



**HAL**  
open science

## Hfq C-terminal region forms a $\beta$ -rich amyloid-like motif without perturbing the N-terminal Sm-like structure

Mélanie Berbon, Denis Martinez, Estelle Morvan, Axelle Grélard, Brice Kauffmann, Jehan Waeytens, Frank Wien, Véronique Arluison, Birgit Habenstein

### ► To cite this version:

Mélanie Berbon, Denis Martinez, Estelle Morvan, Axelle Grélard, Brice Kauffmann, et al.. Hfq C-terminal region forms a  $\beta$ -rich amyloid-like motif without perturbing the N-terminal Sm-like structure. *Communications Biology*, 2023, 6 (1), pp.1075. 10.1038/s42003-023-05462-1 . hal-04284610

**HAL Id: hal-04284610**

**<https://hal.science/hal-04284610>**

Submitted on 14 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## Hfq C-terminal region forms a $\beta$ -rich amyloid-like motif without perturbing the N-terminal Sm-like structure

Mélanie Berbon<sup>1</sup>, Denis Martinez<sup>1</sup>, Estelle Morvan<sup>2</sup>, Axelle Grélard<sup>1</sup>, Brice Kauffmann<sup>2</sup>, Jehan Waeytens <sup>3</sup>, Frank Wien<sup>4</sup>, Véronique Arluison <sup>5,6</sup>✉ & Birgit Habenstein <sup>1</sup>✉

Hfq is a pleiotropic actor that serves as stress response and virulence factor in the bacterial cell. To execute its multiple functions, Hfq assembles into symmetric torus-shaped hexamers. Extending outward from the hexameric core, Hfq presents a C-terminal region, described as intrinsically disordered in solution. Many aspects of the role and the structure of this region remain unclear. For instance, in its truncated form it can promote amyloid-like filament assembly. Here, we show that a minimal 11-residue motif at the C-terminal end of Hfq assembles into filaments with amyloid characteristics. Our data suggest that the full-length Hfq in its filamentous state contains a similar molecular fingerprint than that of the short  $\beta$ -strand peptide, and that the Sm-core structure is not affected by filament formation. Hfq proteins might thus co-exist in two forms in vivo, either as isolated, soluble hexamers or as self-assembled hexamers through amyloid-reminiscent interactions, modulating Hfq cellular functions.

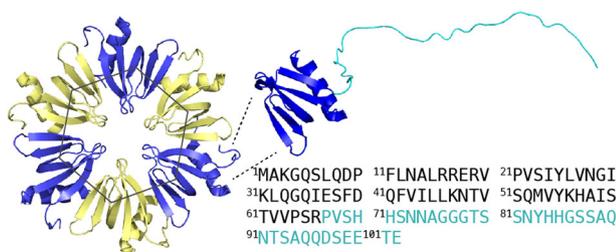
<sup>1</sup>Univ. Bordeaux, CNRS, Bordeaux INP, CBMN, UMR 5248, IECB, Pessac, France. <sup>2</sup>Univ. Bordeaux, CNRS, INSERM, IECB, UAR 3033, Pessac, France.

<sup>3</sup>Structure et Fonction des Membranes Biologiques, Université libre de Bruxelles, Bruxelles, Belgique. <sup>4</sup>Synchrotron SOLEIL, L'Orme des Merisiers, Saint Aubin BP48, 91192 Gif-sur-Yvette, France. <sup>5</sup>Laboratoire Léon Brillouin LLB, UMR12 CEA CNRS, CEA Saclay, 91191 Gif-sur-Yvette, France. <sup>6</sup>Université de Paris Cité, UFR SDV, 75013 Paris, France. ✉email: [veronique.arluison@univ-paris-diderot.fr](mailto:veronique.arluison@univ-paris-diderot.fr); [b.habenstein@cbmn.u-bordeaux.fr](mailto:b.habenstein@cbmn.u-bordeaux.fr)

**H**fq (host factor 1) is a highly conserved protein found in about two-thirds of bacteria<sup>1</sup>. In Gram-negative bacteria, it is implicated in the post-transcriptional stress response through its RNA chaperone function and modulates mRNA translation using small noncoding RNAs (sRNA)<sup>2,3</sup>. Containing an Sm-like globular domain<sup>4</sup>, Hfq assembles into a homo-hexameric ring with structurally unlike surfaces, characterized by an electropositive proximal and a non-polar distal face, both of which can interact with RNA<sup>5–7</sup>, as well as the convex rim<sup>8</sup>. Extending outside this Sm-core, Hfq also comprises a variable C-terminal region (CTR) with an unknown structure in the full protein functional conformation (Fig. 1). These ~40-residue C-terminal in *Escherichia coli* Hfq adopts an intrinsically disordered fold in solution<sup>9</sup>, and none of the numerous Hfq crystal structures reveals its folding<sup>6,10–12</sup>. Nevertheless, it has been shown that numerous functions could be related to the CTR, such as stabilizing the toroidal hexameric structure<sup>13</sup>, sRNA binding<sup>14,15</sup>, single-, double-, and quadruple-stranded DNA binding<sup>16–19</sup>, nucleoid shaping<sup>20</sup>, or membrane poration<sup>21,22</sup>. This can be achieved by adjusting its structure and conferring inter- and intramolecular interactions.

In vitro, Hfq can assemble into filamentous architectures like eukaryotic Sm proteins, including the globular domain in hexameric configuration<sup>23</sup>. In vivo, protein propensity to assemble into higher order polymers containing amyloid structural elements can be associated with deleterious effects, such as in neurodegenerative diseases, or beneficial cellular function as in regulatory or scaffolding processes<sup>24–26</sup>. The Hfq C-terminal conformational adjustments, when confronted with molecular partners, have remained enigmatic, and so has its structure in the Hfq filaments<sup>18,22</sup>. Indeed, the structural determinants located in the C-terminal may contribute to Hfq polymerization and reflect important parameters to regulate the activity of Hfq, possibly also explaining its structural role in nucleic acid interactions, shaping, and competition.

We recently confirmed that a minimal assembling motif of 11-residues (Hfq<sub>87–97</sub>), here denominated Hfq11, is capable of assembling in vitro while fragments excluding the peptide sequence fail to do so<sup>21</sup>. In contrast, Hfq11 does not contribute to the C-terminal nucleic acid binding<sup>27</sup>. We here use magic-angle-spinning (MAS) solid-state nuclear magnetic resonance (NMR) combined with complementary biophysical approaches to shed light on the conformation of the minimal assembling motif Hfq11, isolated or within the context of the full-length Hfq in the assembled filamentous architecture. MAS solid-state NMR constitutes a powerful non-destructive technology to study assembled biomolecules because it provides site-specific information on structures, dynamics, and interactions in the assembled state<sup>28–32</sup>. We complement our analysis using X-ray diffraction, synchrotron



**Fig. 1** *E. coli* Hfq hexamers (PDB code 3QHS) and monomer predicted by AlphaFold<sup>68</sup> (AF-POA6X3-F1) for *E. coli* Hfq including the C-terminal intrinsically disordered region highlighted in cyan on the structure and on the sequence below. The AlphaFold prediction for the C-terminal region indicates a per-residue confidence score (pLDDT) < 90 for all residues with a residue number > Val 68.

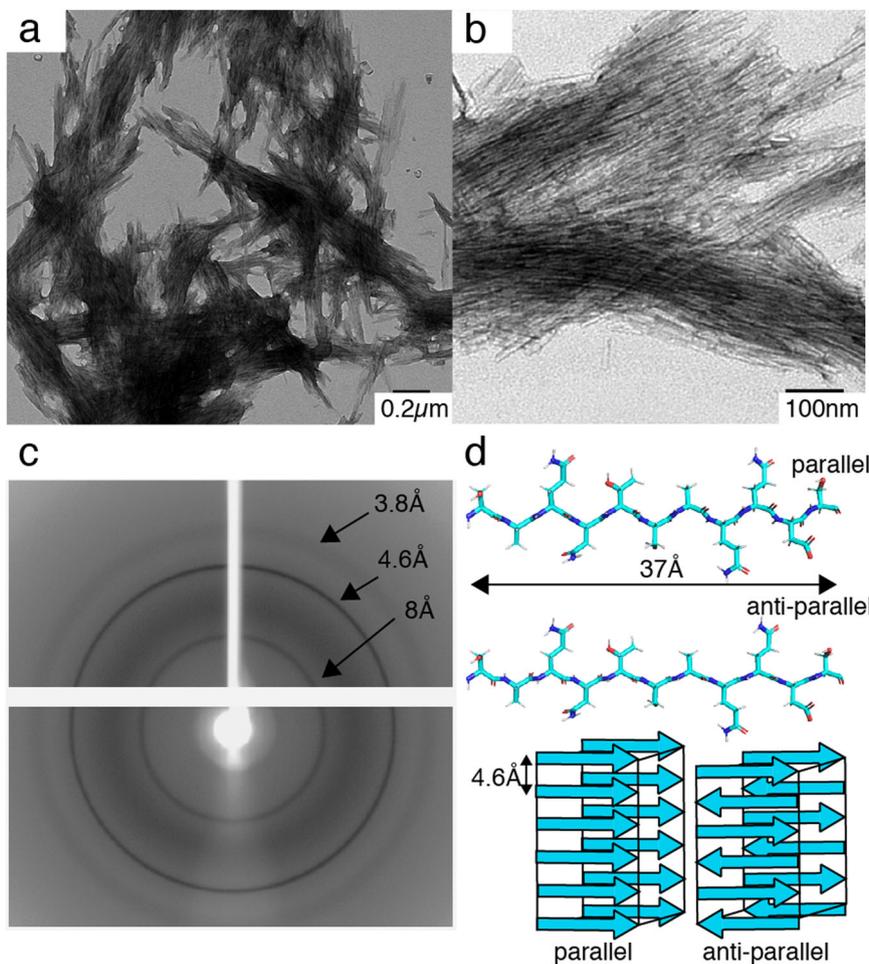
radiation circular dichroism (SRCD), and attenuated total reflectance–Fourier-transform infrared spectroscopy (ATR-FTIR).

Our data indicate the formation of a cross- $\beta$  architecture by Hfq11 in amyloid-like fibers. The peptides are arranged in a well-folded  $\beta$ -strand spine with low flexibility remaining in the primary peptide structure. We further reveal that full-length Hfq monomers assemble into highly ordered structures with a unique conformation in the polymeric assembly. Surprisingly, segments of the CTR are part of the rigid filamentous core in a clearly defined conformation. MAS solid-state NMR data further indicate that the C-terminal segment in the fibrillar core contains the Hfq11 sequence and that this region adopts a similar  $\beta$ -strand fold as in the Hfq11 fragment in isolation.

## Results

**11-residue C-terminal Hfq peptide forms filaments with cross- $\beta$  signature.** We have assembled the minimal Hfq11 motif in water, forming bundles of straight, unbranched fibers of approx. 20 nm (Fig. 2a, b), reminiscent of peptides in amyloid-like fibrils. We then analyzed the fibers by X-ray diffraction, an established technique to detect the presence of a  $\beta$ -amyloid backbone in the powder diffraction pattern of amyloid-like filaments<sup>33–38</sup>. We indeed observe the strongest reflection at 4.6 Å, corresponding to the inter-strand distance between the  $\beta$ -strands along the proto-filament axis (Fig. 2c, d)<sup>39</sup>. The strong hydrogen bonding, typical for the  $\beta$ -amyloid backbone, leads to very repetitive arrangements and explains this intense signal. We observe the typical signal around 7.8–8 Å, which can be assigned to the relatively short inter-sheet distance arising from the tight stacking of the peptide side-chains<sup>36,40</sup>. A more diffuse signal also appears around 3.8 Å. This reflection has been observed multiple times for amyloid-like proteins<sup>36,41,42</sup> and has been attributed to Ca–Ca spacing in the pleated  $\beta$ -strand for fragments of the Amyloid- $\beta$  (A $\beta$ ) protein<sup>35</sup> found in the deposition of amyloid plaques, a hallmark of Alzheimer's disease. In comparison, before washing the sample by centrifugation, the signals at 3.8 Å and 8 Å are doubled (X-ray diffraction shown in Supplementary Fig. 1). The peptide sequence could represent a steric zipper arrangement where multiple interfaces are possible<sup>33,34</sup>, explaining the signal doubling if two peptide arrangements are present (Fig. 2d). Since signal doubling mostly disappears after washing and centrifugation of the peptide, they might arise from non-polymerized oligomeric Hfq11 species. In summary, the data indicate that the 11-residue Hfq peptide can indeed assemble into fibers with cross- $\beta$  architecture.

**11-residue C-terminal Hfq peptides arrange into amyloid-like  $\beta$ -strand filaments.** The technology of MAS solid-state NMR stands out for its power on amyloid-like protein investigations because it can provide site-specific information on the peptide or protein in an assembled filamentous state. We thus chose to apply MAS solid-state NMR on Hfq11 fibers and obtained a highly resolved <sup>13</sup>C-detected cross-polarization (CP) solid-state NMR spectral fingerprint (Fig. 3a). The signal intensity in the <sup>1</sup>H–<sup>13</sup>C CP, relying on a <sup>1</sup>H–<sup>13</sup>C polarization transfer based on dipolar couplings, shows that the peptide residues are immobilized in the amyloid-like fibrils. The resolution and dispersion of the peptide in the CP (Fig. 3a) shows that the peptide is ordered in structural motifs on the atomic level. Ca and C $\beta$  resonances of proteins are exquisite probes to define the secondary structure of the underlying residue<sup>43</sup>. Because only one threonine (Thr) is present in Hfq11 (SAQNT<sub>92</sub>SAQQDS), and the Thr C $\beta$  resonances reveal a unique average chemical shift range in proteins (68–71 ppm)<sup>43</sup>, the analysis of the most downfield shifted resonance in the range of Ca and side-chain resonances reveals atomic information on

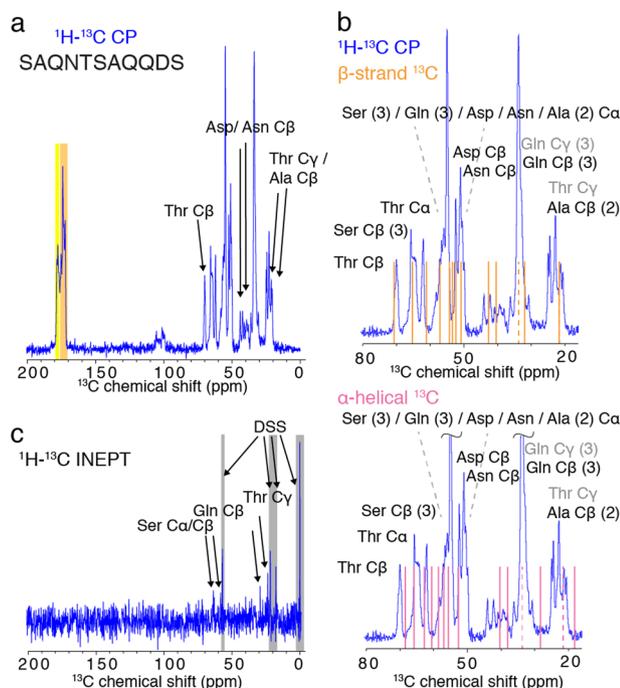


**Fig. 2 Amyloid-like cross- $\beta$  fibers of synthetic Hfq11 filaments. a, b** Transmission electron micrographs **c** X-ray diffraction pattern. **d** Structural model of a monomeric Hfq11 peptide in a parallel or anti-parallel arrangement (PyMol<sup>69</sup>).

Hfq11 Thr C $\beta$ . When comparing the observed Thr C $\beta$  chemical shift (70.1 ppm) to the average Thr chemical shift in  $\alpha$ -helical conformation (68.6 ppm), we can deduce  $\beta$ -strand conformation of the Thr 92 residue (Fig. 3a). Considering the amino acid composition of Hfq11, the C $\beta$  chemical shift of aspartate (Asp) and asparagine (Asn) can further be identified although the secondary structure motif cannot be derived because of the ambiguity to the respective Asp or Asn assignment. The most upfield resonance shifts around 20 ppm represent signals of the alanine (Ala) C $\beta$  in  $\beta$ -strand conformation and Thr C $\gamma$  moieties. No signal is detected below 20 ppm, excluding the presence of Ala C $\beta$  in  $\alpha$ -helical conformation. The downfield signals ranging from 172.3 to 178.3 ppm correspond to resonances of carbonyl ( $^{13}\text{C}=\text{O}$ ) moieties in the amino acid sequence (Fig. 3a and Supplementary 2). Signals centered around 178 ppm, highlighted in yellow in Fig. 3a, should correspond to Gln C $\delta$  (3), Asn C $\gamma$  (1), and Asp C $\gamma$  (1) resonance peaks, whereas backbone carbonyls of a peptide lead to the signals between 172 and 175 ppm. These backbone  $^{13}\text{C}=\text{O}$  resonance shifts clearly indicate a peptide backbone in  $\beta$ -strand conformation because  $^{13}\text{C}=\text{O}$  signals of the amino acids present in Hfq11 should appear upfield of 175.3 ppm (Ala) for a  $\beta$ -strand conformation while arising between 179.5 ppm (Ala) and 176.5 ppm (Ser) if in  $\alpha$ -helical conformation<sup>43</sup>. Peak shape and intensity should contain all signals of 11 backbone  $^{13}\text{C}=\text{O}$  when compared to the signal intensity of the 4 side-chain  $^{13}\text{C}=\text{O}$ . The resonance peaks at 110–100 ppm correspond to spinning sidebands of the  $^{13}\text{C}=\text{O}$  signals (172–178 ppm) due to the MAS frequency of 11 kHz,

corresponding to approx. 73 ppm at a magnetic field strength of 14.1 T.

To further confirm that Hfq11 is assembled into a uniformly  $\beta$ -strand conformation in the entire peptide, we highlighted the expected average chemical shift pattern of the peptide SAQNT-SAQQDS in  $\beta$ -strand conformation (orange lines) and  $\alpha$ -helical conformation (pink lines) on the  $^1\text{H}$ - $^{13}\text{C}$  CP Ca and side-chain Carbon chemical shift range (top and bottom panel in Fig. 3b, respectively). Globally, the  $\beta$ -strand peak predictions (orange) correspond better than the  $\alpha$ -helical peak predictions to the observed signal pattern. When considering the average signal patterns of the Hfq11 primary sequence in pure  $\beta$ -strand vs.  $\alpha$ -helical conformation in detail, several statements can be deduced. (i) In the spectral region of the outmost downfield and upfield resonance shifts, respectively, the Thr C $\beta$  signal is clearly in a chemical shift range, indicating  $\beta$ -strand conformation and no signal of Ala C $\beta$  in  $\alpha$ -helical conformation below 20 ppm is visible. (ii) While the signal cluster around 65 ppm can be explained by C $\beta$  signals of the three serines (Ser) in  $\beta$ -strand conformation, the signals would remain unclear if Ser in  $\alpha$ -helical conformation were present. Only Thr Ca in  $\alpha$ -helical shows a resonance signal at this position and only one Thr is present in the sequence. (iii) The peaks observed around >50 ppm mostly arise from Ala and Asn Ca signals of residues in  $\beta$ -strand conformation and are unexplained if these are in  $\alpha$ -helical conformation. (iv) The same is true for Asp and Asn C $\beta$  signals in 45 ppm. (v) Finally, a signal at chemical shift values <30 ppm is missing for glutamine C $\beta$  moieties if the peptide was in  $\alpha$ -helical



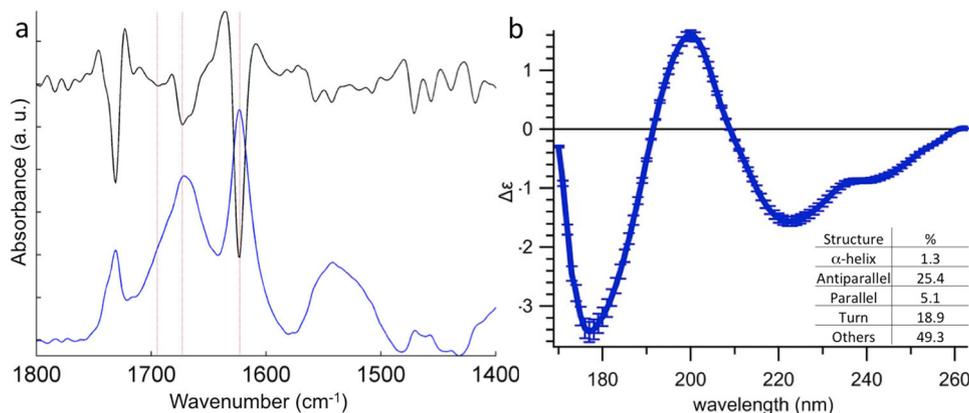
**Fig. 3**  $^{13}\text{C}$ -detected solid-state NMR spectra of Hfq11 on a 14.1 T spectrometer at 11 kHz MAS. **a**  $^1\text{H}$ - $^{13}\text{C}$  cross-polarization (CP) spectrum (blue) detecting signals of atoms in a rigid conformation. In black are annotated unambiguous chemical shift assignments. Highlighted are the spectral regions corresponding to signals of the carbonyl  $^{13}\text{C}=\text{O}$  moieties of, respectively, side-chain  $^{13}\text{C}$  of Gln C $\delta$  (3), Asn C $\gamma$  (1), and Asp C $\gamma$  (yellow) and of the backbone  $^{13}\text{C}=\text{O}$  of amino acids in  $\beta$ -strand conformation<sup>43</sup> (orange), see expanded spectrum in Supplementary Fig. 2. **b**  $^1\text{H}$ - $^{13}\text{C}$  CP spectrum (blue). Orange lines indicate average assignments if the above-annotated corresponding C $\alpha$  or C $\beta$  atom is in  $\beta$ -strand conformation. Pink lines indicate average assignments if the above-annotated corresponding C $\alpha$  or C $\beta$  atom is in  $\alpha$ -helical conformation<sup>43</sup>. Dotted lines and gray annotations correspond to amino acid side-chain assignments that cannot clearly be related to  $\beta$ -strand or  $\alpha$ -helical conformation. **c**  $^1\text{H}$ - $^{13}\text{C}$  INEPT spectrum (blue) detecting signals of atoms with residual flexibility. In black are annotated unambiguous chemical shift assignments. Signals underlined in gray denote assignments to the reference 4-dimethyl-4-silapentane-1-sulphonic acid (DSS).

conformation. In summary, our spectral analysis strongly indicates that the Hfq11 peptide is present in uniformly  $\beta$ -strand conformation. The weak signals in the  $^1\text{H}$ - $^{13}\text{C}$  insensitive nuclei enhanced by polarization transfer INEPT spectrum (Fig. 3c) indicate that only some moieties in the amyloid-like assembly are partly mobile. The averaged chemical shifts can be assigned to Ser C $\alpha$  and C $\beta$ , which are located on both peptide endings, the longer side-chain Thr C $\gamma$  as well as Gln C $\beta$ . Strong signals are detected for the soluble referencing compound 4,4-dimethyl-4-silapentane-1-sulphonic acid (DSS).

We then decided to study Hfq11  $\beta$ -sheet structures by ATR-FTIR and SRCD to determine their arrangements in the amyloid-like structure (Fig. 4). The FTIR spectrum shows three major bands at 1622, 1670 and 1695  $\text{cm}^{-1}$  in amide I region (Fig. 4a). The first FTIR band at 1622  $\text{cm}^{-1}$  indicates the presence of intermolecular  $\beta$ -sheet structure from amyloid fibrils<sup>44,45</sup>. The 1695  $\text{cm}^{-1}$  band is indicative of the antiparallel arrangement of the individual  $\beta$ -strands<sup>46</sup>. As for the amide II region, we also observe a band at 1550  $\text{cm}^{-1}$  characteristic for parallel  $\beta$ -strand arrangement<sup>47</sup>. The band around 1740  $\text{cm}^{-1}$  corresponds to the C=O stretching vibration of an ester. As active esters are often

used for solid-phase peptide synthesis, we suspect this band coming from slight contamination occurring during peptide synthesis. Since this ester band is located outside of the amide I band, used for the characterization of the secondary structure, it does not interfere with our results. The SRCD spectrum (Fig. 4b) indicates a  $\beta$ -sheet structure with a content of 25% antiparallel and only 5% parallel arrangement confirming the FTIR measurements. In both FTIR and SRCD, the signals of random structure might arise from non-polymerized monomeric or smaller oligomeric peptides. This hypothesis is confirmed by ATR-FTIR measurement since we observe a higher contribution of different conformational states when fibers are deposited on the ATR surface without a subsequent washing step, eliminating non-polymerized monomers (Supplementary Fig. 3). Note that these non-polymerized species are likely not detected in solid-state NMR, due to the centrifugation before rotor filling that selects high-molecular-weight species such as fibers.

**Full-length Hfq filaments contain a rigid globular domain and segments of the CTR.** We then decided to extend our structural analysis by solid-state NMR to full-length  $^{13}\text{C}$  isotope-labeled Hfq fibrils. Note that while isolated CTR self-assembly in the absence of cofactor is usually slow and may take a few weeks<sup>48</sup>, the presence of the N-terminal Sm-core accelerates this process. We produced and assembled fully  $^{13}\text{C}$ -labeled Hfq and recorded a  $^{13}\text{C}$ -detected  $^{13}\text{C}$ - $^{13}\text{C}$  proton-driven spin diffusion (PDS) MAS solid-state NMR spectrum. This experiment gives principally access to intra-residual  $^{13}\text{C}$ - $^{13}\text{C}$  correlation signals from  $^{13}\text{C}$  atoms in a rigid conformation (Fig. 5). Signal intensity, resolution, and dispersion of the signals indicate, respectively, an immobile, homogeneously ordered monomer structure containing  $\alpha$ -helical and  $\beta$ -strand structural motifs, in the assembled filaments. Based on previous analysis<sup>23</sup>, we presumed a conserved monomer globular domain in the structural assembly. We thus predicted the C $\alpha$  and C $\beta$  resonances for the monomeric Hfq globular domain structure (PDB code 4RCB) using the SPARTA + algorithm<sup>49</sup> and plotted the  $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$  correlation peaks on the spectrum (Fig. 5a). The predicted resonance peaks correspond to the observed correlation peak pattern in the  $^{13}\text{C}\alpha$ - $^{13}\text{C}\beta$  spectral region, indicating that the globular domain structure is maintained in the filamentous assembly. To evaluate the sensitivity of chemical shift predictions to structural adaptations of the globular domain, we predicted the chemical shifts for the hexameric Hfq arrangement (PDB code 3QHS). Our data show that the limited structural adaptations of the globular structure between the monomeric and the hexameric arrangement lead to significant modifications of the chemical shift predictions such as for Thr 49, located in the loop connecting  $\beta$ -strand 3 and 4, with a very similar structure in the monomer and the hexamer and a chemical shift difference of 1.7 ppm for C $\alpha$  as illustrated in Supplementary Fig. 4. We then assigned spin systems that could clearly identify the globular structure and concentrated therefore on isoleucines (Ile), exclusively present in the globular domain.  $^{13}\text{C}$ - $^{13}\text{C}$  correlation peaks of five Ile spin systems can be assigned (Fig. 5b), colored in a range from red to green with increasing  $\beta$ -strand structural propensity of the underlying residue. Only five Ile are present in full-length Hfq and all detected spin systems display intense signals for all residue atoms, indicating that all five Ile distributed over the Hfq sequence are immobilized in the assembly. The residues are indicated in red on the hexameric crystal structure of the full Hfq (PDB code 3QHS) in Fig. 5c (note that the CTR was not resolved in this structure). The Ile is well-distributed over the primary sequence of the globular domain, indicating that not only parts of the globular structure are visible in the  $^{13}\text{C}$ - $^{13}\text{C}$  PDS NMR spectrum. The side chains of Ile30



**Fig. 4** Secondary structure of the Hfq11 fibrils by ATR-FTIR and SRCD spectra of Hfq11. **a** The ATR-FTIR spectrum in blue and the second derivative in black. The red lines correspond to 1695, 1670, and 1622  $\text{cm}^{-1}$  wavenumbers for, respectively, antiparallel, random, and amyloid  $\beta$ -sheet structures. Note also the band at 1550  $\text{cm}^{-1}$  in amide II, characteristic of the presence of some parallel  $\beta$ -sheets. **b** The SRCD spectrum and the corresponding secondary structure analysis by Betsel indicate the content of 5% of parallel  $\beta$ -sheet and 25% of antiparallel  $\beta$ -sheet with an NMRSD of 0.058 for the spectral fitting.

that are oriented towards the hexameric surface seem to be packed in the filamentous assembly so that they can be detected in the spectrum. Peak positions for 4–5 Ala residues are indicated for the  $^{13}\text{C}\alpha$ – $^{13}\text{C}\beta$  spin systems (blue), thus revealing the presence of structured parts outside of the globular domain in the filamentous assembly of full-length Hfq because the globular domain contains only 3 Ala. This is confirmed by the signals of 4 Thr spin systems, highlighted by cyan lines, with only two Thr in the globular domain, one exposed to the hexameric surface. This corroborates the assumption that even exposed side chains in the globular domain remain immobilized, based on the rigidified side chains of Ile30.

Since Hfq associates with the nucleic acids<sup>50</sup> but also localizes in the vicinity of membranes<sup>51</sup>, we tested the impact of the individual negatively charged lipid component Dioleoylphosphatidylglycerol (DOPG) on C-terminal amyloid stability by FTIR (Supplementary Fig. 5). Over an evolution period of 12 h the fibers disassemble in the presence of the DOPG lipids, indicating the possible modulation of Hfq amyloid assembly by cellular lipid components.

#### Hfq11 fibers and full-length Hfq filaments contain $\beta$ -motifs.

We have shown that the C-terminal domain, encompassing the minimal motif Hfq11, exhibits a propensity to assemble into a cross- $\beta$  spine arrangement. We further revealed that the full-length Hfq assembly contains intact globular domain structures and stretches of the C-terminal domain in a rigid and well-ordered conformation. Because solid-state NMR signals are very sensitive to the local electrochemical environment of the detected nucleus, depending strongly on the secondary protein structure, we compared the  $^{13}\text{C}$  signal pattern obtained for Hfq11 with the two-dimensional  $^{13}\text{C}$ – $^{13}\text{C}$  PDS of the full-length Hfq. All Hfq11 signals could correspond to signals detected in the PDS of full-length Hfq, as clearly exemplified for the spectral regions of Thr, Ser, and Ala (Fig. 6a). Traces reflecting the Thr signals in the full-length Hfq PDS with respect to Hfq11 signals demonstrate the striking solid-state NMR signal comparability. Whereas the signals of the Thr  $\text{C}\alpha$ – $\text{C}\beta$  cross-peak detected at the  $\text{C}\beta$  resonance frequency with a chemical shift of 70 ppm correspond perfectly to the signals detected in Hfq11, this is not the case for the other Thr as exemplified for the Thr  $\text{C}\alpha$ – $\text{C}\beta$  cross-peak at the  $\text{C}\beta$  resonance frequency with a chemical shift of 69 ppm (Fig. 6b). Our data thus indicate that the Hfq11 peptide retains its  $\beta$ -strand conformation in the full-length Hfq assembly.

#### Discussion

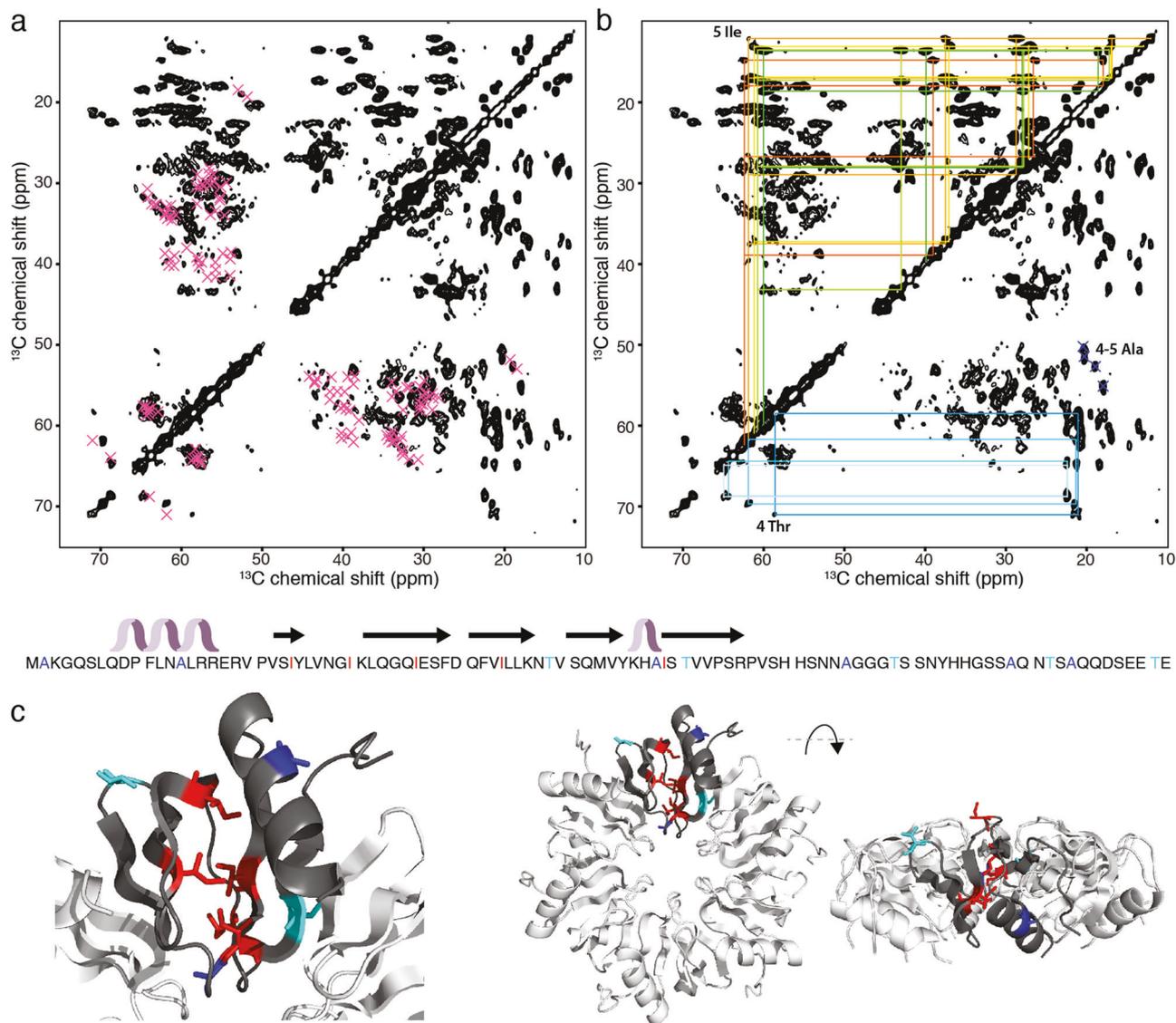
The respective roles of Hfq's structural components, namely the Sm-like globular domain and the C-terminal intrinsically disordered region, are still unclear. Notably, in *E. coli*, the about 40-residue spanning CTR arranges partially disordered in solution<sup>9</sup> but self-assembles when in isolation<sup>48</sup>. As described above, suggested functions of the CTR include sRNA-based regulation<sup>14,15</sup>, DNA binding<sup>16–19</sup>, membrane poration<sup>21,22,52</sup>, or molecular condensation<sup>53,54</sup>.

Interestingly, the evolutionary conservation of the C-terminal amino acid sequence indicates the appearance of two Hfq categories, (1) Hfq proteins of  $\gamma$ - and  $\beta$ -proteobacteria that have an extended C-terminus and (2) other Hfq proteins lacking a CTR, including those of Gram+ bacteria such as *S. aureus* (see sequence alignments<sup>1,27</sup>, Supplementary Fig. 6).

We decided to investigate the determinants of *E. coli* Hfq self-assembly using a divide-and-conquer approach at a site-specific level, focusing on the structural implications of the amyloidogenic CTR.

We report the molecular structure of the minimal assembling C-terminal motif of Hfq, the 11-residue peptide denominated Hfq11, spanning from residue 87–97. Removing Q<sub>90</sub>N<sub>91</sub> and Q<sub>95</sub>Q<sub>96</sub> sequences abolishes the self-assembly, S<sub>65</sub>RPVSHSNNAGGGT<sub>80</sub> by itself is not able to polymerize<sup>21</sup> and mutating the Serines S<sub>88</sub>, S<sub>93</sub> or S<sub>98</sub> in the 11-mer region for an Alanine abolishes the self-assembly<sup>27</sup>. Thus, the length and nature of the sequence are important for the self-assembly<sup>27</sup>. By electron microscopy, we confirm that Hfq11 forms around 20 nm wide filaments. Because our solid-state NMR data indicate that no turn in the monomer structure is present, the 10–20 nm width could reflect the bundling of 4 protofilaments stretching over 37 Å. The solid-state NMR resonance assignment and the X-ray diffraction analysis suggest that Hfq11 assembles into a well-ordered cross- $\beta$  architecture in the assembled filaments. This amyloid cross- $\beta$  structure was further confirmed by SRCD and FTIR and determined with a mixture composed mainly of antiparallel conformation. Considering that this structural segment does not seem to contribute to RNA interactions<sup>55</sup> or DNA binding<sup>27</sup>, its propensity to assemble into  $\beta$ -strand amyloid-like arrangements indicates a potential regulative role through self-interactions.

Solid-state NMR data reveal that full-length Hfq filaments contain the globular Sm-domain in a homogeneously ordered unique structure. Our data reveal segments of the CTR in an equally ordered homogeneous structure in the full-length

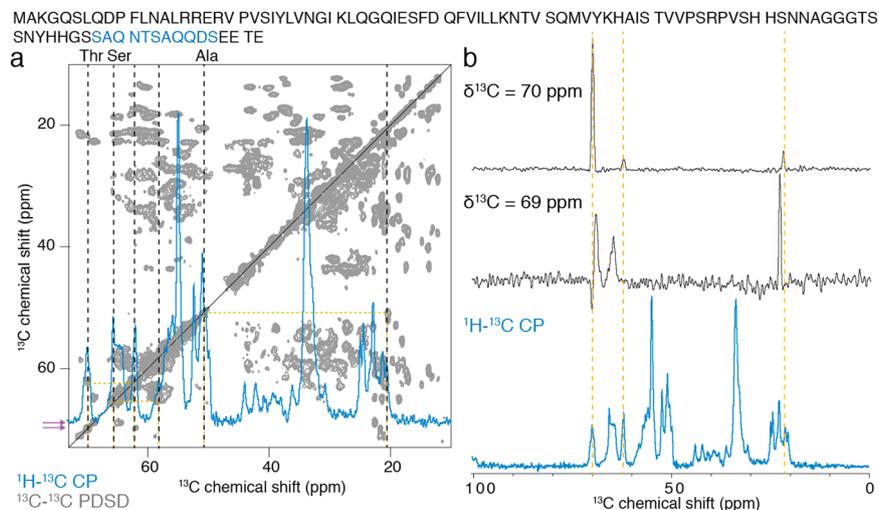


**Fig. 5** Solid-state NMR (ssNMR) analysis of full-length Hfq fibers. **a, b** Aliphatic region of a two-dimensional  $^{13}\text{C}$ -detected proton-driven spin diffusion (PDSF) ssNMR spectra of Hfq fibers at 11 kHz MAS on a 14.1 T spectrometer (50 ms mixing time). **a** Resonance peaks (violet) of Hfq monomer structure (PDB code: 4RCB, chemical shift predictions from Sparta+<sup>49</sup>), lacking the last 30 residues, on Hfq fiber spectrum, **b** on the left side of the diagonal, 5 Ile spin system chemical shift assignments ranging in rainbow colors from green to red (increasing  $\beta$ -strand propensities of the Ile residue) are indicated, on the right side 4 Thr assignments in shades of cyan (increasing  $\beta$ -strand propensities from light to dark cyan) and 4-5 Ala assignments (blue) are shown on the spectral fingerprint. Hfq primary and secondary sequence and amino acids corresponding to the assigned spin systems are shown below in the corresponding color scheme (red—Ile, cyan—Thr, blue—Ala). **c** Amino acids of assigned spin systems are depicted in sticks (color scheme as in (b)) on the hexameric Hfq structure, lacking the last 30 residues, with a view on the proximal face and the rim (PDB code: 3QHS). Sparta+ chemical shift predictions for PDB code 3QHS and 4RCB are provided in Supplementary Data 1 and 2, respectively.

filaments, with  $\beta$ -strand secondary structure propensity. The assignment of the spin system signals arising from the CTR cannot be simply attributed to a specific segment because the underlying amino acids are present in multiple segments of the CTR. When comparing the Hfq11 signal pattern with the full-length Hfq spectral fingerprint, the Hfq11 signals are found in the Hfq filaments pointing to a contribution of intermolecular interactions and immobilization of these residues. Previously reported as unstructured in solution, we here provide evidence for a structuration of the C-terminal domain in full-length Hfq. Solution NMR had shown that the N-terminal globular domain is conserved within soluble Hfq containing the CTR in a partially disordered conformation<sup>9</sup>. However, the soluble form of full-length Hfq shows detectable conformational differences when compared to the globular domain in isolation. This is, for

example, the case for the residues preceding the segment 65–70, which is capable of folding back onto the proximal face of the hexamers (PDB code 1HK9)<sup>56</sup>. Soluble Hfq might therefore contain the globular domain within a similar structure as described in crystalline hexamers. However, also in the soluble state, Hfq adopts an adapted conformation with parts of the CTR in a not fully disordered state and, therefore, undetectable by solution NMR methods. Our data indicate that the structured C-terminal segments of the full-length Hfq might contain residues 87–97, adopting a tightly packed  $\beta$ -strand conformation in assembled filaments.

In conclusion, here we show that Hfq11 assembles into highly structured amyloid filaments that arrange into stacked  $\beta$ -strands on the atomic level. Furthermore, we have evidence that full-length Hfq contains a  $\beta$ -strand motif in its CTR, reminiscent of an



**Fig. 6 Comparison of MAS solid-state NMR analysis of Hfq11 (blue) vs. full-length Hfq (gray).** The full-length Hfq primary sequence is listed in black, with the Hfq11 segment highlighted in blue. **a** Overlay of the  $^1\text{H}$ - $^{13}\text{C}$  CP spectrum of Hfq11 with the aliphatic region of the  $^{13}\text{C}$ - $^{13}\text{C}$  PDSD of full-length Hfq. Dotted black lines indicate  $\text{C}\alpha$  and  $\text{C}\beta$  signals of Thr, Ser, and Ala residues detected in Hfq11, and dotted yellow lines correspond  $\text{C}\alpha$ - $\text{C}\beta$  cross-peaks. Potential assignments are plotted above the spectrum. Pink arrows on the y-axis of the PDSD spectrum indicate the trace position shown in **(b)**. **b** Traces of two Thr signals in the  $^{13}\text{C}$ - $^{13}\text{C}$  PDSD of full-length Hfq plotted above the  $^1\text{H}$ - $^{13}\text{C}$  CP spectrum of Hfq11. Yellow dotted lines indicate the corresponding Thr signals.

amyloid conformation and that the  $\beta$ -strand structure forms without perturbing the N-terminal Sm-core. We previously reported that in vitro, Hfq can assemble into filamentous architectures, including the globular domain in a hexameric configuration, and reveal signals that can possibly be attributed to amyloid formation<sup>23</sup>. Our data now indicate that this assembly contains the 11 amino-acid residues in the CTR, allowing the packing of Sm tori through interactions also driven by the CTR. This goes in line with previously observed FTIR data indicating a potential amyloid-like  $\beta$ -sheet motif in full-length fibers<sup>23,57</sup>. Interestingly, the presence of the N-terminal region accelerates the process of self-assembly compared to isolated CTRs<sup>48</sup>. The in vitro observations of oligomeric assemblies containing the CTR and well-folded hexamers suggest that the interactions promoting the molecular assembly might arise transiently *in cellulo*<sup>58</sup> without disturbing the Sm-core structure and functions while still possibly regulating the activity by steric occlusion. Our data combined with the previous electron microscopy reconstructions and the increase in  $\beta$ -sheet content in Hfq assemblies<sup>23</sup> lead us to suspect a filament architecture containing amyloid-like structural elements and the conserved toroidal rings, schematized in Supplementary Fig. 7. The filament has a helical diameter of 170 Å in which the globular domains of Hfq (illustrated in Supplementary Fig. 7 as blue plates representing the hexagonal monomer arrangements) form repetitive hexameric architectures, leaving a space of approximately 60 Å inside the filament. This interior space could accommodate several arrangements of the CTR, such as a single polymerized intermolecular  $\beta$ -sheet or two complementary  $\beta$ -sheets in an intermolecular amyloid zipper<sup>34,59,60</sup>. We therefore propose the CTRs (here in red) to interact together to provide a backbone architecture, allowing the formation of Hfq filaments. This observation is in agreement with the recent observation that flexible CTR extensions of Hfq form a dense network<sup>61</sup>. These non-covalent CTR amyloid-like interactions could also reflect the conformations adopted in the recently reported single-component and multiple-component elongated Hfq assemblies<sup>53</sup>. Since each hexamer harbors six CTRs (not all represented in the self-assembly of hexamers in Supplementary Fig. 7), the CTRs pointing outside of the filament core cannot contribute to an interior amyloid-like assembly. CTRs, therefore, possibly further stabilize monomer interactions in the

hexamer, for instance, via the acidic tip at the end of each CTR interacting with basic patches inside the Hfq globular domain<sup>61</sup> or interconnecting outside the filament core. The observed intra-hexamer interactions could potentially be maintained while still accommodating amyloid-like self-assembly along the filament axis. Note that Hfq filaments cannot form in the absence of the CTR, illustrating the central role of this region to promote full-length Hfq self-assembly<sup>23</sup>.

Since Hfq has further been found to drive molecular condensation<sup>53,54</sup>, to adopt an amyloid-like structuration in cells<sup>58</sup> and to compact DNA depending on the C-terminal Hfq region<sup>20,53</sup>, we propose that the here-presented amyloid structuration might govern these local inter-molecular interactions. The propensity of Hfq to assemble into amyloid-like structures and the stability of the resulting assemblies could depend on its cellular localization and the local physico-chemical environment. For example, we here show that the presence of lipids such as DOPG modifies the capacity of the CTR to form amyloid structures. Furthermore, the direct interaction of Polyphosphates (PolyP) with the CTR promotes Hfq condensation<sup>53</sup> and could thus induce amyloid formation; possibly by interacting with the two positively charged Histidine clusters preceding Hfq11. Indeed, PolyP's have been associated with amyloid-inducing effects for disease-related and functional amyloids such as  $\alpha$ -synuclein and CsgA<sup>62,63</sup>. The proposed intermolecular interactions between Hfq and PolyP and the here-described amyloid nucleation could provide a structural perspective on the biological regulation of heterochromatin formation in bacteria governed by PolyP-Hfq-DNA<sup>53</sup> interactions.

Moreover, the fact that the eleven-residue sequence contains four residues that can be phosphorylated (three Ser and one Thr) and one residue whose charge state is sensitive to pH (Asp), could further suggest a regulative mechanism of amyloid formation. While Hfq phosphorylation has to our knowledge, not yet been reported, this or other posttranslational modifications, such as acylation<sup>64</sup>, could be relevant to regulate its functions.

## Materials and methods

**Protein expression, purification, and assembly.** *E. coli* BL21-DE3 were transformed with a *pTE607* vector containing the DNA

encoding for Hfq<sup>13</sup> and plated onto LB-agar plates containing 100 µg/mL ampicillin. A pre-culture of 100 mL LB medium was inoculated with a single transformed colony and incubated at 37 °C overnight (about 20 h). 1 L of LB medium is inoculated with the pre-culture at OD<sub>600</sub> = 0.2 and incubated at 37 °C until OD<sub>600</sub> = 0.7–0.8. Protein production was induced with 0.75 mM of IPTG at 30 °C for 20 h. Cells were pelleted at 6000g for 30 min at 4 °C and resuspended in 10 mL lysis buffer (500 mM NaCl, 20 mM Tris-HCl pH 8, 10% Glycerol) containing protease inhibitors (Complete, Roche). Cells were sonicated on ice at 30% magnitude three times (30 s on, 30 s off) and centrifuged at 15,000g for 30 min at 4 °C. The pellet was resuspended in lysis buffer (500 mM NaCl, 20 mM Tris-HCl pH 8, 10% glycerol) containing 1% Triton X-100 and incubated overnight at 4 °C before centrifuging at 250,000g. The supernatant was incubated at 80 °C for 15 min, followed by an incubation on ice for 3 min, and centrifuged at 15,000g for 15 min at 4 °C. The supernatant was incubated at room temperature after adding RNase (10 µg/mL) and DNase (50 U/mL) for 1 h. Purification of Hfq was achieved with an Akta Pure 25 HPLC system (GE Healthcare) on a HisTrap affinity column equilibrated in Buffer A (20 mM Tris-HCl pH 8, 300 mM NaCl, 10 mM imidazole). The supernatant was added to the column and washed with Buffer A to remove non-specifically bound proteins. The protein was eluted by increasing the concentration of imidazole in three steps; 25 mM, 50 mM, and 100 mM. As shown in Supplementary Fig. 8, Hfq hexameric species are visible on the SDS gel after purification and therefore resist SDS denaturation. Once the purification was performed, 5 mM EDTA was immediately added to the solution to avoid precipitation of protein. After purification, proteins were buffer-exchanged against a reconstitution buffer (50 mM Tris-HCl, 50 mM NH<sub>4</sub>Cl, 10% glycerol, pH = 7.4) using dialysis. Proteins were left to spontaneously assemble at a concentration of 4 mg/ml for 1 week at room temperature under mild agitation. Assembled Hfq was centrifuged and filled into the ssNMR rotor. Hfq 11-residues (SAQNTSAQQDS) and Hfq 38-residues (SRPVSHSNNAGGGTSSNYHHGSSAQNTSAQQDSEETE) peptides were chemically synthesized (Proteogenix) and assembled in water at 4 °C at 20 mg/mL for >1 month. When needed for the experiment (FTIR), H<sub>2</sub>O would be replaced by D<sub>2</sub>O.

**X-ray diffraction.** The diffraction patterns were measured at 270 K on a Rigaku FRX rotating anode X-ray generator at the copper wavelength (K $\alpha$ ,  $\lambda$  = 1.54 Å). The source is equipped with Osmic Varimax HF optics and a Rigaku® HyPix6000 detector on a  $2\theta$  arm of a Rigaku partial chi AFC11 goniometer. The samples were mounted in MicroLoops from MiTeGen on a goniometer head under the cryostream nitrogen flux. The diffraction patterns correspond to a 360° rotation along the phi axis (perpendicular to the direct beam with omega and chi axes at the 0 position) with an exposure time of 360 s. Data were integrated with CrysAlisPro (Rigaku Oxford Diffraction, Ltd., Yarnton, Oxfordshire, England). After performing the first X-ray diffraction experiments, we used the remaining sample of Hfq11 filaments, resuspended the filaments in H<sub>2</sub>O, and centrifuged the filaments to discard the water. The procedure was repeated multiple times to record a second set of diffraction data.

**Transmission electron microscopy.** One droplet of Hfq peptide filaments was applied to glow-discharged 300 mesh carbon-coated copper grids for 1 min, washed with water, stained with 2% uranyl acetate (w/v) for 1 min, and dried under dark conditions. Samples were observed using a FEI CM120 transmission electron microscope at an accelerating voltage of 120 kV under

TEM low-dose mode. TEM images were recorded using a Gatan USC1000 2k × 2k camera.

**SRCD.** For SRCD analyses, measurements and data collection were carried out on the DISCO beamline at Synchrotron SOLEIL, as described previously in Partouche et al.<sup>65</sup>. Protein secondary structure was determined with BeStSel software<sup>66</sup>.

**ATR-FTIR.** Attenuated total reflection Fourier-transform infrared spectroscopy (ATR-FTIR) measurements were done on a Bruker Equinox55 spectrophotometer equipped with an MCT detector cooled with liquid nitrogen. 2 µL of the sample was deposited and dried with a nitrogen flow on a Specac diamond and 128 scans were averaged for sample and background. To remove most soluble or small oligomeric Hfq11 peptides (non-polymerized), the deposited sample was washed with a drop of 20 µL of water. The water was removed by pipetting without touching the sample and the sample was finally dried by nitrogen. For measurements of Hfq CTR amyloid structure on lipids, a drop of DOPG lipid solution at 0.5 mg/mL in chloroform (Avanti Polar Lipids, Alabaster, AL, USA) was first deposited on the IRE and dried with a nitrogen flux. 5 µL of D<sub>2</sub>O from Cambridge Isotope was added on top and finally, 0.5 µL of assembled Hfq CTR at 20 mg/mL in D<sub>2</sub>O was added. The system used was a Bruker Equinox55 purged with dry air and a diamond ATR device with a single reflection at an angle of 45° and closed with a golden gate chamber from Specac (Orpington, UK) to avoid evaporation and non-deuterated water vapor exchange. Spectra with a resolution of 4 cm<sup>-1</sup> were acquired every 90 min overnight.

**Solid-state NMR.** Solid-state NMR experiments were performed on a 14.1 T Bruker Biospin spectrometer equipped with 4 mm triple resonance (<sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N) MAS probes. DSS (4,4-dimethyl-4-silapentane-1-sulphonic acid) was added to the sample for internal referencing. The spinning frequency was set to 11 kHz (4 mm rotor). The BCU temperature was set to 273 K for Hfq and 276 K for Hfq11, resulting in sample temperatures of 4 and 11 °C, respectively. All spectra were processed with Bruker TopSpin 4.2.0 and analyzed with the CcpNmr 2.4.2 software. The 2D <sup>13</sup>C–<sup>13</sup>C correlation spectrum was recorded using a <sup>1</sup>H–<sup>13</sup>C CP and a PDSM mixing time of 50 ms. Acquisition times were set to 10 ms and 15 ms in the indirect and the direct <sup>13</sup>C dimension, respectively. 1D <sup>13</sup>C-detected <sup>1</sup>H–<sup>13</sup>C CP and <sup>1</sup>H–<sup>13</sup>C INEPT (insensitive nuclei enhancement by polarization transfer)<sup>67</sup> spectra on Hfq11 and Hfq were both recorded with 20 ms acquisition time.

**Statistics and reproducibility.** Production and purification were performed and optimized for unlabeled and <sup>13</sup>C, <sup>15</sup>N-labeled Hfq samples resulting in protein samples with comparable purity. Synthetic Hfq11 fibril formation has been reported previously<sup>21</sup>. Multi-dimensional solid-state NMR experiments were recorded in experimental building blocks, encompassing approximately one day of experiment time, and added together.

**Reporting summary.** Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

#### Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on request.

Received: 22 May 2023; Accepted: 13 October 2023;

Published online: 21 October 2023

## References

- Sun, X. Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic Acids Res.* **30**, 3662–3671 (2002).
- Gottesman, S. Trouble is coming: signaling pathways that regulate general stress responses in bacteria. *J. Biol. Chem.* **294**, 11685–11700 (2019).
- Valentin-Hansen, P., Eriksen, M. & Udesen, C. The bacterial Sm-like protein Hfq: a key player in RNA transactions: Hfq, an Sm-related protein. *Mol. Microbiol.* **51**, 1525–1533 (2004).
- Wilusz, C. J. & Wilusz, J. Lsm proteins and Hfq: Life at the 3' end. *RNA Biol.* **10**, 592–601 (2013).
- Schumacher, M. A. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J.* **21**, 3546–3556 (2002).
- Link, T. M., Valentin-Hansen, P. & Brennan, R. G. Structure of *Escherichia coli* Hfq bound to polyribadenylate RNA. *Proc. Natl Acad. Sci. USA* **106**, 19292–19297 (2009).
- Zhang, A., Schu, D. J., Tjaden, B. C., Storz, G. & Gottesman, S. Mutations in interaction surfaces differentially impact *E. coli* Hfq association with small RNAs and their mRNA targets. *J. Mol. Biol.* **425**, 3678–3697 (2013).
- Sauer, E., Schmidt, S. & Weichenrieder, O. Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. *Proc. Natl Acad. Sci. USA* **109**, 9396–9401 (2012).
- Beich-Frandsen, M. et al. Structural insights into the dynamics and function of the C-terminus of the *E. coli* RNA chaperone Hfq. *Nucleic Acids Res.* **39**, 4900–4915 (2011).
- Sauer, E. Structure and RNA-binding properties of the bacterial LSm protein Hfq. *RNA Biol.* **10**, 610–618 (2013).
- Orans, J., Kovach, A. R., Hoff, K. E., Horstmann, N. M. & Brennan, R. G. Crystal structure of an *Escherichia coli* Hfq Core (residues 2–69)–DNA complex reveals multifunctional nucleic acid binding sites. *Nucleic Acids Res.* **48**, 3987–3997 (2020).
- Dimastrogiovanni, D. et al. Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. *eLife* **3**, e05375 (2014).
- Arluison, V. et al. The C-terminal domain of *Escherichia coli* Hfq increases the stability of the hexamer: the role of the C-terminal domain on Hfq. *Eur. J. Biochem.* **271**, 1258–1265 (2004).
- Santiago-Frangos, A., Kavita, K., Schu, D. J., Gottesman, S. & Woodson, S. A. C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. *Proc. Natl. Acad. Sci. USA* **113**, E6089–E6096 (2016).
- Turbant, F., Wu, P., Wien, F. & Arluison, V. The amyloid region of Hfq riboregulator promotes DsrA: rpoS RNAs annealing. *Biology* **10**, 900 (2021).
- Parekh, V. J. et al. Crucial role of the C-terminal domain of Hfq protein in genomic instability. *Microorganisms* **8**, 1598 (2020).
- Cech, G. M. et al. The *Escherichia coli* Hfq protein: an unattended DNA-transaction regulator. *Front. Mol. Biosci.* **3**, 36 (2016).
- Malabirade, A. et al. Revised role for Hfq bacterial regulator on DNA topology. *Sci. Rep.* **8**, 16792 (2018).
- Kubiak, K. et al. Amyloid-like Hfq interaction with single-stranded DNA: involvement in recombination and replication in *Escherichia coli*. *QRB Discov.* **3**, e15 (2022).
- Cossa, A. et al. Cryo soft X-ray tomography to explore *Escherichia coli* nucleoid remodeling by Hfq master regulator. *J. Struct. Biol.* **214**, 107912 (2022).
- Malabirade, A. et al. Membrane association of the bacterial riboregulator Hfq and functional perspectives. *Sci. Rep.* **7**, 10724 (2017).
- Turbant, F. et al. Unraveling membrane perturbations caused by the bacterial riboregulator Hfq. *IJMS* **23**, 8739 (2022).
- Arluison, V. et al. Three-dimensional structures of fibrillar Sm proteins: Hfq and other Sm-like proteins. *J. Mol. Biol.* **356**, 86–96 (2006).
- Sawaya, M. R., Hughes, M. P., Rodriguez, J. A., Riek, R. & Eisenberg, D. S. The expanding amyloid family: structure, stability, function, and pathogenesis. *Cell* **184**, 4857–4873 (2021).
- Nguyen, P. H. et al. Amyloid oligomers: a joint experimental/computational perspective on Alzheimer's disease, Parkinson's disease, type II diabetes, and amyotrophic lateral sclerosis. *Chem. Rev.* **121**, 2545–2647 (2021).
- Ke, P. C. et al. Half a century of amyloids: past, present and future. *Chem. Soc. Rev.* **49**, 5473–5509 (2020).
- Turbant, F. et al. Identification and characterization of the Hfq bacterial amyloid region DNA interactions. *BBA Adv.* **1**, 100029 (2021).
- van der Wel, P. C. A. Insights into protein misfolding and aggregation enabled by solid-state NMR spectroscopy. *Solid State Nucl. Magn. Reson.* **88**, 1–14 (2017).
- Loquet, A. et al. 3D structure determination of amyloid fibrils using solid-state NMR spectroscopy. *Methods* **138–139**, 26–38 (2018).
- Habenstein, B. & Loquet, A. Solid-state NMR: an emerging technique in structural biology of self-assemblies. *Biophys. Chem.* **210**, 14–26 (2016).
- Jaroniec, C. P. Two decades of progress in structural and dynamic studies of amyloids by solid-state NMR. *J. Magn. Reson.* **306**, 42–47 (2019).
- Meier, B. H., Riek, R. & Böckmann, A. Emerging structural understanding of amyloid fibrils by solid-state NMR. *Trends Biochem. Sci.* **42**, 777–787 (2017).
- Makin, O. S., Atkins, E., Sikorski, P., Johansson, J. & Serpell, L. C. Molecular basis for amyloid fibril formation and stability. *Proc. Natl Acad. Sci. USA* **102**, 315–320 (2005).
- Nelson, R. et al. Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
- Sunde, M. & Blake, C. C. F. From the globular to the fibrous state: protein structure and structural conversion in amyloid formation. *Quart. Rev. Biophys.* **31**, 1–39 (1998).
- Nguyen, J. T. et al. X-ray diffraction of scrapie prion rods and PrP peptides. *J. Mol. Biol.* **252**, 412–422 (1995).
- Serpell, L. C. Alzheimer's amyloid fibrils: structure and assembly. *Biochim. Biophys. Acta* **1502**, 16–30 (2000).
- Balbirnie, M., Grothe, R. & Eisenberg, D. S. An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated  $\beta$ -sheet structure for amyloid. *Proc. Natl Acad. Sci. USA* **98**, 2375–2380 (2001).
- Eanes, E. D. & Glenner, G. G. X-ray diffraction studies on amyloid filaments. *J. Histochem. Cytochem.* **16**, 673–677 (1968).
- Sikorski, P. & Atkins, E. New model for crystalline polyglutamine assemblies and their connection with amyloid fibrils. *Biomacromolecules* **6**, 425–432 (2005).
- Fraser, P. E. et al. Fibril formation by primate, rodent, and Dutch-hemorrhagic analogs of Alzheimer amyloid  $\beta$ -protein. *Biochemistry* **31**, 10716–10723 (1992).
- Meersman, F., Cabrera, R. Q., McMillan, P. F. & Dmitriev, V. Structural and mechanical properties of TTR105-115 amyloid fibrils from compression experiments. *Biophys. J.* **100**, 193–197 (2011).
- Wang, Y. & Jardetzky, O. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci.* **11**, 852–861 (2002).
- Ruyschaert, J.-M. & Raussens, V. ATR-FTIR Analysis of Amyloid Proteins. in *Peptide Self-Assembly* (eds Nilsson, B. L. & Doran, T. M.) vol. 1777 69–81 (Springer New York, 2018).
- Waeytens, J. et al. Probing amyloid fibril secondary structures by infrared nanospectroscopy: experimental and theoretical considerations. *Analyst* **146**, 132–145 (2021).
- Cerf, E. et al. Antiparallel  $\beta$ -sheet: a signature structure of the oligomeric amyloid  $\beta$ -peptide. *Biochem. J.* **421**, 415–423 (2009).
- Sarroukh, R. et al. Transformation of amyloid  $\beta$ (1–40) oligomers into fibrils is characterized by a major change in secondary structure. *Cell. Mol. Life Sci.* **68**, 1429–1438 (2011).
- Fortas, E. et al. New insight into the structure and function of Hfq C-terminus. *Biosci. Rep.* **35**, e00190 (2015).
- Shen, Y. & Bax, A. Protein backbone chemical shifts predicted from searching a database for torsion angle and sequence homology. *J. Biomol. NMR* **38**, 289–302 (2007).
- Azam, T. A. & Ishihama, A. Twelve species of the nucleoid-associated protein from *Escherichia coli*. *J. Biol. Chem.* **274**, 33105–33113 (1999).
- Diestra, E., Cayrol, B., Arluison, V. & Risco, C. Cellular electron microscopy imaging reveals the localization of the Hfq protein close to the bacterial membrane. *PLoS ONE* **4**, e8301 (2009).
- Turbant, F. et al. Interactions and Insertion of *Escherichia coli* Hfq into outer membrane vesicles as revealed by infrared and orientated circular dichroism spectroscopies. *Int. J. Mol. Sci.* **14**, 11424 (2023).
- Beaufay, F. et al. Polyphosphate drives bacterial heterochromatin formation. *Sci. Adv.* **7**, eabk0233 (2021).
- Goldberger, O., Szoke, T., Nussbaum-Shochat, A. & Amster-Choder, O. Heterotypic phase separation of Hfq is linked to its roles as an RNA chaperone. *Cell Rep.* **41**, 111881 (2022).
- Santiago-Frangos, A., Jeliakzov, J. R., Gray, J. J. & Woodson, S. A. Acidic C-terminal domains autoregulate the RNA chaperone Hfq. *eLife* **6**, e27049 (2017).
- Sauter, C. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic Acids Res.* **31**, 4091–4098 (2003).
- De Meutter, J. & Goormaghtigh, E. Protein structural denaturation evaluated by MCR-ALS of protein microarray FTIR spectra. *Anal. Chem.* **93**, 13441–13449 (2021).
- Partouche, D. et al. In situ characterization of Hfq bacterial amyloid: a Fourier-transform infrared spectroscopy study. *Pathogens* **8**, 36 (2019).
- Schmidt, A., Annamalai, K., Schmidt, M., Grigorieff, N. & Fändrich, M. Cryo-EM reveals the steric zipper structure of a light chain-derived amyloid fibril. *Proc. Natl Acad. Sci. USA* **113**, 6200–6205 (2016).
- Sawaya, M. R. et al. Atomic structures of amyloid cross- $\beta$  spines reveal varied steric zippers. *Nature* **447**, 453–457 (2007).
- Sarni, S. H. et al. Intrinsically disordered interaction network in an RNA chaperone revealed by native mass spectrometry. *Proc. Natl Acad. Sci. USA* **119**, e2208780119 (2022).

62. Yamaguchi, K. et al. Polyphosphates induce amyloid fibril formation of  $\alpha$ -synuclein in concentration-dependent distinct manners. *J. Biol. Chem.* **296**, 100510 (2021).
63. Cremers, C. M. et al. Polyphosphate: a conserved modifier of amyloidogenic processes. *Mol. Cell* **63**, 768–780 (2016).
64. Obregon, K. A., Hoch, C. T. & Sukhodolets, M. V. Sm-like protein Hfq: composition of the native complex, modifications, and interactions. *Biochim. Biophys. Acta* **1854**, 950–966 (2015).
65. Partouche, D. et al. Epigallocatechin gallate remodelling of Hfq amyloid-like region affects *Escherichia coli* survival. *Pathogens* **7**, 95 (2018).
66. Micsonai, A. et al. BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res.* **46**, W315–W322 (2018).
67. Elena, B., Lesage, A., Steuernagel, S., Böckmann, A. & Emsley, L. Proton to carbon-13 INEPT in solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **127**, 17296–17302 (2005).
68. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
69. Molecular Graphics System, Schrödinger, LLC, New York, NY 10036, US.

## Acknowledgements

We acknowledge funding from the Center National de la Recherche Scientifique CNRS Momentum (BH), IdEx University of Bordeaux–Chaire d’installation (BH), the Région Nouvelle Aquitaine (BH, grant convention No 2017-1R10305-00013031), the “Fondation de la Maison de la Chimie” (BH), the French National Research Agency (BH, grant No. ANR-19-CE13-0021), CNRS and CEA (VA). This study contributes to the IdEx Université Paris Cité ANR-18-IDEX-0001 (VA). The work has benefited from the Biophysical and Structural Chemistry Platform at IECB, CNRS UAR 3033, INSERM US001. We are grateful to Vincent Raussens (ULB, Bruxelles) for his help in FTIR analysis. SRCD experiments were performed on the “DISCO” beamline at SOLEIL Synchrotron, France (inhouse project:99230050). Financial support from the IR INFRANALYTICS FR2054 for conducting the research is gratefully acknowledged.

## Author contributions

M.B., D.M., B.K., E.M., A.G., and B.H. conducted the NMR, X-ray diffraction, and EM experiments and analyzed the data. F.W., J.W., and V.A. conducted the FTIR and SRCD

experiments and analyzed the data. M.B. and D.M. performed protein expression and purification. V.A. and B.H. designed the project. J.W., F.W., B.K., V.A., and B.H. wrote the paper. All authors commented on the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05462-1>.

**Correspondence** and requests for materials should be addressed to Véronique Arluison or Birgit Habenstein.

**Peer review information** *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Xiaohui Frank Zhang and Gene Chong. A peer review file is available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023