



**HAL**  
open science

## Prion soft amyloid core driven self-assembly of globular proteins into bioactive nanofibrils

Weiqiang Wang, Susanna Navarro, Rafayel Azizyan, Manuel Baño-Polo, Sebastian Esperante, Andrey Kajava, Salvador Ventura

### ► To cite this version:

Weiqiang Wang, Susanna Navarro, Rafayel Azizyan, Manuel Baño-Polo, Sebastian Esperante, et al.. Prion soft amyloid core driven self-assembly of globular proteins into bioactive nanofibrils. *Nanoscale*, 2019, 11 (26), pp.12680-12694. 10.1039/c9nr01755k . hal-03039021

**HAL Id: hal-03039021**

**<https://hal.science/hal-03039021>**

Submitted on 30 Dec 2020

**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.

# Prion soft amyloid core driven self-assembly of globular proteins into bioactive nanofibrils

*Weiqiang Wang<sup>1</sup>, Susanna Navarro<sup>1</sup>, Rafayel A. Azizyan<sup>2</sup>, Manuel Baño-Polo<sup>1</sup>, Sebastian A. Esperante<sup>1</sup>, Andrey V. Kajava<sup>2</sup> and Salvador Ventura<sup>1\*</sup>*

<sup>1</sup>Institut de Biotecnologia i de Biomedicina and Departament de Bioquímica i Biologia Molecular; Universitat Autònoma de Barcelona; 08193 Bellaterra (Barcelona), Spain.

<sup>2</sup>Centre de Recherche en Biologie cellulaire de Montpellier, UMR 5237 CNRS, Université Montpellier, 1919 Route de Mende, 34293 Montpellier, Cedex 5, France

E-mail: salvador.ventura@uab.es

## **Abstract**

Amyloids have been exploited to build up amazing bioactive materials. In most cases, short synthetic peptides constitute the functional components of such materials. The controlled assembly of globular proteins into active amyloid nanofibrils is still challenging, because the formation of amyloids implies a conformational conversion towards a  $\beta$ -sheet-rich structure, with a concomitant loss of the native fold and the inactivation of the protein. There is, however, a remarkable exception to this rule: the yeast prions. They are singular proteins able to switch between a soluble and an amyloid state. In both states, the structure of their globular domains remains essentially intact. The transit between these two conformations is encoded in prion domains (PrDs): long and disordered sequences to which the active globular domains are appended. PrDs are much larger than typical self-assembling peptides. This seriously limits their use for nanotechnological applications. We have recently shown that these domains contain soft amyloid cores (SACs) that suffice to nucleate their self-assembly reaction. Here we genetically fused a model SAC with different globular proteins. We demonstrate that this very short sequence act as minimalist PrDs, driving the selective and slow assembly of the initially soluble fusions into amyloid fibrils in which the globular proteins keep their native structure and display high activity. Overall, we provide here a novel, modular and straightforward strategy to build up active protein-based nanomaterials at a preparative scale.

**Keywords:** protein self-assembly, amyloid fibrils, prion domain, soft amyloid core, functional amyloids, nanomaterials.

## Introduction

The formation of amyloid fibrils is associated with the onset of a range of protein misfolding diseases.<sup>1</sup> However, amyloid structures are also exploited for functional purposes by different organisms.<sup>2,3</sup> The inner intermolecular  $\beta$ -sheet structure of the fibrils makes these protein assemblies very stable,<sup>4</sup> even in harsh environments. This, together with their tuneable assembly under physiological conditions, make them attractive modules to build up nanomaterials for biomedical and biotechnological applications, including catalysis,<sup>5</sup> biosensors,<sup>6</sup> electronics,<sup>7</sup> tissue engineering<sup>8</sup> or drug delivery.<sup>9</sup> Most efforts so far have been focused on the use of short synthetic peptides as the bioactive components of such materials,<sup>10, 11</sup> while an analogous approach for inducing globular proteins to assemble into functional nanofibres has been much less explored.

The main limitation to create amyloids that display functional proteins comes from the connection between protein function and the attainment and maintenance of a defined folded state. Protein folding and aggregation are two competing reactions which depend on overlapping physicochemical properties.<sup>12</sup> In a large majority of cases, the aggregation of the protein into amyloid fibrils implies a conformational conversion in which it loses the native fold and gains  $\beta$ -sheet structure,<sup>1</sup> becoming thus inactive in the aggregated state.

The design of fibrillar structures containing properly folded domains appears as a challenging task, since the protein of interest should fulfil contradictory properties: it should remain soluble and folded during recombinant expression and subsequent purification and storage, but, at the same time, it should be able to self-assemble into  $\beta$ -

sheet rich amyloid-like structures and, moreover, this must occur without a structural conversion of the globular domains. A way to bypass these limitations consists in the design of hybrid structures in which purified folded domains are chemically linked to preformed fibrillar structures.<sup>13</sup> However, this strategy is deterred by the limited available polypeptide conjugation chemistries, their cross-reactivity and the unavoidable reduction in the proportion of conformationally active molecules in the assembly.

Nevertheless, functional proteins with intrinsic self-assembling properties already exist in nature, among them, the yeast prions. Yeast prions are proteins expressed and stored in the cell in a soluble state, but they are able to self-associate into amyloid structures under certain conditions.<sup>14</sup> A common feature of these proteins is the presence of a prion domain (PrD). PrDs correspond to intrinsically unstructured sequences of low complexity highly enriched in asparagine (N) and/or glutamine (Q) residues<sup>15</sup> and are accompanied by one or more globular domains at their C- or N-terminus.<sup>16</sup> The PrD is both necessary and sufficient for self-assembly, whereas the globular domains account for the protein activity.<sup>17</sup> The nice thing here is that there are evidences that only the PrD is integrated in the core of the fibril, whereas the globular domains hang from the fibril in a folded conformation.<sup>18</sup>

Because yeast prions are modular, one can, in principle, fuse any globular domain to a given PrD and potentially obtain a functionalized fibrillar nanomaterial. In this way, enzymes like alkaline phosphatase (AP)<sup>19</sup> or carbonic anhydrase (CA)<sup>20</sup> have been genetically fused to the PrD of the yeast prion Ure2p and methyl-parathion hydrolase

(MPH) to the PrD of Sup35 prion.<sup>21</sup> The PrDs were shown to drive the association of the correspondent fusion proteins into amyloid fibrils, generating catalytic nanomaterials that displayed folded enzyme moieties.

The required equilibrium between solubility and aggregation propensities explains the absence of highly amyloidogenic sequences in PrDs,<sup>22</sup> fibrillation being thought to rely on the establishment of a large number of weak interactions distributed along the complete low complexity sequence.<sup>23</sup> As a consequence, PrDs are much larger than the majority of self-assembling peptides used for nanotechnological applications.<sup>24</sup> Traditionally, a length of at least 80 residues has been considered necessary for the conversion of PrDs into amyloids.<sup>25</sup> Thus, the PrD would constitute a significant fraction of any protein fusion, especially when fused with small globular domains. In addition, the large size of PrDs results in steric hindrance in the fibrillar state, reducing substrates diffusion and significantly limiting the activity of the adjacent domains, relative to their soluble counterparts.<sup>19</sup> Indeed, in the natural Ure2p yeast prion, the globular GST domain becomes inactive in the fibrils, despite its native-like structure.<sup>20</sup>

We have recently shown that, in addition to a distinctive amino acidic composition, PrDs contain inner weak amyloidogenic sequence stretches that contribute to trigger the initial protein self-assembly reaction.<sup>26</sup> These cryptic amyloids promote conformational conversion in bona fide yeast prions,<sup>27</sup> but they also exist in human prion-like proteins.<sup>28</sup> The soft amyloid cores (SAC) embedded within PrDs can be identified computationally.<sup>29</sup> They differ from the classical amyloid cores of pathogenic proteins in that they are slightly

longer and more polar, in such a way that the amyloid potential is less concentrated, allowing the containing PrD to remain soluble under most physiological conditions, while still keeping a certain amyloid propensity that might facilitate its assembly in certain circumstances.<sup>30</sup>

We hypothesized here, that complete PrDs can be substituted by their SACs, in such a way that these shorter sequences would recapitulate the larger domain self-association properties in the context of modular fusion proteins. It is expected that the polar nature of SAC would prevent aggregation and misfolding of the adjacent globular domain during recombinant expression, purification and storage, while still being able to induce its self-assembly under controlled conditions. Ideally, in the resulting amyloids fibrils, the SAC would form the core of the fibril, whereas the attached globular domain would remain folded and active. As a proof of principle, we selected the canonical yeast prion Sup35, an eukaryotic translation release factor.<sup>31</sup> Its SAC corresponds to a 21-residues long sequence stretch that autonomously self-assembles into highly ordered amyloid fibrils, displaying a typical cross- $\beta$  diffraction pattern. This sequence is able to seed amyloid formation by the entire PrD, *in vitro*, and of the complete protein, *in vivo*,<sup>32</sup> and is indispensable for the induction, propagation and inheritance of the prion state in the mammalian cytosol.<sup>33</sup> We show here, how the properties of this short sequence stretch can be exploited to obtain highly functional amyloid-like nanofibers. We illustrate the wide applicability of the approach by fusing this SAC to three globular proteins displaying different structure;

namely the all- $\alpha$  FF domain, the all- $\beta$  green fluorescent protein and the  $\alpha/\beta$  carbonic anhydrase enzyme.



## Results and discussion

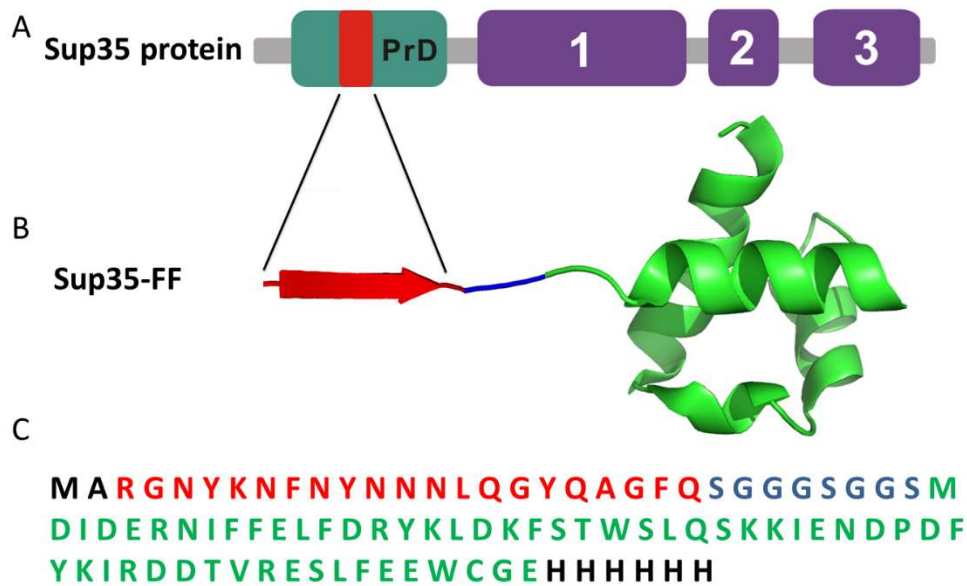
### Design of a self-assembling protein fusion consisting of Sup35-SAC and a small all $\alpha$ -helical domain

The 21 residues-long SAC of Sup35 PrD (Sup35-SAC) corresponds to residues 98-118 of the yeast protein (**Figure 1A**). We decided to exploit the prion-like characteristics of this small segment to design prion-inspired self-assembling fusion proteins, consisting of a chimera of Sup35-SAC and different globular proteins. As a proof of principle, we used the FF domain of the URN1 splicing factor.<sup>34</sup>

FF domains are small protein-protein interaction modules of 50-70 residues, characterized by a fold that consists of three  $\alpha$ -helices arranged as an orthogonal bundle with a  $3_{10}$  helix in the loop connecting the second and the third helices (**Figure 1B**).<sup>35</sup> FF domains are highly soluble and the folding landscape of several of these proteins, including the URN1-FF domain (57 residues), have been characterized in detail.<sup>36,37</sup> Thus, it constitutes a perfect model protein to assess: *i*) whether the fusion of Sup35-SAC to a globular domain impacts its solubility, conformation, thermodynamic stability and/or folding kinetics and, *ii*) if this short segment suffices to drive the self-assembly of a soluble globular protein into amyloid-like structures, allowing to maintain its native structure in the aggregated state.

We fused Sup35-SAC to the N-terminus of the FF domain. A flexible linker consisting of SG<sub>3</sub>SG<sub>2</sub>S was incorporated between both protein moieties (**Figure 1C**) and a His6 tag at the FF domain C-terminus. In this fusion, the ratio between the size of the self-

assembling sequence and the globular domain is 0.4, whereas in the case of the full-length Sup35 PrD (114 residues) instead, this ratio would have increased to 2.0.



**Figure 1. Schematic representation of Sup35-FF fusion protein.** (A) Domain organization of Sup35 protein: soft amyloid core (SAC) (red) within the PrD (green) and the respective elongation factor functional domains (purple) in Sup35 are shown. (B) Cartoon of Sup35 soft amyloid core fused to FF domain shown as ribbon representation (PDB: 2JUC). (C) Sequence of the Sup35-FF fusion protein. Sup35 soft amyloid core, spacer linker, FF domain and His6 tag are shown in red, blue, green and black, respectively.

Previous studies showed that proteins appended with  $\beta$ -sheet forming peptides aggregated into insoluble inclusion bodies during expression.<sup>38,39,40</sup> Thus, we first assessed computationally if the addition of Sup35-SAC to the FF domain would impact its inherent

solubility, compromising thus its expression, using three different aggregation propensity algorithms: AGGRESCAN<sup>41</sup>, FoldAmyloid<sup>42</sup> and PASTA<sup>43</sup>, all them predicting Sup35-SAC as having low aggregation propensity, compared with classical amyloid cores (**Table S1A**). Indeed, 57 % of the residues in Sup35-SAC are polar with only 19 % of them being strictly hydrophobic. The rest correspond to a 14 % of Tyr, sharing both characters, and another 14 % of Gly, a residue with very low aggregation propensity. Accordingly, Sup35-SAC is not expected to impact significantly the solubility of the globular protein in the fusion (**Table S1A**). Therefore, we proceeded to construct the fusion protein (Sup35-FF) and expressed it in *E. coli*. As predicted, the protein was almost completely located in the soluble cell fraction (**Figure S1**), from which it was purified by IMAC and gel-filtration chromatography (**Figure S2A**). The yield was 68 mg of purified fusion per L of culture, much higher than that reported for other fusions between amyloidogenic peptides of similar size and globular proteins, where a significant fraction of the protein remained insoluble.<sup>44</sup>

### **Sup35-SAC does not affect the conformation of the FF domain in the protein fusion**

A first requirement to use Sup35-FF for building up functional nanofibers is that the N-terminal Sup35-SAC does not alter the conformation, stability or folding properties of the globular domain.

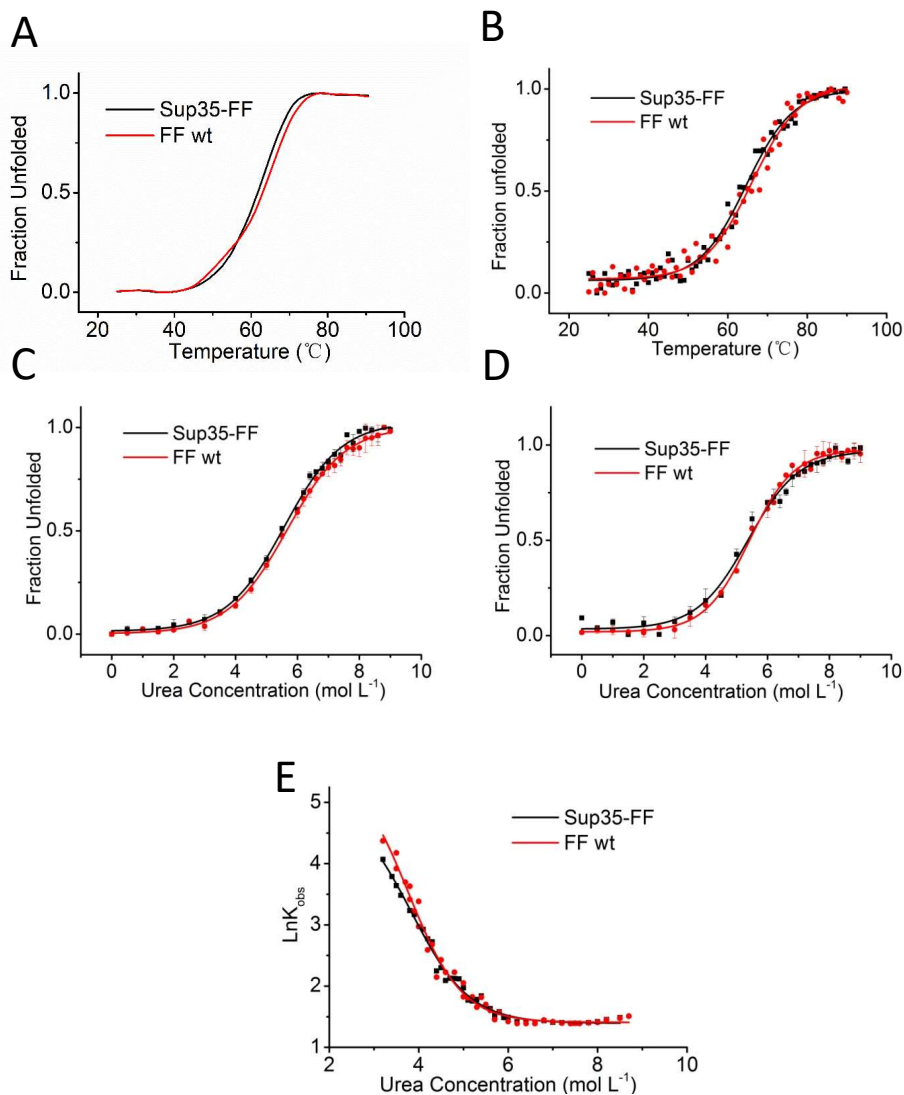
We compared the conformational properties of Sup35-FF and the FF domain alone (FF-wt) at pH 7.4 and 25 °C by monitoring the far-UV CD spectra and Trp intrinsic fluorescence. The far-UV CD spectra of the two proteins at different concentrations closely

resemble and are dominated by  $\alpha$ -helical signals (**Figure S2B**). The intrinsic fluorescence spectrum of FF-wt exhibits an emission maximum at 337 nm. In Sup35-FF, it is shifted to 334.5 nm (**Figure S2C**), suggesting the Trp being in a slightly more protected environment. The FF domain possesses two Trp residues at positions 27 and 56. Trp27 is buried, while Trp56 is structurally adjacent to the N-terminal  $\alpha$ -helix and partially exposed to solvent. Four different disorder prediction algorithms: PONDR (VSL2)<sup>45</sup>, GlobPlot<sup>46</sup>, PASTA<sup>43</sup>, and IUPred<sup>47</sup> suggest that the Sup35-SAC and the linker region will remain disordered in the context of the fusion protein (**Table S1B**). Therefore, it is likely that the structural fluctuations of the disordered N-terminal tail might shield, at least partially, the exposed Trp56 side chain in the FF domain, as we already observed for human SUMO domains.<sup>48</sup> In any case, the shift towards the blue region of the spectrum is indicative of a properly folded domain, since upon destabilization and partial unfolding, the Trp spectrum of the FF domain shifts to the red, displaying a large increase in its fluorescence emission.<sup>36</sup>

### **Sup35-SAC does not affect the thermodynamic stability of the FF domain in the protein fusion**

Despite the native-like conformation of the FF domain in the protein fusion, the adjacent exogenous sequence might compromise the protein stability, promoting partial unfolding. To the best of our knowledge, the stability of globular domains when attached to exogenous PrDs or amyloidogenic sequences has not been yet addressed, likely because these fusions begin to aggregate soon after purification.

The equilibrium unfolding of Sup35-FF and FF-wt proteins was analysed at pH 7.4 and 25 °C. The urea denaturation curves at equilibrium were obtained recording the changes in Trp intrinsic fluorescence at 350 nm (**Figure 2C**) and in molar ellipticity at 222 nm (**Figure 2D**) at increasing denaturant concentrations. Both proteins displayed a single visible transition indicative of a cooperative unfolding process. The main thermodynamic parameters of the unfolding reaction were calculated from the equilibrium curves assuming a two-state model ( $R > 0.99$ ) (**Table S2**). The stabilities of Sup35-FF calculated from fluorescence and CD measurements were similar with  $\Delta G_{U-F}^{H_2O} \approx 4.03$  kcal/mol and a  $[\text{Urea}]_{50\%}$  of  $\approx 5.47$  M. FF-wt shows similar values, with a  $\Delta G_{U-F}^{H_2O} \approx 4.15$  kcal/mol and  $[\text{Urea}]_{50\%}$  of  $\approx 5.57$  M, indicating that the N-terminal extension has a minor impact in the domain's chemical stability (**Table S2**).



**Figure 2. Stability and folding properties of Sup35-FF and FF-wt proteins.** Thermal stabilities were analysed by (A) Trp intrinsic fluorescence emission at 350 nm and, (B) far-UV CD signal at 222 nm. Chemical equilibrium curves with urea were followed at 25 °C by (C) Trp intrinsic fluorescence at 350 nm and, (D) far-UV CD at 222 nm. (E) The kinetics of folding and unfolding were measured by Trp intrinsic fluorescence at 25 °C performing stopped-flow experiments. The rate constants were calculated under conditions of apparent two-state folding.

Thermal unfolding of Sup35-FF and FF-wt at pH 7.4 was followed by Trp intrinsic fluorescence (**Figure 2A**) and far-UV CD (**Figure 2B**), monitoring the changes in molar

ellipticity and fluorescence emission at 222 nm and 350 nm, respectively. Thermal denaturation curves show a single cooperative transition. The melting temperatures obtained for Sup35-FF were  $T_m$  of 64.7 °C and 62.7 °C by far-UV CD and intrinsic fluorescence, respectively. For FF-wt we recorded  $T_m$  of 66.1 °C and 64.6 °C by far-UV CD and intrinsic fluorescence, respectively. Therefore, Sup35-SAC has only a minor effect on the thermal stability of the FF domain (**Table S3**).

### **Sup35-SAC does not affect the folding and unfolding kinetics of the FF domain in the protein fusion**

As far as we know, the folding properties of a globular domain when embedded in a fusion with an aggregation-prone segment have never been assessed. Here, we determined the kinetics of folding and unfolding of FF domains by stopped-flow at pH 7.4 and 25 °C by monitoring the changes in Trp intrinsic fluorescence under a wide range of urea concentrations. For both Sup35-FF and FF-wt, the folding and unfolding fluorescence traces fit well to a single exponential function. Chevron plots appear to be linear in the entire range of denaturant concentrations studied and fit well to a two-state model, indicating the absence of detectable intermediates (**Figure 2E**). The rate constants for folding ( $k_F$ ) and unfolding ( $k_U$ ) and their dependence on the denaturant concentration ( $m_F$  and  $m_U$ ) are shown in **Table 1**. The folding and unfolding rates of Sup35-FF are only a 24% slower and an 8% faster than that of FF-wt, respectively. According to this kinetic

rates, Sup35-FF is only about 0.2 kcal/mol more stable alone than in the fusion, in good agreement with the equilibrium data.

**Table 1. Folding kinetics parameters of Sup35-FF and FF-wt proteins.**

	$\Delta G_{U-F}^{U \rightarrow F}$ <sup>a</sup> (kcal mol <sup>-1</sup> )	$m_{U-F}$ <sup>b</sup> (kcal mol <sup>-1</sup> M <sup>-1</sup> )	$C_m$ <sup>c</sup> (M)	$k_F$ (s <sup>-1</sup> )	$k_U$ (s <sup>-1</sup> )	$RTm_U$ (kcal mol <sup>-1</sup> M <sup>-1</sup> )	$RTm_F$ (kcal mol <sup>-1</sup> M <sup>-1</sup> )
Sup35-FF	4.32±0.13	0.79±0.08	5.50±0.56	2835±260	1.93±0.25	0.057±0.009	0.728±0.083
FF-wt	4.53±0.16	0.78±0.04	5.74±0.40	3714±733	1.78±0.12	0.063±0.004	0.719±0.043

<sup>a</sup> Gibbs energy of unfolding at [Urea] = 0.

<sup>b</sup> m value, dependence of free energy of unfolding with denaturing agent.

<sup>c</sup> The urea concentration required to unfold 50% of the protein molecules.

Overall, the data presented in this and the previous sections clearly indicate that, in the soluble state, the FF domain maintains its fold, stability and folding properties almost intact when fused to Sup35-SAC.

### **Sup35-SAC promotes the assembly of the Sup35-FF fusion into amyloid fibrils under mild conditions**

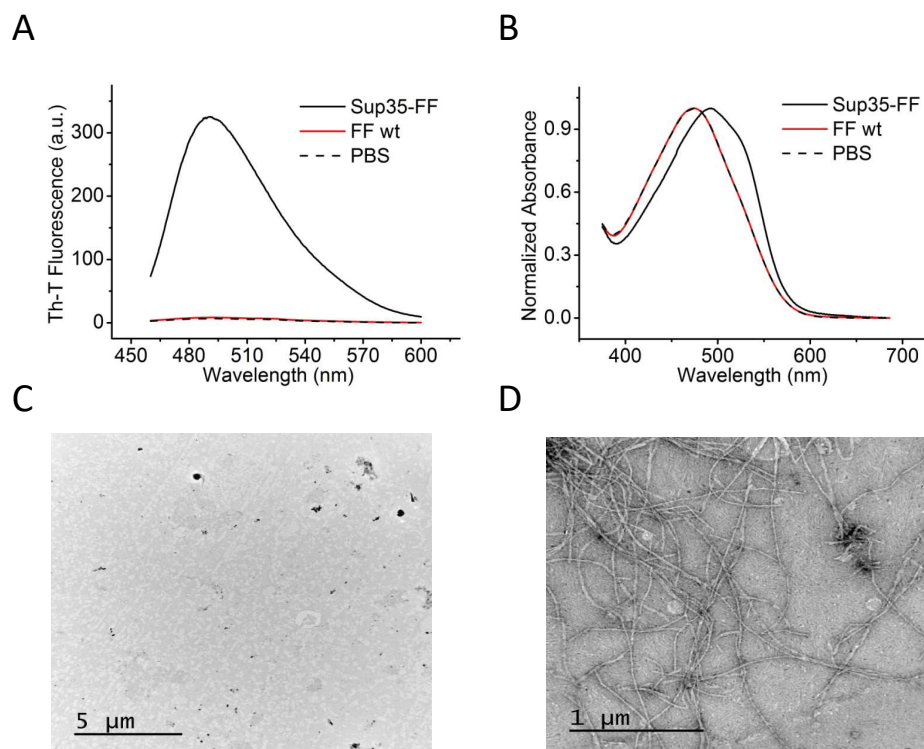
We used the amyloid-specific dyes Thioflavin-T (Th-T) and Congo Red (CR) to assess if the Sup35-FF protein fusion self-assembles with time as amyloid-like structures under mild conditions. To this aim, Sup35-FF and FF-wt were incubated at pH 7.4 and 37 °C for one week. Th-T is a dye which fluorescence emission maximum at 488 nm increases in the presence of amyloid-like structures.<sup>49</sup> The presence of incubated Sup35-FF promoted a large increase of Th-T fluorescence emission signal, whereas FF-wt had a negligible



effect (**Figure 3A**). In agreement with these results, CR binding was observed for Sup35-FF, resulting in a clear red-shift of CR absorption spectrum, indicative of the dye binding to an amyloid structure<sup>50</sup>, whereas FF-wt did not promote any spectral shift (**Figure 3B**). The morphological analysis of the two protein solutions by negative-staining and transmission electron microscopy (TEM) confirmed the presence of characteristic long and unbranched amyloid fibrils of  $16.0 \pm 1.7$  nm in width for Sup35-FF (**Figure 3D**), whereas the FF-wt solution did not show any detectable ordered aggregate (**Figure 3C**).

We further analysed the assembly of Sup35-FF into amyloid fibrils by monitoring the changes in Th-T signal with time. To this aim, the reaction was set up in 96 well plates with constant agitation in the presence of Teflon beads, a condition that accelerates aggregation and allows a continuous reading of the dye signal.<sup>51</sup> The kinetics of amyloid fibril formation can be usually adjusted to a sigmoidal curve, reflecting the existence of a nucleation-dependent growth reaction.<sup>52</sup> The aggregation of Sup35-FF under native conditions follows this kinetic scheme, exhibiting a lag phase of  $\approx 6$  h and being completed after  $\approx 17$  h (**Figure S3**). A characteristic of most amyloid assemblies is that they are able to seed and accelerate the reaction of their soluble counterparts. Importantly, the presence of 2% Sup35-FF preformed fibrils was enough to dramatically accelerate the aggregation reaction, reducing the lag phase to  $\approx 1.5$  h and allowing it to complete in  $\approx 7$  h (**Figure S3**).

Overall, the data converge to indicate that Sup35-SAC is able to specifically promote the self-assembly of the Sup35-FF fusion, containing the otherwise soluble FF domain, into ordered nanofibrils.



**Figure 3. Characterization of Sup35-FF fibrils.** Sup35-FF and FF-wt solutions were incubated for one week and analysed by measuring (A) Th-T fluorescence emission and (B) Congo red absorbance. FF-wt and Sup35-FF are shown in red and black, respectively. PBS without protein was included as a control (dashed line). Representative TEM micrographs of incubated proteins upon negative staining: (C) FF-wt and, (D) Sup35-FF. The scale bar represents 5  $\mu\text{m}$  and 1  $\mu\text{m}$ , respectively.

### Secondary structure of Sup35-FF amyloid fibrils

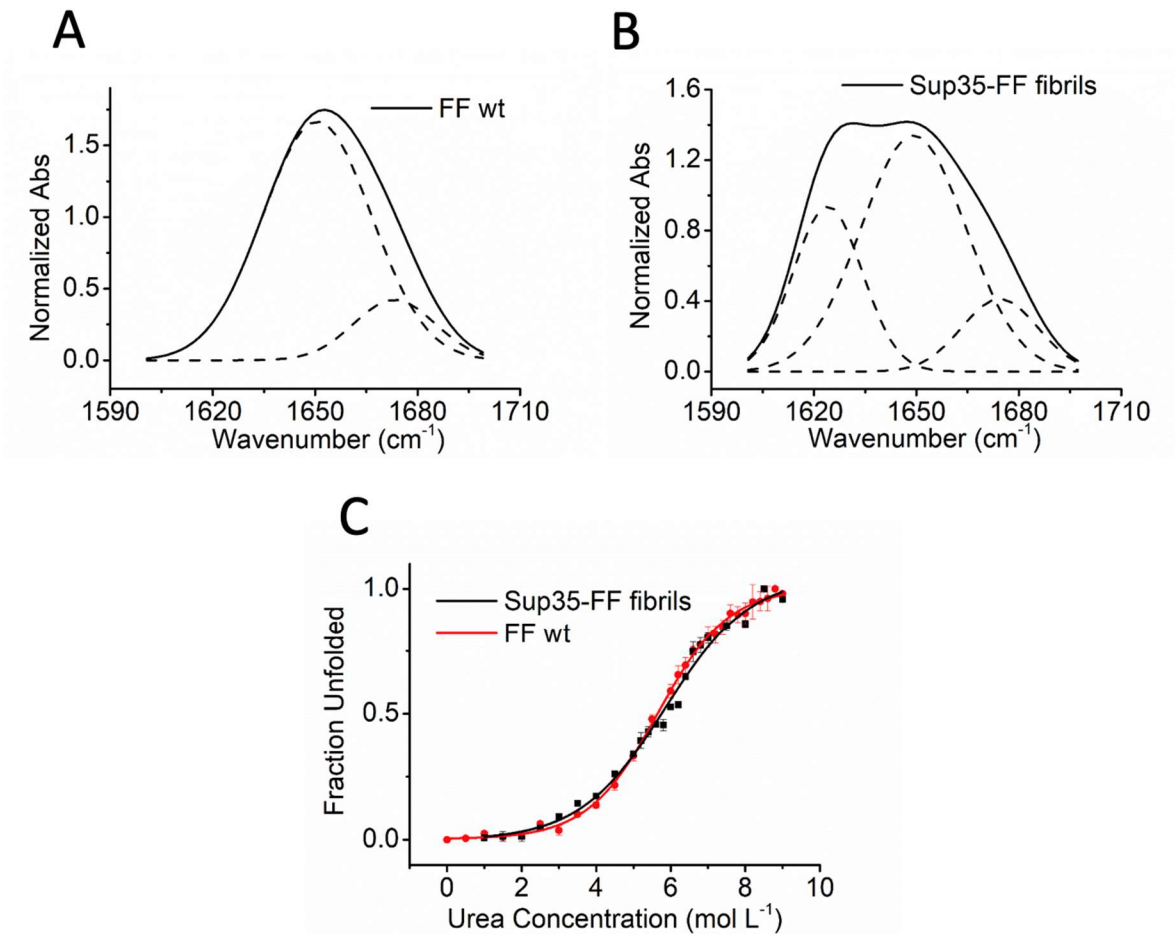
The modular nature of yeast prions accounts for the fact that globular domains retain its original fold in the amyloid state.<sup>53</sup> We wondered whether Sup35-SAC might act as a mimic of the 5-fold longer Sup35 PrD, driving the self-assembly of protein fusions into amyloid structures, but allowing to maintain the native structure of the adjacent globular domain. Because the conversion of a globular protein into an amyloid implies a

conformational conversion into a  $\beta$ -sheet rich structure, independently of the initial 3D-conformation, we took profit of the all  $\alpha$ -helical nature of the FF domain to monitor if it remains folded in the amyloid assembly, using Attenuated Total Reflectance Fourier Transform Infrared spectroscopy (ATR-FTIR).

We recorded the infrared spectra of Sup35-FF amyloid fibrils in the amide I region of the spectrum (1700-1600  $\text{cm}^{-1}$ ) (**Figure 4A**). This region corresponds to the absorption of the carbonyl peptide bond group of the protein main chain and is sensitive to the peptide conformation. Deconvolution of the spectra allows to assign the secondary structure elements and their relative contribution to the main signal (**Table S4**). The spectra of Sup35-FF displayed signals indicative of the formation of  $\beta$ -sheet structure, coming from the intermolecular  $\beta$ -sheet region at 1624  $\text{cm}^{-1}$  contributing to 27.25 % of the area. However, the largest contribution to the spectra, accounting a 59.5 % of the area, comes the band at 1649  $\text{cm}^{-1}$ , compatible with the presence of native-like  $\alpha$ -helices,<sup>54</sup> in the Sup35-FF amyloid fibrils. This putative helical band fits well with the major band at 1650  $\text{cm}^{-1}$  observed in the ATR-FTIR spectra of the soluble FF-wt domain (**Figure 4A and 4B** and **Table S4**).

In the Sup35-FF fusion the two Trp residues are located in the FF domain  $\alpha$ -helices; because the core of the amyloid fibrils is expected to be resistant to urea denaturation, or at least more resistant than the helices, following urea denaturation by intrinsic fluorescence should allow us to assess if the Trp residues remain in a native-like context in the fibrillar state. As it can be seen in **Figure 4C** the denaturation curve of aggregated

Sup35-FF is cooperative and resembles that of soluble FF-wt, despite the calculated  $[Urea]_{50\%}$  is slightly higher for the fibrils ( $5.86 \pm 0.11$  M) than for the soluble wild type protein ( $5.68 \pm 0.04$  M), both calculated from the changes in intrinsic fluorescence. These data support the globular domain remaining folded in the amyloid assembly.

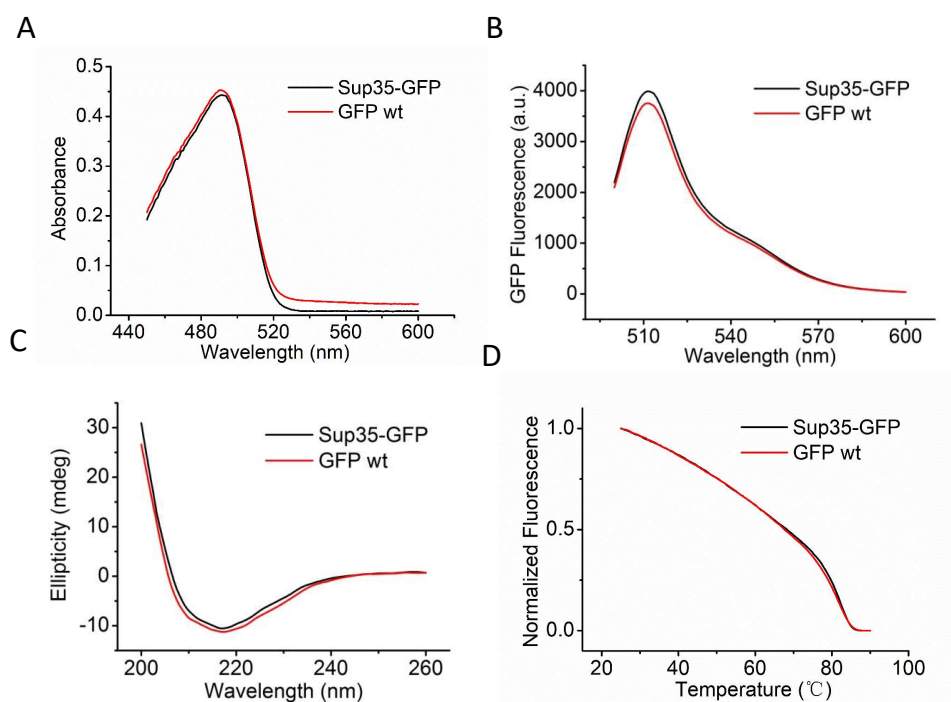


**Figure 4. Conformational properties of Sup35-FF fusion protein fibrils.** Sup35-FF proteins solutions were incubated for one week. The absorbance spectra of (A) Soluble FF-wt and (B) Sup35-FF fibrils in the amide I region (solid line) and the components bands (dashed lines) are shown. (C) The chemical equilibrium curve with urea for aggregated Sup35-FF was followed at 25 °C by Trp intrinsic fluorescence at 350 nm. Soluble FF-wt was measured in the same conditions as a control.

### **Sup35-SAC drives the formation of GFP fluorescent amyloid fibrils**

The above described data indicates that the short Sup35-SAC allows the expression and purification of soluble fusions in which the adjacent C-terminal globular domain remains properly folded as well as the self-assembly of the protein into nanofibers that retain detectable native structure. On this basis, we explored whether we can obtain fluorescent amyloid fibrils by fusing Sup35-SAC to GFP (Sup35-GFP) (**Figure S4A**).

Sup35-GFP was purified from the soluble cellular fraction (**Figure S5**) with a yield of 85 mg/L culture. We characterized the conformational features of the GFP moiety in the fusion by monitoring GFP absorption and fluorescence emission spectra (**Figure 5A and 5B**) as well as by far-UV CD (**Figure 5C**) at pH 7.4 and 25 °C. Sup35-GFP showed spectral properties indistinguishable from that of GFP alone, with an absorption maximum at ~ 490 nm, maximum fluorescence emission at ~511 nm, and a far UV-CD spectrum exhibiting a single minimum at ~218 nm characteristic of  $\beta$ -sheet proteins.



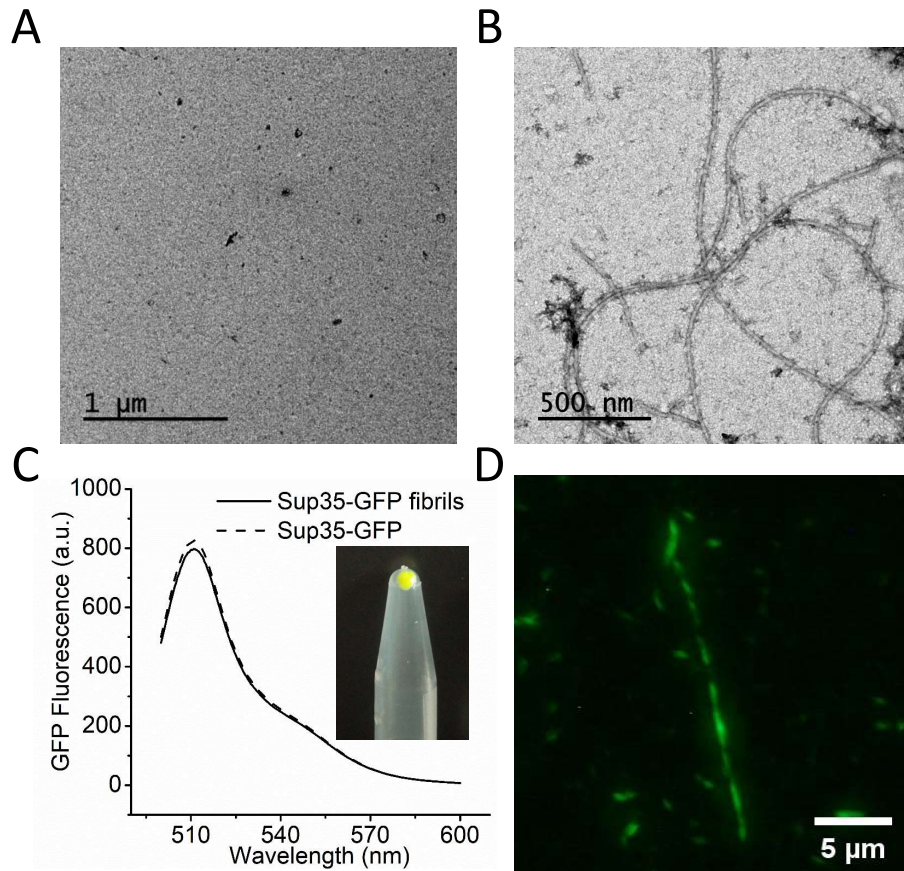
**Figure 5. Conformational properties and thermal stabilities of Sup35-GFP.** Emission (A) excitation (B) and far-UV CD (C) spectra. (D) Thermal unfolding curves followed by monitoring GFP fluorescence at 515 nm from 25 °C to 90 °C with a heating rate of 1 °C/min. Sup35-GFP and GFP are shown in black and red, respectively.

Next, we monitored the thermal stability of Sup35-GFP and GFP by following the changes in GFP fluorescence emission at 515 nm with the temperature (**Figure 5D**). We obtained cooperative, superimposable, denaturation curves with  $T_m = 81.0 \pm 0.2$  °C and  $81.9 \pm 0.2$  °C for Sup35-GFP and GFP, respectively. Altogether the data indicate that, as in the case of the FF domain, Sup35-SAC does not alter the fold or the stability of GFP.

Finally, we incubated Sup35-GFP and GFP at pH 7.4 and 37 °C for one week. Because Th-T and CR cannot be used to monitor amyloid structure when GFP is present, we directly characterized the morphology of the potential assemblies using TEM, confirming the presence of long fibrillar assemblies with a diameter of  $20.1 \pm 1.1$  nm for

Sup35-GFP (**Figure 6B**) and the absence of any ordered aggregate for GFP (**Figure 6A**), indicating that Sup35-SAC drives specifically the amyloid assembly in the context of the Sup35-GFP fusion.

To assess if GFP maintains a functional conformation when embedded in the fibrils, we recorded the GFP fluorescence emission spectra of the aggregated material. It exhibited, the characteristic maximum at ~511 nm (**Figure 6C**) and, indeed, when this material was imaged in a fluorescence microscope using an FITC filter (excitation at 465-495 nm) the presence of large green fluorescent fibrillar structures was directly observed (**Figure 6D**). A quantitative analysis of the fluorescence emission of Sup35-GFP in the soluble and fibrillar states indicated that the specific fluorescence emission of the fibrils was 2.6 times lower (**Figure 6C**). Whether this decrease in activity responds to a fraction of the domains being inactive in the assembly or it results instead from self-quenching, owing to the expected dense packing of the domains and close proximity of the fluorophores in the fibrils, should be further examined.



**Figure 6. Morphology and functionality of Sup35-GFP fibrils.** Sup35-GFP and GFP solutions were incubated for one week. (A) and (B) show representative TEM micrographs of GFP and Sup35-GFP, respectively. Scale bar represents 1  $\mu\text{m}$  and 500 nm. (C) Emission fluorescence spectra of Sup35-GFP fibrils at 5  $\mu\text{M}$  and photograph of precipitated fibrils. The fluorescence emission of soluble Sup35-GFP at 2  $\mu\text{M}$  is shown for comparison. (D) Fluorescence microscopy image of Sup35-GFP fibrils. Scale bar represents 5  $\mu\text{m}$ .

### Sup35-SAC drives the formation of catalytically active amyloid fibrils

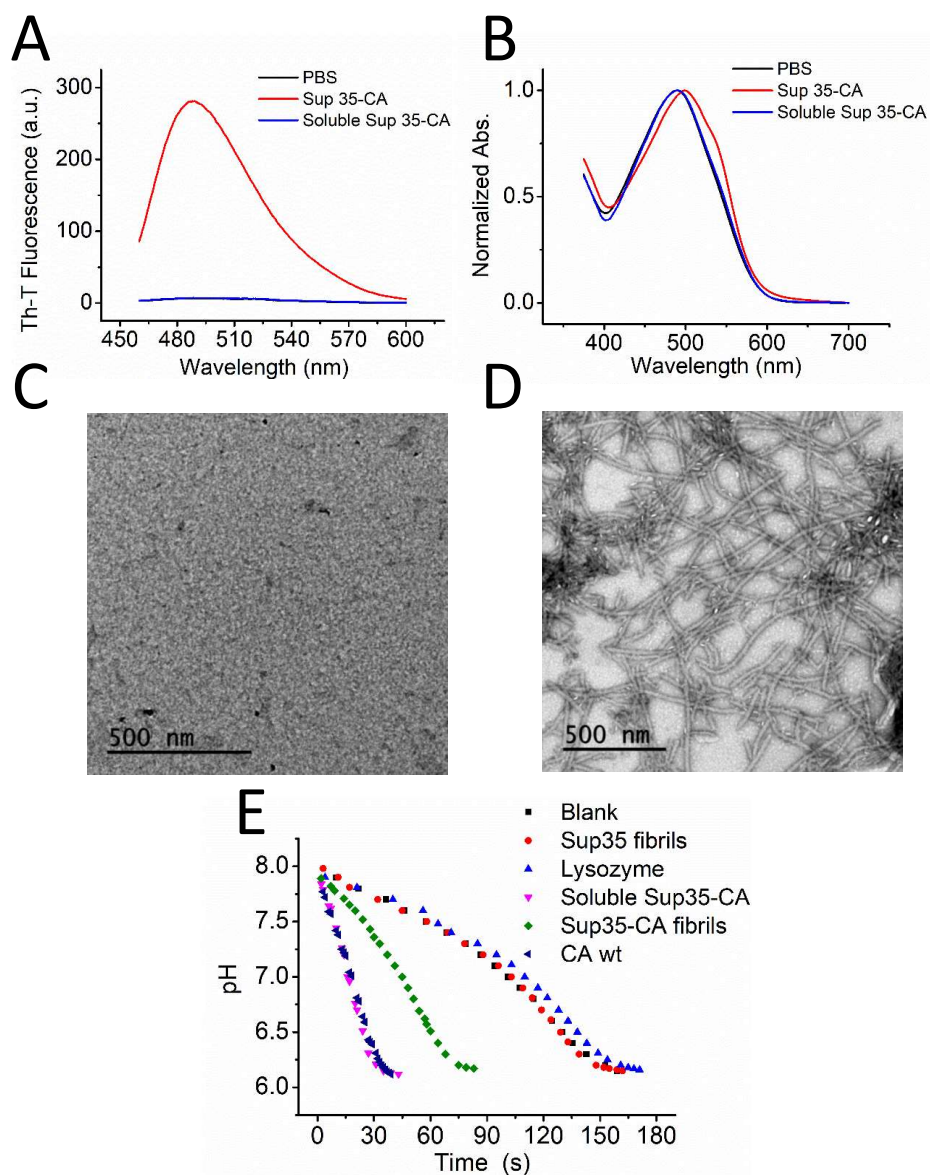
To further confirm the idea that Sup35-SAC can be exploited to obtain modular and functional nanofibrillar assemblies, this short sequence stretch was fused to the carbonic anhydrase enzyme (Sup35-CA) (**Figure S4B**).

Again, Sup35-CA was purified from the soluble cell fraction (**Figure S6**), with a yield of 98 mg/L culture. As in the case of the FF and GFP proteins, the comparison of the



conformational properties of Sup35-CA with those of wild type CA (CA-wt) indicated that the fusion of the 21-residues peptide does not alter significantly the adjacent globular structure (**Figure S7A and S7B**), neither its stability, since CA-wt and Sup35-CA exhibit similar cooperative denaturation curves with  $[Urea]_{50\%}$  of  $4.24 \pm 0.03$  M and  $4.40 \pm 0.04$  M, respectively, as monitored by intrinsic fluorescence (**Figure S7C**).

To analyse the self-assembling capacities of Sup35-CA, we incubated the protein at pH 7.4 and 25 °C for one week and analysed the presence of amyloid-like material using Th-T (**Figure 7A**) and CR assays (**Figure 7B**). In the presence of incubated Sup35-CA, Th-T exhibits a high increase in fluorescence emission; likewise, the protein promotes a red shift of the CR spectra, both proves indicating an amyloid-like nature. The morphology of the Sup35-CA assemblies was analysed by TEM (**Figure 7D**), which allowed to observe the presence of long amyloid fibrillar structures of  $23.0 \pm 1.5$  nm in diameter. No aggregate was detected in a fresh Sup35-CA protein solution (**Figure 7C**), consistent with its inability to bind Th-T and CR (**Figure 7A and 7B**).



**Figure 7. Tinctorial properties, morphology and activity of Sup35-CA fibrils.** Sup35-CA was incubated for one week. (A) Th-T and (B) CR binding assays in the absence (black) or presence (red) of incubated Sup35-CA protein. Soluble Sup35-CA was used as a control (blue). Panels (C) and (D) show representative TEM micrographs of soluble and incubated Sup35-CA protein, respectively. Scale bar represents 500 nm. (E) CO<sub>2</sub> hydration activity of Sup35-CA fibrils. Buffer alone (black), Sup35-SAC fibrils (red) and lysozyme (blue) were used as negative controls while CA-wt (dark blue) was used as positive control. pH measurements showed that carbonic anhydrase moieties embedded in the Sup35-CA fibrils (green) or in the soluble fusion (pink) were active.

Finally, we characterized the catalytic activity of Sup35-CA fibrils by modified Wilbur-Anderson method<sup>55</sup> (**Figure 7E**). In this method, the activity of CA is measured by monitoring the solution pH acidification caused by the CA catalysed conversion of carbon dioxide and water into bicarbonate and hydrogen ions, resulting in a decrease of pH. The kinetics of the reaction in the pH range 8.1- 6.3 were recorded. Before measuring the activity of fibrils, soluble protein was removed by repeated centrifugation and washing steps. The activity of the protein solutions was calculated using the equation:  $WAU = (t_0 - t)/t$  where  $t$  is the time required for the pH change when the protein is present and  $t_0$  is the time required in buffer alone. Bromothymol blue (BTB) was used as a pH indicator, to confirm the solution acidification. When the reaction was performed in the presence of Sup35-CA fibres, an activity of 1.2 WAU was recorded, and the solution colour in the presence of BTB changed from blue to yellow, confirming that at least a fraction of the CA domains in the fibrils was functional. No activity was detected when lysozyme or fibrils formed by the Sup35-SAC peptide alone, therefore devoid of CA, were used; indicative of the specificity of the reaction. The activity of CA in the fibrils was 3.4 times lower than the one exhibited by an equivalent amount of either soluble Sup35-CA (4.1 WAU) or CA-wt (4.0 WAU). This reduction of CA activity in the fibrillar state has been already observed for a fusion between the complete PrD of the yeast Ure2p and CA,<sup>20</sup> despite in that case the reduction of activity in the filaments relative to the soluble form was higher, about 10-fold. Two explanations may account for these differences in activity: *i*) when folded, CA

displays the same activity in both fibrils, but the amount of properly folded domains is higher in Sup35-CA fibrils. *ii*) CA is more active in Sup35-CA fibrils.

CA is a near-diffusion-limited enzyme.<sup>56</sup> This implies that the reaction rate depends on the speed at which substrate diffuses to the active site. Aggregation might decrease the apparent activity of CA because the fibrillar structure reduces the diffusion rate of the substrate. In this case, it is expected that the fibrils formed by smaller self-assembling domains, such as the Sup35-SAC, would impose lower restrictions to diffusion than those of the larger PrDs. The same rationale would apply in case the reduction in activity is caused by steric hindrance. Indeed, an inverse relationship between the size of the fibril-forming component and the activity of the adjacent enzyme in the aggregated state was reported for a fusion of the Ure2p PrD with horseradish peroxidase (HRP).<sup>19</sup> The use of the 93-residue full-length PrD rendered very low HRP activity, whereas shortening it to 80 residues permitted to double the catalytic activity of the enzyme in the fibrillar form.

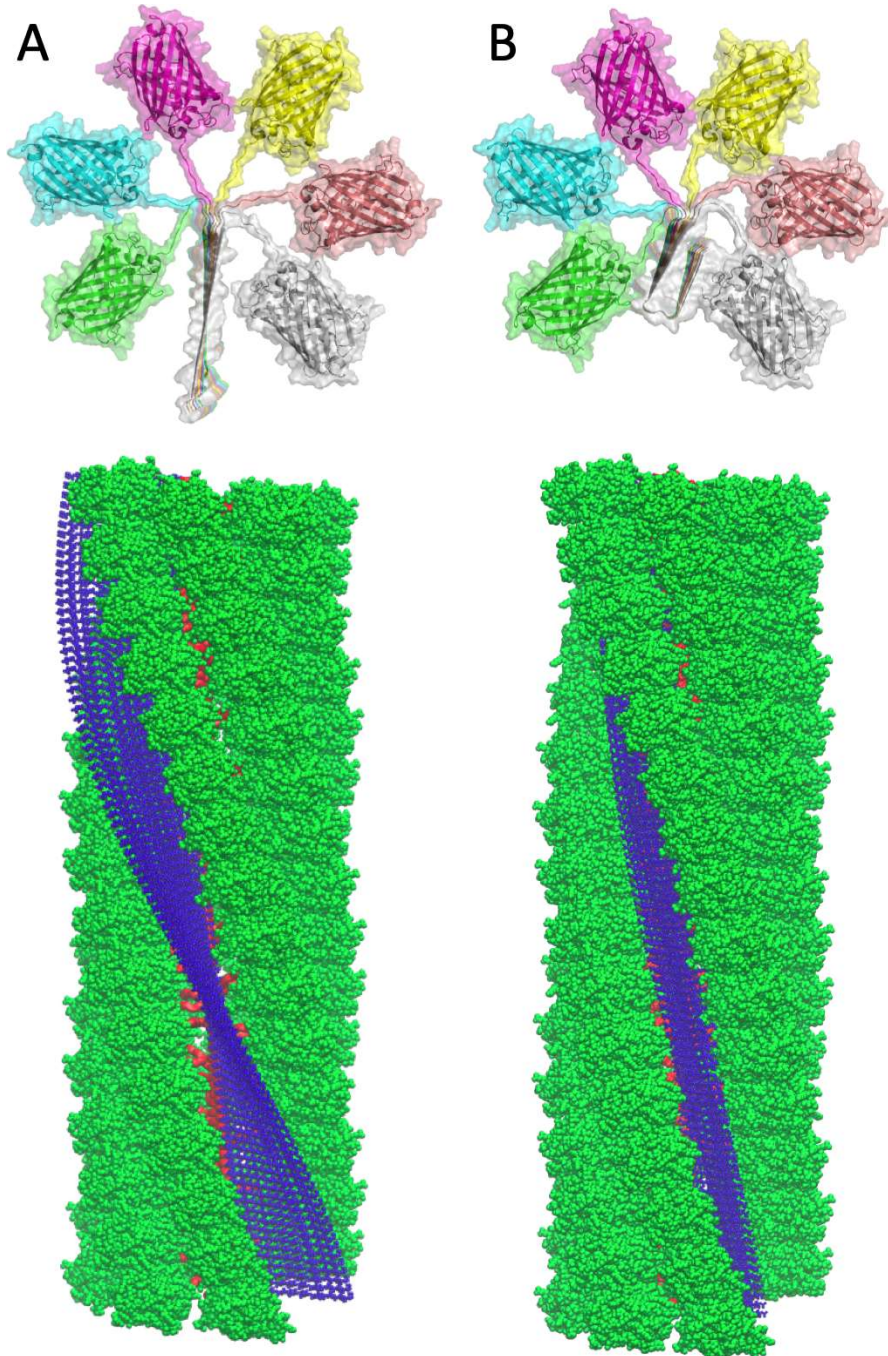
In any case, it seems clear from our experiments that the short soft amyloid core of Sup35 PrD can be employed to generate catalytically active fibrils.

### **Modelling of amyloid fibrils containing Sup35-SAC linked to a globular domain.**

The distance between  $\beta$ -strands in a typical amyloid  $\beta$ -sheet is  $\approx 4.8 \text{ \AA}$ , whereas the size of the globular domains is typically ranged between 30 and 40  $\text{\AA}$ . Therefore, when located in close vicinity to amyloid regions along the chain, globular domains might prevent the formation of ordered amyloids because of the steric repulsion. This is especially true for

the parallel arrangement because of the bigger difference of the axial distances between the  $\beta$ -strands (4.8 Å) and between the globular domains (30 Å). To confirm our conclusion that a Sup35-SAC amyloid core surrounded by globular domains can be formed without significant steric tension, we applied a molecular modelling approach to the Sup35-SAC amyloid with the parallel and in-register arrangement. Furthermore, a number of experimental data suggest that the prion domain of Sup35 and its fragments form amyloid fibrils with the parallel and in-register cross- $\beta$  structure.<sup>58-60</sup> The molecular modelling was performed as previously described.<sup>57</sup> In this work, as a model, we used a hybrid protein containing the amyloid-forming Sup35-SAC peptide and GFP linked by 8-residues linker, as in our Sup35-GFP construct. Furthermore, it is known that the amyloid protofibrils usually have a slight left-handed twist. Therefore, we constructed these amyloid fibrils with an axial twist angle of 2° per 4.8-Å step (**Figure 8**). For the parallel and in-register cross- $\beta$  amyloid core we used two structural arrangements: stacks of linear  $\beta$ -strands and of  $\beta$ -arches. The 21-residues Sup35-SAC is long enough to form  $\beta$ -strand- $\beta$ -arc- $\beta$ -strand elements called ‘ $\beta$ -arches’ that have been found in a significant number of disease-related and functional amyloid fibrils. The possibility of a  $\beta$ -arch arrangement for Sup35-SAC was correctly predicted by the ArchCandy algorithm.<sup>61</sup> Our molecular modelling confirmed that the hybrid molecules experimentally tested in this work can form parallel and in-register cross- $\beta$  amyloids (with Sup35-SAC either in the linear or  $\beta$ -arch conformations) decorated by globular domains (**Figure 8**). The structural models were energy-minimized using steepest descent algorithm of GROMACS package version 4.6.7

(Hess, B.; Kutzner, C.; Van Der Spoel, D.; Lindahl, E. GRGMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* 2008, 4 (3), 435–447.) and their steric tension and overall stereochemistry were evaluated by using PROCHECK package (Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* 1993, 26(2), 283–291). In the final structures (Figure 8), the GFP domains do not have steric clashes. The conclusion of the molecular modelling can be extended to the antiparallel cross- $\beta$  amyloids of the hybrid molecules because these fibrils would have even more space to accommodate the globular domains. The diameter of the modelled parallel and in-register fibrils are 17.7 nm and 17 nm for the linear and  $\beta$ -arch conformations, respectively; in very good agreement with the 20 nm we measured for Sup35-GFP amyloid fibrils.



**Figure 8. Structural models of amyloid fibrils formed by Sup35-GFP** (A) Fibrils formed by stacking of linear  $\beta$ -strands of Sup35-SAC. Axial projection of the repetitive element (top). Lateral projection of the 90-mer fibril that corresponds to a half helix turn ( $180^\circ$ ). The fibril was generated by the  $28.8 \text{ \AA}$  translation of the hexameric element along the fibril axis (bottom). B. Fibrils formed by stacking of  $\beta$ -arches of Sup35-SAC. Axial projection (top) and lateral projections (bottom). Images were generated by using PyMol software.<sup>62</sup>

## Conclusions

We have developed different functional nanofibrils exploiting our previous discovery that the self-assembling properties of PrDs can be mimicked by their much shorter SACs. In particular, we have demonstrated that the conjugation of globular proteins to Sup35-SAC induces a controlled self-assembly into amyloid-like supramolecular structures. In contrast to other fusions between  $\beta$ -sheet forming peptides and globular domains, our proteins could be obtained at high yield in a soluble state, without any need for refolding procedures. This is likely because Sup35-SAC remains largely disordered when soluble, despite its ability to adopt a  $\beta$ -sheet conformation in the fibrillar state. Moreover, in Sup35-SAC fusions the adjacent globular proteins display conformational, stability and folding properties almost indistinguishable of those of the original proteins. Importantly, the globular domains within our fusion proteins seem to maintain their structure preserved in the fibrillar state. Indeed, the diameter of the individual fibrils correlates well with that of the conjugated globular domain: 16, 20 and 23 nm for the FF domain (8 kDa), GFP (27 kDa) and CA (30 kDa) fibrils, respectively. This trend suggests that the packing of FF and CA domains in the fibrils might be similar to the one observed in the Sup35-GFP fibrils derived models. Therefore, Sup35-SAC appears as a module that can be readily used to immobilize bioactive proteins of different sizes and structures. The results agree well with the previous molecular modelling study that allowed to establish a more general relationship between the size of the globular domains and the length of the linkers in the parallel and in-register cross- $\beta$  fibrils<sup>57</sup>.



The modular genetic fusion approach described here can be applied to decorate fibrils with different functionalities, including active enzymes. Altogether, in addition to validate the prion-like proteins of short SACs, the present work illustrates a straightforward strategy to obtain novel bionanomaterials displaying immobilized functional proteins of biological or chemical interest.

## Materials and methods

**Reagents and enzymes.** Reagents and enzymes were purchased from Sigma-Aldrich (UK), unless otherwise stated. Carbon grid (400 square mesh copper) were purchased from Micro to Nano (Netherlands) and the uranyl acetate solution were provided by the microscopy service (Universitat Autònoma de Barcelona). Sup35-SAC 21-residues peptides were purchased from CASLO ApS (Scion Denmark Technical University).

**Expression and Purification of Proteins.** The FF domain, corresponding to residues 212-266 of yeast URN1, was cloned into a pET-28(a) plasmid (Addgene, USA), that previously contained the soft amyloid core Sup35 sequence, resulting in a plasmid encoding for a chimeric protein (Sup35-linker-FF domain) with an His6 tag. The cDNAs of Sup35-GFP (folding reporter green fluorescent protein<sup>63</sup>) and Sup35-CA (Carbonic anhydrase) cloned in the plasmid pET28(a) with a His6 tag were acquired from GenScript (USA). *E. coli* BL21 (DE3) competent cells were transformed with the correspondent plasmids. Then, transformed cells were grown in 10 mL LB medium containing 50 µg/mL kanamycin, overnight at 37 °C, and transferred into 1 L fresh LB media containing 50 µg/mL kanamycin. After reaching an OD<sub>600</sub> of 0.6, the culture was induced with 0.4 mM IPTG and grown either at 20 °C for 16 h or at 37 °C for 4h. At both temperatures, the proteins were found mainly in the soluble cell fraction; however, 20 °C was selected for all preparative productions, since the soluble/insoluble ratio was slightly higher. Cells were collected by centrifugation at 5000 rpm for 15 min at 4 °C. The collected pellet was

resuspended into 20 mL PBS pH 7.4 containing 20 mM imidazole, 1 mg/mL lysozyme and 1 mM PMSF. The solution was incubated on ice, followed by sonication for 20 min. The supernatant was collected by centrifugation at 15000 rpm for 30 min at 4 °C and, purified in an His-tag column, according to the manufacturer's protocol, followed by a gel filtration onto a HiLoad™ Superdex™ 75 prepgrade column (GE Healthcare, USA). The purified proteins were frozen with liquid nitrogen and stored at -80 °C. The purity of the sample was confirmed by SDS-PAGE. The concentration of the proteins: Sup35-FF, Sup35-GFP and Sup35-CA was determined by UV absorption using a  $\epsilon$  value of 18450 L·mol<sup>-1</sup>·cm<sup>-1</sup>, 24995 L·mol<sup>-1</sup>·cm<sup>-1</sup> and 54890 L·mol<sup>-1</sup>·cm<sup>-1</sup>, respectively. The protein was found mainly in the soluble cell fraction, independently if the expression was done at 37 °C or at 20 °C; however, this last temperature was used for all preparative protein productions since the soluble/insoluble ratio was slightly higher.

**Prediction of Aggregation and Disorder.** The protein sequence of FF domain, Sup35-FF, Sup35-SAC, and Sup35-linker were used to predict aggregation and disorder. The aggregation propensity was predicted by three different methods: Aggrescan<sup>41</sup>, FoldAmyloid<sup>42</sup> and PASTA 2.0<sup>43</sup>. The disordered regions were predicted using four different algorithms: PONDR (VSL2)<sup>45</sup>, GlobPlot<sup>46</sup>, PASTA 2.0<sup>43</sup> and IUPred<sup>47</sup>. The parameters used for the predictions were the following: Aggrescan and PONDR (VSL2) were run with default parameters. For FoldAmyloid, the scale option was “expected number of contacts 8Å” the “averaging frame” was 5 and threshold was 21.4. For PASTA

2.0, the “custom” mode was used and the “top pairing energies” was 20, the “energy threshold” was -5 (1.0 Pasta Energy Unit corresponding to 1.192 Kcal/mol), and the TPR (sensitivity) and FPR (1- specificity) were 40.5% and 4.7%, respectively. For GlobPlot, the mode of propensities was “Russell/Linding (recommended P=RC-SS)”, for disorder prediction, minimum peak width was 5 and maximum join distance was 4; the globular domain hunting was performed using domain prediction “SMART/Pfam” parameter with minimum peak width 74 and maximum join distance 15, and the plot was smoothed using 1st and 2nd derivative Savitzky-Golay with frame 10. For IUPred, the “short disorder of prediction type” was used and the output type was “generate plot” with window size 500.

**Conformational Characterization.** Proteins were dissolved at a final concentration of 20  $\mu$ M for Sup35-FF, Sup35-GFP and Sup35-CA in PBS pH 7.4 buffer, then samples were filtered through a 0.22  $\mu$ m Millipore filter and immediately analysed. Far-UV CD spectra were recorded from 260 nm to 200 nm at 1 nm bandwidth, response time of 1 second, and a scan speed of 100 nm/min in a Jasco-710 spectropolarimeter (Jasco Corporation, Japan), thermostated at 25 °C. Ten accumulations were averaged for each spectrum. Trp intrinsic fluorescence and GFP intrinsic fluorescence spectra were measured at 25 °C on a Jasco FP-8200 Spectrofluorometer (Jasco Corporation, Japan). Three averaged spectra were accumulated using an excitation wavelength of 280 nm and 485 nm and recording emission from 300 to 400 nm and 500 to 600 nm for Trp and GFP emission fluorescence, respectively, with slit widths of 5 nm. Three averaged absorbance spectra of Sup35-GFP

were recorded from 450 to 600 nm on a SPECORD 200 plus spectrophotometer (Analytik Jena, Germany), with scan speed of 20 nm/s. As controls, 20  $\mu$ M FF-wt, GFP and CA-wt in PBS buffer, pH 7.4, were measured in all the assays under the same conditions. The fluorescence emission spectra of Sup35-GFP fibrils was recorded as described for the soluble fusion at 5  $\mu$ M final concentration. The amount of protein in the fibrillar solution was titrated after denaturation by SDS-PAGE against a concentration standard of soluble Sup35-GFP.

**Thermal and chemical denaturation.** FF-wt and Sup35-FF were dissolved at 20  $\mu$ M in PBS, pH 7.4. Trp intrinsic fluorescence was monitored in a Jasco FP-8200 Spectrofluorometer from 25  $^{\circ}$ C to 90  $^{\circ}$ C, with an excitation wavelength of 280 nm and recording emission at 340 nm, with an interval of 0.5  $^{\circ}$ C and a heating rate of 0.5  $^{\circ}$ C/min. Ellipticity was recorded at 222 nm each 1  $^{\circ}$ C with a heating rate 0.5  $^{\circ}$ C/min from 25  $^{\circ}$ C to 90  $^{\circ}$ C, using a Jasco-710 spectropolarimeter. In the case of Sup35-GFP and GFP, soluble proteins were prepared at 10  $\mu$ M in PBS, pH 7.4. GFP intrinsic fluorescence was recorded from 25  $^{\circ}$ C to 90  $^{\circ}$ C, with excitation wavelength of 485 nm and emission recorded at 515 nm each 0.5  $^{\circ}$ C with a heating rate of 0.5  $^{\circ}$ C/min.

For chemical denaturation, samples were prepared at 10  $\mu$ M in PBS, pH 7.4 in the presence of different concentrations of denaturing agent (0-9 M of urea), and the reaction was left to equilibrate for 20 h at room temperature. Each sample was prepared in duplicate. Trp intrinsic fluorescence spectra and far-UV CD spectra were collected at 280/360 nm

(excitation/emission) and 222 nm, respectively. One-week aggregated Sup35-FF was analysed in the same manner at 5  $\mu\text{M}$  final concentration. For Sup35-CA and CA-wt the shift in the center of mass (CM) of the Trp emission spectrum was calculated and plotted against the urea concentration.

The results were fitted to a two-state transition curve where the signals of the folded and unfolded states are dependent on the temperature/denaturant concentration using the nonlinear least squares algorithm provided by Origin 8.5 (OriginLab Corporation).

**Folding Kinetics.** Kinetics of unfolding and refolding were analysed by recording intrinsic Trp fluorescence changes upon transition at 25  $^{\circ}\text{C}$  in a Bio-Logic SFM-3 stopped-flow instrument (Bio-Logic Science Instruments), employing an excitation wavelength at 280 nm and a 320 nm cut-off filter. For unfolding reactions, the protein solution was prepared in PBS pH 7.4, at 10  $\mu\text{M}$  and appropriate volumes of the same buffer containing concentrated urea were added to initiate the reaction. For the refolding reaction, selected volumes of denaturant free buffer were added to the initial protein solution at 10  $\mu\text{M}$  in 9M urea. All fluorescence traces were fitted to an exponential function, and kinetic parameters were calculated considering a two-state folding model, using Origin 8.5 program (OriginLab Corporation). Kinetic and free energy values were determined using the following equations:

$$\Delta G_{U-F}^{H_2O} = -RT \cdot \ln(k_U^{H_2O} / k_F^{H_2O})$$

$$m_{U-F} = RT(m_F + m_U)$$

where  $k_F$  and  $k_U$  are the rate constants for folding and unfolding, respectively, and the  $m_F$  and  $m_U$  values correspond to the slopes of the respective folding and unfolding regions.

**Fibril Formation.** Sup35-FF, Sup35-GFP and Sup35-CA proteins were prepared at 50  $\mu$ M, 1.5 mM and 1.0 mM in PBS pH 7.4, and filtered through a 0.22  $\mu$ m filter. The Sup35-FF and Sup35-GFP samples were incubated at 37 °C, and Sup35-CA at 25 °C, with agitation at 300 rpm for one week. FF-wt and GFP were incubated at the same concentrations and conditions as controls.

### **Amyloid dyes binding**

Thioflavin T (Th-T) and Congo red (CR) was used to determine the formation of amyloid fibrils. For the binding assay, incubated proteins were diluted to a final concentration of 20  $\mu$ M in PBS pH7.4, in the presence of 25  $\mu$ M Th-T. Emission fluorescence was recorded in the 460-600 nm range, using an excitation wavelength of 445 nm and emission bandwidth of 5 nm on a Jasco FP-8200 Spectrofluorometer (Jasco Corporation, Japan). For the CR binding assay, incubated proteins were prepared at final concentration of 20  $\mu$ M and, CR was mixed to a final concentration of 20  $\mu$ M. Optical absorption spectra were recorded in the range from 375 to 700 nm in a Specord 200 Plus spectrophotometer (Analytik Jena, Germany). Spectra of protein alone and buffer were acquired to subtract protein scattering.

**Transmission Electron Microscopy (TEM).** For TEM samples preparation, 10 $\mu$ L of the incubated proteins were deposited on a carbon-coated copper grid for 10 min and the excess liquid was removed with filter paper, followed by a negative stain with 10  $\mu$ L of 2 %(w/v) uranyl acetate for 1min. Grids were exhaustively scanned using a JEM 1400 transmission electron microscope (JEOL ltd, Japan) operating at 80 kV, and images were acquired with a CCD GATAN ES1000W Erlangshen camera (Gatan Inc., USA). The width of fibrils was analysed by Image J (National Health Institute), averaging the measures of 10 individual fibrils for each fusion protein.

**Fourier Transform Infrared Spectroscopy (FTIR).**

FF-wt was dissolved at 180  $\mu$ M in PBS, pH 7.4, and filtered through a 0.22  $\mu$ m filter. 30  $\mu$ L of the prepared Sup35-FF fibrils at 50  $\mu$ M were centrifuged at 12000g for 30 min and resuspended in 10  $\mu$ L of water. Samples were placed on the ATR crystal and dried out under N<sub>2</sub> flow. The experiments were carried out in a Bruker Tensor 27 FTIR (Bruker Optics, USA) supplied with a Specac Golden Gate MKII ATR accessory. Each spectrum consists of 32 acquisitions measured at a resolution of 1 cm<sup>-1</sup> using the three-term Blackman-Harris Window apodization function. Data were acquired and normalized, using the OPUS MIR Tensor 27 software (Bruker Optics, USA), with the Min/Max normalization method, which scales spectrum intensities to the effect that the minimum absorbance unit will be 0 and the maximum 2. The analysis of the IR spectra was performed with the PeakFit program (SeaSolve Software Inc.). Obtained data, in the amide I region



(1700 to 1600  $\text{cm}^{-1}$ ), were first pre-smoothed using non-parametric smoother Loess ((locally weighted smoothing) procedure at 5% level, and the lineal baseline was subtracted. IR spectra were fitted employing the residuals method for finding hidden peaks, which consists on finding the difference in y-value between a data point and the sum of component peaks evaluated at the data point's x-value. From the second derivative plot for each sample absorbance spectrum, peaks and local minima were identified and the number and positions of them manually placed to deconvolute the absorbance spectra. Afterwards, automated peak fitting was done using the "AutoFit Peaks I Residuals" option with the "vary widths" condition for the autoscan procedure, until reaching iteration 7 and a  $r^2 > 0.997$ . The resulting area, amplitude and center values of the fitted bands were exported as a table and plotted.

**Aggregation and Seeding Kinetics.** Sup35-FF solutions were prepared at 50  $\mu\text{M}$  in PBS pH 7.4 containing 20  $\mu\text{M}$  Th-T and transferred into 96 well plates in the absence and presence of 2% (v/v) Sup35-FF amyloid seeds. Each sample was prepared by triplicate. The Th-T fluorescence at 485 nm was recorded every 3 min at 37 °C during 1000 min, using a Victor III Multilabel Plate Reader (Perkin Elmer, USA), equipped with P450 CW-lamp filter and 486/10nm emission filter. The obtained data were analyzed and plotted by following a Boltzmann function in Origin 8.5 (OriginLab Corporation).

**Fluorescence Microscopy.** 50  $\mu\text{L}$  aggregated Sup35-GFP was centrifuged at 12000g for 30 min. Supernatant was removed and fibrils were resuspended in PBS pH 7.4. 5  $\mu\text{L}$  of the resuspended fibres were mixed with 5  $\mu\text{L}$  mounting medium solution and dropped on a clean glass slide (Deltalab, 26 $\times$ 76 mm) and covered by a cover slide (Deltalab, 22 $\times$ 22mm). Fluorescence imaging of nanofibers was carried out on an Eclipse 90i epifluorescence optical microscopy equipped with a Nikon DXM1200F (Nikon,Japan) camera and ACT-1 software. Images were acquired with an excitation filter of 465-495 nm and detecting fluorescence emission in a range of 515-555 nm.

**Catalytic activity of Sup35-CA fibrils.** Catalytic activity of Carbonic Anhydrase (CA) was assayed by using a modified Wilbur-Anderson method.<sup>55</sup> Specifically, 2 mL of 50 mM Tris-HCl, 100 mM NaCl, pH 8.1 buffer were mixed with 200  $\mu\text{L}$  0.05% BTB (Bromothymol blue) on ice. 10  $\mu\text{L}$  of 20  $\mu\text{M}$  protein samples were added and the reaction was initiated by addition of 1.0 mL of ice-cold  $\text{CO}_2$  saturated water. To determine the catalytic activity, the time required for the pH to drop from 8.1 to 6.3 was recorded. BTB was used as a pH indicator. The activity of the tested sample was calculated using the equation:  $\text{WAU} = (t_0 - t) / t$  where  $t$  is the time required for the pH change when the test sample is present and  $t_0$  is the time required for the pH change when the buffer is substituted for the test sample. CA wt was used as positive control, to evaluate the effect of Sup35 fusion on the CA activity. Sup35-SAC fibrils and lysozyme were used as negative controls to show the specificity of CA activity in the fibrils.

**Molecular modelling of fibril with a pseudo-helical packing of GFPs.** The left-handed twisted fibrils were built with  $2^\circ$  angle twist per axial  $4.8 \text{ \AA}$  translation between the monomers. The twisting operator was imposed around a Z-axis located at the center of mass of the main-chain atoms from the two  $\beta$ -strands of the Sup35 fragment. To build models of hybrid protein we used the crystal structure of GFP protein (pdb entry 1GFL).<sup>64</sup> To position the GFP structures around the amyloid core and evaluate their packing, we built a model of two successive hexameric layers of the subunits.<sup>57</sup> The first GFP subunit was manually docked on one side of the fibril stack while keeping the axis of the GFP  $\beta$ -barrel perpendicular to the Z-axis. The five other subunits were positioned by applying the successive rotation of  $\sim 51^\circ$  around an axis that intersects the X-Y plane at the  $C\alpha$  atom of the last Sup35 residue of the first monomer and their successive translations of  $4.8 \text{ \AA}$  along the Z-axis. This geometry was chosen so that it provides the most optimal close packing of the GFP molecules within the hexamer (**Figure 8**). To evaluate the packing of GFP domains between the hexameric units, an adjacent hexameric layer was generated by a  $28.8 \text{ \AA}$  translation along the Z-axis. The  $28.8 \text{ \AA}$  distance agrees well with the corresponding distance between the GFPs from *Aequorea victoria* in the crystal packing (pdb entry 1GFL).<sup>64</sup> CHIMERA<sup>65</sup> program was used for the symmetrical positioning of GFPs and molecular modelling procedures. The linkers connecting Sup35 peptides to GFPs were generated using MODLOOP.<sup>66</sup> The quality and consistency of the generated models were checked with the PROCHECK program.<sup>67</sup>

## **Conflicts of interest**

There are no conflicts to declare.

## **Acknowledgements**

We thank Drs. Anita Carija and Ricardo Sant'Anna for experimental help. This work was funded by the Spanish Ministry of Economy and Competitiveness BIO2016-78310-R to S.V and by ICREA, ICREA-Academia 2015 to S.V. Weiqiang Wang acknowledges financial support from the China Scholarship Council (CSC): NO. 201606500007.

## References

1. F. Chiti and C. M. Dobson, *Annual review of biochemistry*, 2017, **86**, 27-68.
2. K. Si, *Annual review of cell and developmental biology*, 2015, **31**, 149-169.
3. D. Otzen, *Prion*, 2010, **4**, 256-264.
4. R. Sabaté and S. Ventura, in *Protein Supersecondary Structures*, Springer, 2012, pp. 237-257.
5. M. Diaz Caballero, S. Navarro, I. Fuentes, F. Teixidor and S. Ventura, *ACS nano*, 2018.
6. C. Li, S. Bolisetty and R. Mezzenga, *Advanced Materials*, 2013, **25**, 3694-3700.
7. T. Scheibel, R. Parthasarathy, G. Sawicki, X.-M. Lin, H. Jaeger and S. L. Lindquist, *Proceedings of the National Academy of Sciences*, 2003, **100**, 4527-4532.
8. T. C. Holmes, S. de Lacalle, X. Su, G. Liu, A. Rich and S. Zhang, *Proceedings of the National Academy of Sciences*, 2000, **97**, 6728-6733.
9. S. Bolisetty, C. S. Boddupalli, S. Handschin, K. Chaitanya, J. Adamcik, Y. Saito, M. G. Manz and R. Mezzenga, *Biomacromolecules*, 2014, **15**, 2793-2799.
10. C. M. Rufo, Y. S. Moroz, O. V. Moroz, J. Stöhr, T. A. Smith, X. Hu, W. F. DeGrado and I. V. Korendovych, *Nature chemistry*, 2014, **6**, 303.
11. Z. Al-Garawi, B. McIntosh, D. Neill-Hall, A. Hatimy, S. Sweet, M. Bagley and L. Serpell, *Nanoscale*, 2017, **9**, 10773-10783.
12. P. Marinelli, S. Navarro, M. Baño-Polo, B. Morel, R. Graña-Montes, A. Sabe, F. Canals, M. R. Fernandez, F. Conejero-Lara and S. Ventura, *ACS chemical biology*, 2018, **13**, 2094-2105.
13. R. Wakabayashi, A. Suehiro, M. Goto and N. Kamiya, *Chemical Communications*, 2019, **55**, 640-643.
14. P. Chien, J. S. Weissman and A. H. DePace, *Annual review of biochemistry*, 2004, **73**, 617-656.
15. S. M. Uptain and S. Lindquist, *Annual Reviews in Microbiology*, 2002, **56**, 703-741.
16. E. D. Ross, A. Minton and R. B. Wickner, *Nature cell biology*, 2005, **7**, 1039.
17. I. Hafner-Bratkovic, R. Bester, P. Pristovsek, L. Gaedtke, P. Veranic, J. Gaspersic, M. Mancek-Keber, M. Avbelj, M. Polymenidou and C. Julius, *Journal of Biological Chemistry*, 2011, jbc. M110. 213926.
18. P. K. Baral, M. Swayampakula, A. Aguzzi and M. N. James, *Journal of structural biology*, 2015, **192**, 37-47.
19. X. M. Zhou, A. Entwistle, H. Zhang, A. P. Jackson, T. O. Mason, U. Shimanovich, T. P. Knowles, A. T. Smith, E. B. Sawyer and S. Perrett, *ChemCatChem*, 2014, **6**, 1961-1968.
20. U. Baxa, V. Speransky, A. C. Steven and R. B. Wickner, *Proceedings of the National Academy of Sciences*, 2002, **99**, 5253-5260.

21. D. Men, Y.-C. Guo, Z.-P. Zhang, H.-p. Wei, Y.-F. Zhou, Z.-Q. Cui, X.-S. Liang, K. Li, Y. Leng and X.-Y. You, *Nano letters*, 2009, **9**, 2246-2250.
22. J. A. Toombs, M. Petri, K. R. Paul, G. Y. Kan, A. Ben-Hur and E. D. Ross, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 6519-6524.
23. E. D. Ross, H. K. Edskes, M. J. Terry and R. B. Wickner, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 12825-12830.
24. G. Wei, Z. Su, N. P. Reynolds, P. Arosio, I. W. Hamley, E. Gazit and R. Mezzenga, *Chem. Soc. Rev.*, 2017, **46**, 4661-4708.
25. S. Alberti, R. Halfmann, O. King, A. Kapila and S. Lindquist, *Cell*, 2009, **137**, 146-158.
26. R. Sabate, F. Rousseau, J. Schymkowitz, C. Batlle and S. Ventura, *Prion*, 2015, **9**, 200-206.
27. R. Sant'Anna, M. R. Fernández, C. Batlle, S. Navarro, N. S. De Groot, L. Serpell and S. Ventura, *Scientific reports*, 2016, **6**, 34274.
28. C. Batlle, N. S. de Groot, V. Iglesias, S. Navarro and S. Ventura, *Scientific reports*, 2017, **7**, 12134.
29. R. Zambrano, O. Conchillo-Sole, V. Iglesias, R. Illa, F. Rousseau, J. Schymkowitz, R. Sabate, X. Daura and S. Ventura, *Nucleic acids research*, 2015, **43**, W331-W337.
30. J. A. Toombs, M. Petri, K. R. Paul, G. Y. Kan, A. Ben-Hur and E. D. Ross, *Proceedings of the National Academy of Sciences*, 2012, **109**, 6519-6524.
31. S. Alberti, R. Halfmann, O. King, A. Kapila and S. Lindquist, *Cell*, 2009, **137**, 146-158.
32. S. Kawai-Noma, C.-G. Pack, T. Kojidani, H. Asakawa, Y. Hiraoka, M. Kinjo, T. Haraguchi, H. Taguchi and A. Hirata, *The Journal of cell biology*, 2010, jcb. 201002149.
33. Y. Duernberger, S. Liu, K. Riemschoss, L. Paulsen, R. Bester, P.-H. Kuhn, M. Schölling, S. F. Lichtenthaler and I. Vorberg, *Molecular and cellular biology*, 2018, MCB. 00111-00118.
34. R. Bonet, X. Ramirez - Espain and M. J. Macias, *Proteins: Structure, Function, and Bioinformatics*, 2008, **73**, 1001-1009.
35. M. Allen, A. Friedler, O. Schon and M. Bycroft, *Journal of molecular biology*, 2002, **323**, 411-416.
36. V. Castillo, F. Chiti and S. Ventura, *PLoS One*, 2013, **8**, e58297.
37. P. Jemth, R. Day, S. Gianni, F. Khan, M. Allen, V. Daggett and A. R. Fersht, *Journal of molecular biology*, 2005, **350**, 363-378.
38. W. Kim, Y. Kim, J. Min, D. J. Kim, Y.-T. Chang and M. H. Hecht, *ACS chemical biology*, 2006, **1**, 461-469.
39. X. Wang, B. Zhou, W. Hu, Q. Zhao and Z. Lin, *Microbial cell factories*, 2015, **14**, 88.
40. E. García-Fruitós, N. González-Montalbán, M. Morell, A. Vera, R. M. Ferraz, A. Arís, S. Ventura and A. Villaverde, *Microbial cell factories*, 2005, **4**, 27.

41. O. Conchillo-Solé, N. S. de Groot, F. X. Avilés, J. Vendrell, X. Daura and S. Ventura, *BMC bioinformatics*, 2007, **8**, 65.
42. S. O. Garbuzynskiy, M. Y. Lobanov and O. V. Galzitskaya, *Bioinformatics*, 2009, **26**, 326-332.
43. A. Trovato, F. Seno and S. C. E. Tosatto, *Protein Engineering, Design and Selection*, 2007, **20**, 521-523.
44. G. A. Hudalla, T. Sun, J. Z. Gasiorowski, H. Han, Y. F. Tian, A. S. Chong and J. H. Collier, *Nature materials*, 2014, **13**, 829.
45. K. Peng, P. Radivojac, S. Vucetic, A. K. Dunker and Z. Obradovic, *BMC Bioinformatics*, 2006, **7**, 208.
46. R. Linding, R. B. Russell, V. Neduva and T. J. Gibson, *Nucleic acids research*, 2003, **31**, 3701-3708.
47. Z. Dosztányi, V. Csizmok, P. Tompa and I. Simon, *Bioinformatics*, 2005, **21**, 3433-3434.
48. R. Graña-Montes, P. Marinelli, D. Reverter and S. Ventura, *Biomacromolecules*, 2014, **15**, 1194-1203.
49. H. Levine and J. D. Scholten, in *Methods in enzymology*, Elsevier, 1999, vol. 309, pp. 467-476.
50. W. E. Klunk, J. Pettegrew and D. J. Abraham, *Journal of Histochemistry & Cytochemistry*, 1989, **37**, 1273-1281.
51. J. Pujols, S. Peña-Díaz, M. Conde-Giménez, F. Pinheiro, S. Navarro, J. Sancho and S. Ventura, *International journal of molecular sciences*, 2017, **18**, 478.
52. W.-F. Xue, S. W. Homans and S. E. Radford, *Proceedings of the National Academy of Sciences*, 2008, **105**, 8926-8931.
53. R. B. Wickner, F. P. Shewmaker, D. A. Bateman, H. K. Edskes, A. Gorkovskiy, Y. Dayani and E. E. Bezsonov, *Microbiology and Molecular Biology Reviews*, 2015, **79**, 1-17.
54. E. Goormaghtigh, V. Cabiaux and J. M. RUYSSCHAERT, *European Journal of Biochemistry*, 1990, **193**, 409-420.
55. K. M. Wilbur and N. G. Anderson, *Journal of biological chemistry*, 1948, **176**, 147-154.
56. B. Jönsson and H. Wennerström, *Biophysical chemistry*, 1978, **7**, 285-292.
57. R. A. Azizyan, A. Garro, Z. Radkova, A. Anikeenko, A. Bakulina, C. Dumas and A. V. Kajava, *Journal of molecular biology*, 2018, **430**, 3835-3846.
58. U. Baxa, T. Cassese, A. V. Kajava and A. C. Steven, *Advances in protein chemistry*, 2006, **73**, 125-180.
59. F. Shewmaker, D. Kryndushkin, B. Chen, R. Tycko and R. B. Wickner, *Biochemistry*, 2009, **48**, 5074-5082.
60. N. Luckgei, A. K. Schütz, L. Bousset, B. Habenstein, Y. Sourigues, C. Gardiennet, B. H. Meier, R. Melki and A. Böckmann, *Angewandte Chemie International Edition*, 2013, **52**, 12741-12744.

61. A. B. Ahmed, N. Znassi, M.-T. Château and A. V. Kajava, *Alzheimer's & Dementia*, 2015, **11**, 681-690.
62. W. L. DeLano, *CCP4 Newsletter On Protein Crystallography*, 2002, **40**, 82-92.
63. G. S. Waldo, B. M. Standish, J. Berendzen and T. C. Terwilliger, *Nature biotechnology*, 1999, **17**, 691.
64. F. Yang, L. G. Moss and G. N. Phillips, *Nature biotechnology*, 1996, **14**, 1246.
65. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *Journal of computational chemistry*, 2004, **25**, 1605-1612.
66. A. Fiser and A. Sali, *Bioinformatics*, 2003, **19**, 2500-2501.
67. R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *Journal of applied crystallography*, 1993, **26**, 283-291.