Initial Steps of Chaperone-Aided Fibrillation of *Pseudomonas aeruginosa* Biofilm Forming Functional Amyloid FapC

Chang-Hyeock Byeon[1], Kasper Holst Hansen[2], Jasper Jeffrey[1], Maria Andreasen[2], Ümit Akbey[1]*

[1]Department of Structural Biology, School of Medicine, University of Pittsburgh, Biomedical Science Tower 3, 3501 Fifth Avenue, Pittsburgh, 15261, United Stated. [2]Deparment of Biomedicine, Aarhus University, Wilhelm Meyers Alle 3, Building 1234, DK-8000, Arhus C, Denmark

Abstract

Functional bacterial amyloids (FuBA) are involved in biofilm formation in *Pseudomonas aeruginosa*. Biofilm forming bacterial colonies are difficult to treat, cause most of the chronic bacterial infections and are relevant in the context of antimicrobial resistance. Here in this work by focusing on the major biofilm forming FapC protein we demonstrate the initial steps of the chaperone aided amyloid formation. Solution NMR spectroscopy gave atomic resolution information on the prefibrillar soluble monomeric FapC that is an intrinsically disordered protein (IPD). Moreover, we show that the chaperone FapA within the same biofilm forming protein operon interacts with FapC. FapA is itself an IDP and remarkably slows down the FapC fibril formation without modifying the final fibril morphology. Monitoring of the small but significant chemical shift perturbations allowed the determination of the interaction interface which was not observed previously. To the best of our knowledge, our work provides the first atomic-level identification of a rather specific interaction of an IDP chaperone, FapA, to modulate and slow-down the functional amyloid protein FapC towards its path to mature fibril formation. Our results establish a better understanding of the functional amyloid formation in bacterial biofilms and its molecular level tuning or prevention by chaperones. We foresee that future therapeutic treatments of biofilms and relevant infections could emerged from these structural and mechanistic insights.

Keywords: Solution-state NMR Spectroscopy, functional bacterial amyloid (FuBA), bacterial biofilm, antimicrobial resistance, *Pseudomonas aeruginosa* FapC, prefibrillar assembly, chaperone FapA

Introduction

Functional bacterial amyloids (FuBA) are unique class of amyloid fibrils that have function in living organisms, and do not directly cause diseases as their pathologic relatives.[1-6] They are involved in many biological processes such as biofilm formation as a survival mechanism of bacteria to protect themselves under stress inducing hostile conditions. Although biofilm forming FuBAs do not directly cause diseases, they are indirectly mediating pathology through chronic bacterial infections, and leading to increased antimicrobial resistance. Most of all chronic infections, e.g., are due to biofilm forming bacteria.[7] So far, several biofilm forming FuBAs have been studied in detail so far, such as CsgA from *E. coli*, FapC from *P. aeruginosa*, TasA from *B. subtilis* and PSMs from *S. aureus*, and others.[8-11] For a detailed review on this topic, we refer the reader to recent reviews.[6, 12]

In recent years, there has been large interest in the functional amyloids, their biological roles and biophysical characterization. Despite abundant biophysical information available on functional amyloids, structural insights have been elusive due to the difficulties in studying either monomeric intrinsically disordered protein (IDP) forms or amyloids fibril forming macroscopic filaments.[13] The lack of structural understanding of amyloid formation from the monomeric to fibrillar forms in biofilms, hampers development of therapeutic strategies. To make a breakthrough in the fight against biofilm-mediated chronic infections and/or biofilm-mediated antimicrobial resistance, novel approaches directed by structural insights are necessary.

Most of the functional and pathological amyloid forming proteins are IDPs in their monomeric forms, except TasA being one of the known exception with a unique globular fold.[6, 14] These unstructured monomers common to almost all of the functional and pathologic amyloids can be studied by only a few techniques at atomic resolution, with high-resolution solution state NMR (solution NMR) spectroscopy being a powerful choice.[15-17] For small fragments of amyloid proteins that can be crystallized, X-Ray crystallography established the first atomic models of different systems,[3, 18, 19] however, longer fragments of amyloids are not crystalline. For the full-length insoluble filamentous amyloid fibrils, solid-state magic-angle spinning (MAS) NMR (ssNMR) had been the only available technique to study these non-crystalline insoluble proteins until the resolution revolution by cryo electron microscopy (Cryo-EM) in the last decade.[20-24] Nowadays, ssNMR and Cryo-EM are used for structural studies on amyloids, both having stronger and weaker sides in terms of the information they can provide, and they can be combined for a more complete understanding of amyloids.[25-29]

To address the lack of structural information on functional amyloids, in this work we present insights towards better understanding of molecular structure and organization of functional amyloids. We report results on prefibrillar soluble monomeric form of the functional amyloid protein FapC from *Pseudomonas aeruginosa* by using solution NMR spectroscopy.[30] Solution NMR spectroscopy is a powerful technique to study IDPs, in simple homogeneous pure *in vitro* context, as well as in complex heterogeneous native *in*

vivo environments.[31] It may be the only atomic resolution method to study the first steps of oligomerization of prefibrillar species (soluble monomers, soluble oligomers, and any other high molecular weight soluble proteins).[32, 33] In contrast, Thioflavin T (ThT) assays, a common method to monitor fibrillization, can only detect the bulk fibrillation and do not provide atomistic information.[34] Solution NMR has been utilized successfully to characterize NMR-visible soluble monomeric species as well as the NMR-invisible high molecular-weight oligomeric/protofilament and sparsely populated oligomeric species of several different amyloid fibrils from both pathologic and functional amyloids.[32, 35-39, 40-44] These methods are also summarized in an excellent recent review.[32]

A few very important points to note here concerning fibril formation are the sample concentration, temperature, and the freshness of the prepared samples to be in an absolute monomeric state in the NMR experiments. NMR measurements benefits from higher concentrations to increase the overall observed signal, however, fibrillation rate increases significantly with concentration and observation of this process timely becomes more difficult. Temperature is another important factor, due to the fact that the rate of fibrillation is strongly temperature dependent.[45] Finally, the existence of any high molecular weight species at the beginning of fibrillation, such as seeds or high molecular weight soluble/insoluble oligomers, will affect the rate of fibrillation and may cause bias if not properly addressed, e.g. by freshly preparing the monomeric samples. All of these factors were considered carefully in this work.

Chaperones can alter the fibrillation of amyloid fibrils. Examples from both pathologic and functional amyloid perspectives cover a wide range of chaperone effect on amyloid fibrils, ranging from slowing down to the prevention or even to their reversal.[17, 35, 37, 46-48] CsgE was shown to slow down the fibril formation of CsgA functional amyloid. Similarly, pathologic amyloid fibril formation of αSynuclein was slowed down by HSP90 family chaperones.[46] FapA is a *Pseudomonas aeruginosa* chaperone and is one of the six proteins expressed in the relevant operon. Compared to the HSP family or other folded chaperones, FapA belongs to a unique class of proteins that are intrinsically unfolded, but still can act as a chaperone.[17] Understanding details of the molecular mechanism of FapA chaperone is crucial to develop strategies to prevent amyloid formation in biofilm context and to prevent related chronic infections.

Here in this work, we present a time-resolved high-resolution solution NMR approach to characterize FapC from its prefibrillar monomeric form towards amyloid fibril formation. As a first milestone towards complete molecular understanding of biofilm forming FuBAs, we monitored changes in the NMR spectrum of monomeric FapC. By utilizing our recent NMR resonance assignment,[30] we determined the time-dependent NMR chemical shift and intensity changes, which was then correlated to the FapC secondary structure to understand the first nucleation interface. Moreover, the molecular basis of chaperone-aided fibrillation prevention mechanism is monitored and quantified, towards a general chaperone based amyloid fibril prevention mechanism. We used the *Pseudomonas aeruginosa* chaperone FapA and monitored its effect on FapC. To facilitate the correctness of our time-resolved experiments, we used low protein

concentration and freshly prepared FapC sample to prevent seeding/oligomerization effects. Moreover, we performed the NMR experiments at a lower temperature of 274 K in order to slow down the fibrillation kinetics to our favor and also to observe all the FapC NMR resonances that partially disappeared at elevated temperatures at our experimental conditions.

Results

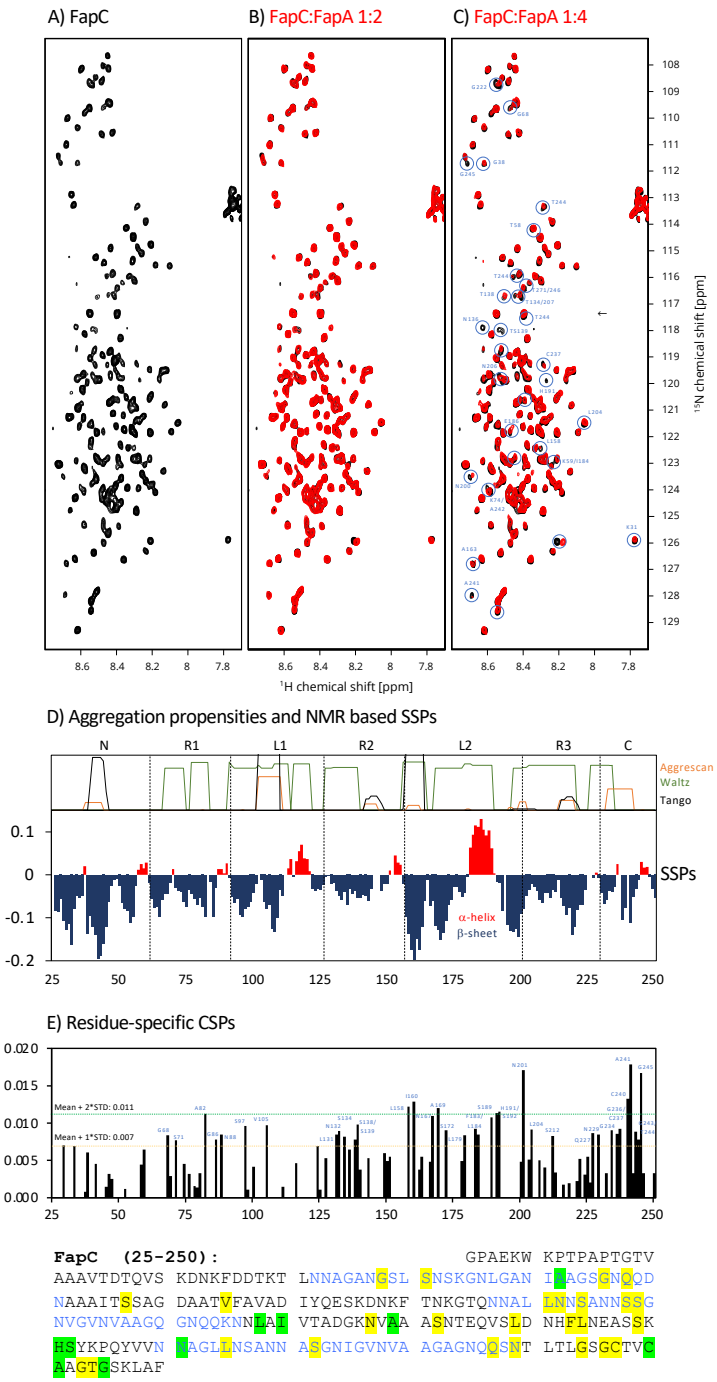


Figure 1. 2D fingerprint ^{1}H-^{15}N HSQC NMR spectra of **A)** pure full-length FapC (black spectrum), **B)** FapC (black) versus FapC mixed with FapA (red spectrum) at 1:2 molar ratio and **C)** FapC (black) versus FapC mixed with FapA (red spectrum) at 1:4 molar ratio recorded at 274 K. The spectra are shown at the identical contour levels for direct comparison. The NMR resonance assignment for FapC is presented previously.[30] FapC concentration is 50 μM and FapA concentration is adjusted accordingly for each sample. The significant chemical shift perturbations due to FapA addition to the FapC sample are shown with blue circles in C with the corresponding resonance assignments. **D)** Representation of amyloid forming regions predicted by three different algorithms, along with the NMR chemical shift based secondary neighbor-corrected structure propensities (ncSSPs) for the full-length FapC. Y-axis value of 1 refers to 100% for ncSSPs. **E)** Representation of the chemical shift perturbations (CSPs) in the FapC NMR spectrum upon addition of FapA at the FapC:FapA 1:4 sample compared to the pure FapC. The weighted cumulative CSPs given in terms of ppm (y-axis) and are calculated based on the method presented earlier, by $\Delta\delta = [(\delta^{1H}_{FapC:FapA} - \delta^{1H}_{FapC})^2 + (0.14*(\delta^{15N}_{FapC:FapA} - \delta^{15N}_{FapC}))^2]^{0.5}$.[49] Additionally, the FapC sequence (between 25-250 amino acids) is shown to visualize the CSPs, the repeat regions are colored in blue. Yellow highlight indicates significant changes, and the green highlight indicates much larger changes as depicted with two dashed lines in the CSPs plot.

FapC is an intrinsically disordered protein. By utilizing solution NMR spectroscopy, we recently presented the complete NMR resonance assignment of soluble monomeric FapC.[30] NMR spectra and secondary structure propensities (SSPs) determined using NMR chemical shifts indicate that FapC is an intrinsically disordered protein (IDP) with still significant α-helix and β-sheet structural propensities,[6] see **Figure 1A/D**. The 2D ^{1}H-^{15}N HSQC NMR spectra shown in **Figure 1 A-C** of both the pure FapC and the FapC complexed with its chaperone FapA show a narrow proton chemical shift dispersion of ~1 ppm, due to its IDP nature. The alignment of the SPPs with the predicted amyloidogenic segments as well as with the proposed domain organization of FapC (L: loop, R: repeat, C: c-terminus and N: n-terminus) based on the sequence properties are shown in the FELLS analysis in **Figure SI1**.[50, 51] **Figure 1D** depicts the overlay of the NMR determined SSPs along with the β-aggregation potential predicted using three algorithms TANGO, WALTS and AGGRESCAN.[52-54] These *in silico* methods indicate different aggregation prone regions (APRs), however, the regions between amino acids 39-45 (~5% propensity by TANGO algorithm), 102-109 (~46%), 142-148 (~1.2%), 156-163 (~18%), 197-205 (~0.1%), 215-222 (~1%) are covered by all three algorithms though with varying propensities (See **Figure SI2** for details). These APRs correlate with the three repeat regions of FapC (R1,R2 and R3), but also falls in the unexpected areas outside the repeats in the N, L1 and L2 of FapC. This observation of APRs in the non-repeating FapC regions fits to our previous observation that L1/L2 loop regions increase amyloid formation rates determined by ThT assays.[51] The NMR determined SSPs of the FapC indicates predominantly random coil with up to ~40% polyproline 2 (PP2) helix conformations. This large PP2 propensity could be important for structural transition towards the correct amyloid fibril fold in contrast to amorphous aggregates.[55] Nevertheless, sizeable β-sheet propensities between 5-25% are observed throughout the protein, along with a few α-helix indicating areas 181-189 with ~15% and 113-121 with ~7% propensity, based on the ncSSP analysis shown in **Figure 1D**.

There are areas that indicate turns or small helices, such as residues at around 37, 57-60, 70, 87-90, 151-155, 228, 236 and 245-247. The predicted APRs fall within the NMR determined β-sheet propensity areas, indicating a good correlation between the NMR and *in silico* methods. The structural transition from the monomeric to amyloid FapC during primary nucleation could be initiated at/around these APRs. Overall, even though the NMR determined SSPs are relatively small propensities, they correlate well with the secondary structure elements predicted by the Alphafold model, shown in **Figure 2A**. This highlighting a possible less hindered conformational transition from monomeric IDP fold of FapC into its amyloid fold.

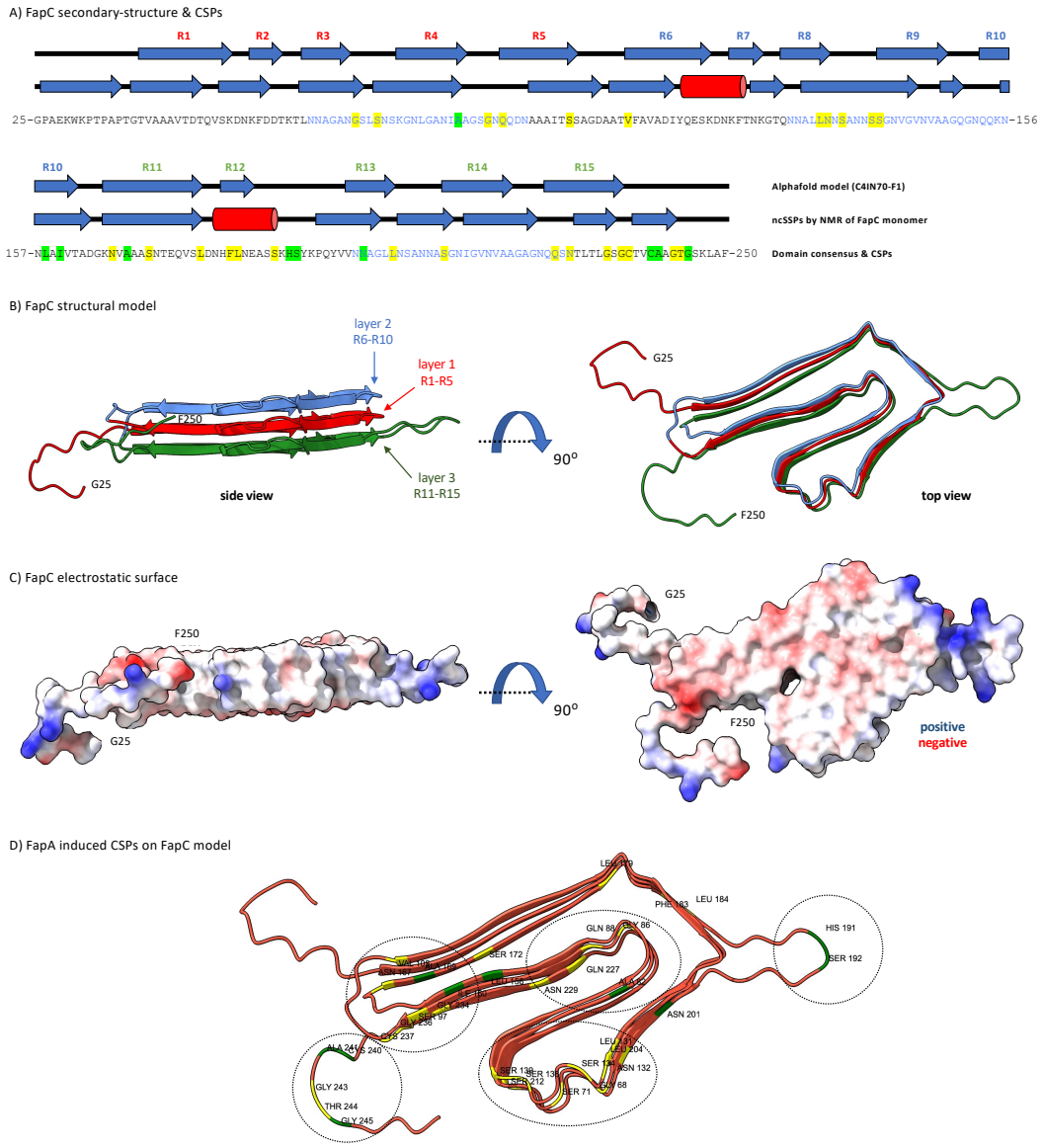


Figure 2. A) The secondary structure comparison of the amyloid FapC structure based on Alphafold model (AF-C4IN70-F1) and the neighbor-corrected secondary structure propensities (ncSSPs) based on our experimental solution NMR chemical shifts of the monomeric FapC with the BMRB entry #51793. The amino acid sequence of FapC (25-250 residues) are shown at the bottom. The black color represents the N-terminus, loop regions and C-terminus in the sequence, whereas the blue color represents the generally accepted repeat regions (R1-R3). The Alphafold

secondary structure elements are depicted by the model. We used a threshold of 4-residues for depicting the α-helices (red cylinders) or β-sheets (numbered and blue arrows) as a secondary structure element for the NMR ncSSPs data. **B)** Representation of Alphafold amyloid FapC structural model (should be visualized as a model of the amyloid fibril structure of FapC) in two orientations, side, and top view. The repeats forming three different layers are indicated with color coding: beta-strands R1-5 form layer 1 (red), beta-strands R6-10 form layer 2 (blue) and the beta-strands R11-15 form the layer 3 (green). Note than the nomenclature used for the FapC protein segments (repeat region 1-3) are different from the beta-strand numbering here. **C)** Representation of the electrostatic surface for the FapC fibrillar structure model shown in B. **D)** Locations of the chemical shift perturbations (CSPs) on the structural model, that are observed in the solution NMR spectra of isotope labeled monomeric FapC as a soluble IDP upon addition of four-fold unlabeled FapA chaperone. The green and yellow colors are placed according to the representation in Figure 1E, represent the large (green) and medium (yellow) CSPs. The chemical shift perturbation should be understood with care, since they were recorded for the soluble monomeric FapC, however shown here at the structural model representing the amyloid fibrillar conformation.

FapA is an intrinsically disordered chaperone, strongly interacts with FapC monomer and prevents fibrillation. FapA is a chaperone and is one of the six proteins expressed within the functional amyloid related operon (fapA-F) of *Pseudomonas aeruginosa*.[56] FapA is an IDP similar to FapC based on the 2D HSQC fingerprint spectrum from the unlabeled FapA by our work here, **Figure SI3**, as well as recent work by others.[17] Recently it was shown that FapA dramatically affects FapC fibrilization monitored by ThT assay, but speculated that this effect was through an "NMR invisible oligomeric species" which did not affect the NMR spectra of FapC. Even though it is not straightforward to think of an IDP acting as a chaperone, the data we present here clearly shows that FapA acts as a chaperone for monomeric FapC *in vitro* though specific interactions during functional amyloid formation and hence suppresses FapC fibrillation. Since one of the major differences between the functional and pathologic amyloids is the tight control of functional amyloid formation, the FapA chaperone is a candidate to interact and therefore regulate FapC.[56] This mechanism was proposed to prevent self-harm to bacteria by eliminating harmful off-pathway fibril formation.[6, 57] Conditional disorder of a chaperone for its action is known as well as IDPs acting as a shield and slowing down fibrillation,[58, 59] however, a complete IDP acting as a chaperone and having significant effect on the monomeric prefibrillar species has not been reported previously in the literature for other systems. Nevertheless, non-chaperone IPD proteins were shown to slow down the aggregation of amyloids, such as Aβ by interfering with the nucleation phase, which can be correlated to FapA interaction with FapC.[59]

Here, we show direct evidence of IDP FapA interacting with prefibrillar monomeric FapC by using high-resolution solution NMR. We monitored FapA/FapC interaction from the FapC perspective by using isotope labeled FapC and unlabeled FapA. 2D ^{1}H-^{15}N HSQC NMR spectra of pure ^{15}N isotope-labeled FapC recorded with and without unlabeled FapA (at 1:2 and 1:4 mole ratios) under identical conditions and complexed, shown in **Figure 1A-C**. Remarkably, the incubation of FapA with FapC induces spectral changes through chemical shift perturbations and signal intensity changes, particularly at the 1:4 molar ratio as marked in the spectrum in **Figure 1C**. The zoomed areas with selected regions are shown in **Figure SI4**.

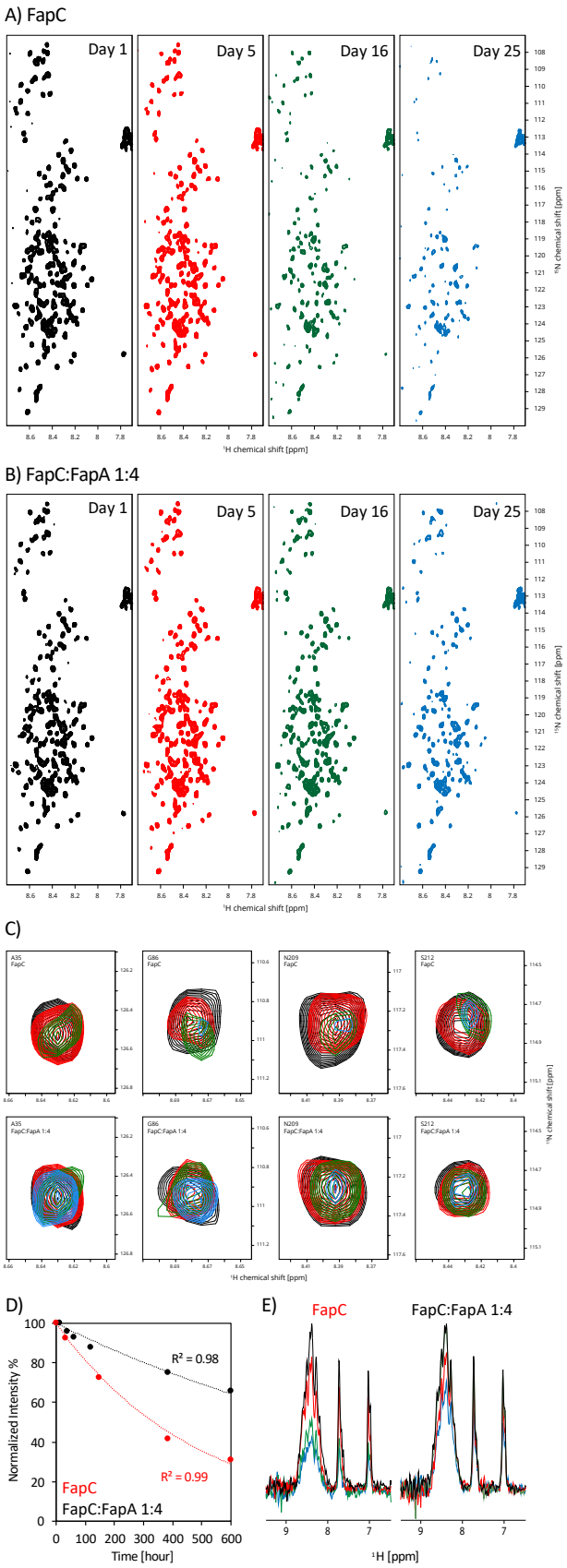

A) FapC
Day 1
Day 5
Day 16
Day 25
^{15}N chemical shift [ppm]
^{1}H chemical shift [ppm]
B) FapC:FapA 1:4
Day 1
Day 5
Day 16
Day 25
^{15}N chemical shift [ppm]
^{1}H chemical shift [ppm]
C)
A35 FapC
G86 FapC
N209 FapC
S212 FapC
A35 FapC:FapA 1:4
G86 FapC:FapA 1:4
N209 FapC:FapA 1:4
S212 FapC:FapA 1:4
^{15}N chemical shift [ppm]
^{1}H chemical shift [ppm]
D)
Normalized Intensity %
R^2 = 0.98
R^2 = 0.99
FapC
FapC:FapA 1:4
Time [hour]
E)
FapC
FapC:FapA 1:4
^{1}H [ppm]

Figure 3. Time-dependent 2D ^{1}H-^{15}N HSQC NMR spectra for pure **A)** FapC and **B)** FapC:FapA 1:4 samples. The spectra were recorded over 25 days at Day 1, 5, 16 and 25 for both samples under identical conditions and plotted with the same contour levels for direct comparison. **C)** Zoomed regions for selected amino acids are shown for pure FapC (top) and FapC:FapA 1:4 (bottom), with the same color coding as in A/B. **D)** Time-dependent NMR signal intensity reduction for pure FapC (red) and FapC:FapA1:4 (black) samples. The first spectrum of each 2D ^{1}H-^{15}N HSQC spectra is used to quantify the signal amount plotted here by signal integration at the amide proton region, that are shown in **E)**. The R^2 values of the fits are depicted on the plot to show the quality of the fits.

Despite the very small chemical shift dispersion in the NMR spectra of FapC, we observed significant chemical shift perturbations (CSPs greater than "mean + standard-deviation") in the NMR spectra of FapC upon FapA addition. This is a direct sign of interaction, **Figure 1E**. Previous interaction studies on IDP proteins suggest chemical shift changes of similar magnitude as significant shifts.[15] The FapC and FapA interaction was not observed previously by NMR probably due to different experimental conditions.[17] Overall, 36 residues show significant CSP spread over the whole FapC sequence as highlighted in the protein sequence except C-terminus, **Figure 1E**.

These results are also visualized in **Figure 2A-D** in combination with the FapC amyloid structural model based on Alphafold's prediction.[60] Strikingly, the largest changes, with values greater than the "mean + 2*standard-deviation", are observed predominantly at the L2, R3 and C-terminus section of FapC. These residues, L158, I160, A169, H191, S192, N201, C240, A241 and G245, could indicate the hotspots of fibrillation inhibition/control interface between FapA and FapC. Particularly, the N201, C240, A241 and G245 have the largest CSPs, and this may indicate a specific interaction of FapA at the L2/R3 crossing interface area and at the C-terminus. Previous ThT assays performed on R3C, L2R3C, L1L2R3C and FL FapC constructs indicate that the L1 and L2 loop regions increase the aggregation propensity without changing the fibril morphology.[51] The CSPs we observed shown in **Figure 1E** indicate the largest overall changes in the L2 and C-terminus region, which can be correlated to the increased ThT response by the flexible loop regions. **Figure 2A/D** represents the location of these residues showing CSPs at the structural model. It is surprising to observe that the chemical shift changes due to FapC-FapA interaction are clustered close to each other in the structural model at the C-terminus and linkers between the β-sheets. It should be noted that the model represents the structure of the folded fibrillar FapC, whereas the CSPs were observed for the monomeric protein having predominantly random coil conformation with mentioned sizable SSPs. Nevertheless, the localization of FapC affected areas may indicate structural basis of chaperone-based prevention of functional amyloid formation. The electrostatic potential of the structural model shown in **Figure 2C** indicates that the CSPs are not at the largest charged areas of the structural model. This indicates that the FapA and FapC interaction maybe structurally specific, rather than an electrostatic interaction.

Another observation supporting FapA interaction with FapC is the changes observed in NMR signal intensities in the HSQC spectrum of the FapC upon addition of FapA at 1:4 mole ratio, see **Figure 3** and **4A**. We observed strong time-dependent intensity changes at the NMR resonances for freshly prepared FapC and FapC:FapA samples. Surprisingly, the FapC:FapA sample at 1:4 ratio has an overall slightly reduced NMR signal intensity compared to the pure FapC (average intensity FapC:FapA/FapC = ~0.9), for samples freshly prepared and incubated for one day at 277 K and recorded at 274 K otherwise under identical conditions. Similar experiments were done on CsgA functional amyloid from *E. coli* incubated with its CsgE chaperone partner.[15] The result from this work showed the addition of the chaperone increased the overall NMR signal compared to the pure CsgA sample at one day incubation time at 4 °C but measured at 284 K, a different trend compared to our measurements at day one incubation.[15] Nevertheless, our NMR measurements performed at longer incubation times (at Day 5, 16 and 25) show that chaperone containing FapC has larger overall NMR signal intensity compared to the pure FapC up to a factor of ~2, a result similar to the CsgA/CsgE, **Figure 3**. At Day 5 and later time points, the FapC sample incubated with FapA resulted in larger overall NMR intensities compared to pure FapC, indicating the fibrillation-prevention effect of FapA as discussed below in more detail. The spectra recorded at lower temperatures slows down the fibrillation rate dramatically compared to ambient temperature or 37 °C, and this approach allowed us to timely monitor spectral changes by NMR. The kinetics and thermodynamics of the interaction of a chaperone to a monomeric functional amyloid protein strongly depends on the conditions of the experiments performed, such as concentration, pH, temperature, incubation time and others. Since the temperature and pH of our and other's data are different (hence the overall kinetics of the process), we rely on and discuss results shown here that are consistently reproducible.

Residue-resolved monitoring of FapA - FapC interaction. We performed time-resolved solution NMR spectroscopy to record 2D ^{1}H-^{15}N HSQC spectra of pure FapC and FapC:FapA (1:4 ratio) samples, **Figure 3A/B**. All the spectra were recorded at 274 K and the samples were stored at 4 °C in between the NMR measurements to keep them at the same conditions throughout the experimental time span of 25 days. Lowered temperatures allowed us first, to observe all the signals in FapC and FapC:FapA samples at the current sample conditions and second, more importantly to slow down the fibrillation so that we can monitor changes correctly and without transient effects. Temperature dependent slowing down of fibrillation were reported similarly before by using ThT assays.[45] This simple yet powerful experiment can reveal the residue-specific aggregation propensities by quantifying the individual intensity losses over time in the HSQC spectra. One can think of this experiment as a ThT assay that can monitor changes at the atomic level for each residue in the protein. The structural transition or conformation change of the soluble FapC monomer into either a soluble high molecular-weight species (oligomers) or insoluble fibril/oligomer result in solution NMR signal intensity decay or a complete signal loss. Since solution NMR predominantly

monitors the low molecular weight species, i.e., monomeric FapC, the sites that are interacting/in-exchange with higher molecular weight species and/or forming fibrils are "invisible" to solution NMR and will disappear from the solution NMR spectrum.[61]

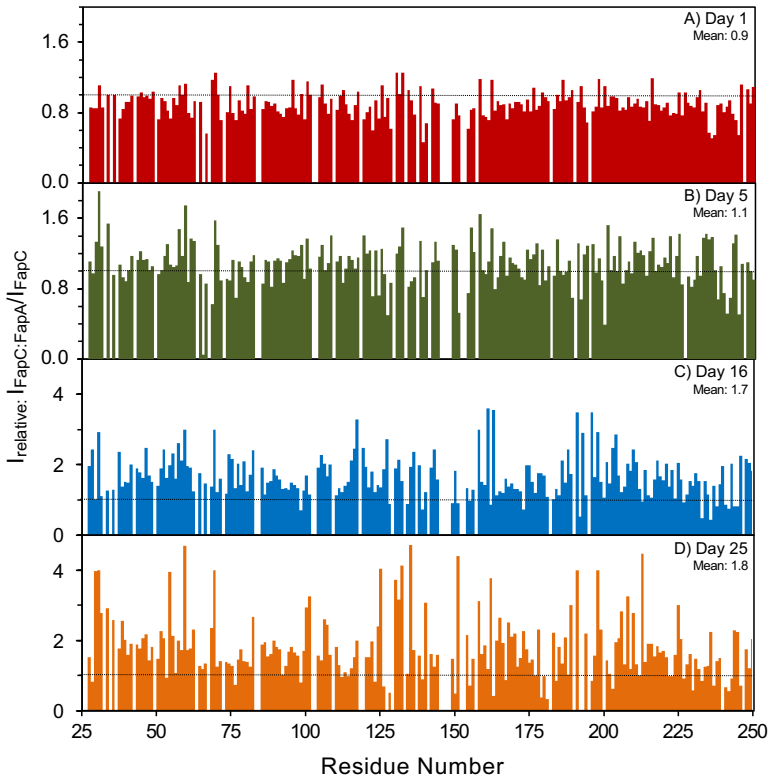


Figure 4. Representation of the residue-resolved NMR signal intensity ratios for FapC and FapC:FapA samples. The signal intensity of the NMR signal from each residue of the FapC:FapA 1:4 sample is divided with the signal intensity of the same residue of the pure FapC sample. These plots show an estimate of the time-dependent effect of FapA chaperone on the overall signal amount observed in FapC NMR spectrum at **A)** Day 1, **B)** 5, **C)** 16 and **D)** 25. It is a residue-specific representation of the bulk signal intensity decay results shown in Figure 3C. The average values of the relative intensities are given at each plot, as well as the dashed line at y-value=1. Overall, it is clear that FapA significantly reduces the signal intensity decrease due to fibrillation compared to the pure FapC. Note different scales on the y-axes for A-D.

The 2D HSQC spectra are shown in **Figure 3A-B** for pure FapC and incubated with FapA (FapC:FapA 1:4). The spectra are plotted at the same contour levels and directly comparable. The overlay of the 2D HSQC recorded at Day 1 and Day 25 for FapC and FapC:FapA 1:4 samples are shown in **Figure SI5**. The reduction in the overall intensity and number of observed peaks demonstrate that the soluble monomer to high molecular weight species transition occurs slowly at 274 K. NMR signals from the A35, G86, N209 and S212 are shown in **Figure 3C** as zoomed regions for visualization of the time-dependent signal intensity reduction for pure FapC and FapC:FapA 1:4 samples. Small chemical shift changes are observed for some residues over time, however, we did not attempt to quantify these due to error-prone lower signal to noise

ratios at longer incubation times. A qualitative analysis is performed by the integration of the protein backbone amide signal intensities of the first slice of each 2D HSQC experiments, **Figure 3D/E**. This data shows that FapA slows down the bulk FapC signal intensity decay at the backbone amide region correlating to the signals of the protein backbone. The single exponential decays fitted to the experimental data are of high quality for both curves. The signal reduction over 25 days of total experimental time determined by the 1D NMR indicate that at our experimental conditions ~31% of the signal remains for pure FapC and to 66% for FapC:FapA 1:4 samples. Similar 1D NMR signal intensity monitoring of amyloid fibril fibrillation was done on pathological Aβ and Huntington fibrils.[40, 43]

The quantitative residue-specific signal intensity reduction in the 2D spectra follows the same trend and is shown in **Figure 4**. As mentioned above, the average of the signal intensity ratio of the FapC:FapA spectrum relative to the FapC spectrum (NMR signal intensity FapC:FapA / FapC) is close to unity but slightly less, ~0.9 at Day 1, **Figure 4A**. This indicates that FapA is interacting with FapC, as also seen from the chemical shift perturbations shown in **Figure 1C/E**. Upon repeating the same set of 2D NMR experiments at day 5, 16 and 25 to quantify the relative time-dependent intensity changes, the FapC sample incubated with FapA chaperone shows higher overall signal intensity compared to the pure FapC. At day 5, this ratio is ~1.1, and further increases to ~1.7 and 1.8 at day 16 and 25, respectively. This is a clear indication that FapA is not only interacting with FapC, but also keeping it more monomeric to a higher degree by preventing its fibrillation, as observed similarly for the CsgA/CsgE.[15] The data from day 1 and 5 are more uniform with similar overall signal intensity ratios, however, becomes less uniform and show differential values for later time points recorded at day 16 and 25 without any easy-to-understand clear domain motif. As a speculation, the N-terminus and R1-L2-R3 regions maybe differentiating in **Figure 4C/D** compared to the rest of the sequence.

Initiation of fibrillation is mediated by the entire FapC sequence. To further quantify the time-dependent NMR signal decay characteristics of NMR samples, a different analysis of the data was performed, shown in **Figure 3/4**. To be able to understand the time-dependent and residue-specific signal intensity decay for FapC and FapC:FapA samples, the signal intensity of each residue for each sample is normalized individually to its day 1 intensity and plotted for data recorded at longer time-points. This is similar to the data shown in **Figure 3C** for the bulk NMR signal, but now shows residue-specific representation. This presentation and normalization ensure that the signal decays can be quantified and compared for each residue independently and correctly, because the signal intensities among different residues are not uniform within the spectra (at day 1 as well as at later time points). There are large signal intensities differences in the 2D HSQC NMR spectrum shown in **Figure 1**, due to possible different dynamics and relaxation properties of each residue. As a result, a consistent way to compare the intensity

reduction at different time points is to normalize the signal intensities individually to the first time point for each residue/sample.

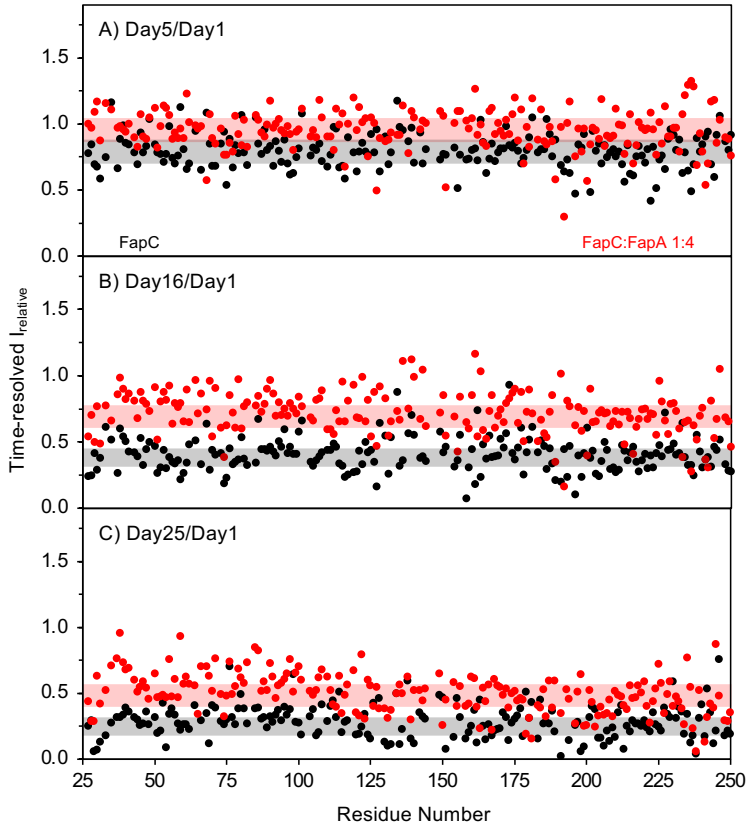


Figure 5. Representation of the time-dependent intensity reduction of each residue for pure FapC and FapC:FapA 1:4 samples. The NMR signal intensity recorded at the 2D HSQC spectra for **A)** Day5/Day1, **B)** Day16/Day1 and **C)** Day25/Day1 ratios are plotted for FapC (black) and FapC:FapA 1:4 (red) samples. These plots show if some residues within the same preparation have a different intensity decrease compared to the other regions. The averages of the intensity ratio distributions are 0.81 (A), 0.41 (B), 0.28 (C) for FapC sample, and 0.97 (A), 0.72 (B), 0.50 (C) for FapC:FapA sample, with a standard deviation of ~0.12-0.17. The average plus/minus the standard deviation are represented as black and red transparent areas for FapC and FapC:FapA samples, respectively.

The residue-specific NMR signal intensity ratios for day5/day1, day16/day1 and day25/day1 are shown in **Figure 5A-C** for FapC (black) and FapC:FapA 1:4 (red) samples. For the FapC sample, even though some irregularities exist throughout the sequence, we did not identify any areas that lose intensity significantly different relative to the rest of the protein sequence that fall outside the confidence interval (mean + standard deviation). This indicates that the entire FapC sequence is involved in the initial nucleation step towards fibril formation, or alternatively towards formation of higher molecular weight insoluble species. The FapC:FapA 1:4 sample behaves similar, however, the residues between 25-125 (NR1L1 region) may seem to represent a deviation outside the confidence interval for the day25/Day1 comparison, **Figure 5C**. This

observation may indicate that the N-terminal half of FapC behaves different from the C-terminal half upon interacting with FapA chaperone. Considering these observations, here we hypothesize that molecular initiation of FapC fibrillation is mediated by its whole sequence, and FapA maybe altering fibrillation initiation mechanism of FapC. More advanced NMR experiments such as relaxation measurements are required to shed light on these preliminary observations, which will also be beneficial to quantitatively determine the mechanism of fibrillation-initiation as well as the rates of FapC exchange between free-monomeric and bound-monomeric, or free-monomeric and bound-fibrillar states in a residue-specific manner.

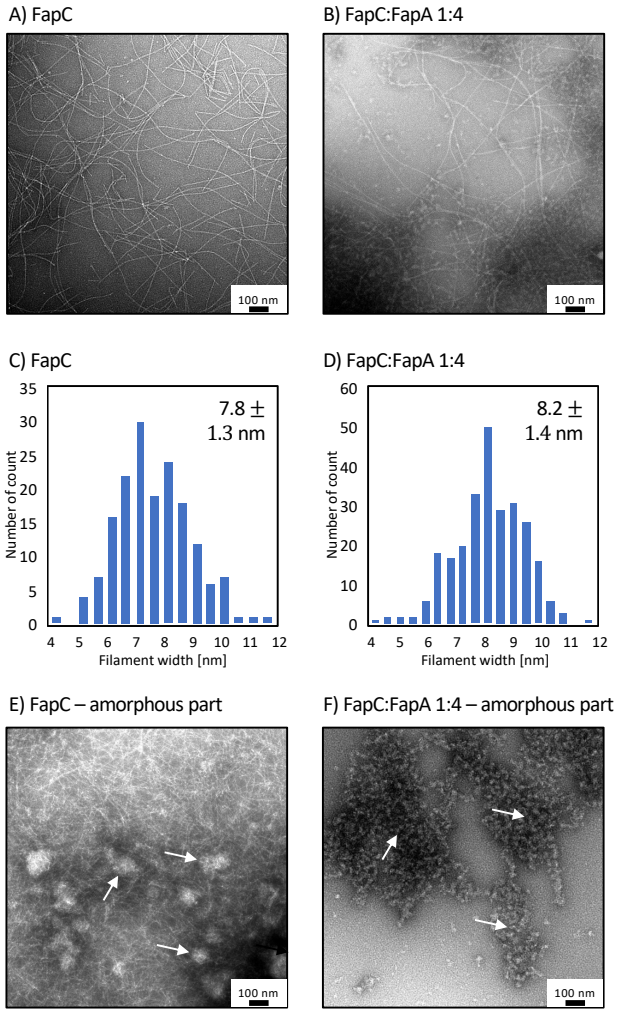


Figure 6. The negative-staining transmission electron microscopy (TEM) images of the fibrils of **A)** pure FapC and **B)** FapC:FapA 1:4 samples. The averages of ~200 filament width measurements are shown as histograms for **C)** pure FapC and **D)** FapC:FapA 1:4. Similar filament widths are observed for both samples. Note that the 100 nm scale bars are slightly different in length. The FapC contains predominantly pure filaments with a small fraction of amorphous fraction. Larger amorphous aggregate fraction is observed for FapC:FapA 1:4 sample along with still high-quality filaments. The amorphous parts are shown in **E)** for pure FapC and **F)** for FapC:FapA 1:4 samples, respectively, indicated with white arrows.

FapA does not alter the morphology of FapC. Negative-staining transmission electron microscopy (TEM) was performed to identify the morphology of the fibrils formed by FapC samples with and without FapA. The NMR samples are used right after the 25 days of NMR measurement time cycle to prepare the EM samples and to record the micrographs shown in **Figure 6A** and **B** for the pure FapC and FapC:FapA 1:4 samples. Overall, the pure FapC sample shows very high quality throughout the micrographs, with long and mostly single filaments with an averaged diameter of 7.8 ± 1.3 nm, **Figure 6A/C**, that is calculated from ~200 measurements on different filaments from different EM grids. This value is consistent with previous reports.[17, 51] Some amorphous aggregates were found for the pure FapC sample in the grids and observed consistently for different EM grid preparations.

The FapC:FapA 1:4 sample also contains long filaments similar to the ones observed for FapC with an averaged filament diameter of 8.2 ± 1.4 nm, **Figure 6B/D**. The filament widths for the pure FapC and FapC:FapA samples are similar. However, the micrographs prepared from the FapC:FapA 1:4 sample contain a much larger amount of amorphous aggregates as standalone moieties or attached to the FapC filaments (see **Figure 6E/F**), which was similarly observed previously.[17] We note that the amorphous aggregates have different appearance and ratio for both samples, and are due to FapC in the pure FapC sample, but due to excess FapA in the FapC:FapA sample. The amorphous aggregates observed for FapC:FapA sample shown in **Figure 6F** were never detected for pure FapC sample, and same type of aggregates were observed for the micrographs prepared by pure FapA (data not shown).

Discussion and Conclusions

We used high-resolution solution NMR spectroscopy to study biofilm forming FuBA FapC from *P. aeruginosa* in its prefibrillar monomeric form. We determined the secondary structural propensity of the intrinsically disordered FapC, which indicates that FapC is predominantly in random coil or PP2 helix conformation with significant α-helical and β-sheet propensities. The PP2 and the pre-structural α-helical and β-sheet propensities are key towards structural transition to the correct amyloid fibrillar fold. These SSPs and the linker regions separating them seem to align well with the secondary structure motifs from the Alphafold model of amyloid FapC. Moreover, the *in silico* predicted APRs fall within these β-sheet predicted regions in the structure. Two α-helical regions indicated by the ncSSPs do not align with the structure or APR predictions, highlighting differences between the IDP structural elements of FapC determined by NMR from its amyloid fold. Overall, the agreement of the IDP SSPs with the predicted model structure may indicate an easy and unhindered structural transition towards the amyloid fold without large structural rearrangements.

We investigated the effect of FapA chaperone on FapC fibrillation. FapA is the component of the functional amyloid operon that regulates FapC *in vivo* and prevents its unwanted fibrillation.[9] By using solution NMR,

we show that FapA is also an IDP, similar to FapC. Time-dependent NMR spectroscopy show direct evidence that FapA slows down FapC fibrillation and indeed acts as a chaperone. Starting at the first incubation time point of FapC with FapA, we observed CSPs in the FapC spectrum as well as intensity changes for the excess chaperone preparation at 1:4 mole ratio, which were not observed previously by NMR.[17] Total of 36 residues show significant shifts, which are surprisingly clustered at a few locations at the predicted FapC fibrillar structural model. The N201, C240, A241 and G245 residues have the largest CSPs, and may indicate a specific interaction of FapA with FapC at the crossing interface of L2 and R3, as well as at the C-terminus. This observation is motivating to perform more in-depth experiments to structurally quantify chaperone aided amyloid interaction/prevention mechanisms on the fibrillar FapC fold. Time-resolved NMR signal intensity changes were monitored for pure FapC and FapC:FapA 1:4 samples, which clearly proves that the FapA chaperone significantly slows the amyloid fibril formation within the 25 days of experimental time span. Lower NMR measurement temperatures was crucial for the timely and correct observation of these changes and allowed us to monitored time-dependent resonance intensity changes on these two preparations. When the FapC:FapA sample was measured at day 1 of incubation time, we observed an overall reduced signal intensity compared to the pure FapC, indicating that FapA interacts with FapC. At longer incubation times of day 5 to day 25, the averaged intensity of the FapC samples incubated with FapA becomes almost two-fold larger compared to the pure FapC. Overall, these results demonstrate the chaperone effect of the intrinsically disordered protein FapA on the functional amyloid FapC. FapA not only interacts with monomeric FapC, but also slows down its fibrillation. Moreover, hints for the structural basis of the chaperone-based prevention of functional amyloids are identified for residue-resolved interaction hotspots. FapA has chaperone properties similar to the well folded CsgE nonameric chaperone from *E. coli* functional amyloid system, despite their very different structural properties, folded versus unfolded.

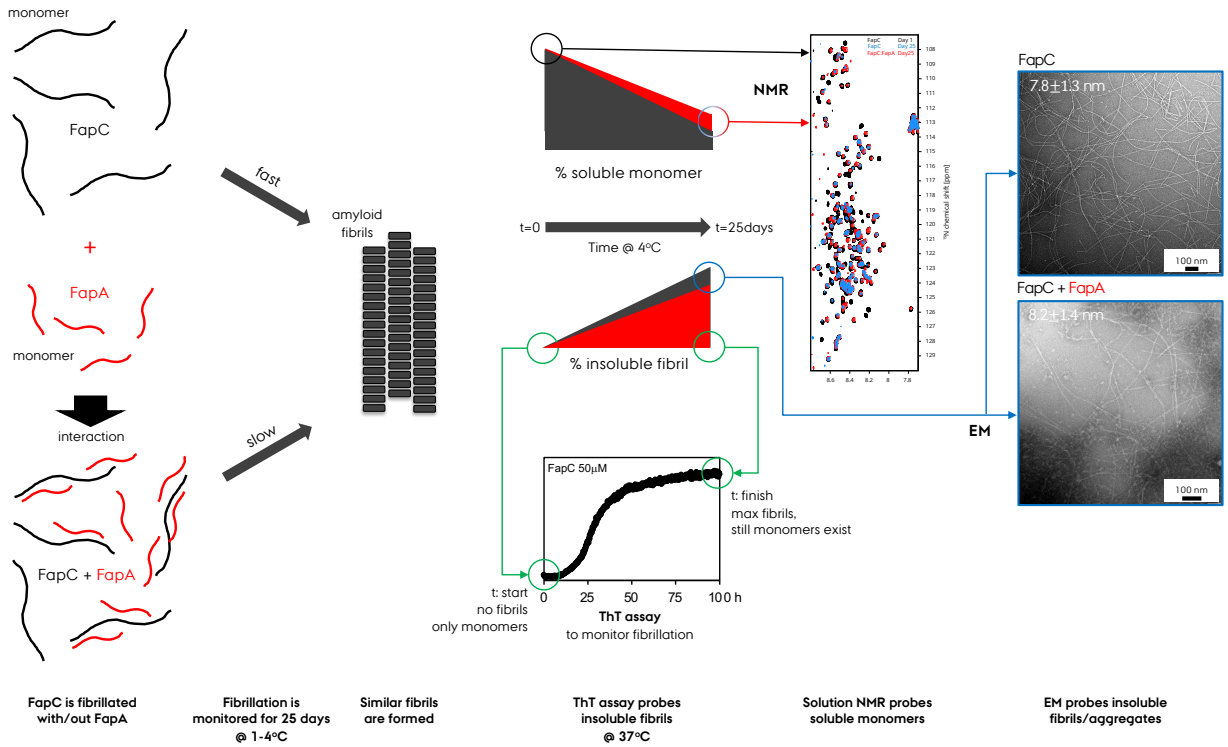


Figure 7. Schematic representation of the effect of FapA chaperone on the FapC functional amyloid fibrillation and on the FapC prefibrillar soluble monomers.

The residue-specific NMR resonance intensity ratios for pure FapC and FapC:FapA 1:4 samples are uniform at day 1 to day 5 incubation times, however, becomes less uniform and shows variations among the sequence at day 16 and 25. As a speculation, the N-terminus and R1-L2-R3 regions may have larger overall NMR resonance intensity compared to the rest of the FapC, which could indicate more specific and/or preferential action of FapA on FapC over time. Along these lines, a different analysis of the time-dependent NMR signal intensity loss data represented by a normalization approach quantified that the intensities of residues decreasing with a similar rate. This indicates that the initial nucleation step towards FapC fibril formation, or alternatively towards formation of higher molecular weight insoluble species, is mediated by the entire FapC sequence. However, one can still speculate that the FapA may be altering this behavior slightly, since the FapC:FapA sample compared to the pure FapC signal intensity ratio, shows a noticeable difference at the day 25/day 1 signal intensity ratios for the NR1L1 part of the protein compared to the second half.

The TEM results indicate that the FapC fibril morphology was not altered upon its interaction with FapA, since we observed similar filaments widths, ~8 nm. However, we observed a marked difference in the amorphous part in the EM grids, where the FapC:FapA sample has much larger amount of amorphous fraction observed throughout the grids that also seem to stick to FapC filaments as well. Even though FapC

also formed small amount of amorphous aggregates, the much larger amount of it in the FapC:FapA sample is due to the excess FapA that was precipitated during the grip preparations.

In summary, we presented a solution-state NMR based study to get atomic resolution information on FapC functional amyloid. The NMR determined secondary structure elements align very well with the predicted Alphafold fibrillar structural model. We provided insights into the first steps of amyloid formation, by monitoring the changes in the high-resolution 2D ^{1}H-^{15}N HSQC NMR spectra of soluble monomeric FapC. The effect of intrinsically disordered chaperone FapA on intrinsically disordered functional amyloid FapC fibrillation is quantified and clear evidence is provided showing that FapA indeed slows down the FapC fibrillation. However, FapA did not alter the final FapC filament morphology as seen by EM. Structural insights into the FapC-FapA interaction are provided, with correlations on to the predicted structural model. The key findings are summarized schematically in **Figure 7**. Overall, we suggest a model where FapA slows down FapC fibrillation by specifically interacting with the monomeric FapC at the flexible loops and C-terminus based on the predicted Alphafold structural model.

Methods

Recombinant protein production. Details of FapC protein expression and purification is described in detail previously,[30] we give a short summary here. The full-length FapC (residues 25-250, 226 aa) was cloned into a pET28a vector system. There are nine additional amino acids in the produced protein samples due to the construct design, one methionine on the N-terminus and a histidine purification tag of LEHHHHHH at the C-terminus. Bacterial plasmids were transformed into the BL21(DE3) *E. coli* bacteria. Minimal media containing ^{15}N-ammonium chloride is used to uniformly ^{15}N isotope label FapC. The transformed BL21(DE3) cells were plated onto LB agar and grown overnight in 37 °C incubator. The colonies were resuspended and inoculated into large volume media. These cultures were grown in shaker incubator at 37 °C until the OD600 reached a value between 0.8-1.0 OD. For induction, 1 mM IPTG was added to the culture and bacteria was grown for 4 hours at 37 °C. Cells were harvested by 6000 RCF centrifugation for 20 min at 4 °C. The cells were lysed using a sonicator in denaturing buffer. His-tag affinity column was used to purify the protein by step-eluting the protein. FapC containing fractions were pooled together and spun down by 20800 RCF, 10 min at RT. The supernatant was stored at -20 °C.

FapA gene block was ordered from Twist Bioscience and primers (IDT DNA) were used to clone into pET32 vector using In-fusion cloning mechanism (Takara Bio). The FapA construct without the signal sequence (residues 29-152, 124aa) was expressed with an N-terminal Thioredoxin (Trx) followed by a 10 histidine (His10) tag and thrombin cleavage site. Due to the lack of a tryptophan, one was engineered into the C-terminal end of the protein for accurate protein quantification. The final cleaved FapA protein

contains five additional amino acids, four at the N-terminus as cloning and cleavage artifacts (GSGT) and one at the C-terminus (W).

FapA plasmid was transformed into BL21(DE3) e. coli bacteria. The transformed BL21(DE3) cells were plated onto LB agar with ampicillin and grown overnight in 37 °C incubator. The colonies were resuspended in LB media with ampicillin antibiotic and inoculated into large volume media. These cultures were grown in shaker incubator at 37 °C until the OD600 reached a value between 0.8-1.0 OD. For induction, IPTG was added to each culture for a final concentration of 1 mM and grown for another 4 hours at 37 °C. Cells were harvested by centrifugation (6000 RCF, for 20 min at 4 °C). Cells were resuspended in resuspension buffer (50 mM Tris, 300 mM NaCl, 2 mM PMSF, pH 8.0) and lysed using a sonicator. Cell lysate was centrifuged (29,000 RCF, for 20 min at 4 °C), and the soluble fraction was decanted and collected. Trx-FapA was purified from the soluble fraction using a His-Tag affinity column using a gradient elution from 0-500mM imidazole (in addition to 50 mM Tris, 300 mM NaCl, 2 mM PMSF, pH 8.0). The eluted Trx-FapA was pooled and dialyzed into 20 mM Tris pH 8.7 at 4 °C. Thrombin was added to the dialysis bag for the last dialysis at 4U per mg of Trx-FapA, and cleavage was done overnight at 4 °C. The cleaved FapA was purified using a SourceQ column over a 0-500 mM NaCl gradient (with 20 mM Tris, 2 mM PMSF, pH 8.7). Pure FapA was pooled and stored at 4 °C.

NMR Spectroscopy. 2D ^{1}H-^{15}N HSQC NMR spectra of uniformly ^{15}N isotope labeled FapC FL were recorded at a Bruker Avance III 600 spectrometers equipped with a 5 mm triple resonance TCI cryoprobes. 3 mm NMR sample tubes were used. A total volume of 160 µl of 50 µM protein was prepared in 50 mM sodium phosphate, 30 mM DTT, 1 mM D6-DSS, 10% D_2O and at pH 7.4. The NMR samples were prepared freshly prior to performing the NMR experiment. All the 2D ^{1}H-^{15}N HSQC NMR experiments were recorded at 274 K. We previously described the details of the NMR resonance assignment, and the data were recently deposited to BMRB (51792 and 51793).[30] The ^{1}H chemical shifts were referenced directly to 0 ppm by using DSS as an internal standard added to the NMR samples, and the ^{15}N chemical shifts were indirectly referenced by using the ^{1}H frequency.[62]

2D ^{1}H-^{15}N HSQC NMR spectrum of natural abundance FapA was recorded at a Bruker Avance III 600 spectrometers equipped with a 5 mm triple resonance TCI cryoprobes. 5 mm NMR sample tube was used. A total volume of 500 µl ~0.16 mM protein was prepared in 50 mM sodium phosphate, 1 mM D6-DSS, 10% D_2O and at pH 7.5. The NMR sample was prepared freshly prior performing the NMR experiment. A total of 1024 scans were recorded for 128 increments, with 0.9 s recycle delay.

The FapC-FapA interaction assays were performed by preparing NMR samples with a constant concentration of ^{15}N isotope labeled FapC at 50 µM, and the corresponding unlabeled FapA concentrations of 50 µM for 1:1 sample, 100 µM for 1:2 sample and 200 µM for 1:4 sample in 50 mM sodium phosphate, 30 mM DTT, 1 mM D6-DSS, 10% D_2O and at pH 7.4.

Negative-staining TEM Microscopy: 400 mesh carbon film copper grids were glow discharged for 90 seconds. ~2.5 μL of pure FapC and FapC:FapA 1:4 samples were applied to grids and left for 10 seconds before side blotting on filter paper. Grids were negative stained with 2% (weight/volume) uranyl acetate for 10 seconds before side blotting on filter paper. EM micrographs were recorded at a FEI Tecnai TF20 microscope with a field emission gun operating at 200 kV equipped with a TVIPS XF416 CMOS camera. No further image processing was done to the images shown.

Declaration of Competing Interest

Authors have no conflict of interest to declare.

Data Availability

The datasets generated and/or analyzed during the current study are available in the BMRB repository, accession number #51793 (https://bmrb.io/data_library/summary/index.php?bmrbId=51793).

Author Contribution

Conceptualization: UA; Methodology: CHB, KH, JJ, MA, UA; Formal analysis and investigation: CHB, KH, JJ, MA, UA; Writing - original draft preparation: CHB, MA, UA; Funding acquisition: UA; Resources: MA, UA; Supervision: UA. All authors reviewed the manuscript. Correspondence: umitakbey@pitt.edu

Acknowledgement

UA acknowledges financial support from University of Pittsburgh startup funding and the high-field NMR infrastructure at the Structural Biology Department, School of Medicine, University of Pittsburgh.

References

(1) Chiti, F.; Dobson, C. M. Protein Misfolding, Amyloid Formation, and Human Disease: A Summary of Progress Over the Last Decade. In Annual Review of Biochemistry, Vol 86, Kornberg, R. D. Ed.; Annual Review of Biochemistry, Vol. 86; 2017; pp 27-68.

(2) Rubel, M. S.; Fedotov, S. A.; Grizel, A. V.; Sopova, J. V.; Malikova, O. A.; Chernoff, Y. O.; Rubel, A. A. Functional Mammalian Amyloids and Amyloid-Like Proteins. Life-Basel 2020, 10 (9). DOI: 10.3390/life10090156.

(3) Sawaya, M. R.; Hughes, M. P.; Rodriguez, J. A.; Riek, R.; Eisenberg, D. S. The expanding amyloid family: Structure, stability, function, and pathogenesis. Cell 2021, 184 (19), 4857-4873. DOI: 10.1016/j.cell.2021.08.013.

(4) Salinas, N.; Povolotsky, T. L.; Landau, M.; Kolodkin-Gal, I. Emerging Roles of Functional Bacterial Amyloids in Gene Regulation, Toxicity, and Immunomodulation. Microbiology and Molecular Biology Reviews 2021, 85 (1). DOI: 10.1128/mmbr.00062-20.

(5) Levkovich, S. A.; Gazit, E.; Bar-Yosef, D. L. Two Decades of Studying Functional Amyloids in Microorganisms. Trends in Microbiology 2021, 29 (3), 251-265. DOI: 10.1016/j.tim.2020.09.005.

(6) Akbey, U.; Andreasen, M. Functional amyloids from bacterial biofilms - structural properties and interaction partners. Chemical Science 2022, 13 (22), 6457-6477. DOI: 10.1039/d2sc00645f.

(7) Bjarnsholt, T. The role of bacterial biofilms in chronic infections. Apmis 2013, 121, 1-58. DOI: 10.1111/apm.12099.

(8) Chapman, M. R.; Robinson, L. S.; Pinkner, J. S.; Roth, R.; Heuser, J.; Hammar, M.; Normark, S.; Hultgren, S. J. Role of Escherichia coli curli operons in directing amyloid fiber formation. Science 2002, 295 (5556), 851-855. DOI: 10.1126/science.1067484.

(9) Dueholm, M. S.; Petersen, S. V.; Sonderkaer, M.; Larsen, P.; Christiansen, G.; Hein, K. L.; Enghild, J. J.; Nielsen, J. L.; Nielsen, K. L.; Nielsen, P. H.; et al. Functional amyloid in Pseudomonas. Molecular Microbiology 2010, 77 (4), 1009-1020. DOI: 10.1111/j.1365-2958.2010.07269.x.

(10) Romero, D.; Aguilar, C.; Losick, R.; Kolter, R. Amyloid fibers provide structural integrity to Bacillus subtilis biofilms. Proceedings of the National Academy of Sciences of the United States of America 2010, 107 (5), 2230-2234. DOI: 10.1073/pnas.0910560107.

(11) Periasamy, S.; Joo, H. S.; Duong, A. C.; Bach, T. H. L.; Tan, V. Y.; Chatterjee, S. S.; Cheung, G. Y. C.; Otto, M. How Staphylococcus aureus biofilms develop their characteristic structure. Proceedings of the National Academy of Sciences of the United States of America 2012, 109 (4), 1281-1286. DOI: 10.1073/pnas.1115006109.

(12) Khambhati, K.; Patel, J.; Saxena, V.; Parvathy, A.; Jain, N. Gene Regulation of Biofilm-Associated Functional Amyloids. Pathogens 2021, 10 (4). DOI: 10.3390/pathogens10040490.

(13) Turoverov, K. K.; Kuznetsova, I. M.; Uversky, V. N. The protein kingdom extended: Ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. Progress in Biophysics & Molecular Biology 2010, 102 (2-3), 73-84. DOI: 10.1016/j.pbiomolbio.2010.01.003.

(14) Jain, N.; Chapman, M. R. Bacterial functional amyloids: Order from disorder. Biochimica Et Biophysica Acta-Proteins and Proteomics 2019, 1867 (10), 954-960. DOI: 10.1016/j.bbapap.2019.05.010.

(15) Sewell, L.; Stylianou, F.; Xu, Y. Q.; Taylor, J.; Sefer, L.; Matthews, S. NMR insights into the pre-amyloid ensemble and secretion targeting of the curli subunit CsgA. Scientific Reports 2020, 10 (1). DOI: 10.1038/s41598-020-64135-9.

(16) Stelzl, L. S.; Pietrek, L. M.; Holla, A.; Oroz, J.; Sikora, M.; Kofinger, J.; Schuler, B.; Zweckstetter, M.; Hummer, G. Global Structure of the Intrinsically Disordered Protein Tau Emerges from Its Local Structure. Jacs Au 2022, 2 (3), 673-686. DOI: 10.1021/jacsau.1c00536.

(17) Rasmussen, H. Ø.; Kumar, A.; Shin, B.; Stylianou, F.; Sewell, L.; Xu, Y.; Otzen, D. E.; Pedersen, J. S.; Matthews, S. J. FapA is an Intrinsically Disordered Chaperone for Pseudomonas Functional Amyloid FapC. Journal of Molecular Biology 2023, 435 (2), 167878. DOI: https://doi.org/10.1016/j.jmb.2022.167878.

(18) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; et al. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. Nature 2007, 447 (7143), 453-457. DOI: 10.1038/nature05695.

(19) Ke, P. C.; Zhou, R. H.; Serpell, L. C.; Riek, R.; Knowles, T. P. J.; Lashuel, H. A.; Gazit, E.; Hamley, I. W.; Davis, T. P.; Fandrich, M.; et al. Half a century of amyloids: past, present and future. Chemical Society Reviews 2020, 49 (15), 5473-5509. DOI: 10.1039/c9cs00199a.

(20) Tycko, R. Solid-state NMR as a probe of amyloid fibril structure. Current Opinion in Chemical Biology 2000, 4 (5), 500-506. DOI: 10.1016/s1367-5931(00)00123-x.

(21) Tycko, R. Solid-State NMR Studies of Amyloid Fibril Structure. In Annual Review of Physical Chemistry, Vol 62, Leone, S. R., Cremer, P. S., Groves, J. T., Johnson, M. A. Eds.; Annual Review of Physical Chemistry, Vol. 62; 2011; pp 279-299.

(22) Bai, X. C.; McMullan, G.; Scheres, S. H. W. How cryo-EM is revolutionizing structural biology. Trends in Biochemical Sciences 2015, 40 (1), 49-57. DOI: 10.1016/j.tibs.2014.10.005.

(23) Qiang, W.; Yau, W. M.; Lu, J. X.; Collinge, J.; Tycko, R. Structural variation in amyloid-beta fibrils from Alzheimer's disease clinical subtypes. Nature 2017, 541 (7636), 217-+. DOI: 10.1038/nature20814.

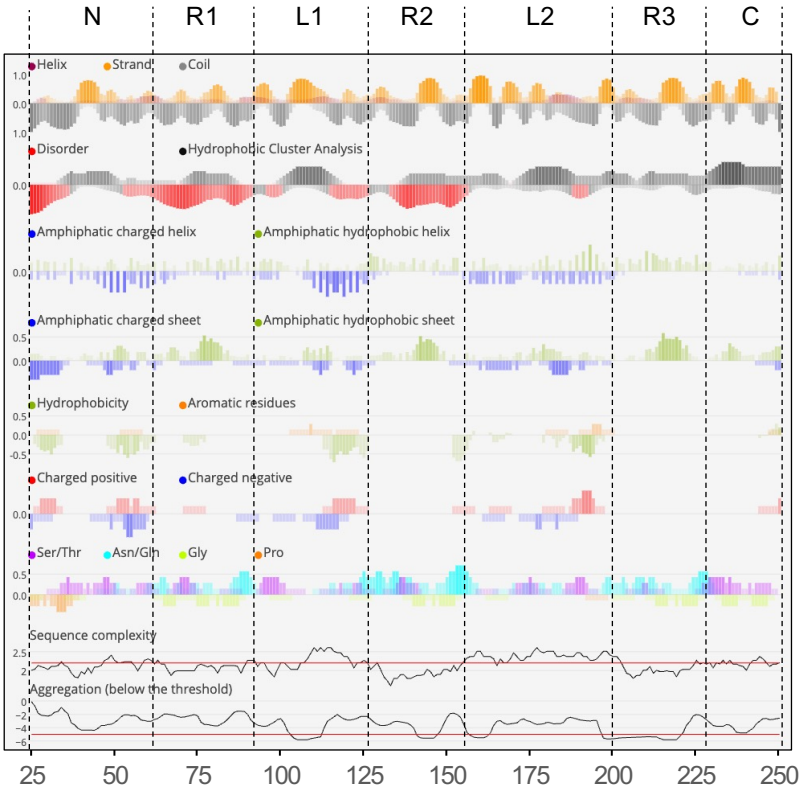
(24) Hervas, R.; Rau, M. J.; Park, Y.; Zhang, W. J.; Murzin, A. G.; Fitzpatrick, J. A. J.; Scheres, S. H. W.; Si, K. Cryo-EM structure of a neuronal functional amyloid implicated in memory persistence in Drosophila. Science 2020, 367 (6483), 1230-+. DOI: 10.1126/science.aba3526.

(25) Loquet, A.; Habenstein, B.; Lange, A. Structural Investigations of Molecular Machines by Solid-State NMR. Accounts of Chemical Research 2013, 46 (9), 2070-2079. DOI: 10.1021/ar300320p.
(26) Shimada, I.; Ueda, T.; Kofuku, Y.; Eddy, M. T.; Wuthrich, K. GPCR drug discovery: integrating solution NMR data with crystal and cryo-EM structures. Nature Reviews Drug Discovery 2019, 18 (1), 59-82. DOI: 10.1038/nrd.2018.180.
(27) Gauto, D. F.; Estrozi, L. F.; Schwieters, C. D.; Effantin, G.; Macek, P.; Sounier, R.; Sivertsen, A. C.; Schmidt, E.; Kerfah, R.; Mas, G.; et al. Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex. Nature Communications 2019, 10. DOI: 10.1038/s41467-019-10490-9.
(28) Geraets, J. A.; Pothula, K. R.; Schroder, G. F. Integrating cryo-EM and NMR data. Current Opinion in Structural Biology 2020, 61, 173-181. DOI: 10.1016/j.sbi.2020.01.008.
(29) Daskalov, A.; El Mammeri, N.; Lends, A.; Shenoy, J.; Lamon, G.; Fichou, Y.; Saad, A.; Martinez, D.; Morvan, E.; Berbon, M.; et al. Structures of Pathological and Functional Amyloids and Prions, a Solid-State NMR Perspective. Frontiers in Molecular Neuroscience 2021, 14. DOI: 10.3389/fnmol.2021.670513.
(30) Chang-Hyeock Byeon, P. C. W., In-Ja L. Byeon, Umit Akbey. Solution-state NMR Assignment and Secondary Structural Propensities of the Full-Length and Minimalistic-Truncated Prefibrillar Monomeric Form of Biofilm-Forming Functional-Amyloid FapC from Pseudomonas aeruginosa. bioRxiv. DOI: 10.1101/2023.01.22.525107.
(31) Selenko, P.; Wagner, G. Looking into live cells with in-cell NMR spectroscopy. Journal of Structural Biology 2007, 158 (2), 244-253. DOI: 10.1016/j.jsb.2007.04.001.
(32) Karamanos, T. K.; Kalverda, A. P.; Thompson, G. S.; Radford, S. E. Mechanisms of amyloid formation revealed by solution NMR. Progress in Nuclear Magnetic Resonance Spectroscopy 2015, 88-89, 86-104. DOI: 10.1016/j.pnmrs.2015.05.002.
(33) Gallardo, R.; Ranson, N. A.; Radford, S. E. Amyloid structures: much more than just a cross-beta fold. Current Opinion in Structural Biology 2020, 60, 7-16. DOI: 10.1016/j.sbi.2019.09.001.
(34) Levine, H. THIOFLAVINE-T INTERACTION WITH SYNTHETIC ALZHEIMERS-DISEASE BETA-AMYLOID PEPTIDES - DETECTION OF AMYLOID AGGREGATION IN SOLUTION. Protein Science 1993, 2 (3), 404-410. DOI: 10.1002/pro.5560020312.
(35) Ulamec, S. M.; Brockwell, D. J.; Radford, S. E. Looking Beyond the Core: The Role of Flanking Regions in the Aggregation of Amyloidogenic Peptides and Proteins. Frontiers in Neuroscience 2020, 14. DOI: 10.3389/fnins.2020.611285.
(36) Okamura, E.; Aki, K. Real-time in-situ H-1 NMR of reactions in peptide solution: preaggregation of amyloid-beta fragments prior to fibril formation. Pure and Applied Chemistry 2020, 92 (10), 1575-1583. DOI: 10.1515/pac-2019-1201.
(37) Sewell, L.; Stylianou, F.; Xu, Y.; Taylor, J.; Sefer, L.; Matthews, S. NMR insights into the pre-amyloid ensemble and secretion targeting of the curli subunit CsgA. Sci Rep 2020, 10 (1), 7896. DOI: 10.1038/s41598-020-64135-9.
(38) Santoro, A.; Grimaldi, M.; Buonocore, M.; Stillitano, I.; D'Ursi, A. M. Exploring the Early Stages of the Amyloid A beta(1-42) Peptide Aggregation Process: An NMR Study. Pharmaceuticals 2021, 14 (8). DOI: 10.3390/ph14080732.
(39) Clore, G. M. NMR spectroscopy, excited states and relevance to problems in cell biology - transient pre-nucleation tetramerization of huntingtin and insights into Huntington's disease. Journal of Cell Science 2022, 135 (12). DOI: 10.1242/jcs.258695.
(40) Fawzi, N. L.; Ying, J. F.; Torchia, D. A.; Clore, G. M. Kinetics of Amyloid beta Monomer-to-Oligomer Exchange by NMR Relaxation. Journal of the American Chemical Society 2010, 132 (29), 9948-9951. DOI: 10.1021/ja1048253.
(41) Fawzi, N. L.; Ying, J. F.; Ghirlando, R.; Torchia, D. A.; Clore, G. M. Atomic-resolution dynamics on the surface of amyloid-beta protofibrils probed by solution NMR. Nature 2011, 480 (7376), 268-U161. DOI: 10.1038/nature10577.
(42) Davies, H. A.; Rigden, D. J.; Phelan, M. M.; Madine, J. Probing Medin Monomer Structure and its Amyloid Nucleation Using 13C-Direct Detection NMR in Combination with Structural Bioinformatics. Scientific Reports 2017, 7 (1), 45224. DOI: 10.1038/srep45224.
(43) Kotler, S. A.; Tugarinov, V.; Schmidt, T.; Ceccon, A.; Libich, D. S.; Ghirlando, R.; Schwieters, C. D.; Clore, G. M. Probing initial transient oligomerization events facilitating Huntingtin fibril nucleation at atomic resolution by relaxation-based NMR. Proceedings of the National Academy of Sciences of the United States of America 2019, 116 (9), 3562-3571. DOI: 10.1073/pnas.1821216116.
(44) Schwarze, B.; Korn, A.; Höfling, C.; Zeitschel, U.; Krueger, M.; Roßner, S.; Huster, D. Peptide backbone modifications of amyloid β (1–40) impact fibrillation behavior and neuronal toxicity. Scientific Reports 2021, 11 (1), 23767. DOI: 10.1038/s41598-021-03091-4.
(45) Norrild, R. K.; Vettore, N.; Coden, A.; Xue, W. F.; Buell, A. K. Thermodynamics of amyloid fibril formation from non-equilibrium experiments of growth and dissociation. Biophysical Chemistry 2021, 271. DOI: 10.1016/j.bpc.2021.106549.
(46) Burmann, B. M.; Gerez, J. A.; Matecko-Burmann, I.; Campioni, S.; Kumari, P.; Ghosh, D.; Mazur, A.; Aspholm, E. E.; Sulskis, D.; Wawrzyniuk, M.; et al. Regulation of alpha-synuclein by chaperones in mammalian cells. Nature 2020, 577 (7788), 127-+. DOI: 10.1038/s41586-019-1808-9.
(47) Wentink, A. S.; Nillegoda, N. B.; Feufel, J.; Ubartaite, G.; Schneider, C. P.; De Los Rios, P.; Hennig, J.; Barducci, A.; Bukau, B. Molecular dissection of amyloid disaggregation by human HSP70. Nature 2020, 587 (7834), 483-+. DOI: 10.1038/s41586-020-2904-6.
(48) Williams, D. M.; Thorn, D. C.; Dobson, C. M.; Meehan, S.; Jackson, S. E.; Woodcock, J. M.; Carver, J. A. The Amyloid Fibril-Forming beta-Sheet Regions of Amyloid beta and alpha-Synuclein Preferentially Interact with the Molecular Chaperone 14-3-3 zeta. Molecules 2021, 26 (20). DOI: 10.3390/molecules26206120.
(49) Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. Progress in Nuclear Magnetic Resonance Spectroscopy 2013, 73, 1-16. DOI: https://doi.org/10.1016/j.pnmrs.2013.02.001.

(50) Piovesan, D.; Walsh, I.; Minervini, G.; Tosatto, S. C. E. FELLS: fast estimator of latent local structure. Bioinformatics 2017, 33 (12), 1889-1891. DOI: 10.1093/bioinformatics/btx085 (acccessed 2/10/2023).
(51) Nagaraj, M.; Ahmed, M.; Lyngso, J.; Vad, B. S.; Boggild, A.; Fillipsen, A.; Pedersen, J. S.; Otzen, D. E.; Akbey, U. Predicted Loop Regions Promote Aggregation: A Study of Amyloidogenic Domains in the Functional Amyloid FapC. Journal of Molecular Biology 2020, 432 (7), 2232-2252. DOI: 10.1016/j.jmb.2020.01.044.
(52) Linding, R.; Schymkowitz, J.; Rousseau, F.; Diella, F.; Serrano, L. A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. Journal of Molecular Biology 2004, 342 (1), 345-353. DOI: 10.1016/j.jmb.2004.06.088.
(53) Conchillo-Sole, O.; de Groot, N. S.; Aviles, F. X.; Vendrell, J.; Daura, X.; Ventura, S. AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. Bmc Bioinformatics 2007, 8. DOI: 10.1186/1471-2105-8-65.
(54) Louros, N.; Konstantoulea, K.; De Vleeschouwer, M.; Ramakers, M.; Schymkowitz, J.; Rousseau, F. WALTZ-DB 2.0: an updated database containing structural information of experimentally determined amyloid-forming peptides. Nucleic Acids Research 2020, 48 (D1), D389-D393. DOI: 10.1093/nar/gkz758.
(55) Blanch, E. W.; Morozova-Roche, L. A.; Cochran, D. A. E.; Doig, A. J.; Hecht, L.; Barron, L. D. Is polyproline II helix the killer conformation? A Raman optical activity study of the amyloidogenic prefibrillar intermediate of human lysozyme. Journal of Molecular Biology 2000, 301 (2), 553-563. DOI: 10.1006/jmbi.2000.3981.
(56) Dueholm, M. S.; Otzen, D.; Nielsen, P. H. Evolutionary Insight into the Functional Amyloids of the Pseudomonads. Plos One 2013, 8 (10). DOI: 10.1371/journal.pone.0076630.
(57) Andreasen, M.; Meisl, G.; Taylor, J. D.; Michaels, T. C. T.; Levin, A.; Otzen, D. E.; Chapman, M. R.; Dobson, C. M.; Matthews, S. J.; Knowles, T. P. J. Physical Determinants of Amyloid Assembly in Biofilm Formation. Mbio 2019, 10 (1). DOI: 10.1128/mBio.02279-18.
(58) Bardwell, J. C. A.; Jakob, U. Conditional disorder in chaperone action. Trends in Biochemical Sciences 2012, 37 (12), 517-525. DOI: 10.1016/j.tibs.2012.08.006.
(59) Ikeda, K.; Suzuki, S.; Shigemitsu, Y.; Tenno, T.; Goda, N.; Oshima, A.; Hiroaki, H. Presence of intrinsically disordered proteins can inhibit the nucleation phase of amyloid fibril formation of A beta(1-42) in amino acid sequence independent manner. Scientific Reports 2020, 10 (1). DOI: 10.1038/s41598-020-69129-1.
(60) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021, 596 (7873), 583-+. DOI: 10.1038/s41586-021-03819-2.
(61) Salmon, L.; Nodet, G.; Ozenne, V.; Yin, G. W.; Jensen, M. R.; Zweckstetter, M.; Blackledge, M. NMR Characterization of Long-Range Order in Intrinsically Disordered Proteins. Journal of the American Chemical Society 2010, 132 (24), 8407-8418. DOI: 10.1021/ja101645g.
(62) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. H-1, C-13 AND N-15 RANDOM COIL NMR CHEMICAL-SHIFTS OF THE COMMON AMINO-ACIDS .1. INVESTIGATIONS OF NEAREST-NEIGHBOR EFFECTS. Journal of Biomolecular Nmr 1995, 5 (1), 67-81. DOI: 10.1007/bf00227471.

Supplementary Information

Figure SI1: FELLS sequence analysis of FapC 25-250 amino acids, as described in the study. FapC amino acid sequence are given in Figure 1 in the manuscript. Different sequence specific protein properties are shown in the plot and summarized in the table below. See more information on the nomenclature at: http://old.protein.bio.unipd.it/fells



Feature	Blobs	%
Hydrophobic Cluster Analysis (HCA)	36	15.93
Hydrophobic	16	7.08
Charged positive	3	1.33
Charged negative	2	0.88
Aromatic (F/Y/W)	0	0.00
Glycine	0	0.00
Proline	3	1.33
Serine/Threonine	0	0.00
Asparagine/Glutamine	23	10.18
Amphi Helix Charged	10	4.42
Amphi Helix Hydrophobic	3	1.33
Amphi Sheet Charged	18	7.96
Amphi Sheet Hydrophobic	28	12.39

Feature	Residues	%
Sequence	226	100.00
Helix	0	0.00
Sheet	71	31.42
Coil	127	56.19
Disorder	97	42.92
Low complexity	145	64.16
Aggregation	71	31.42

Figure SI2: TANGO, WALTZ, AGGRESCAN results showing the aggregation prone regions of FapC. For comparison the NMR determined ncSSPs are shown on the top.

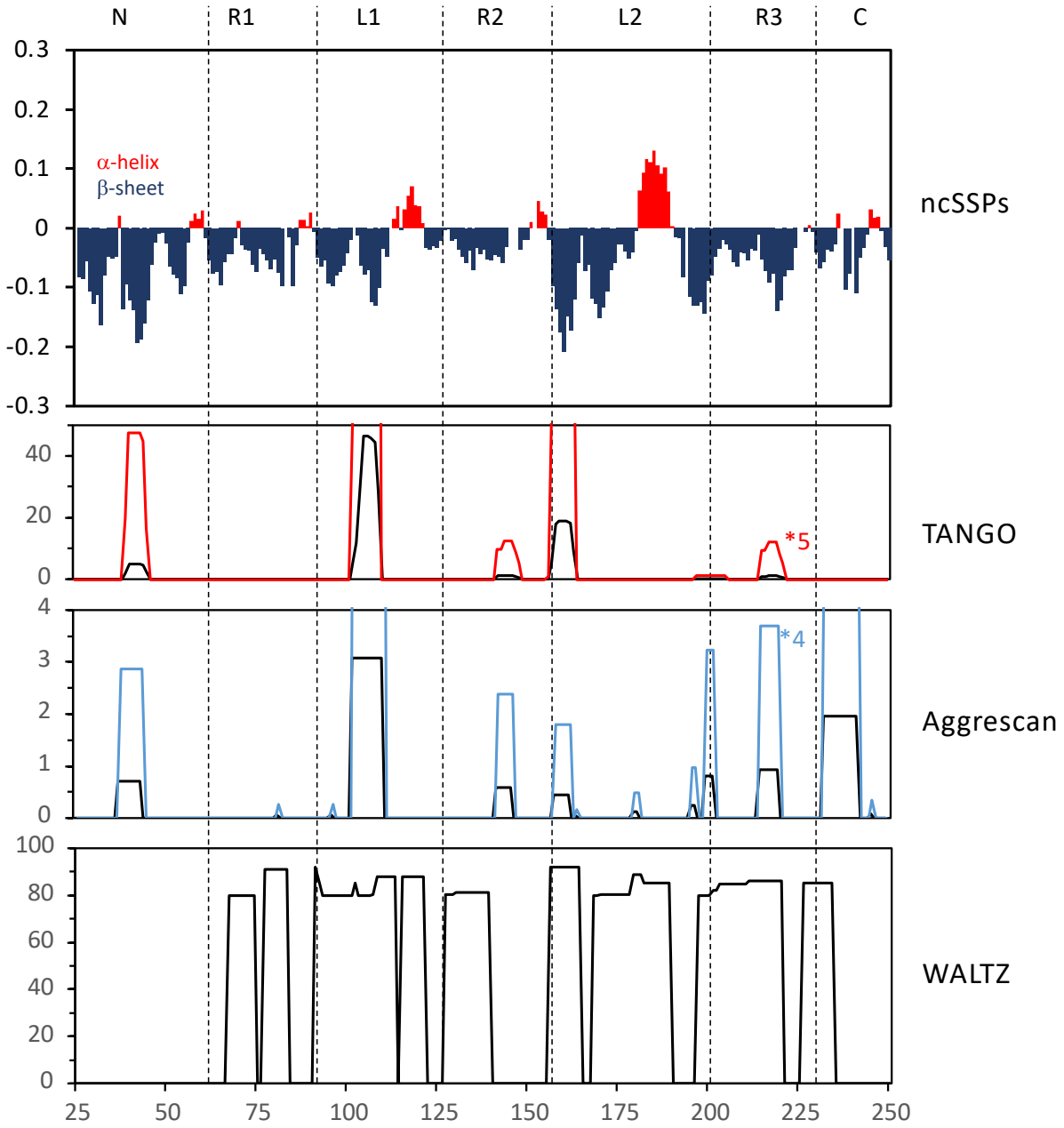
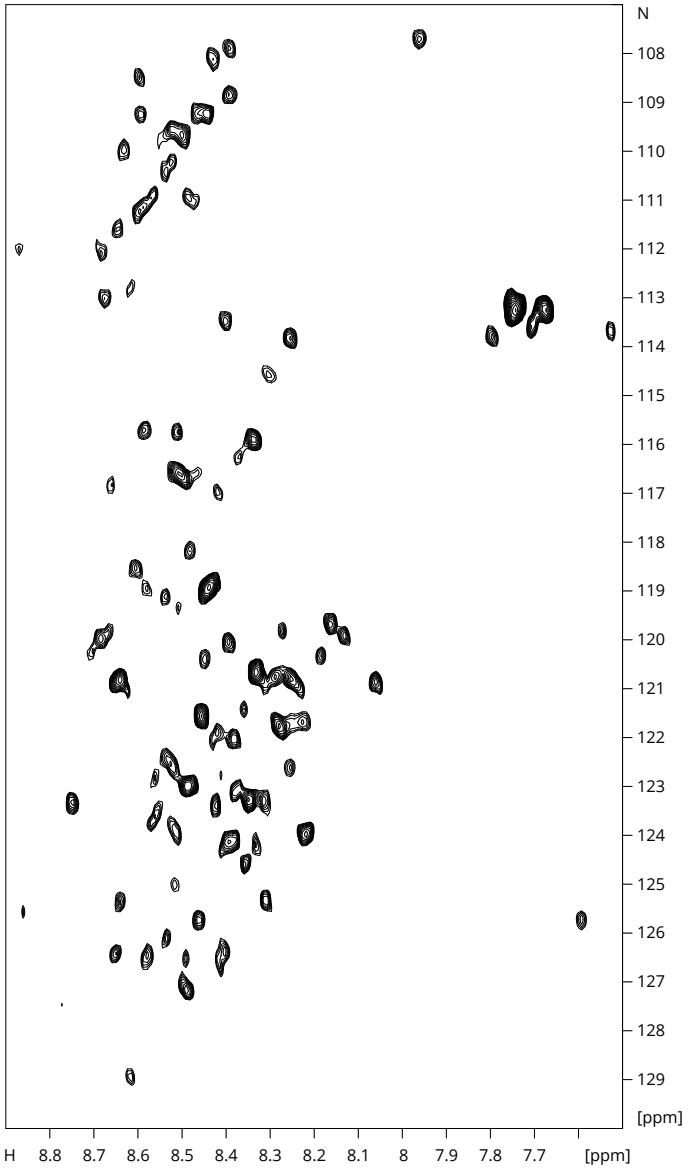


Figure SI3: 2D ^{1}H-^{15}N HSQC NMR spectrum of natural abundance FapA (29-152 amino acids, without the signal peptide), along with amino acid sequence and statistics given below. The NMR spectrum was recorded at a Bruker Avance III 600 spectrometers equipped with a 5 mm triple resonance TCI cryoprobes. 5 mm NMR sample tube was used. A total volume of 500 µl ~0.16 mM protein was prepared in 50 mM sodium phosphate, 1 mM D6-DSS, 10% D_2O and at pH 7.5. The NMR sample was prepared freshly prior to performing the NMR experiment. A total of 1024 scans were recorded for 128 increments, with 0.9 s recycle delay. The number of glycine cross peaks in the HSQC spectrum matches to the number of glycine in the FapA sequence.



FapA sequence (C4IN68: 1-28 signal sequence / 29-152 amino acid):
MATLNQGSLS LLLISCVLGA SVSVPASAED GIIVTGRNVQ GFMVGRPSFG KDPYPSTANA NPSQQILRAT GGELSDNDFA GVSSGSTITR AILPNGDLPG LSNTLGNNSA GLGAGAAAGH ASGNSLGGQI SGSIERGLAP LNNIGAMMGV QP

Amino acid statistics for FapA 29-152 (124 amino acids) construct used to record the 2D spectrum above: Ala 13, Arg 5, Asn 11, Asp 5, Cys 0, Gln 5, Glu 3, Gly 23, His 1, Ile 8, Leu 10, Lys 1, Met 3, Phe 3, Pro 8, Ser 13, Thr 6, Trp 0, Tyr 1, Val 5

Figure SI4: Zoomed regions of individual resonances from the 2D ^{1}H-^{15}N HSQC NMR spectra of the FapC (black) and FapC:FapA 1:4 (red) samples showing chemical shift perturbations at Day 1 incubation time.

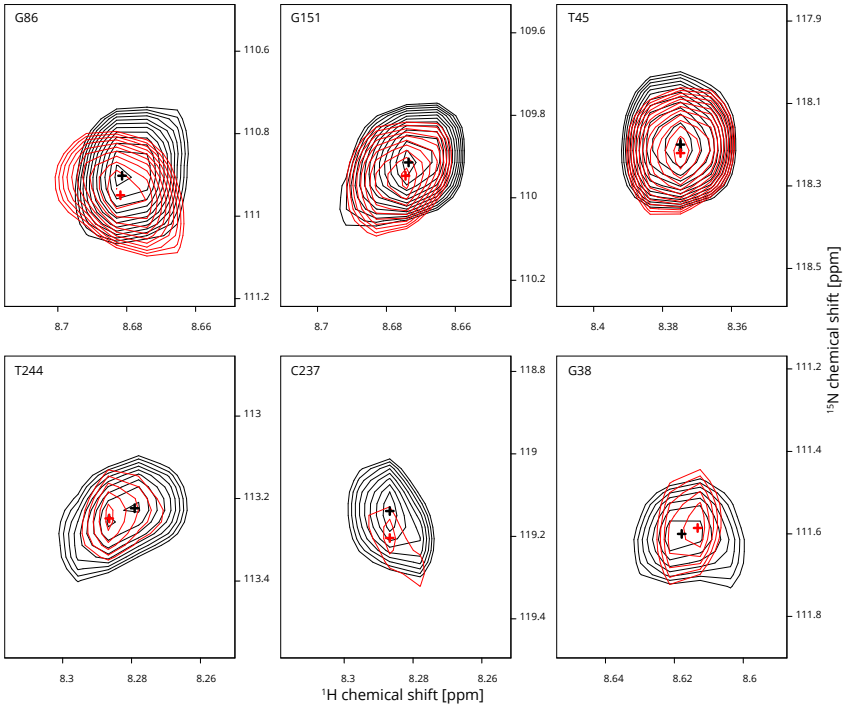
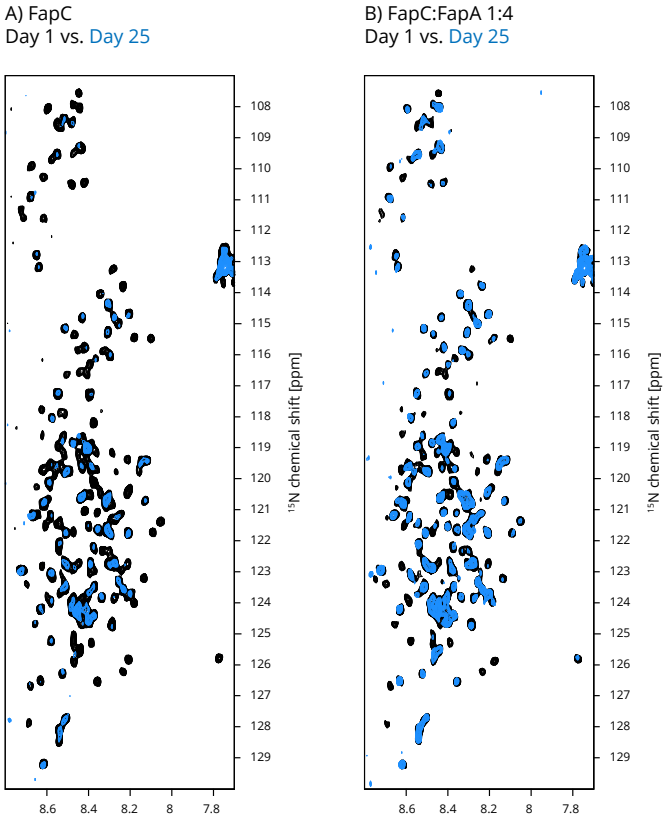



Figure SI5: Overlay of the 2D ^{1}H-^{15}N HSQC NMR spectra for pure FapC and FapC:FapA 1:4 samples recorded at Day 1 (black) and Day 25 (blue).