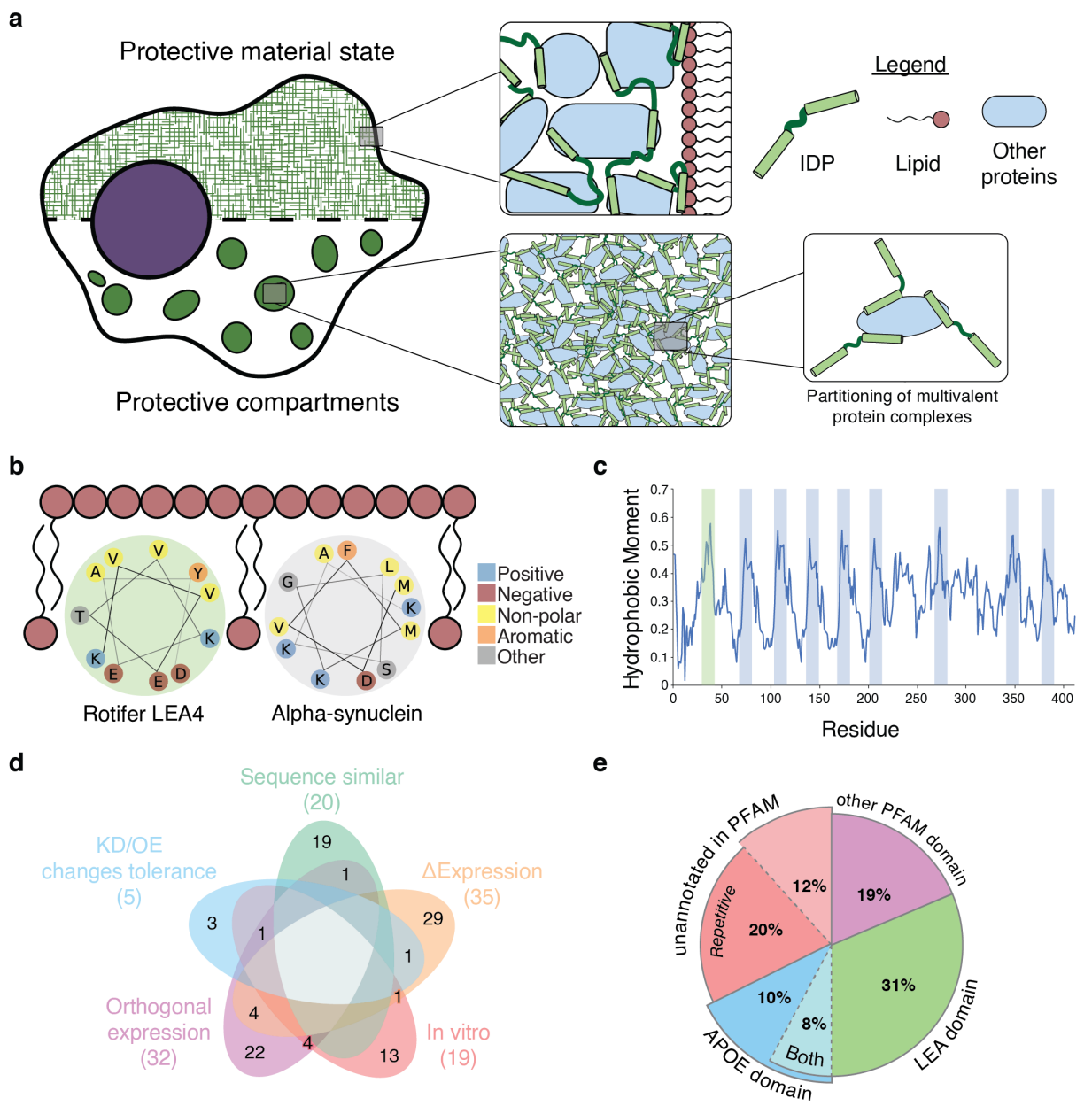


Figure 1: Selection of candidate extremotolerance-associated proteins (a) Illustration of different protective models for cellular stress tolerance. **(b)** Illustration of amphipathic membrane interaction for rotifer LEA4 protein and Alpha-synuclein, thus highlighting a feature indicative of extremotolerance-associated proteins. **(c)** Hydrophobic moment analysis of rotifer LEA4, showing the repetitive nature of the amphipathic signature. Highlighted green segment corresponds to the LEA4 helix in (b). **(d)** Venn diagram of the lines of evidence used to select targets in the final library. Abbreviations: KD, knockdown; OE, overexpression; Δ, change in expression levels. **(e)** Domain analysis of included targets reveals LEA and ApoE domains in almost half of the targets tested. The other half contains proteins with other PFAM domains (19%) or without any affiliated PFAM entry (32%). Of the proteins lacking a known PFAM annotation, many exhibit repetitive sequence (20% of all targets).

67

68 The resurgence of interest in biological phase separation, extensively reviewed in [12-15],
69 has shown that human cells can employ similar strategies to survive stressful conditions.
70 While the reversible formation of stress granules in healthy cells is the archetype of stress
71 tolerance, recent studies have revealed phase separation as a key adaptation that cancer
72 cells use to avoid apoptosis and cell death [16-17]. Phase-separated condensates have been
73 implicated in sequestering key tumor repressors[17], enhancing or repressing gene
74 expression in response to stress [18], and regulating mRNA or protein function through
75 sequestration [19]. These processes parallel the mechanisms that ExTol proteins are
76 thought to use to protect complex biological systems, suggesting there could be some
77 functional overlap. That is, sequence features and properties within ExTol proteins could
78 be implicated in mammalian stress tolerance pathways, including apoptosis resistance.

79

80 A key sequence feature of some ExTol proteins is the presence of amphipathic helices
81 [20]. While disordered in most contexts, ExTol proteins can transiently form amphipathic
82 helices in certain molecular environments, notably when interacting with and stabilizing
83 membranes [6]. The helices often exhibit a hydrophobic face, which interacts with core
84 lipids within the membrane, flanked by positively charged residues that might interact with
85 the negatively charged lipid head groups (Figure 1B). This amphipathic moment often
86 repeats across the length of the protein (Figure 1C). This feature is a hallmark of the late
87 embryonic abundant (LEA) class of ExTol proteins [21]. Human membrane interacting
88 proteins, like alpha-synuclein, also exhibit a similar repeat motif [22]. Such parallels in

89 sequence features suggest that orthogonal expression of ExTol proteins may confer their
90 protective properties to other biological systems, like human cells.
91
92 Indeed, previous reports have shown that some of these ExTol proteins provide similar
93 protective properties when expressed heterologously. Expressing tardigrade-specific
94 IDPs (TDPs) in bacteria provides protection against desiccation [3]. Anhydrin, an IDP
95 upregulated by *Aphelenchus avenae* during dehydration, limits PABPN1-A17
96 aggregation, a hallmark of oculopharyngeal muscular dystrophy, when expressed in a
97 human cell line [7]. Similar reports demonstrate protection of human cells provided by
98 expression of LEA proteins originating from the African midge [23] and by TDPs from
99 tardigrades [24]. These findings reinforce the promise of transferring protective phenotypes
100 from their native contexts to human cells.
101
102 Taken together, these data indicate a potential avenue for using this class of proteins to
103 protect biological systems in a controlled manner. To test this hypothesis, we generated
104 approximately 300 sequences inspired by ExTol proteins and examined their ability to
105 protect against apoptotic cell death. Our results show that human proteins containing
106 sequence hallmarks of ExTol proteins as well as *de novo* designed condensate forming
107 proteins can protect human cells from chemically-induced apoptosis. We also
108 demonstrate that condensate induction using synthetic modifications can strengthen
109 apoptotic protection. Finally, we show that key components of the apoptotic pathway can
110 localize to inducible condensates, suggesting enzyme sequestration could be a
111 mechanism of action.

112

113 **RESULTS**

114 **Assembling a library of stress tolerance proteins**

115 Based on the hallmark sequence features of ExTol proteins, we searched the published
116 literature to identify proteins that not only contain these features but also exhibit some
117 evidence for protection of biological structures. The library was generated based on over
118 one hundred naturally-occurring proteins based on five lines of evidence for stress
119 tolerance association, including: change in expression during extremotolerance;
120 knockdown or overexpression changing the tolerant state; *in vitro* protection; orthogonal
121 expression transferring tolerance; and sequence similarity to other proteins with these
122 features (Figure 1D and Supplemental table 1). We also examined our sequences for the
123 presence of repetitive elements and intrinsic disorder (Supplemental Figure 1).

124

125 A PFAM domain analysis of the identified proteins reveals almost half of the identified
126 targets have two key domains: Late Embryonic Abundant (LEA) and Apolipoprotein E
127 (ApoE) domains (Figure 1E, Supplemental table 1). LEA proteins are well known to be
128 upregulated during the last stages of plant seed maturation [25]. This domain is thought to
129 protect the seed embryo during transport to its new destination. There is also evidence
130 that this domain has been co-opted by bacteria as protection from desiccation and
131 radiation [26]. ApoE, in contrast, is not commonly discussed in the context of
132 extremotolerance. ApoE proteins interact with lipid particles to allow for lipid reuptake
133 into cells. Mutations within human ApoE have been identified as strong genetic predictors

of Alzheimer's disease in human populations [27]. While seemingly unrelated to other proteins in our study, we chose to include human ApoE and its various mutations as ExTol proteins due to the prominence of the ApoE domain in the PFAM analysis and since they exhibit similar features as other selected proteins (e.g. their lipid interaction properties and phase separation tendencies [28]).

Designing novel extremotolerance-associated proteins

To test hypotheses about sequence features and their relationship to phase separation and apoptosis, we designed novel proteins based on the set of ExTol proteins. These new designs take 3 approaches (Figure 2): truncation of proteins to focus on repetitive, intrinsically disordered, and amphipathic regions; isolation of individual helices with desired sequence features; and *de novo* design of alpha helix-containing proteins that self-assemble into higher order structures. Through these designs, we sought to identify the minimal

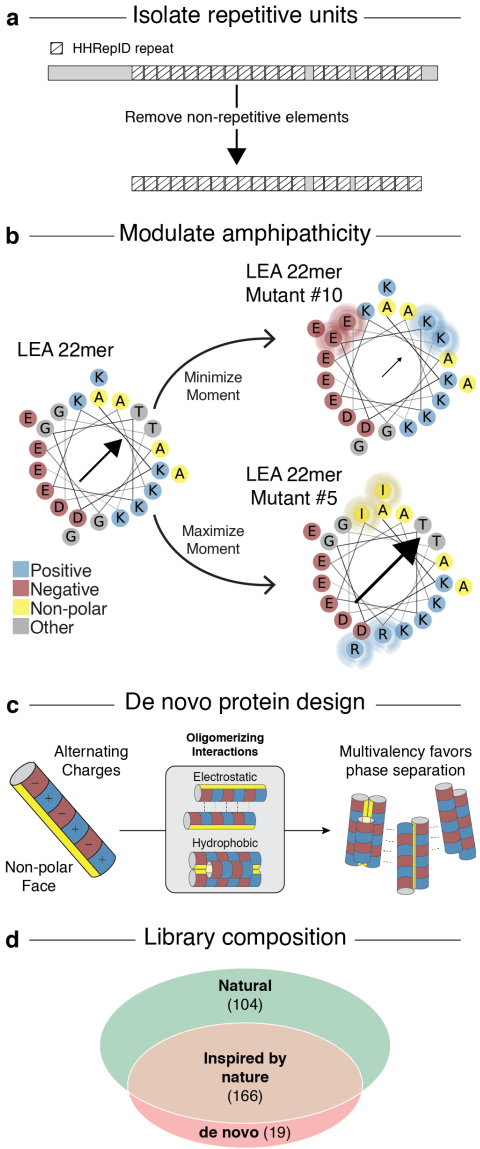


Figure 2: Design based on candidate extremotolerance-associated proteins (a) Truncations were made to isolate regions of repetitive and amphipathic sequence. **(b)** Small helices were selected and altered based on their hydrophobic moment. **(c)** Synthetic alpha helix-containing proteins designed to oligomerize. **(d)** Venn diagram highlighting the library that was tested.

156 elements required for protection and phase separation.

157

158 First, we identified regions within natural proteins that contain intrinsic disorder [29],
159 repetitive sequences [30], and amphipathic helices [31]. Figure 2A illustrates the use of
160 HHRepID on the sequence of full length human ApoE. Based on this analysis, we
161 identified the repeating domain within ApoE for further testing. Other ExTol proteins were
162 similarly modified and examined.

163

164 We also included short peptides based on the consensus sequence of repetitive regions
165 that have previously been shown to be protective [32]. To test if any of the sequence
166 features we identified are important for apoptosis protection, we altered these features
167 within these short peptides. For example, we altered the amphipathic character of these
168 helices (Figure 2B) to determine if apoptosis protection was changed.

169

170 To examine whether generic condensate formation is sufficient to confer apoptosis
171 protection, we also examined completely synthetic alpha helix-containing proteins that
172 were designed to form repetitive structure with sequences orthogonal to naturally-
173 occurring proteins [33]. These designed helical repeat (DHR) proteins interact along the
174 axis of the helix, forming long spiraling fibrils, but may also potentially self-interact along
175 the face of the protein based on a pattern of positive and negative charges (Figure 2C).
176 This ability to oligomerize and form multivalent complexes, which may then interact via
177 weak non-covalent interactions, is an established determinant of condensate formation
178 [15].

179

180 Together, our library consists of 104 naturally-occurring proteins, 166 truncations or
181 isolations of helical regions, and 19 *de novo* designs (Figure 2D). These proteins
182 interrogate several hypotheses about how disordered proteins may protect human cells
183 against apoptosis.

184

185 **Apoptosis assay identifies extremotolerance-associated proteins that protect**
186 **human cells**

187 To model stress in a mammalian cell system, we induced cellular apoptosis by treatment
188 with the chemotherapeutic camptothecin (CPT), a DNA topoisomerase inhibitor well
189 known to trigger apoptosis [34-35]. Under these conditions, protection against apoptosis
190 could be achieved through a number of avenues, including protection against the DNA
191 damage caused by CPT-induced topoisomerase inhibition [34], interruption of the signaling
192 cascade, or even protection of the mitochondrial membrane whose rupture ultimately
193 triggers apoptosis [36].

194

195 To test our designs for protection against apoptosis, we used a human cell-based
196 apoptosis assay (Figure 3A). This assay uses HT-1080 fibrosarcoma cells from ATCC [37].
197 These cells are seeded in a 96-well plate and allowed to adhere overnight. The next day,
198 the cells are co-transfected with a plasmid encoding the ExTol candidate proteins of
199 interest as well as a transfection marker. The following day, the media is swapped to one
200 containing 13.2 μM CPT. After addition of CPT, caspase-3 and -7 activation is measured
201 by microscopy every 45 minutes for 1 day using a fluorophore quencher pair attached to

202 either end of a peptide with the sequence “DEVD”. As caspase-3 and -7 are activated,
203 the DEVD peptide is cleaved, releasing the fluorophore from the quencher. The
204 unquenched fluorophore localizes to the nucleus where it can be quantified by imaging
205 on a single-cell level. The single-cell resolution allows for the separation of transfected
206 and non-transfected cells based on the transfection marker. Only single cells with the
207 transfection marker were subjected to further analysis; Figure 3B shows sample images
208 highlighting the type of data being analyzed.
209
210 On every plate, an empty vector negative control is run as well as a positive control
211 containing the X-linked inhibitor of apoptosis (XIAP). This protein is known to bind to
212 caspase-3 and -7 to block their activity [38] . Figure 3C shows example data comparing
213 median DEVD peptide signal intensity over time after induction with CPT of four wells on
214 the same plate transfected with XIAP and empty vector. The results show a reduction in
215 DEVD peptide signal intensity when the XIAP-containing vector is transfected,
216 demonstrating that the assay can identify modulators of apoptosis.
217
218 To convert these time-resolved data to a measure of protection against apoptosis, the
219 maximum DEVD peptide signal intensity is calculated and normalized to a plate-matched
220 empty vector control (see methods for further details). The normalized apoptosis induction
221 intensity is shown in Figure 3D left. The data are slightly skewed towards apoptosis-
222 protection, which supports our hypothesis that the targeted proteins can protect against
223 apoptosis.

224

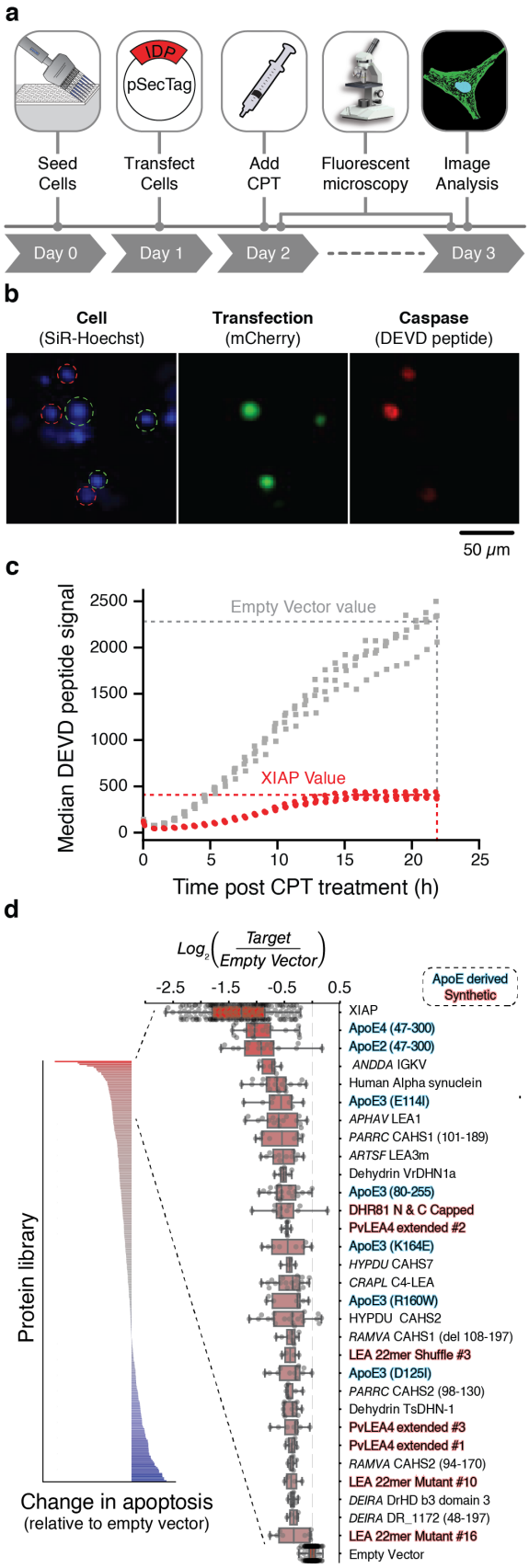


Figure 3: Screen for CPT resistance by extremotolerance-associated proteins (a) Flow diagram of the assay. **(b)** Representative images showing cells during the assay. Cells are identified using SiR-Hoechst marking the nucleus (*blue*). Transfected cells are identified by the presence of the mCherry marker (*green*). Finally, the degree of apoptosis is identified by the presence of the cleaved DEVD peptide (*red*). **(c)** Positive control showing apoptosis reduction upon expression of XIAP. **(d)** Bar graph highlighting the apoptosis resistance provided by different constructs, including zoomed in area to highlight the top hits in the assay. Source species of a sequence are abbreviated and shown in italics; abbreviations are drawn from Uniprot, including *RAMVA* (*Ramazzottius varieornatus*), *HYPDU* (*Hypsibius dujardini* or *Hypsibius exemplaris* historical naming), *PARRC* (Paramacrobiotus richtersi), *ANDDA* (*Andrias davidianus*), *APHAV* (*Aphelenchus avenae*), *ARTSF* (*Artemia franciscana*), *CRAPL* (*Craterostigma plantagineum*), and *DEIRA* (*Deinococcus radiodurans*).

225

226 To highlight the top-performing designs, the top 10% of the targets are shown in Figure
227 3D right. As expected, the XIAP positive control showed the strongest anti-apoptotic
228 effect. The two top-performing hits were ApoE designs. Human alpha-synuclein is the 5th
229 hit, which is expected as it is known to be both intrinsically disordered, and anti-apoptotic
230 [39]. Additionally, several ApoE designs and short LEA helices are top hits. The data used
231 to generate these figures is included in supplemental table 2.

232

233 **Known extremotolerance-associated proteins affect apoptosis**

234 Among the top hits in apoptosis reduction are an IGKV protein highly expressed in mucus
235 of the Chinese giant salamander *Andrias davidianus* [40] that is used as a medical glue in
236 traditional Chinese medicine, an LEA protein from the nematode *Aphelenchus avenae*
237 previously shown to have the ability to reduce a form of protein aggregation associated
238 with oculopharyngeal muscular dystrophy in human cells [7], and a mitochondrial LEA
239 protein from the brine shrimp *Artemia franciscana* shown to protect human cells during
240 desiccation [23] (Figure 3D). The prominence of these ExTol proteins suggests the assay
241 can help identify other proteins that can protect cells from more general stresses beyond
242 CPT-induced apoptosis.

243

244 **Short repetitive peptides and *de novo* designs modulate apoptosis**

245 LEA repeat-containing peptides are among the top-performing hits in the screen. While
246 the original short LEA designs from Furuki et al. 2019 had a minor effect reducing
247 apoptosis, several of the designs tested here performed better (Supplemental Figure 2A)

248 [32]. As 22 amino acids is small and could lead to peptide stability issues, longer 50-amino
249 acid designs inspired by the same LEA proteins utilized by Furuki et al. were also
250 developed (Supplemental Figure 2B). Some of these designs were also able to reduce
251 apoptosis, including the PvLEA4 with extended repeats (notably PvLEA4 repeats
252 extended #2), which fell into the top hits of the entire library (Figure 3D and Supplemental
253 Figure 2B). These results suggest that optimization of synthetic designs yields proteins
254 with stronger function than naturally-occurring proteins.
255
256 While these proteins were designed to interrogate the relationship between sequence
257 and apoptosis protection in short peptides, little relationship was found with respect to the
258 sequence features we analyzed (Supplemental Figure 2C). Several designs were made
259 that would alter the hydrophobic moments of putative helices these peptides are predicted
260 to form. To examine if this sequence property relates to the stress-tolerance function of
261 these proteins, we attempted to correlate this property with assay performance. However,
262 no significant correlation was found, suggesting (at least in these cases and experimental
263 conditions) that degree of amphipathicity on its own is not the key determinant of stress-
264 tolerance function of these proteins.
265
266 **The repeat region of ApoE comprises its minimal anti-apoptotic unit**
267 The top two hits in the assay were ApoE4 (47-300) and ApoE2 (47-300). These truncated
268 forms of ApoE maintain the repeat domains shown in Figure 4A, while removing the signal
269 peptide and terminal regions. This suggests that the repeat region of ApoE accounts for
270 its core anti-apoptotic functionality. Interestingly, the full length ApoE2 and E4 do not

271 show up as top hits, suggesting that these truncated forms improve on the apoptosis
272 protection provided by the full-length variants (Figure 4B).
273
274 Based on the location of the signal peptide, it is expected that ApoE4 should localize to
275 the secretory pathway whereas the truncated form may be cytoplasmic. To test this
276 hypothesis, cells were subjected to Western blot and immunofluorescence analysis using
277 commercially available antibodies (Abcam ab24139; Figure 4C and Supplemental Figure
278 3A). The Western blot results show that both the full-length and truncated forms of ApoE4
279 are expressed at the expected size (Supplemental Figure 3A). The immunofluorescence
280 shows co-localization between the ApoE4 full length protein and the endoplasmic
281 reticulum (marked by Thermo's PDI antibody MA3-019, Figure 4C). However, the
282 truncated form did not show the same co-localization, suggesting as expected that it may
283 not enter the secretory pathway as the full-length version does.
284
285 Strikingly, the truncated version appeared to form more pronounced puncta throughout
286 the cytoplasm of the cell, suggesting the formation of potentially phase-separated
287 condensates. In screening through the point mutants of ApoE also tested here, we found
288 a similar condensate phenotype with the full length ApoE3 D125I mutation (Supplemental
289 Figure 3B). Together, these data suggest that ApoE may be able to support condensate
290 formation and confer apoptosis protection in part via the protective compartment
291 mechanism (Figure 1A bottom).
292
293

294

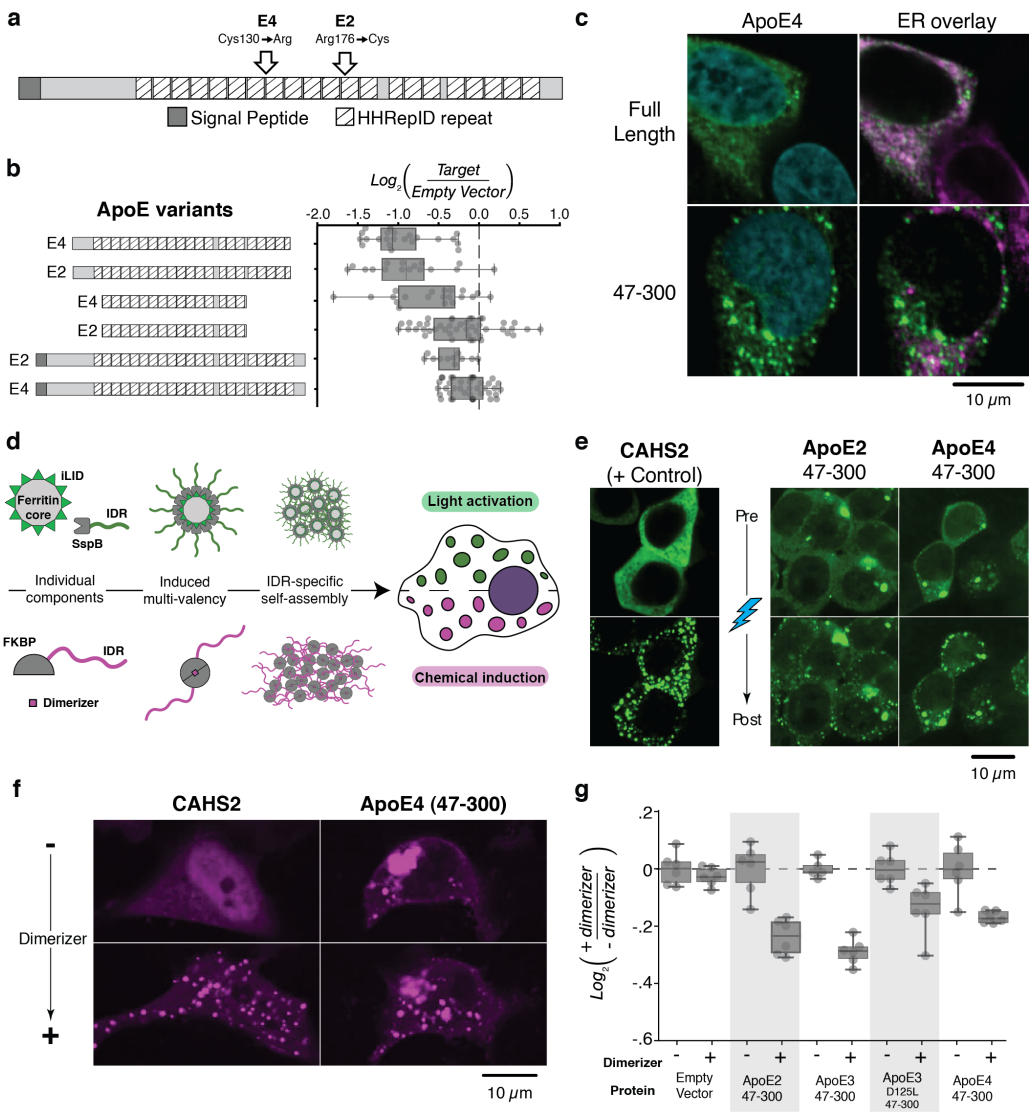


Fig. 4. ApoE truncations and synthetic condensates further reduce apoptosis.
(a) Gene diagram highlighting the repetitive region of ApoE. **(b)** Boxplots highlighting the apoptosis assay performance of the full length and truncated forms of ApoE2 and E4. **(c)** Immunofluorescence of ApoE4, full length and truncated forms. *White* indicates colocalization, which is more apparent with the full-length protein. The truncated form appears to exhibit larger and more distinguishable puncta. **(d)** Diagram of the induced multi-valency systems implemented here, including the light-inducible system (*top*) and the chemically-inducible system (*bottom*). For the chemically induced system, we used two FKBP domains. **(e)** Condensate formation assay using the light-inducible system, including the *PARRC* CAHS2 positive control showing condensates only upon light induction and several ApoE examples that form condensates both before and after induction. **(f)** Condensate formation assay using the chemically-inducible system, including the *PARRC* CAHS2 positive control showing condensates only upon added dimerizer and ApoE4 (47-300) exhibiting condensates both before and after induction. **(g)** Apoptosis data showing reduction in apoptosis upon condensate induction with dimerizer. Data are normalized to mean apoptosis level of uninduced (i.e. no dimerizer) sample for each construct.

295 **Induced multivalency to probe the relation between condensates and apoptosis**

296 We made use of established condensate engineering tools [41] to test for latent phase
297 separation propensity among a selection of ExTol proteins representing the range of
298 apoptosis effects. We used both light- and chemically-inducible systems of multivalency.
299 In the light system, a ferritin core (comprised of self-assembled monomers, FTH1 domain-
300 eGFP-iLID) becomes decorated with SspB-modified proteins upon exposure to blue light
301 [42] (Figure 4D top). In the chemical system, multivalency (i.e. multiple arms of the ExTol
302 protein of interest) is induced using a dimerization agent [43] that brings together FKBP-
303 modified proteins [41] (Figure 4D bottom). These inducible methods increase the valency
304 of a system, and thus work to favor phase separation. In this way, they can help reveal
305 any latent ability to form condensates that is not immediately observable without
306 modification.

307

308 Condensates were indeed observed in a number of ExTol proteins (Supplemental Figure
309 4), including ApoE, *ARTSF* LEA3m, DHR81, *HYPDU* CAHS3, and *PARRC* CAHS2 using
310 the light-inducible system. For those proteins with no visible condensates, the
311 experimental conditions may be insufficient to support phase separation. *PARRC* CAHS2
312 exhibited the most dramatic phenotype, which is consistent with recent work describing
313 the gel forming capabilities of CAHS proteins [44-45] (Figure 4E&F). ApoE variants were of
314 interest, based on the puncta observed in the top performing ApoE constructs. All
315 versions of ApoE tested formed puncta without activation, consistent with our previous
316 observation using untagged native ApoE (Figure 4C). Upon light- or chemical-activation,

317 there is a noticeable increase in condensates, exemplifying how phase separation can be
318 enhanced with increased multivalency (Figure 4E&F).
319
320 To examine how enhanced phase separation may affect apoptosis, we performed the
321 apoptosis assay with ExTol proteins that were modified for chemically-induced
322 multivalency. Chemical induction was more compatible with the apoptosis assay, while
323 enabling the same increase in condensates seen with the light-induced system (Figure
324 4F). We observed that the presence of the chemical dimerizer had little to no effect on
325 apoptosis in cells transfected with the empty vector, with a construct encoding the
326 dimerization domains alone, or with constructs encoding an ExTol protein that was not
327 strongly anti-apoptotic (e.g. *PARRC* CAHS2) (Figure 4G; Supplemental Figure 5). In
328 contrast, with constructs encoding a strongly anti-apoptotic ExTol protein – including
329 ApoE, *ARTSF* LEA3m, and DHR81 – added dimerizer corresponded with reduced
330 apoptosis in transfected cells (Figure 4G; Supplemental Figure 5). These data suggest
331 that condensate formation plays a role in the anti-apoptotic effect of these proteins.
332
333 **Synthetic DHR81 condensates sequester caspase-7**
334 To begin elucidating the mechanism by which ExTol proteins protect against apoptosis,
335 particularly the hypothesis of membraneless compartments that sequester essential
336 effectors, we examined the ability for the various condensates to partition apoptotic
337 signaling molecules, identifying the synthetic protein DHR81 as the only ExTol protein
338 capable of doing so. DHR81 is a synthetic alpha helical-rich protein designed to promote
339 self-interaction. It is formed from four helical hairpin repeat regions; the core of the protein

is hydrophobic to ensure consistent folding, while its outside is decorated with checkerboard positive and negative charges (Figure 5A) [33]. This structure ensures the protein is well folded, while also enabling the potential for higher order structure formation via electrostatic interactions. Not only did DHR81 support condensates in the chemically-inducible condensate system (Figure 5B), it also exhibited a strong increase in apoptotic protection after induced-multivalency (Figure 5C; Supplemental Figure 5&6A).

As such, we chose DHR81, along with a few other ExTol proteins with notable phase separation phenotypes (i.e. ApoE4 and *PARRC* CAHS2), to further explore how their condensates may confer an anti-apoptotic effect. Specifically, we used immunofluorescence to examine whether they sequester major components of apoptotic signaling, namely caspase-3 and -7. The results showed that activated caspase-7 localized to DHR81 condensates but not *PARRC* CAHS2 or ApoE4 condensates (Figure 5D), while caspase-3

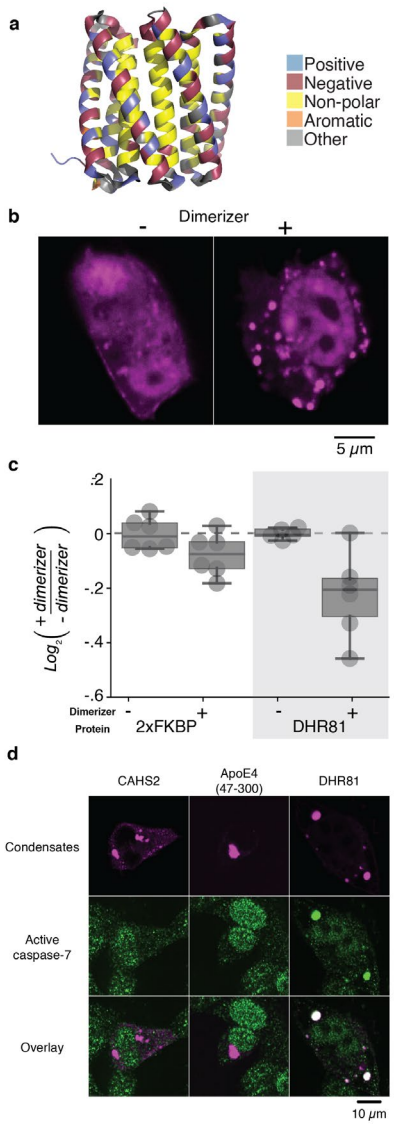


Fig. 5. DHR81 condensate formation and apoptosis reduction. (a) Diagram of DHR81, illustrating a charge distribution that can support oligomerization. **(b)** Sample images highlighting how the synthetic protein DHR81 supports chemically-induced condensates. **(c)** Apoptosis is further reduced in the presence of dimerizer (i.e. with induced DHR81 condensates). The FKBP-modifications alone, which enable chemically-induced multi-valency, have a minimal effect on apoptosis. **(d)** Activated caspase-7 (*green*) co-localizes with DHR81 condensates but not *PARRC* CAHS2 or ApoE4 (47-300) condensates (*magenta*). *White* indicates colocalization.

363 did not localize to any of the examined condensates (Supplemental Figure 6B). Even
364 though condensate properties may confound (i.e. physically occlude) antibody binding
365 during immunofluorescence, the positive DHR81 result suggests sequestration of key
366 apoptotic factors can indeed be a potential mechanism by which these condensates
367 reduce apoptosis.
368

369 **Discussion**

370 To broadly evaluate the ability of ExTol proteins to protect orthogonal systems from
371 environmental stress, we studied over one hundred different ExTol proteins. We used
372 these as a basis to design approximately 300 variants, including truncations and fusion
373 proteins. These designs were evaluated on their ability to protect human HT-1080 cells
374 against CPT-induced apoptosis. The results showed that known ExTol proteins, ApoE
375 variants, and the synthetic DHR81 protein could modulate apoptosis levels when
376 transferred to human cells.
377

378 Protection against apoptosis is a key feature of many ExTol proteins. As apoptosis is a
379 common cellular response to stress, the ability to mitigate this response is important in
380 providing cell-autonomous protection in stressful environments. While the mechanism
381 resilient organisms use to survive is still not fully understood, it is becoming increasingly
382 clear that phase separation may play an important role. Some ExTol proteins are
383 proposed to form expansive higher order structures (e.g. gels or glassy systems) that
384 reduce the viscosity of the cytoplasm, especially in desiccated states [3]. Other reports
385 show that these proteins could mediate formation of protective phase-separated

386 compartments (i.e. condensates) that sequester damage-inducing proteins and/or protect
387 important targets [11]. Based on our data, it appears the latter may be at play for proteins
388 like ApoE, *ARTSF* LEA3m, and DHR81, which exhibit anti-apoptotic effects while also
389 being condensate prone.
390
391 Condensates alone are not sufficient for apoptosis protection. While we observe
392 condensate formation in many of our top performing hits, we also observe condensate
393 formation in targets like *HYPDU* CAHS3 and *PARRC* CAHS2, which do not reduce
394 apoptosis very much. Additionally, some proteins like *PARRC* CAHS1 (101-189) are
395 truncated in ways predicted to reduce condensate formation[44, 46-47], but they still perform
396 well in apoptosis protection. These observations suggest yet-to-be-discovered
397 mechanisms are likely responsible for apoptosis protection in these cases. However, for
398 those proteins that exhibit both anti-apoptotic effects as well as elevated phase separation
399 propensities, increasing the amount of condensates via an engineered inducible system
400 led to stronger apoptosis protection (Supplemental Figure 5). In these cases,
401 condensates could be a key part of the protective mechanism.
402
403 Specifically, we observed that activated caspase-7 partitions into DHR81 condensates,
404 suggesting caspase sequestration as a potential anti-apoptotic mechanism. Phase-
405 separated condensates and higher order assemblies are already implicated in mediating
406 apoptosis: stress granules inhibit apoptosis by suppressing MAPK pathways [48]; the
407 apoptosome relies on Apaf-1 multimerization, which has been reported to undergo
408 amyloid fibrillation in specific conditions [49]; and stress-dependent phase separation

409 modulates proteasome formation, whose function affects apoptosis [50-51]. It is interesting
410 to note that we only observe clear caspase-7 partitioning into DHR81 condensates,
411 suggesting the possibility of a specific binding interaction in this case. ApoE and *ARTSF*
412 LEA3m condensates may be sequestering other apoptotic factors or using a different
413 mechanism to reduce apoptosis.
414
415 It is also possible that these ExTol proteins achieve their anti-apoptotic effect by
416 implementing a more general physical mechanism, namely slowing the intracellular
417 mobility of biological effectors. There are various reports describing how the effective
418 viscosity within the cytoplasm can be increased in response to external stresses [4-5] or
419 mediated by protein composition, specifically intrinsically disordered proteins [52]. Notably,
420 a pervasive cell-spanning gel or network is not necessary to achieve such an effect; it is
421 possible that the presence of condensates, coupled with their ability to selectively partition
422 biomolecules (i.e. control what biomolecules can enter and diffuse through them), could
423 broadly slow biomolecular movement. That is, by forcing different proteins to either be
424 trapped in them or diffuse around them, condensates can decrease protein mobility and
425 perturb e.g. signaling pathways. Our results showing a lack of caspase-3 and -7
426 sequestration by most of our protein condensates could thus support this more general
427 mechanism of condensates perturbing biology, though they certainly do not disqualify
428 sequestration of other effectors not examined in this study.
429
430 Condensates are becoming increasingly implicated in disease [13], particularly when their
431 structural metastability and reversibility is dysregulated (e.g. through mutation). Notably,

432 drugs that block the formation of amyloid fibrils and plaques, which are products of phase
433 separation dysregulation, are promising yet strongly debated [53] therapies for many
434 neurological diseases. As such, our observations concerning ApoE proteins, while
435 unexpected from an extremotolerance perspective, are particularly salient. We found that
436 expression of a cytoplasmic form of ApoE is protective against apoptosis. In particular,
437 the truncated repeat region of ApoE is sufficient and shows the strongest anti-apoptotic
438 activity of all designs we tested. Others have shown similar effects for exogenously added
439 ApoE-containing liposomes protecting rat neurons against apoptosis [54], though we show
440 the effect through overexpression of ApoE inside of human cells. When combined with
441 our condensate results and the clinical relevance of ApoE in Alzheimer's disease [55], these
442 observations suggest the anti-apoptotic effect observed here warrants further study, to
443 further elucidate its role in Alzheimer's disease progression.
444
445 Altogether, our work expands upon the understanding of the relationship between
446 condensate formation and cell physiology, demonstrating how condensate accumulation
447 can be beneficial in reducing apoptosis while also identifying diverse new protein domains
448 that confer protection. The proteins and concepts presented here demonstrate the utility
449 of this class of proteins for engineering new properties into cells and ultimately higher
450 organisms.
451
452 **Methods**
453 **Cell culturing and maintenance**

454 HT-1080 cells (ATCC, CCL-121) were maintained in EMEM medium supplemented with
455 10% FBS and Penicillin-Streptomycin according to ATCC's protocol. For additional details
456 see supplemental methods.

457 **DNA Preparation**

458 Apoptosis constructs were synthesized from TWIST and transformed into DH5α cells for
459 amplification and plasmid purification. Condensate assay constructs were subcloned from
460 these synthesized constructs into vectors containing the appropriate chemically or light
461 inducible condensate elements. DNA was prepared using the QIAGEN Plasmid Plus Midi
462 Kit (Qiagen 12943) for transfection grade DNA purification. See supplemental methods
463 for additional details.

464 **Apoptosis assay**

465 200 µL of HT-1080 cells at to 7.5×10^4 cells per mL (15000 cells per well) were added to
466 each well of a microscope grade 96 well plate (Corning 3603). Plates were then
467 centrifuged at 330 rpm for 3 mins (Eppendorf, Centrifuge 5810). Plates were then
468 incubated at 37°C for 24 hours. Cells were then transfected with appropriate DNA
469 constructs with the lipofectamine 3000 kit (Thermo Fisher Scientific, L3000015) as
470 specified by the manufacturer. 24 hours after transfection, the transfection media was
471 removed and replaced with FluoroBrite DMEM Media (Thermo A1896702) supplemented
472 with 1x GlutaMAX (Thermo 35-050-061) and 10% FBS (Thermo Fisher Scientific,
473 10082147). This media also contained 13.2 µM (S)-(+)-Camptothecin (Sigma C9911-
474 100MG), 0.3 µM SiR-DNA stain (SpiroChrome CY-SC007), and 1.25 µM of the DEVD
475 peptide for detecting caspase activation (Sartorius 4440). The plates were sealed with
476 sterile Breathe-Easy film (USA scientific 9123-6100) and incubated at 37°C with 5% CO_2

477 for 30 minutes. After 30 min incubation, the plates were placed in an ImageXpress-
478 confocal plate imager with CO_2, humidity, and temperature control set to 37°C with 5%
479 CO_2 (Molecular Devices). Images of all wells were taken every 45 minutes and stored on
480 a remote server for later analysis. See supplemental methods for further details on this
481 method as well as the image analysis pipeline used.

482 **Confocal microscopy**

483 Confocal microscopy was done at the Nikon Imaging Center at Harvard Medical School
484 using a Nikon Ti fluorescence microscope attached to a Yokogawa W1 spinning disk
485 confocal. See supplemental methods for further details on the equipment and settings
486 used.

487 **Light inducible condensate assay**

488 HEK293-T cells were seeded on 96 well glass bottom plates from Cellvis (P96-1.5H-N).
489 Cells within each well were transfected with 50 ng of the LLPS1-iLID::EGFP::FTH1
490 (Addgene plasmid #122149) construct containing the core FTH1 scaffolding domain
491 attached to iLID for light inducible interactions. The cells were also transfected with 50 ng
492 of a construct containing the target protein attached to mCherry and SspB. The
493 transfection was allowed to continue overnight before the media was swapped to
494 FluoroBrite media supplemented with 10% FBS and 1x GlutaMAX. The cells were then
495 imaged using the settings listed above for mCherry and eGFP for the pre-induction
496 images. The cells were then exposed to 1 second of the DAPI imaging protocol described
497 in supplemental methods before new images were captured on the mCherry and eGFP
498 settings.

499 **Chemically inducible condensate assay**

500 HT-1080 cells were seeded on 96 well glass bottom plates from Cellvis (P96-1.5H-N).
501 Cells within each well were transfected with 50 ng of target construct fused to a mCherry-
502 2xFKBP domain. The transfection was allowed to continue overnight before the media
503 was swapped to FluoroBrite media supplemented with 10% FBS, 1x GlutaMAX, and 300
504 nM SiR-Hoechst. The cells were then imaged using the settings listed in the supplemental
505 methods for mCherry and SiR-Hoechst for the pre-induction images. Then, B/B dimerizer
506 (AP20187 ligand; TaKaRa 635058) was added to a final concentration of 500 nM in the
507 well. Imaging continued for 15 minutes with new images every 45 seconds.

508 **Caspase assay with chemically induced condensates**

509 These caspase assays are performed similarly to the standard caspase assays with a
510 few key differences. Briefly, HT-1080 cells were seeded as normal on to 96 well plates
511 and allowed to adhere overnight. The next day, 12.5 ng of nuclear mCherry marker and
512 37.5 ng of the target construct fused to mCherry-2xFKBP was transfected into the cells
513 as described above. The next day, 500 nM of the B/B dimerizer (TaKaRa 635058) was
514 added 2 hours before the start of the assay (or an Ethanol vehicle control was added).
515 The media was then removed and replaced with the apoptosis assay media described
516 above with the addition of 500 nM B/B dimerizer or a vehicle control. See supplemental
517 methods for further details.

518

519 **Abbreviations**

520 ExTol: extremotolerance-associated

521 IDP: Intrinsically Disordered Protein

522 LEA: Late Embryonic Abundant (protein)

523 ApoE: Apolipoprotein E

524 DHR: Designed helical repeat (protein)

525 CPT: Camptothecin

526 ATCC: American Type Culture Collection

527 XIAP: X-linked inhibitor of apoptosis

528 CAHS: Cytosolic Abundant Heat Soluble (protein)

529 FTH1: Ferritin Heavy Chain 1

530 iLID: light inducible dimer

531 SspB: Stringent Starvation Protein B

532 eGFP: Enhanced Green Fluorescent Protein

533 FKBP: FK506 binding protein

534 PvLEA: *Polypedilum vanderplanki* Late Embryonic Abundant (Protein)

535 MAPK: Mitogen-activated protein kinase

536 Apaf: Apoptotic protease activating factor

537 EMEM: Eagle's Minimum Essential Medium

538 FBS: Fetal bovine serum

539 DMEM: Dulbecco's Modified Eagle Medium

540

541 **Author Information.**

542 Correspondences should be sent to Pamela Silver, Harvard Medical School Department

543 of Systems Biology 200 Longwood Avenue Warren Alpert Building Boston, MA 02115.

544

545 **Author Contributions**

M.T.V. and D.T.N. designed and performed experiments, analyzed the data, and wrote the paper. N.N.T. analyzed the data and wrote the paper. M.E.O provided technical support. N.J.R., K.B.B., N.P.B., and S.L., designed constructs. D.B. and J.C.W. managed the teams and provided key insight. D.S.M., R.L.C., and P.A.S., designed constructs, managed teams, and wrote the paper.

Conflict of interest statement

No conflicts of interest

Supporting information

Supporting information contains supplemental figures providing additional context to the figures in the main text. This includes: Supplemental Figures 1 through 6 supporting the data presented in the main text.

Acknowledgements

Research was sponsored by the NSF, DARPA, the Wyss Institute for Biologically inspired engineering, and the Army Research Office and was accomplished under Cooperative Agreement Number W911NF-19-2-0017. M.T.V. was sponsored by the National Science Foundation under Grant No. 2010370. D.T.N., N.N.T., M.E.O., N.J.R., K.P.B., J.C.W, D.S.M., and P.A.S. were supported by the Harvard Medical School Department of Systems Biology. J.C.W. was further supported by the Harvard Medical School Laboratory of Systems Pharmacology. K.P.B. and D.S.M. were further supported by CZI Neurodegenerative Challenge Ben Barres Early Career Award (CZI2018-191853). N.P.B

569 was further supported by the HHMI Hanna Gray fellowship (grant GT11817). We would
570 like to acknowledge Clarence Yapp for help with microscopy as well as the NCI U54-
571 CA225088 grant and a gift from the Ludwig Institute for Cancer Research to the
572 Laboratory for Systems Pharmacology. We would like to acknowledge the Nikon Imaging
573 Center at Harvard Medical Center for providing microscope support. The views and
574 conclusions contained in this document are those of the authors and should not be
575 interpreted as representing the official policies, either expressed or implied, of any funding
576 agency including the NSF, DARPA, the Army Research Office, or the U.S. Government.
577 The U.S. Government is authorized to reproduce and distribute reprints for Government
578 purposes notwithstanding any copyright notation herein. Visual table of contents entry
579 was created with support from BioRender.com.

580

581 **Bibliography & References**

582

583 1. Jonsson, K. I.; Rabbow, E.; Schill, R. O.; Harms-Ringdahl, M.; Rettberg, P., Tardigrades
584 survive exposure to space in low Earth orbit. *Curr Biol* **2008,** *18* (17), R729-R731.
585 2. Thevelein, J. M., Regulation of trehalose mobilization in fungi. *Microbiol Rev* **1984,** *48*
586 (1), 42-59.
587 3. Boothby, T. C.; Tapia, H.; Brozena, A. H.; Piszkiewicz, S.; Smith, A. E.; Giovannini, I.;
588 Rebecchi, L.; Pielak, G. J.; Koshland, D.; Goldstein, B., Tardigrades Use Intrinsically Disordered
589 Proteins to Survive Desiccation. *Mol Cell* **2017,** *65* (6), 975-984 e5.
590 4. Dijksterhuis, J.; Nijsse, J.; Hoekstra, F. A.; Golovina, E. A., High viscosity and anisotropy
591 characterize the cytoplasm of fungal dormant stress-resistant spores. *Eukaryot Cell* **2007,** *6* (2),
592 157-70.
593 5. Parry, B. R.; Surovtsev, I. V.; Cabeen, M. T.; O'Hern, C. S.; Dufresne, E. R.; Jacobs-Wagner,
594 C., The bacterial cytoplasm has glass-like properties and is fluidized by metabolic activity. *Cell*
595 **2014,** *156* (1-2), 183-94.
596 6. Hand, S. C.; Menze, M. A.; Toner, M.; Boswell, L.; Moore, D., LEA proteins during water
597 stress: not just for plants anymore. *Annu Rev Physiol* **2011,** *73*, 115-34.
598 7. Chakrabortee, S.; Meersman, F.; Kaminski Schierle, G. S.; Bertoncini, C. W.; McGee, B.;
599 Kaminski, C. F.; Tunnacliffe, A., Catalytic and chaperone-like functions in an intrinsically
600 disordered protein associated with desiccation tolerance. *Proc Natl Acad Sci U S A* **2010,** *107*
601 (37), 16084-9.

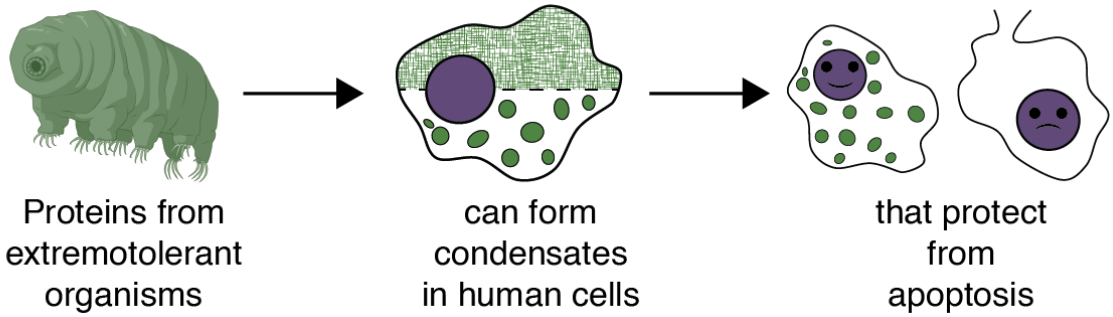
602 8. Boothby, T. C., Water content influences the vitrified properties of CAHS proteins. *Mol*
603 *Cell* **2021,** *81* (3), 411-413.
604 9. Arakawa, K.; Numata, K., Reconsidering the "glass transition" hypothesis of intrinsically
605 unstructured CAHS proteins in desiccation tolerance of tardigrades. *Mol Cell* **2021,** *81* (3), 409-
606 410.
607 10. Crilly, C. J.; Brom, J. A.; Warmuth, O.; Esterly, H. J.; Pielak, G. J., Protection by
608 desiccation-tolerance proteins probed at the residue level. *Protein Sci* **2021**.
609 11. Belott, C.; Janis, B.; Menze, M. A., Liquid-liquid phase separation promotes animal
610 desiccation tolerance. *Proc Natl Acad Sci U S A* **2020,** *117* (44), 27676-27684.
611 12. Hyman, A. A.; Weber, C. A.; Julicher, F., Liquid-liquid phase separation in biology. *Annu*
612 *Rev Cell Dev Biol* **2014,** *30*, 39-58.
613 13. Shin, Y.; Brangwynne, C. P., Liquid phase condensation in cell physiology and disease.
614 *Science* **2017,** *357* (6357).
615 14. Boeynaems, S.; Alberti, S.; Fawzi, N. L.; Mittag, T.; Polymenidou, M.; Rousseau, F.;
616 Schymkowitz, J.; Shorter, J.; Wolozin, B.; Van Den Bosch, L.; Tompa, P.; Fuxreiter, M., Protein
617 Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* **2018,** *28* (6), 420-435.
618 15. Choi, J. M.; Holehouse, A. S.; Pappu, R. V., Physical Principles Underlying the Complex
619 Biology of Intracellular Phase Transitions. *Annu Rev Biophys* **2020,** *49*, 107-133.
620 16. Jiang, S.; Fagman, J. B.; Chen, C.; Alberti, S.; Liu, B., Protein phase separation and its role
621 in tumorigenesis. *Elife* **2020,** *9*.
622 17. Boija, A.; Klein, I. A.; Young, R. A., Biomolecular Condensates and Cancer. *Cancer Cell*
623 **2021,** *39* (2), 174-192.
624 18. Boija, A.; Klein, I. A.; Sabari, B. R.; Dall'Agnese, A.; Coffey, E. L.; Zamudio, A. V.; Li, C. H.;
625 Shrinivas, K.; Manteiga, J. C.; Hannett, N. M.; Abraham, B. J.; Afeyan, L. K.; Guo, Y. E.; Rimel, J.
626 K.; Fant, C. B.; Schuijers, J.; Lee, T. I.; Taatjes, D. J.; Young, R. A., Transcription Factors Activate
627 Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **2018,** *175* (7),
628 1842-1855 e16.
629 19. Hubstenberger, A.; Courel, M.; Benard, M.; Souquere, S.; Ernoult-Lange, M.; Chouaib, R.;
630 Yi, Z.; Morlot, J. B.; Munier, A.; Fradet, M.; Daunesse, M.; Bertrand, E.; Pierron, G.;
631 Mozziconacci, J.; Kress, M.; Weil, D., P-Body Purification Reveals the Condensation of Repressed
632 mRNA Regulons. *Mol Cell* **2017,** *68* (1), 144-157 e5.
633 20. Tolleter, D.; Jaquinod, M.; Mangavel, C.; Passirani, C.; Saulnier, P.; Manon, S.; Teyssier,
634 E.; Payet, N.; Avelange-Macherel, M. H.; Macherel, D., Structure and function of a mitochondrial
635 late embryogenesis abundant protein are revealed by desiccation. *Plant Cell* **2007,** *19* (5), 1580-
636 9.
637 21. Leon, D., III, Occurrence of a Repeating 11-Mer Amino Acid Sequence Motif in Diverse
638 Organisms. *Protein & Peptide Letters* **2001,** *8* (2), 115-122.
639 22. Bussell, R., Jr.; Eliezer, D., A structural and functional role for 11-mer repeats in alpha-
640 synuclein and other exchangeable lipid binding proteins. *J Mol Biol* **2003,** *329* (4), 763-78.
641 23. Li, S.; Chakraborty, N.; Borcar, A.; Menze, M. A.; Toner, M.; Hand, S. C., Late
642 embryogenesis abundant proteins protect human hepatoma cells during acute desiccation.
643 *Proc Natl Acad Sci U S A* **2012,** *109* (51), 20859-64.
644 24. Tanaka, S.; Tanaka, J.; Miwa, Y.; Horikawa, D. D.; Katayama, T.; Arakawa, K.; Toyoda, A.;
645 Kubo, T.; Kunieda, T., Novel mitochondria-targeted heat-soluble proteins identified in the

646 anhydrobiotic Tardigrade improve osmotic tolerance of human cells. *PLoS One* **2015,** *10* (2),
647 e0118272.
648 25. Mertens, J.; Aliyu, H.; Cowan, D. A., LEA Proteins and the Evolution of the WHy Domain.
649 *Appl Environ Microbiol* **2018,** *84* (15).
650 26. Makarova, K. S.; Aravind, L.; Wolf, Y. I.; Tatusov, R. L.; Minton, K. W.; Koonin, E. V.; Daly,
651 M. J., Genome of the extremely radiation-resistant bacterium Deinococcus radiodurans viewed
652 from the perspective of comparative genomics. *Microbiol Mol Biol Rev* **2001,** *65* (1), 44-79.
653 27. Yamazaki, Y.; Zhao, N.; Caulfield, T. R.; Liu, C. C.; Bu, G., Apolipoprotein E and Alzheimer
654 disease: pathobiology and targeting strategies. *Nat Rev Neurol* **2019,** *15* (9), 501-518.
655 28. Raulin, A. C.; Kraft, L.; Al-Hilaly, Y. K.; Xue, W. F.; McGeehan, J. E.; Atack, J. R.; Serpell, L.,
656 The Molecular Basis for Apolipoprotein E4 as the Major Risk Factor for Late-Onset Alzheimer's
657 Disease. *J Mol Biol* **2019,** *431* (12), 2248-2265.
658 29. Erdos, G.; Dosztanyi, Z., Analyzing Protein Disorder with IUPred2A. *Curr Protoc*
659 *Bioinformatics* **2020,** *70* (1), e99.
660 30. Biegert, A.; Soding, J., De novo identification of highly diverged protein repeats by
661 probabilistic consistency. *Bioinformatics* **2008,** *24* (6), 807-14.
662 31. Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C., The helical hydrophobic moment: a
663 measure of the amphiphilicity of a helix. *Nature* **1982,** *299* (5881), 371-4.
664 32. Furuki, T.; Niwa, T.; Taguchi, H.; Hatanaka, R.; Kikawada, T.; Sakurai, M., A LEA model
665 peptide protects the function of a red fluorescent protein in the dry state. *Biochem Biophys Rep*
666 **2019,** *17*, 27-31.
667 33. Brunette, T. J.; Parmeggiani, F.; Huang, P. S.; Bhabha, G.; Ekiert, D. C.; Tsutakawa, S. E.;
668 Hura, G. L.; Tainer, J. A.; Baker, D., Exploring the repeat protein universe through computational
669 protein design. *Nature* **2015,** *528* (7583), 580-4.
670 34. Li, F.; Jiang, T.; Li, Q.; Ling, X., Camptothecin (CPT) and its derivatives are known to target
671 topoisomerase I (Top1) as their mechanism of action: did we miss something in CPT analogue
672 molecular targets for treating human disease such as cancer? *Am J Cancer Res* **2017,** *7* (12),
673 2350-2394.
674 35. Wall, M. E.; Wani, M. C.; Cook, C. E.; Palmer, K. H.; McPhail, A. T.; Sim, G. A., Plant
675 Antitumor Agents. I. The Isolation and Structure of Camptothecin, a Novel Alkaloidal Leukemia
676 and Tumor Inhibitor from Camptotheca acuminata1,2. *Journal of the American Chemical Society*
677 **1966,** *88* (16), 3888-3890.
678 36. Bortner, C. D.; Cidlowski, J. A., Cellular mechanisms for the repression of apoptosis.
679 *Annu Rev Pharmacol Toxicol* **2002,** *42*, 259-81.
680 37. Rasheed, S.; Nelson-Rees, W. A.; Toth, E. M.; Arnstein, P.; Gardner, M. B.,
681 Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer* **1974,** *33* (4),
682 1027-33.
683 38. Obexer, P.; Ausserlechner, M. J., X-linked inhibitor of apoptosis protein - a critical death
684 resistance regulator and therapeutic target for personalized cancer therapy. *Front Oncol* **2014,**
685 *4*, 197.
686 39. Li, W.; Lee, M. K., Antiapoptotic property of human alpha-synuclein in neuronal cell lines
687 is associated with the inhibition of caspase-3 but not caspase-9 activity. *J Neurochem* **2005,** *93*
688 (6), 1542-50.

689 40. Geng, X.; Guo, J.; Zang, X.; Chang, C.; Shang, H.; Wei, H.; Xu, C., Proteomic analysis of
690 eleven tissues in the Chinese giant salamander (Andrias davidianus). *Sci Rep* **2019,** *9* (1), 16415.
691 41. Bracha, D.; Walls, M. T.; Brangwynne, C. P., Probing and engineering liquid-phase
692 organelles. *Nat Biotechnol* **2019,** *37* (12), 1435-1445.
693 42. Bracha, D.; Walls, M. T.; Wei, M. T.; Zhu, L.; Kurian, M.; Avalos, J. L.; Toettcher, J. E.;
694 Brangwynne, C. P., Mapping Local and Global Liquid Phase Behavior in Living Cells Using Photo-
695 Oligomerizable Seeds. *Cell* **2018,** *175* (6), 1467-1480 e13.
696 43. Stanton, B. Z.; Chory, E. J.; Crabtree, G. R., Chemically induced proximity in biology and
697 medicine. *Science* **2018,** *359* (6380).
698 44. Yagi-Utsumi, M.; Aoki, K.; Watanabe, H.; Song, C.; Nishimura, S.; Satoh, T.; Yanaka, S.;
699 Ganser, C.; Tanaka, S.; Schnapka, V.; Goh, E. W.; Furutani, Y.; Murata, K.; Uchihashi, T.; Arakawa,
700 K.; Kato, K., Desiccation-induced fibrous condensation of CAHS protein from an anhydrobiotic
701 tardigrade. *Sci Rep* **2021,** *11* (1), 21328.
702 45. Malki, A.; Teulon, J. M.; Camacho-Zarco, A. R.; Chen, S. W.; Adamski, W.; Maurin, D.;
703 Salvi, N.; Pellequer, J. L.; Blackledge, M., Intrinsically Disordered Tardigrade Proteins Self-
704 Assemble into Fibrous Gels in Response to Environmental Stress. *Angew Chem Int Ed Engl* **2022,**
705 *61* (1), e202109961.
706 46. Hesgrove, C. S.; Nguyen, K. H.; Biswas, S.; Childs, C. A.; Shraddha, K.; Medina, B. X.;
707 Alvarado, V.; Yu, F.; Sukenik, S.; Malferrari, M.; Francia, F.; Venturoli, G.; Martin, E. W.;
708 Holehouse, A. S.; Boothby, T. C., Tardigrade CAHS Proteins Act as Molecular Swiss Army Knives
709 to Mediate Desiccation Tolerance Through Multiple Mechanisms. *bioRxiv* **2021**,
710 2021.08.16.456555.
711 47. Tanaka, A.; Nakano, T.; Watanabe, K.; Masuda, K.; Kamata, S.; Yasui, R.; Kozuka-Hata, H.;
712 Watanabe, C.; Chinen, T.; Kitagawa, D.; Oyama, M.; Yanagisawa, M.; Kunieda, T., Stress-
713 dependent dynamic and reversible formation of cytoskeleton-like filaments and gel-transition
714 by tardigrade tolerance proteins. *bioRxiv* **2021**, 2021.10.02.462891.
715 48. Arimoto, K.; Fukuda, H.; Imajoh-Ohmi, S.; Saito, H.; Takekawa, M., Formation of stress
716 granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. *Nat Cell Biol*
717 **2008,** *10* (11), 1324-32.
718 49. Rao, P. N.; Reddy, K. S.; Bhuyan, A. K., Amyloid fibrillation of human Apaf-1 CARD.
719 *Biochemistry* **2009,** *48* (32), 7656-64.
720 50. Yasuda, S.; Tsuchiya, H.; Kaiho, A.; Guo, Q.; Ikeuchi, K.; Endo, A.; Arai, N.; Ohtake, F.;
721 Murata, S.; Inada, T.; Baumeister, W.; Fernandez-Busnadiego, R.; Tanaka, K.; Saeki, Y., Stress-
722 and ubiquitylation-dependent phase separation of the proteasome. *Nature* **2020,** *578* (7794),
723 296-300.
724 51. Gupta, I.; Singh, K.; Varshney, N. K.; Khan, S., Delineating Crosstalk Mechanisms of the
725 Ubiquitin Proteasome System That Regulate Apoptosis. *Front Cell Dev Biol* **2018,** *6*, 11.
726 52. Theillet, F. X.; Binolfi, A.; Frembgen-Kesner, T.; Hingorani, K.; Sarkar, M.; Kyne, C.; Li, C.;
727 Crowley, P. B.; Gierasch, L.; Pielak, G. J.; Elcock, A. H.; Gershenson, A.; Selenko, P.,
728 Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs).
729 *Chem Rev* **2014,** *114* (13), 6661-714.
730 53. Lalli, G.; Schott, J. M.; Hardy, J.; De Strooper, B., Aducanumab: a new phase in
731 therapeutic development for Alzheimer's disease? *EMBO Mol Med* **2021,** *13* (8), e14781.

54. Hayashi, H.; Campenot, R. B.; Vance, D. E.; Vance, J. E., Protection of neurons from apoptosis by apolipoprotein E-containing lipoproteins does not require lipoprotein uptake and involves activation of phospholipase Cgamma1 and inhibition of calcineurin. *J Biol Chem* **2009,** *284* (43), 29605-13.
55. Schellenberg, G. D.; D'Souza, I.; Poorkaj, P., The genetics of Alzheimer's disease. *Curr Psychiatry Rep* **2000,** *2* (2), 158-64.

Visual Table of Contents entry:



Natural and designed proteins inspired by extremotolerant organisms can form condensates and attenuate apoptosis in human cells

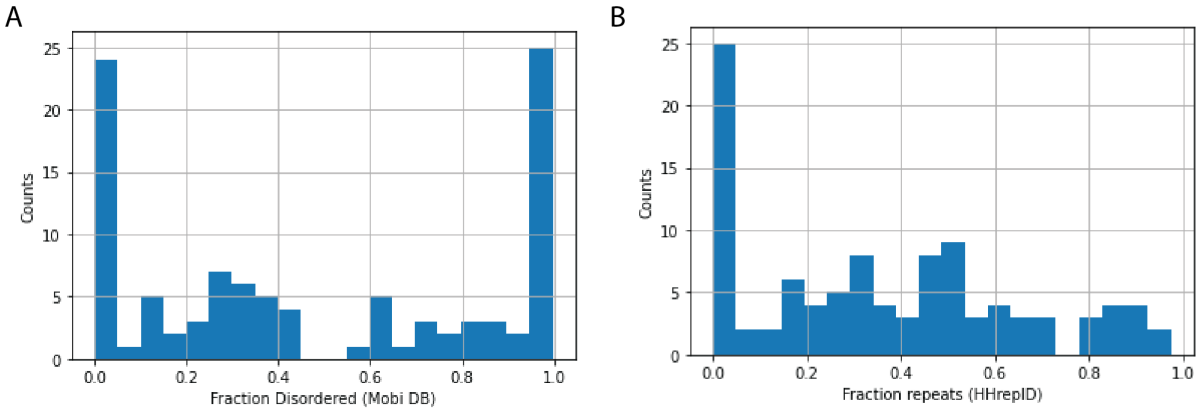
Mike T. Veling[1,2,8], Dan T. Nguyen[1,2,8], Nicole N. Thadani[1], Michela E. Oster[1,2], Nathan J. Rollins[1,2], Kelly P. Brock[1], Neville P. Bethel[3,4], Samuel Lim[1,2], David Baker[3,4,5], Jeffrey C. Way[1,2], Debora S. Marks[1,6], Roger L. Chang[1,2,7,*], Pamela A. Silver[1,2,*]

Affiliations

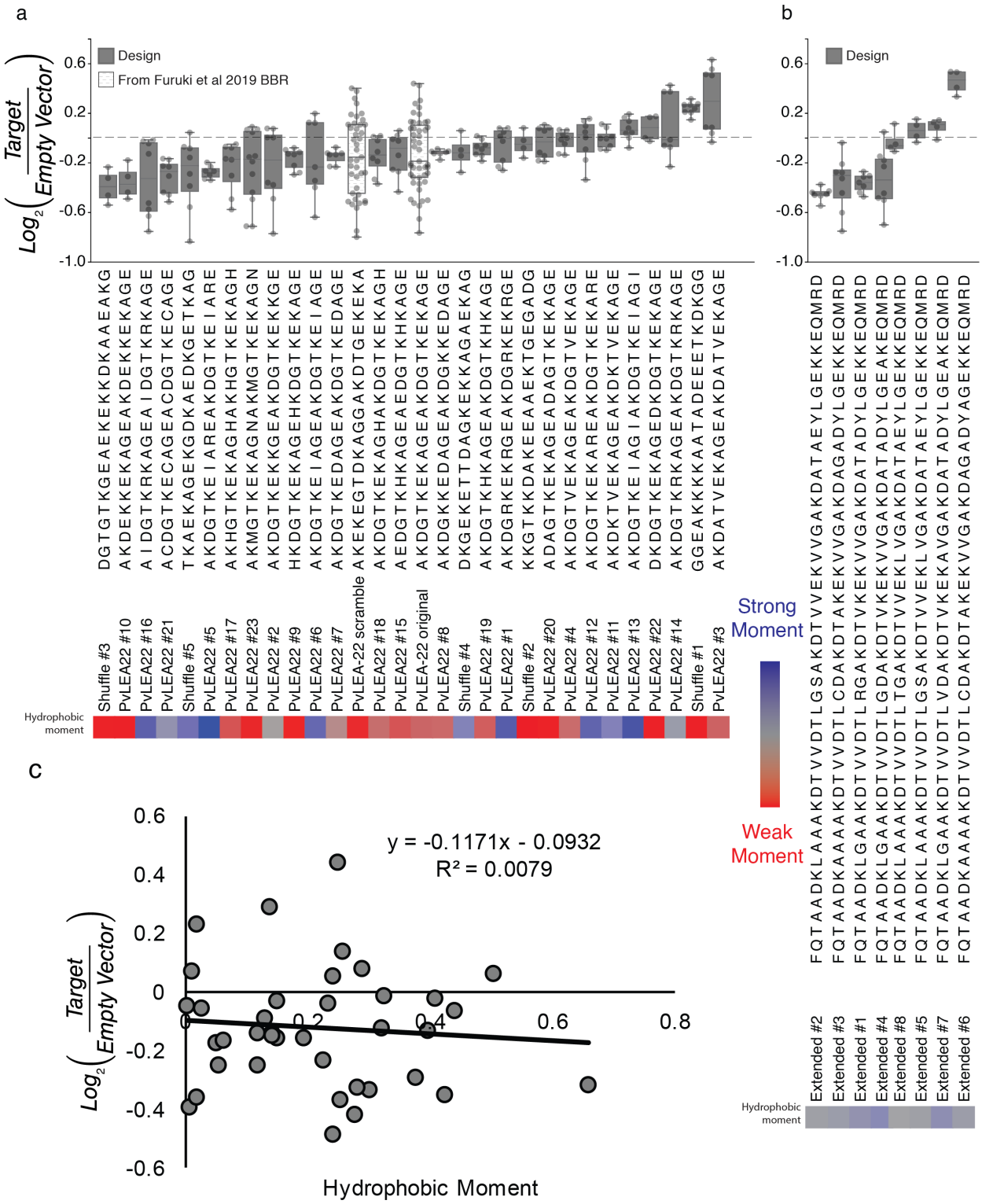
1. Department of Systems Biology, Harvard Medical School, Boston, MA, 02115, USA;
2. Wyss Institute for Biologically Inspired Engineering, Boston, MA, 02115, USA;
3. Institute for Protein Design, University of Washington, Seattle, WA, 98195, USA;
4. Department of Biochemistry, University of Washington, Seattle, WA, 98195, USA;
5. Howard Hughes Medical Institute, University of Washington, Seattle, WA, 98195, USA;
6. Broad Institute of Harvard and MIT, Cambridge, MA, 02142, USA;
7. Department of Systems & Computational Biology, Albert Einstein College of Medicine, Bronx, NY, 10461
8. Co-first authors

*. Co-corresponding authors (pamela_silver@hms.harvard.edu, roger.chang@einsteinmed.edu)

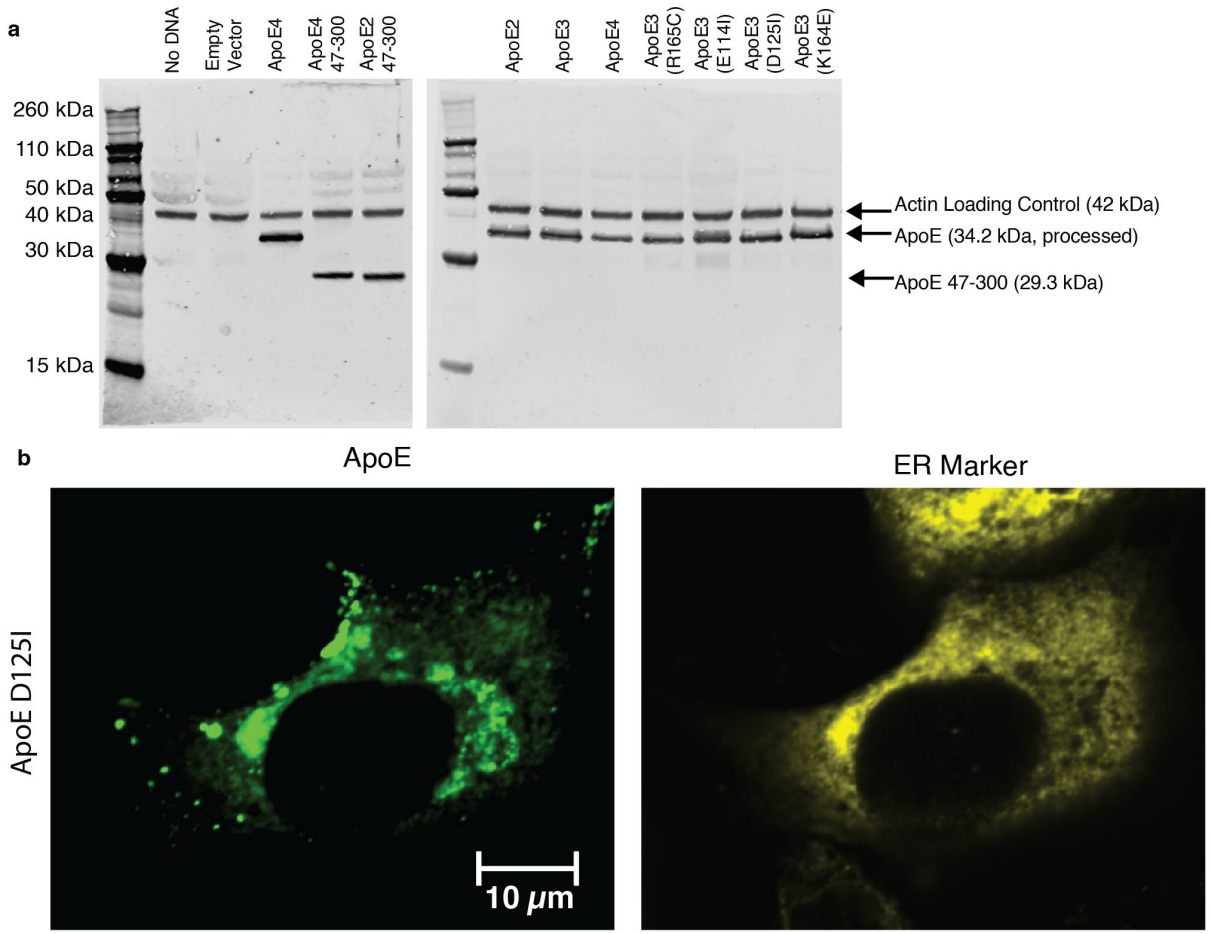
S2



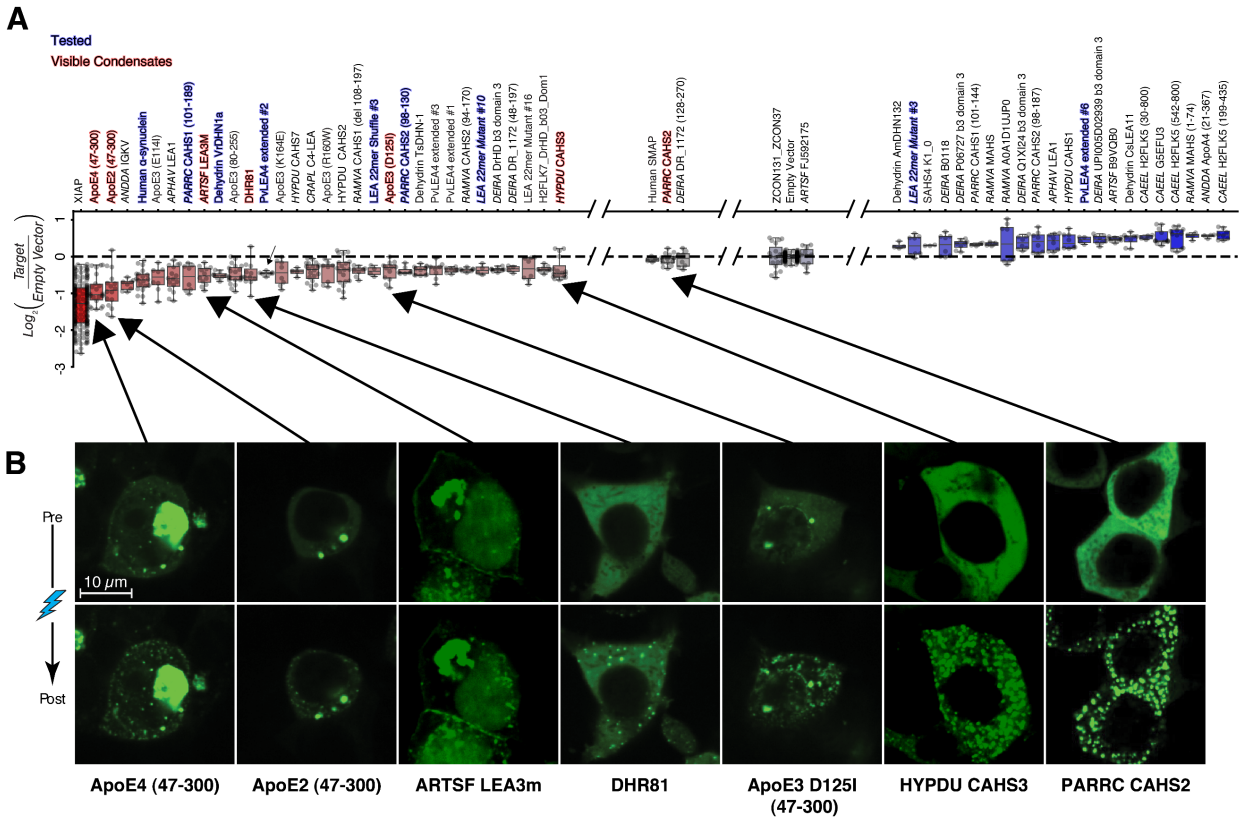
Supplemental Figure 1 (a) Distribution of fractional disorder within the sequences of the selected ExTol proteins. **(b)** Distribution of fractional repeats within the sequences of the selected ExTol proteins.



Supplemental Figure 2 (a) Graph and sequences of various LEA repeats tested, with full sequences and performance. **(b)** Graph and sequences of extended LEA repeats tested. **(c)** Correlation analysis for relationship between hydrophobic moment with apoptosis assay performance.

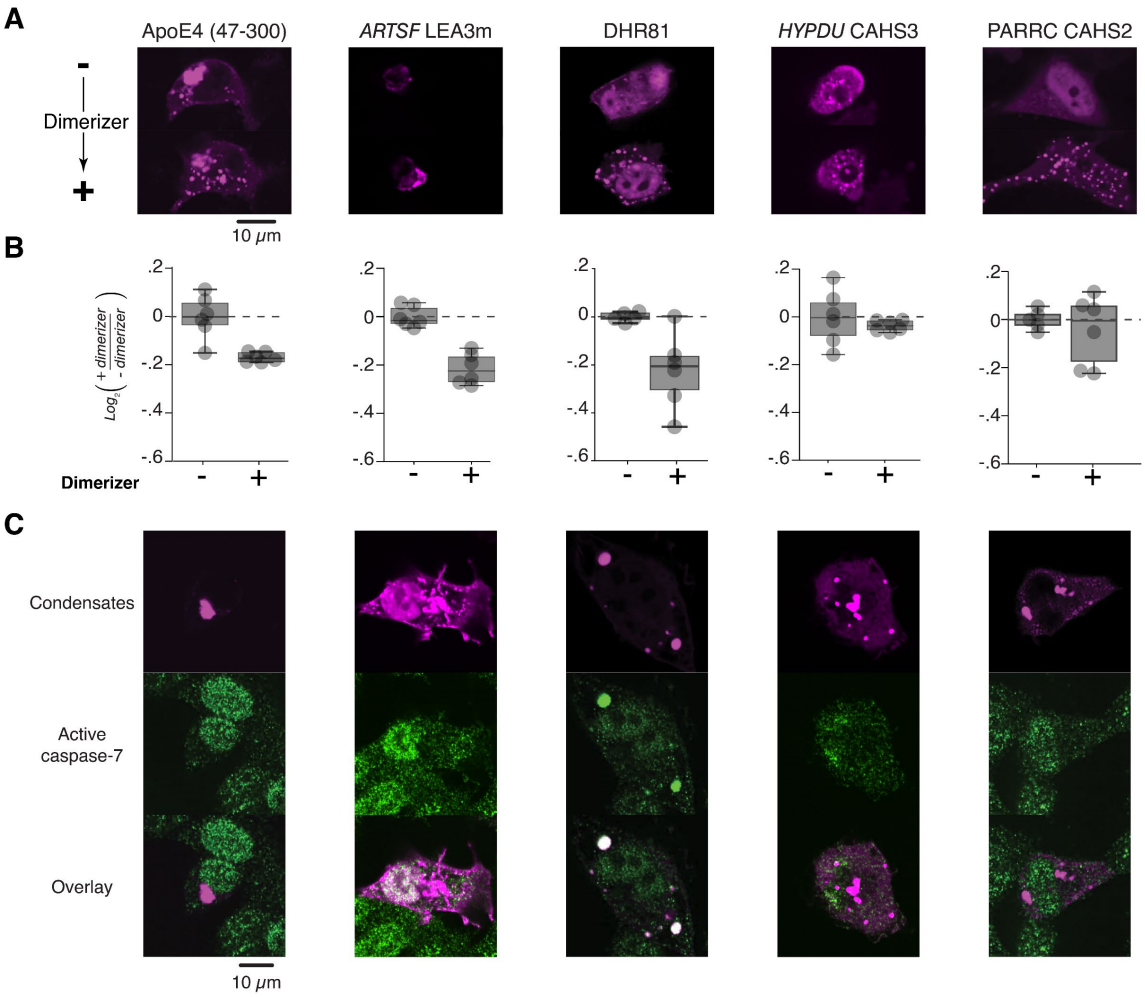


Supplemental Figure 3 (a) Western blot data of some ApoE constructs showing expression at the expected molecular weight. **(b)** Sample images showing condensates forming in the full length ApoE3 D125I mutation.

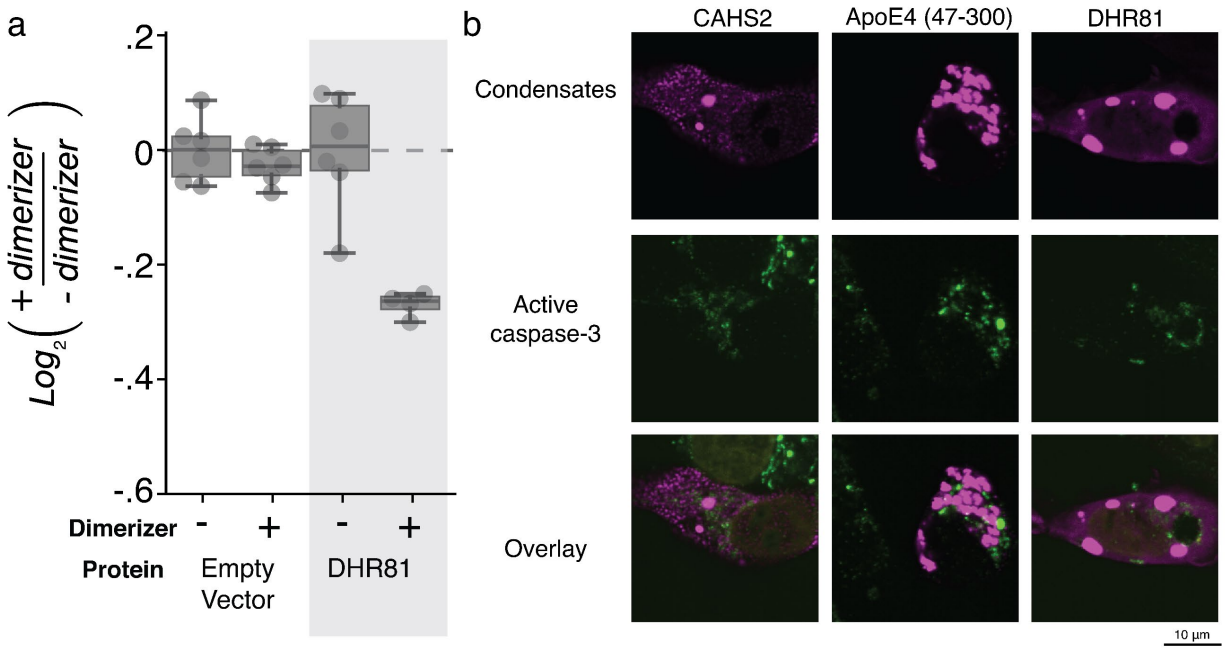


Supplemental Figure 4 (a) Apoptosis assay data highlighting all the ExTol proteins tested and their apoptosis assay performance. Note that red and blue colored font refer to whether condensates were visible or not, respectively, via our microscopy. **(b)** Light inducible condensate formation assays for selected ExTol proteins. Data shown includes all tested proteins with visible condensates.

S6


A
ApoE4 (47-300)
ARTSF LEA3m
DHR81
HYPDU CAHS3
PARRC CAHS2
-
Dimerizer
+
10 μm
B
Log2(+ dimerizer / - dimerizer)
.2
0
-.2
-.4
-.6
Dimerizer
-
+
C
Condensates
Active caspase-7
Overlay
10 μm

47
48



Supplemental Figure 6 (**a**) DHR81 inducible condensates reduce apoptosis. (**b**) Sample images showing no co-localization between activated caspase 3 and condensates.

53 **Table S1:** Supplemental table containing constructs and designs used in this study
54
55 **Table S2:** Data values obtained for apoptosis experiments

Supplemental Methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Pamela Silver.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Escherichia coli strain DH5α (NEB) was used for all DNA synthesis applications and grown at 37°C in LB media with antibiotics. Mammalian cell culture work was done in *Homo sapiens* HT-1080. See methods below for more details.

METHOD DETAILS

Cell culturing and maintenance

Eagle's Minimum Essential Medium (EMEM) (ATCC 20-2003), Heat-inactivated Fetal Bovine Serum (FBS) (Thermo Fisher Scientific 10082147), and Penicillin-Streptomycin (Corning 30-002-CI) were all vacuum sterilized by filtration (0.22 µm pore size, Corning, 430767) and used for maintaining cell growth. Cells were maintained at 37°C. *Homo sapiens* HT-1080 cells (ATCC, CCL-121) were thawed in a 37°C water bath and immediately transferred to a cell culture flask. Cells were then grown for 48 hours. At this point, the media was gently aspirated and replaced and cells were allowed to grow for another 24 hours. Cells were then split into new flasks at one third of the density (Trypsin EDTA 1X 0.25% Trypsin/ 2.21mM EDTA in HBSS without sodium bicarbonate, calcium and magnesium, VWR 45000-664) and continued to grow in EMEM supplemented by FBS and Penicillin-Streptomycin.

Cloning apoptosis constructs

The empty vector used in the apoptosis assay was pSecTag2 A ordered from Thermo (V90020). A GFP sequence was cloned into this vector to provide a replacement sequence for Twist Biosciences to clone over. The resulting vector (YK 02, vector map can be found here https://benchling.com/s/seq-gjUPUvEDiLQT6K6CyGPl) was onboarded at Twist Biosciences. The cloning location was specified using the "replace GFP" feature in the YK 02 vector map described above. Column C of Table S1 specifies

the insert that was used to replace the GFP. The resulting vectors are identical to the YK 02 map shown above with the sequence from column C overwriting the “replace GFP” feature in the map.

Cloning condensate constructs

To generate the light inducible constructs, the same YK 02 backbone was used. The target sequence was appended before a short linker, an mCherry sequence, another short linker, and the SspB protein (Addgene plasmid #122148). The chemically inducible condensate sequences were generated by attaching the target protein up stream of an mCherry and 2x FKBP (F36V) sequences. The below table includes vector maps of the vectors created

Vector ID	Name	Sequence map
pBS1043	CAHS2parrc-mCh-SspB	https://benchling.com/s/seq-mwhbZwTC9pJF1uR8owV2
pBS1068	ApoE2_47-300_mCh-SspB	https://benchling.com/s/seq-MCLamVEadIFGtsvDoeS0
pBS1069	ApoE4_47-300_mCh-SspB	https://benchling.com/s/seq-v0A5J1H806zWv64cogkG
pBS1072	DHR81_Nccapped_mCh-SspB	https://benchling.com/s/seq-Yyqj5GsaPVjLGjLw7wUY
pBS1109	CAHS2parrc-mCh-2xFKBP	https://benchling.com/s/seq-JahUmtrgF84VBOBCvEEj
pBS1111	ApoE2_47-300_mCh-2xFKBP	https://benchling.com/s/seq-Nh0mlTXD5ciGEgfp7gCz
pBS1112	ApoE4_47-300_mCh-2xFKBP	https://benchling.com/s/seq-geutb277CV3mCzZfKgOv
pBS1113	ApoE3_47-300_mCh-2xFKBP	https://benchling.com/s/seq-vV2W6H7pDo4Itr6c85vW
pBS1114	ApoE3_D125I_47-300_mCh-2xFKBP	https://benchling.com/s/seq-2ej7AZbyfB2FaIneJxdp
pBS1115	DHR81_Nccapped_mCh-2xFKBP	https://benchling.com/s/seq-JFzThIQ8s83yDljKC0O6

DNA Preparation

All DNA was synthesized from TWIST and transformed into DH5α cells for amplification and plasmid purification using the QIAGEN Plasmid Plus Midi Kit for transfection grade DNA purification.

Cell seeding

HT-1080 cells were washed with Phosphate-Buffered Saline (PBS) (Thermo Fisher Scientific, AM9625) and trypsinized. Cell density was determined by optical density with an automated cell counter (BIO RAD TC20), cell counting slides (BIO RAD, 1450015) and 0.4% solution Trypan Blue (BIO RAD, 1450021). After live cell count was obtained, cells were diluted to 7.5×10^4 cells per mL in Eagle's Minimum Essential Medium (EMEM) (ATCC, 20-2003) supplemented with 10% Heat-inactivated Fetal Bovine Serum (FBS) (Thermo Fisher Scientific, 10082147), that had been previously vacuum sterilized by filtration (0.22 µm pore size, Corning). 200 µL of cells were added to each well of a microscope grade 96 well plate (Corning 3603). Plates were then centrifuged at 330 rpm for 3 mins (Eppendorf, Centrifuge 5810). Plates were then incubated at 37°C for 24 hours.

Transfection

HT-1080 cells were transfected 24 hours after plating on 96-well plates and were carried out in EMEM (ATCC, 20-2003) supplemented with 10% FBS (Thermo Fisher Scientific, 10082147). Prior to starting the transfection, the seeding media was removed and fresh media was added to the plate. Per well, 75 ng of target DNA and 25 ng of an mCherry transfection marker were transfected. This was achieved by mixing 360 ng of target DNA and 120 ng of mCherry transfection marker DNA together in a strip tube. Separately, a mix was prepared containing 750 µL of Opti-MEM (Thermo 31985062) and 30 µL of the p3000 reagent from the lipofectamine 3000 kit (Thermo Fisher Scientific, L3000015). 24 µL of this mix was added to the well containing the target DNA and transfection marker. Again, in a separate tube, a mix was prepared containing 750 µL of Opti-MEM and 45 µL of the the lipofectamine reagent. 24 µL of this mix was also added to the well containing the target DNA and transfection marker. This mix was briefly vortexed and allowed to stand for 5 minutes before 10 µL of the mix was added to 4 wells on the plate.

CPT treatment for caspase assay.

24 hours after transfection, the transfection media was removed and replaced with FluoroBrite DMEM Media (Thermo A1896702) supplemented with 1x GlutaMAX (Thermo 35-050-061) and 10% FBS (Thermo Fisher Scientific, 10082147). This media also contained 13.2 µM (S)-(+)-Camptothecin (Sigma C9911-100MG), 0.3 µM SiR-DNA stain (SpiroChrome CY-SC007), and 1.25 µM of the DEVD peptide for detecting caspase

activation (Sartorius 4440). The plates were sealed with sterile Breathe-Easy film (USA scientific 9123-6100) and incubated at 37°C with 5% CO_2 for 30 minutes.

Imaging for caspase assay without condensate induction

After 30 min incubation, the plates were placed in an ImageXpress-confocal plate imager with CO_2, humidity, and temperature control set to 37°C with 5% CO_2 (Molecular Devices). Images of all wells were taken every 45 minutes for 22.5 hours and stored on a remote server for later analysis.

Imaging for caspase assay with condensate induction

The imaging process was very similar to the caspase assay without condensate induction with the exception that most assays (those on plates 20210715, 20210914, and 20210924) were collected for 30 hours instead of 22.5. The images captured on the 20211226 plate were only captured for 22.5 hours.

Western blotting

Cells were cultured and transfected in 6 well dishes as described above but scaled up 30-fold to match the difference in surface area. The day after transfection, cells were scraped and collected into a 1.5 mL conical tube. Pellets were resuspended in about 1x pellet volume of RIPA buffer (10 mM Tris-Cl pH 8.0, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, 140 mM NaCl). Cells were allowed to lyse on ice for 30 min before being spun at 20,000xg for 1 min at 4°C. The resulting clarified lysate was run on 12% bis tris in MES buffer at 200 v for about 1 h. The gel was transferred to a membrane using the iBlot-2 membrane transfer system. The blot was then blocked for 1 hour at room temperature with agitation in 5% non-fat dried milk in TBS before adding the primary antibody and incubating overnight at 4°C. Below is a table of primary antibodies used and their dilution factors. All antibodies were used in TBST containing 1% non-fat dried milk.

Supplier	Order Number	Target	Dilution Factor
Sigma	F1804	FLAG	1:100
Abcam	ab24139	ApoE	1:200

Abcam	ab8227	Actin	1:1000
Abcam	ab8224	Actin	1:1000

The next day, the blot was washed 3x in TBST before being stained with secondary antibody for 1 hour at room temperature. Below is a table of the secondary antibodies used.

Supplier	Order Number	Target	Dilution Factor	Fluorophore
VWR	102971-154	Mouse	1:15000	800 CW
VWR	103011-498	Rabbit	1:15000	680 RD

Membranes were then scanned using a LICOR CLx instrument using auto exposure.

Immunofluorescence

Cells were grown on 96 well glass bottom plates from Cellvis (P96-1.5H-N). Cells were cultured and transfected as described above. Cells were fixed in 4% formaldehyde in culture medium. The cells were then washed 3x with PBS. Cells were then permeabilized with 100% cold MeOH for 10 min while being stored at-20°C. Permeabilized cells were blocked with PBS containing 1% BSA and 0.15% Triton X-100 for 1 hour. Blocking solution was removed and fresh blocking solution with a dilution of the primary antibody was added back to the well. Below is a table of the antibodies used and their dilution factors. Antibodies were left to bind overnight in the cold room.

Supplier	Order Number	Target	Dilution Factor
Sigma	F1804	FLAG	1:500
Thermo	MA3-019	PDI	1:400
Abcam	ab24139	ApoE	1:500
Cell Signaling Technologies	9507S	Caspase-9	1:200
Cell Signaling Technologies	8438T	Caspase-7	1:500
Cell Signaling Technologies	9664T	Caspase-3	1:500
Cell Signaling Technologies	12963S	Cytochrome C	1:200

Cell Signaling Technologies	89477S	Bax	1:1000
Cell Signaling Technologies	2933S	Bim	1:100
Abcam	ab32445	Bad	1:50

The next day, the primary containing solution was removed, and the cells were washed with PBS 3x. PBS containing 1% BSA, 0.15% Triton X-100, and 2 µg/mL of DAPI was prepared. Depending on the source of the primary antibody and the desired fluorescent channel, the appropriate secondary antibody was added to stain the cells. This mix was added to the cells and allowed to stain for 1 hour in the dark at room temperature. Below is a table of the secondary antibodies used here.

Supplier	**Order Number**	**Target**	**Dilution Factor**	**Fluorophore**
Thermo	A-11029	Mouse	1:2000	Alexa Fluor 488
Thermo	A-11008	Rabbit	1:2000	Alexa Fluor 488
Thermo	A-11031	Mouse	1:2000	Alexa Fluor 568
Thermo	A-11036	Rabbit	1:2000	Alexa Fluor 568
Thermo	A-21235	Mouse	1:2000	Alexa Fluor 647
Thermo	A-21245	Rabbit	1:2000	Alexa Fluor 647

After staining, the cells were washed 3x with PBS before being stored in PBS for imaging.

Confocal microscopy

Confocal microscopy was done at the Nikon Imaging Center at Harvard Medical School using a Nikon Ti fluorescence microscope attached to a Yokogawa W1 spinning disk confocal. The microscope was equipped with a Nikon Perfect Focus System for focus assistance. A 60x objective was used for image capture (60x Nikon Plan Apo 60x 1.4 NA DCI N2). Live cell imaging was done with a Okolab Stage Top Incubator at 37°C 5% CO_2 mounted on top of a Prior Proscan III motorized stage and shutters. Fixed cells were imaged in the PBS whereas live cells were imaged in FluoroBrite supplemented with 10% FBS and 1x GlutaMAX. The fluorophores used include the Alexa Fluor series (488,568, and 647) as well as mCherry, DAPI, SiR-Hoechst (SpiroChrome, CHF260.00), and eGFP

depending on the experiment. Below is a table of the filters used to detect each fluorophore.

Fluorophore	Excitation Laser	Emission Filter
Alexa Fluor 488	488	Chroma ET525/50m
Alexa Fluor 568	561	Chroma ET605/52m
Alexa Fluor 647	640	Chroma ET705/72m
DAPI	405	Chroma ET455/50m
SiR-Hoechst	640	Chroma ET705/72m
mCherry	561	Chroma ET705/72m
eGFP	488	Chroma ET525/50m

Excitation was done with a Nikon LUN-F XL solid state laser. Images were captured on a Hamamatsu ORCA-Fusion BT CMOS camera (6.5 μm^2 photodiode). Captured images were saved using NIS-Elements.

QUANTIFICATION AND STATISTICAL ANALYSIS

Image analysis for caspase assay

The proprietary MetaXpress software was used to analyze the images. Within this software is a “find round objects” option. This was used to identify the initial nuclei of the cells. These nuclei were modified with the “expand without touching” option set to default settings. The area already defined by the nucleus was taken out using the area algebra feature to define the local area that is not the nucleus of the cell. The software was then set to quantify the FITC (DEVD peptide) and TRITC (Transfection marker) signals within the defined nuclear area and the local background. Each cell was assigned nuclear and background signals for TRITC and FITC values. These values were used to separate transfected from non-transfected cells and to quantify caspase activation.

Quantification of apoptosis signal in transfected cells.

Based on the image analysis described above, each cell was assigned a background subtracted transfection (TRITC) and apoptosis (FITC) value. This value was calculated by subtracting the local background signal from the nuclear signal to get the background

subtracted signal. If the background subtracted transfection signal was above 300, then the cell was considered "transfected." Otherwise, the cell was not considered for this analysis. The median apoptosis signal intensity was calculated by taking the median background subtracted FITC signal of all transfected cells in a particular well. This signal will heretofore be referred to as the apoptosis signal for the well.

Data quality control.

To ensure high quality data, an unequal variances two tailed t-test was performed comparing the XIAP positive control to the empty vector on each plate. If the t-test failed to reject the null hypothesis that the 4 XIAP wells had a lower apoptosis signal than the 4 empty vector wells, then the plate as a whole was discarded.

Data Normalization and processing

Each plate contained 4 replicates of each construct. The apoptosis signal for each well was calculated as described above. The 25th and 75th percentile was calculated. If the top or bottom points were more than 2 interquartile ranges away from these points, that point was identified as an outlier and removed. This value was then divided by an equivalent value from the empty vector run on the same plate (averaged across the 4 wells). The Log_2 of this value was calculated and will heretofore be referred to as the Log_2 Fold Change. Below is an equation showing how this value was calculated:

$$Log_2\ Fold\ Change\ =\ Log_2\left(\frac{Target\ Apoptosis\ signal}{Empty\ Vector\ Apoptosis\ Signal}\right)$$

Based on this analysis, the Log_2 Fold Change was calculated for each construct tested on each plate. This analysis amounts to 3 to 4 datapoints per date a construct was run.

Data aggregation across days

As each plate can only test a maximum of 21 constructs per plate with 4 wells per construct, data across dates were aggregated and analyzed for outliers. For a given construct, if a well had a Log_2 Fold Change more than 2 interquartile ranges away from the 25th or 75th percentile, then that well was removed. The resulting set of Log_2 Fold Changes for each construct are reported in Supplemental Table 2.

Hydrophobic moment calculation

Hydrophobic moment was calculated by assigning residue hydrophobicity based on the Fauchere-Pliska scale and then computing hydrophobic moment according to the formula defined in Eisenberg, 1982 [1-2]. Calculations assume a side-chain angular separation of 100°, typical of alpha helices. Python code to compute the hydrophobic moment was adapted from this script:
https://gist.github.com/JoaoRodrigues/568c845915aea3efa3578babfd72423c.

PFAM domain analysis

Protein sequences were analyzed for domain content using hmmscan (hmmer.org) to search sequences for PFAM-A domains [3]. Domain hits with full-sequence E-values < 0.01 were included.

HHRepID analysis

Protein sequences were analyzed for repeat content using HHRepID [4]. An alignment was built for each protein using 3 iterations of HHblits against uniclust30 [5]. Then HHrepID was used with default parameters to determine repeats coverage. This analysis was used to highlight repetitive sequences that lack pfam domains in Figure 1.

Disorder analysis

Protein sequences were analyzed for disorder content using MobiDB-lite with default parameters [6].

Bibliography & References

1. Fauchere, J.; Pliska, V., Hydrophobic parameters II of amino acid side-chains from the partitioning of N-acetyl-amino acid amides. *Eur. J. Med. Chem.* **1983,** *18*.
2. Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C., The helical hydrophobic moment: a measure of the amphiphilicity of a helix. *Nature* **1982,** *299* (5881), 371-4.
3. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G. A.; Sonnhammer, E. L. L.; Tosatto, S. C. E.; Paladin, L.; Raj, S.; Richardson, L. J.; Finn, R. D.; Bateman, A., Pfam: The protein families database in 2021. *Nucleic Acids Res* **2021,** *49* (D1), D412-D419.

4. Biegert, A.; Soding, J., De novo identification of highly diverged protein repeats by probabilistic consistency. *Bioinformatics* **2008,** *24* (6), 807-14.
5. Steinegger, M.; Meier, M.; Mirdita, M.; Vohringer, H.; Haunsberger, S. J.; Soding, J., HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinformatics* **2019,** *20* (1), 473.
6. Piovesan, D.; Necci, M.; Escobedo, N.; Monzon, A. M.; Hatos, A.; Micetic, I.; Quaglia, F.; Paladin, L.; Ramasamy, P.; Dosztanyi, Z.; Vranken, W. F.; Davey, N. E.; Parisi, G.; Fuxreiter, M.; Tosatto, S. C. E., MobiDB: intrinsically disordered proteins in 2021. *Nucleic Acids Res* **2021,** *49* (D1), D361-D367.