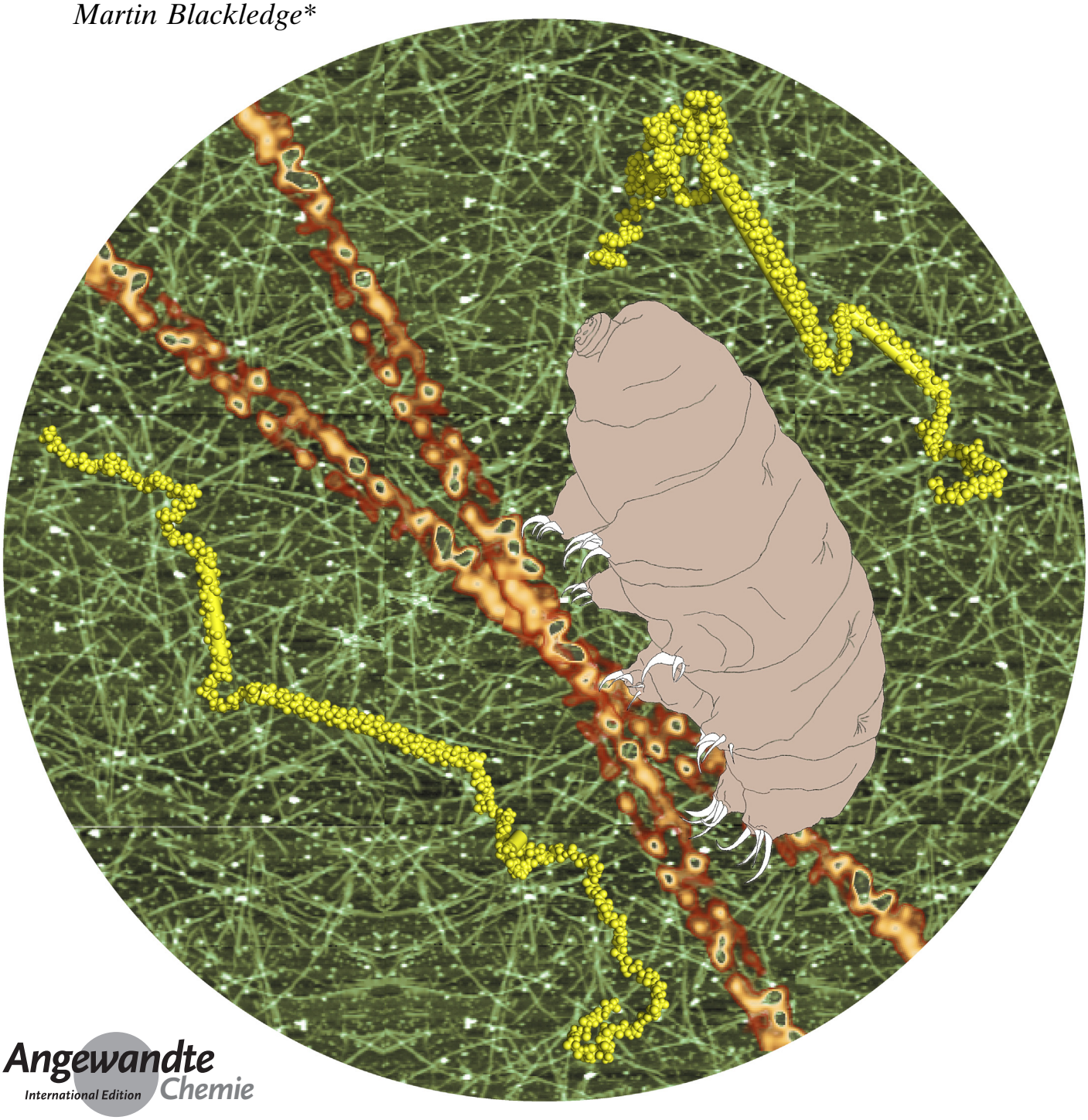
How to cite: *Angew. Chem. Int. Ed.* **2022**, *61*, e202109961
International Edition: doi.org/10.1002/anie.202109961
German Edition: doi.org/10.1002/ange.202109961

Intrinsically Disordered Proteins | Hot Paper

Intrinsically Disordered Tardigrade Proteins Self-Assemble into Fibrous Gels in Response to Environmental Stress

*Anas Malki, Jean-Marie Teulon, Aldo R. Camacho-Zarco, Shu-wen W. Chen, Wiktor Adamski, Damien Maurin, Nicola Salvi, Jean-Luc Pellequer, and Martin Blackledge**



Abstract: *Tardigrades are remarkable for their ability to survive harsh stress conditions as diverse as extreme temperature and desiccation. The molecular mechanisms that confer this unusual resistance to physical stress remain unknown. Recently, tardigrade-unique intrinsically disordered proteins have been shown to play an essential role in tardigrade anhydrobiosis. Here, we characterize the conformational and physical behaviour of CAHS-8 from Hypsibius exemplaris. NMR spectroscopy reveals that the protein comprises an extended central helical domain flanked by disordered termini. Upon concentration, the protein is shown to successively form oligomers, long fibres, and finally gels constituted of fibres in a strongly temperature-dependent manner. The helical domain forms the core of the fibrillar structure, with the disordered termini remaining highly dynamic within the gel. Soluble proteins can be encapsulated within cavities in the gel, maintaining their functional form. The ability to reversibly form fibrous gels may be associated with the enhanced protective properties of these proteins.*

Tardigrades are microscopic aquatic animals[1] that exhibit remarkable resistance to environmental stresses such as extreme temperature, radiation, high pressure, or absence of oxygen or water, a characteristic that has remained mysterious for centuries.[2–4] In response to such stress, tardigrades transform into a characteristic *tun* state that can survive for long periods of time.[5] Although the molecular mechanisms that confer this protection against extreme physiological stress remains poorly understood,[6] recent studies have identified the importance of two families of tardigrade-unique intrinsically disordered proteins (cytosolic abundant heat-soluble (CAHS) proteins and secreted abundant heat-soluble (SAHS) proteins).[7–10] These proteins are collectively known as tardigrade-specific disordered proteins (TDPs).[9] RNA interference demonstrated that TDPs were required for desiccation tolerance in tardigrades, while heterologous expression of certain TDPs demonstrated their ability to increase desiccation tolerance in both yeast and *E. coli*.[9]

[*] A. Malki, J.-M. Teulon, A. R. Camacho-Zarco, W. Adamski, D. Maurin, N. Salvi, J.-L. Pellequer, M. Blackledge
Univ. Grenoble Alpes, CNRS, CEA
Institut de Biologie Structurale
Grenoble (France)
E-mail: martin.blackledge@ibs.fr

S.-w. W. Chen
niChe Lab for Stem Cell and Regenerative Medicine
Department of Biochemical Science and Technology
National (Taiwan) University
Taipei 10617 (Taiwan)

Supporting information and the ORCID identification number for some of the authors of this article can be found under: https://doi.org/10.1002/anie.202109961.

The protective role of TDPs remains under discussion in the literature.[11–13] The observation of the glass transition temperature of TDPs by differential scanning calorimetry was interpreted as evidence that TDPs can reversibly vitrify the cellular environment.[9] Similar phenomena were observed for the intrinsically disordered late embryogenesis abundant (LEA) proteins,[14] through the stabilization and apparent vitrification of co-solutes such as trehalose. This sugar is known to play an important role in anhydrobiosis in a wide range of organisms,[15] although trehalose is not generally abundant in tardigrades.[16]

To better understand how TDPs alone can confer protective properties in response to desiccation and other environmental stresses, we investigated the conformational behaviour of a CAHS from *Hypsibius exemplaris* at atomic resolution. By using high-field NMR spectroscopy, we reveal that the protein comprises an extended central helical domain, with apparent coiled-coil propensity, flanked by long, highly disordered N- and C-terminal domains. Small-angle X-ray scattering (SAXS), atomic force microscopy (AFM), and dynamic light scattering (DLS) were used to follow the successive formation of oligomers, long fibres, and finally gels constituted of fibres. We were able to probe the behaviour of soluble proteins that are encapsulated within the gel matrix, and propose possible mechanisms by which the formation of such fibrous gels can enable anhydrobiosis.

The ^{15}N-^{1}H HSQC spectrum of CAHS-8 (Figure 1 A) is typical of an intrinsically disordered protein, with low dispersion in the $^{1}H^{N}$ dimension. Nevertheless, >98% of the backbone resonances could be assigned (see Methods in the Supporting Information), revealing the presence of a helical domain between residues 95 and 195 (Figure 1 B).

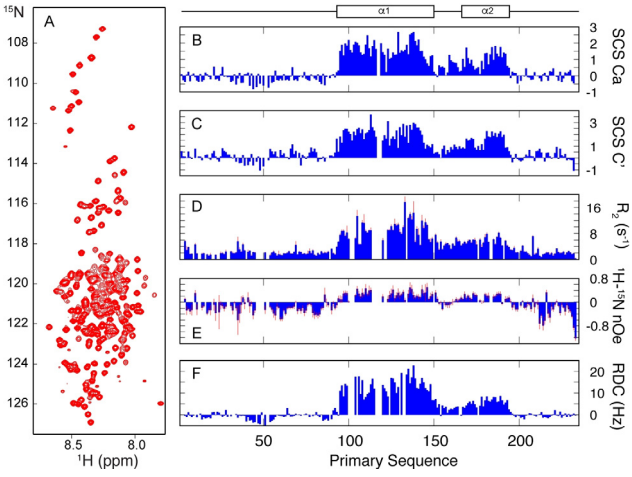


Figure 1. Characterisation of the structural and dynamic properties of CAHS-8 in solution. A) The ^{15}N-^{1}H correlation spectrum is characteristic of an intrinsically disordered protein. B, C) Secondary ^{13}Cα and ^{13}C′ chemical shifts showing strong α-helical propensities from 95–195, divided into two distinct helices, α_1 (E95-F150) and α_2 (V167-L194). ^{15}N relaxation (D; 850 MHz, E; 600 MHz), ^{15}N-^{1}H RDCs (F), and PREs (see Figure S1) suggest the protein acts as a disordered chain comprising two highly populated α-helices (303 K) and no persistent tertiary structure. The NMR spectrum was recorded at 350 μM (9.3 mg mL^{-1}). The NMR assignments are deposited in BMRB (accession code 51115).

This region is divided into two segments, α_1 (E95–F150) and α_2 (V167–L194), with α_2 showing a lower level of helical propensity (Figure 1 B,C; estimated from ^{13}C secondary shifts to be around 60 and 30 %, respectively). The linker between α_1 and α_2 (12 consecutive sites from 152–163, with ^{1}H-^{15}N nOe effects <0.2 and ^{13}C shifts showing a secondary structure population of <0.25) exhibits reduced helical propensity (Figure 1 B,C), and enhanced flexibility (Figure 1 D,E).

The tertiary folding of CAHS-8 was probed using paramagnetic relaxation enhancements, with TEMPO-maleimide labelling of positions A134C and A185C. The absence of significant enhancements beyond immediate neighbours indicates an absence of persistent long-range structure (see Figure S1).

The presence of two helices connected by a more flexible linker is supported by the measurement of ^{15}N-^{1}H residual dipolar couplings (RDCs; see Methods in the Supporting Information). The RDCs show a profile typical of IDPs with helical elements that do not specifically interact with each other, as predicted by simulation and experimentally validated in numerous systems. The RDCs exhibit positive values in the helical regions and lower or negative values in the unstructured parts of the chain (Figure 1 f).[17,18] A model of CAHS-8, comprising a flexible chain with two consecutive helices connected by a short linking peptide and flanked by intrinsically disordered termini was calculated using the disordered protein ensemble generator *flexible-meccano*[19,20] and compared to SAXS measurements. Ensemble selection using the ASTEROIDS approach suggests that the protein samples a largely extended conformation in the monomeric state with an average radius of gyration of 58 ± 7 Å (Figure 2 A). The average radius of gyration $R_{g,av}$ of an ensemble of fully disordered conformers is 47 ± 10 Å, while an ensemble of conformers containing the helical elements ($R_{g,av} = 59 \pm 11$ Å; see Figure S2) coincides with the experimental radius of gyration, thereby supporting the nature of the ensemble of unfolded chains comprising significant populations of helices α_1 and α_2.

CAHS-8 shows a strongly concentration-dependent propensity to oligomerize, as illustrated by SAXS, where scattering curves that are characteristic of fibril formation are observed at higher concentrations (Figure 2 A). The ^{13}C chemical shifts show strong temperature dependence, revealing a significantly higher helical content at 278 K for helix α_2 (see Figure S3). The ^{15}N relaxation is also temperature dependent (Figure 2 C and see also Figure S4), notably higher than that normally observed in intrinsically disordered proteins containing helical elements[21,22] (see Figure S2), thus supporting the formation of higher order oligomers at lower temperature. Dynamic light scattering (DLS) measurements were used to map the phase diagram of CAHS-8 by observing the dependence of the decay in the autocorrelation of scattered intensity as a function of temperature and concentration (Figure 2 B). We observe the formation of an apparent physical hydrogel phase at low temperatures and higher concentration, in qualitative agreement with observations from the NMR and SAXS studies.

We used atomic force microscopy (AFM) in liquid and air modes to observe the apparent formation of fibrils that eventually form a dense mesh that resembles a gel and to investigate the nature of these gels. Extension often follows a twisted helix structure comprising paired filaments, exhibiting a pitch of approximately 600 nm (Figure 3 A). AFM was used to follow gel formation, and revealed the generation of

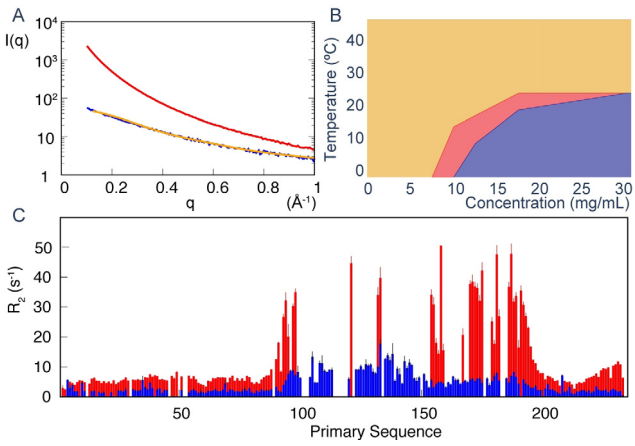


Figure 2. Monomeric CAHS-8 forms higher order oligomers as a function of concentration and temperature. A) SAXS measured at 2 mg mL^{-1} (blue) and 5 mg mL^{-1} (red). At the higher concentration, the curve shows characteristics associated with fibril formation. The orange curve shows the predicted SAXS curve from an ASTEROIDS-selected ensemble of 7 disordered conformers calculated using *flexible-meccano* (average R_g of best-fitting ensembles $= 58 \pm 7$ Å). B) Phase diagram derived from DLS. Orange represents the isotropic phase, red represents increased oligomerization associated with slight gel formation, and blue represents gel formation. C) Transverse ^{15}N relaxation (R_2) measured at 303 K (blue) and 278 K (red) at 850 MHz, showing a very large increase in molecular tumbling at the lower temperature.

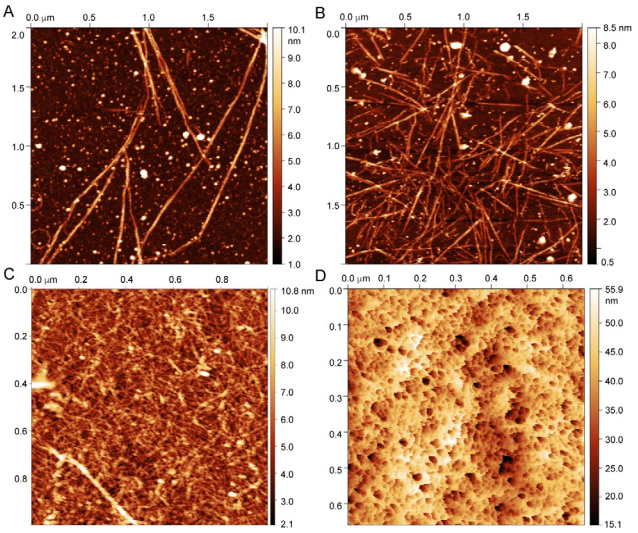


Figure 3. Different states of CAHS-8 imaged by AFM in the liquid (A,B) and air mode (C,D). A) Dissolving the gel on mica revealed fibrils and allowed a kinetic series to be recorded (see Figure 3). B) Allowing fibrillation to continue whilst the sample is drying shows an increasingly dense matrix of fibrils. C) After the kinetic series from (A) was recorded and the sample fully dehydrated, air imaging displays the entanglement of the fibrils forming a gel. D) Placing a 30 mg mL^{-1} gel on glass reveals porous structures, comprising cavities with diameters of tens of nanometers.

fibrillar structures as a function of time (Figure 3B, see also Figure S3), eventually leading to a dense mesh of fibres (Figure 3C). Circular fibrillar structures with a radius of about 150 nm were also formed (see Figure S5). Application of a Laplacian-weight method[23,24] highlights the twist, and again suggests the presence of paired strands forming single fibrils (see Figure S6). Placing concentrated CAHS-8 on glass and cooling resulted in the formation of a porous structure comprising large cavities with diameters of 10–30 nm (Figure 3D).

NMR spectroscopic analysis of the gels formed from ^{15}N-labelled CAHS-8 demonstrates that the intrinsically disordered domains (excluding helix α_1) remain flexible in the context of the fibrillar gel (Figure 4A). No signals from helix α_1 were detected in the ^{15}N-^{1}H HSQC spectrum of a gel formed from ^{15}N-labelled CAHS-8, and ^{15}N spin relaxation rates reveal an overall similar dynamic behaviour of the N-terminal intrinsically disordered region at 303 K as the protein in free solution. The 62–72 strand, comprising both hydrophobic and acidic residues, shows elevated R_2 values, which suggests increased correlation times, possibly arising from interaction with more rigid elements of the gel. The C-terminal domain shows a similar dynamic behaviour to the monomeric phase, with the exception of the terminus, where the non-native TEV-cleavage site (228–233) also shows enhanced R_2 values. Signals from helix α_2 are visible in the spectrum, but show considerably higher relaxation rates than those in free solution (25 s^{-1} compared to 10 s^{-1}), which suggests that this domain is tethered to a much larger, or immobile, object. These observations are compatible with the involvement of α_1 in the formation of fibrils that constitute the gel, with α_2 and the disordered domains remaining flexible within the fibril cavities.

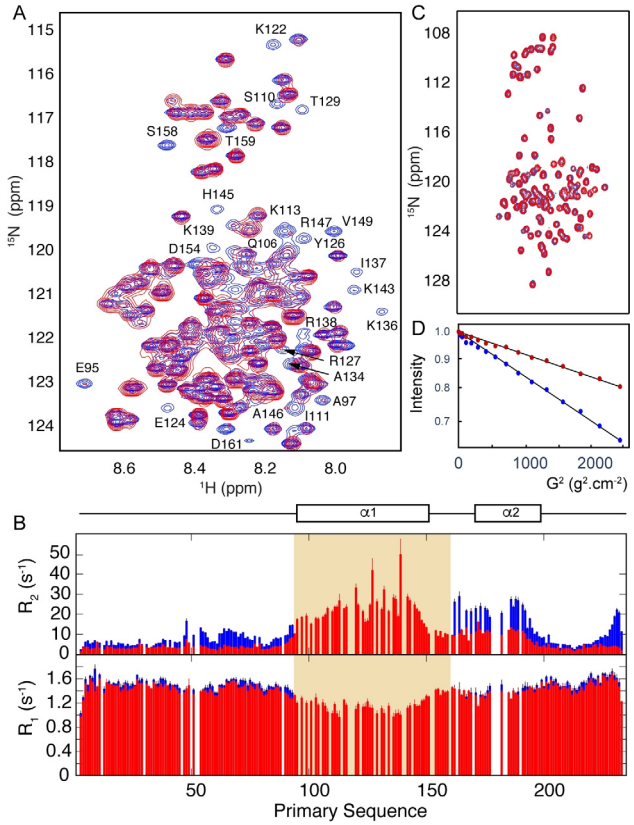


Figure 4. CAHS-8 disordered domains remain flexible in gels that confine soluble proteins. A) Comparison of the ^{15}N-^{1}H correlation spectra (293 K, 850 MHz) of CAHS-8 in monomeric (blue) and gel (red) forms. Residues in helix α1 that disappear in the gel spectrum are annotated. B) Transverse (R_2) and longitudinal (R_1) ^{15}N relaxation (293 K, 850 MHz) of monomeric (red) and gel (blue) CAHS-8. Signals from the shaded region were not visible in the spectrum. C) ^{15}N-^{1}H correlation spectra of the disordered domain of transcription factor ANP32a in buffer (blue) and gel (red) reveal that conformational sampling of ANP32a is unperturbed in gel. D) DOSY measurements of ANP32a in buffer (blue) and gel (red) reveal that translational diffusion of ANP32a is significantly reduced (4.8 compared to 9.6 10^{-11} m^2s^{-1}) in the gel.

To investigate the possible role of gel formation in the establishment of protective properties of CAHS-8, gels were formed in mixtures of the TDP and soluble proteins that were isotopically labelled to allow observation by NMR spectroscopy. The results (illustrated in Figure 4C) demonstrate that the ^{15}N-^{1}H HSQC spectrum of the intrinsically disordered domain of ANP32a,[25] a transcription factor, is essentially unchanged in the gel. The translational diffusion of the protein is reduced by approximately a factor of two compared to standard buffer conditions (Figure 4D). This level of restriction is similar to that measured for a small, folded protein in 6% polyacrylamide gels.[26] Despite this slow diffusion, the conformational sampling of the sequestered protein is maintained in the CAHS-8 gel cavity. Similar results were measured on the intrinsically disordered C-terminal domain of Sendai virus nucleoprotein (N_{TAIL}).[27] Similarly, a folded SH3 domain[28] exhibited an essentially identical HSQC spectrum in gel and isotropic solution (see Figure S7). The intensity of the spectrum is very weak, which suggests a significant fraction is partially immobilised in the gel.

Comparison of the sequence of α_1 with related CAHS sequences reveals highest conservation in the helical domain, which exhibits an irregular coiled coil propensity comprising both canonical heptads and stutters,[29,30] as illustrated for the segment A133–R152 (Figure 5A,B, see also Figure S8). Formation of a coiled-coil conformation has also been associated with desiccation-protectant LEA proteins upon dehydration.[14,31,32]

On the basis of these observations, we propose that the long helical domains of the partially disordered protein associate through coiled-coil-like interactions under certain environmental conditions (low temperature, high concentration) to form oligomers, then fibrils (Figure 5C), which eventually form a gel-like matrix. Both charged and hydrophobic amino acids in this region are strongly conserved over CAHS proteins from different species. NMR relaxation data also show strong evidence that this region is central to the intermolecular interaction, with signals in this region disappearing at lower temperatures, where higher order oligomerization is favoured, and relaxation rates increasing more in this region than elsewhere. Although the molecular structure of the twisted fibril is not known, helix α_1 likely contributes to the structured core, possibly through a super-helical[33] or alternatively a cross-α[34] configuration, while the remainder

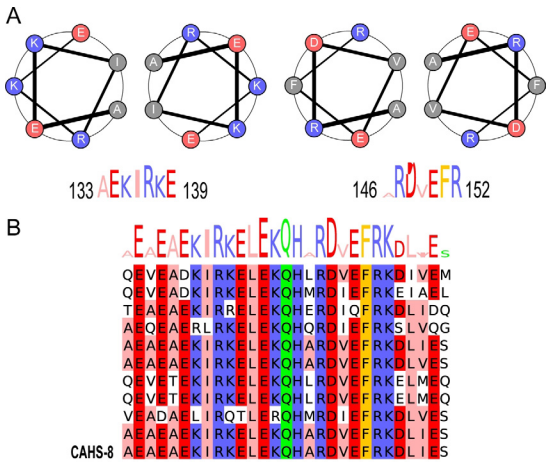


Figure 5. A) CAHS-8 comprises coiled-coil-like sequences located specifically in helix α_1. Two sequences are shown: A133–E139 and A146–R152, which exhibit a typical heptad repeat. B) This propensity is conserved in related sequences (see Figure S8).

of the molecule remains flexible, as shown by NMR spectroscopy.

The exact mechanism by which the formation of such gels protects the organism remains unknown, although the total water volume required to maintain biomolecules in their functional state is clearly reduced in such a matrix. The disordered termini of CAHS-8 remain flexible in the gel and may contribute to the solvation of sequestered biomolecules. In this respect, it is interesting to note that the C-terminal IDR, in particular, is rich in polar amino acids. Cryoprotection may also be facilitated by a reduction in the volume of bulk water, thereby reducing the formation of ice crystals.[35] Under our conditions, we see no evidence of vitrification during the formation of fibrillar hydrogels, as has been observed for trehalose, sucrose, and LEA proteins and proposed as a mechanism of TDP protection.[12,13]

In conclusion, we present a detailed characterization of the transformation of the physical properties of CAHS-8 as a function of environmental conditions. We describe the monomeric behaviour of the protein as partially disordered, with a long two-helical domain flanked by long flexible tails. We propose that this domain mediates homo-dimerization through a coiled-coil interaction, and have experimentally observed the successive formation of oligomers, fibrils, and eventually gels in response to concentration and temperature change. Importantly the gels can sequester soluble proteins in their native state, maintaining their functional form. We propose that the observed transformation of this important protein from soluble monomer to intracellular matrix may be responsible for its protective role in the phenomenal response of tardigrades to environmental stress.

Acknowledgements

This work used the platforms of the Grenoble Instruct-ERIC center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-05-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). We thank Caroline Mas for assistance and access to the Biophysical platform. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG CEA). This work acknowledges the AFM platform at the IBS and the ESRF. M.B. acknowledges support from ERC advanced grant DynamicAssemblies.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: atomic force microscopy · gels · intrinsically disordered proteins · NMR spectroscopy · tardigrades

[1] H. Greven, in *Water Bears: The Biology of Tardigrades* (Ed.: R. O. Schill), Springer International Publishing, **2018**, pp. 1–55.
[2] L. Spallanzani, J. G. Dalyell, M. Baillie, *Tracts on the Nature of Animals and Vegetables.*, William Creech, Edinburgh, **1804**.
[3] W. Wełnicz, M. A. Grohme, L. Kaczmarek, R. O. Schill, M. Frohme, *J. Insect Physiol.* **2011**, *57*, 577–583.
[4] L. Somme, *EJE* **2013**, *93*, 349–357.
[5] R. Bertolani, R. Guidetti, I. K. Jönsson, T. Altiero, D. Boschini, L. Rebecchi, *J. Limnol.* **2004**, *63*, 16–25.
[6] R. O. Schill, B. Mali, T. Dandekar, M. Schnölzer, D. Reuter, M. Frohme, *Biotechnol. Adv.* **2009**, *27*, 348–352.
[7] A. Yamaguchi, S. Tanaka, S. Yamaguchi, H. Kuwahara, C. Takamura, S. Imajoh-Ohmi, D. D. Horikawa, A. Toyoda, T. Katayama, K. Arakawa, A. Fujiyama, T. Kubo, T. Kunieda, *PLoS One* **2012**, *7*, e44209.
[8] T. Hashimoto, D. D. Horikawa, Y. Saito, H. Kuwahara, H. Kozuka-Hata, T. Shin-I, Y. Minakuchi, K. Ohishi, A. Motoyama, T. Aizu, A. Enomoto, K. Kondo, S. Tanaka, Y. Hara, S. Koshikawa, H. Sagara, T. Miura, S.-I. Yokobori, K. Miyagawa, Y. Suzuki, T. Kubo, M. Oyama, Y. Kohara, A. Fujiyama, K. Arakawa, T. Katayama, A. Toyoda, T. Kunieda, *Nat. Commun.* **2016**, *7*, 12808.
[9] T. C. Boothby, H. Tapia, A. H. Brozena, S. Piszkiewicz, A. E. Smith, I. Giovannini, L. Rebecchi, G. J. Pielak, D. Koshland, B. Goldstein, *Mol. Cell* **2017**, *65*, 975–984.
[10] S. Piszkiewicz, K. H. Gunn, O. Warmuth, A. Propst, A. Mehta, K. H. Nguyen, E. Kuhlman, A. J. Guseman, S. S. Stadmiller, T. C. Boothby, S. B. Neher, G. J. Pielak, *Protein Sci.* **2019**, *28*, 941–951.
[11] T. C. Boothby, G. J. Pielak, *BioEssays* **2017**, *39*, 1700119.
[12] K. Arakawa, K. Numata, *Mol. Cell* **2021**, *81*, 409–410.
[13] T. C. Boothby, *Mol. Cell* **2021**, *81*, 411–413.
[14] J. Ingram, D. Bartels, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1996**, *47*, 377–403.
[15] L. M. Crowe, *Comp. Biochem. Physiol. Part A* **2002**, *131*, 505–513.
[16] S. Hengherr, A. G. Heyer, H.-R. Köhler, R. O. Schill, *FEBS J.* **2008**, *275*, 281–288.
[17] L. Salmon, G. Nodet, V. Ozenne, G. Yin, M. R. Jensen, M. Zweckstetter, M. Blackledge, *J. Am. Chem. Soc.* **2010**, *132*, 8407–8418.
[18] M. R. Jensen, M. Zweckstetter, J. Huang, M. Backledge, *Chem. Rev.* **2014**, *114*, 6632–6660.

[19] P. Bernado, L. Blanchard, P. Timmins, D. Marion, R. W. H. Ruigrok, M. Blackledge, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17002 – 17007.
[20] V. Ozenne, F. Bauer, L. Salmon, J.-R. Huang, M. R. Jensen, S. Segard, P. Bernadó, C. Charavay, M. Blackledge, *Bioinformatics* **2012**, *28*, 1463 – 1470.
[21] A. Abyzov, N. Salvi, R. Schneider, D. Maurin, R. W. H. Ruigrok, M. R. Jensen, M. Blackledge, *J. Am. Chem. Soc.* **2016**, *138*, 6240 – 6251.
[22] W. Adamski, N. Salvi, D. Maurin, J. Magnat, S. Milles, M. R. Jensen, A. Abyzov, C. J. Moreau, M. Blackledge, *J. Am. Chem. Soc.* **2019**, *141*, 17817 – 17829.
[23] S.-W. W. Chen, J.-M. Teulon, C. Godon, J.-L. Pellequer, *J. Mol. Recognit.* **2016**, *29*, 51 – 55.
[24] S.-W. W. Chen, A.-S. Banneville, J.-M. Teulon, J. Timmins, J.-L. Pellequer, *Nanoscale* **2020**, *12*, 22628 – 22638.
[25] A. R. Camacho-Zarco, S. Kalayil, D. Maurin, N. Salvi, E. Delaforge, S. Milles, M. R. Jensen, D. J. Hart, S. Cusack, M. Blackledge, *Nat. Commun.* **2020**, *11*, 3656.
[26] H.-J. Sass, G. Musco, S. J. Stahl, P. T. Wingfield, S. Grzesiek, *J. Biomol. NMR* **2000**, *18*, 303 – 309.
[27] M. R. Jensen, K. Houben, E. Lescop, L. Blanchard, R. W. H. Ruigrok, M. Blackledge, *J. Am. Chem. Soc.* **2008**, *130*, 8055 – 8061.
[28] J. L. O. Roldan, M. Blackledge, N. A. J. van Nuland, A. I. Azuaga, *J. Biomol. NMR* **2011**, *50*, 103 – 117.
[29] J. H. Brown, C. Cohen, D. A. Parry, *Proteins* **1996**, *26*, 134 – 145.
[30] A. N. Lupas, M. Gruber, *Adv. Protein Chem.* **2005**, *70*, 37 – 38.
[31] J. Browne, A. Tunnacliffe, A. Burnell, *Nature* **2002**, *416*, 38.
[32] K. Goyal, L. Tisi, A. Basran, J. Browne, A. Burnell, J. Zurdo, A. Tunnacliffe, *J. Biol. Chem.* **2003**, *278*, 12977 – 12984.
[33] S. Mondal, L. Adler-Abramovich, A. Lampel, Y. Bram, S. Lipstman, E. Gazit, *Nat. Commun.* **2015**, *6*, 8615.
[34] E. Tayeb-Fligelman, O. Tabachnikov, A. Moshe, O. Goldshmidt-Tran, M. R. Sawaya, N. Coquelle, J.-P. Colletier, M. Landau, *Science* **2017**, *355*, 831 – 833.
[35] P. Mazur, *Am. J. Physiol. Cell Physiol.* **1984**, *247*, C125 – C142.

Manuscript received: July 26, 2021
Revised manuscript received: November 3, 2021
Accepted manuscript online: November 9, 2021
Version of record online: November 25, 2021