

OPEN

# High-yield Production of Amyloid- $\beta$ Peptide Enabled by a Customized Spider Silk Domain

Axel Abelein<sup>1</sup>, Gefei Chen<sup>1</sup>, Kristīne Kitoka<sup>2</sup>, Rihards Aleksis<sup>2</sup>, Filips Oleskovs<sup>2</sup>, Médoune Sarr<sup>1</sup>, Michael Landreh<sup>3</sup>, Jens Pahnke<sup>4,5,6</sup>, Kerstin Nordling<sup>1</sup>, Nina Kronqvist<sup>1</sup>, Kristaps Jaudzems<sup>2</sup>, Anna Rising<sup>1,7</sup>, Jan Johansson<sup>1</sup> & Henrik Biverstål<sup>1,2\*</sup>

During storage in the silk gland, the N-terminal domain (NT) of spider silk proteins (spidroins) keeps the aggregation-prone repetitive region in solution at extreme concentrations. We observe that NTs from different spidroins have co-evolved with their respective repeat region, and now use an NT that is distantly related to previously used NTs, for efficient recombinant production of the amyloid- $\beta$  peptide (A $\beta$ ) implicated in Alzheimer's disease. A designed variant of NT from *Nephila clavipes* flagelliform spidroin, which in nature allows production and storage of  $\beta$ -hairpin repeat segments, gives exceptionally high yields of different human A $\beta$  variants as a solubility tag. This tool enables efficient production of target peptides also in minimal medium and gives up to 10 times more isotope-labeled monomeric A $\beta$  peptides per liter bacterial culture than previously reported.

Orb-weaving spiders manufacture up to seven different silks, e.g. dragline silk derived from major ampullate silk proteins (spidroins, MaSp) and flagelliform silk derived from flagelliform spidroins (FlSp). The various spidroins share a common architecture - a large core repetitive region capped by globular N- and C-terminal domains (NT and CT)<sup>1</sup>. The divergent and large aggregation-prone repetitive regions of the spidroins determine the mechanical properties of the respective spider silks, while the terminal domains regulate silk fiber formation<sup>2,3</sup>. Despite their high aggregation propensity the spidroins can be stored at extremely high concentrations (30–50% w/v) in the spider silk gland, solubilized by the NT domain<sup>1,4</sup>.

The NT dimerizes upon a drop in pH, which is crucial for silk fiber formation<sup>1,5</sup>. To ensure solubility also at low pH and widen the applicability of NT as a solubility enhancing fusion partner, a charged-reversed mutant has been designed (referred to as NT\*<sub>MaSp</sub>)<sup>6</sup>. The previously reported NT\*<sub>MaSp</sub> tag is derived from the NT domain of *Euprostheno australis* MaSp1 and folds as a five-helix bundle<sup>6,7</sup>. NT\*<sub>MaSp</sub> is a pH insensitive constitutive monomer, highly stable and extremely soluble, and has been successfully applied for efficient production and purification of, among others, lung surfactant protein analogs, cholecystokinin-58, human antimicrobial cathelicidin and a designed  $\beta$ -sheet protein<sup>6,8</sup>.

Aggregation-prone proteins and peptides are associated with several neurodegenerative disorders, e.g. Alzheimer's disease (AD), the most prevalent form of dementia<sup>9,10</sup>. These proteins/peptides often exhibit high  $\beta$ -sheet propensity, which make them prone to aggregate and form insoluble amyloid fibrils<sup>11</sup>. These intrinsic properties of amyloid-forming proteins make high-yield biochemical production challenging, yet the availability of pure protein samples is crucial for studying protein self-assembly and its associated neurotoxicity *in vitro* and *in vivo*. This is probably one important reason behind the fact that, despite immense efforts, the exact mechanisms of A $\beta$  self-assembly are still unknown<sup>9–11</sup>. Recent advances have however revealed new insights into the nucleation mechanism of A $\beta$  *in vitro*<sup>12,13</sup>, structural details of the fibril morphology<sup>14</sup> and biological mechanisms implicated in the AD etiopathology<sup>15,16</sup>. These experiments typically require access to very pure and homogeneous A $\beta$  peptides as small impurities or preformed seeds have a great impact on the aggregation behavior<sup>17</sup>. In particular

<sup>1</sup>Department of Neurobiology, Care Sciences and Society, Center for Alzheimer Research, Division of Neurogeriatrics, Karolinska Institutet, 141 52, Huddinge, Sweden. <sup>2</sup>Department of Physical Organic Chemistry, Latvian Institute of Organic Synthesis, Riga, LV-1006, Latvia. <sup>3</sup>Department of Molecular Tumor and Cell Biology, Karolinska Institutet, 171 65, Solna, Sweden. <sup>4</sup>Department of Pathology, University of Oslo/Oslo University Hospital, N-0424, Oslo, Norway. <sup>5</sup>LIED, University of Lübeck, D-23538, Lübeck, Germany. <sup>6</sup>Department of Pharmacology, Medical Faculty, University of Latvia, Riga, LV-1004, Latvia. <sup>7</sup>Department of Anatomy, Physiology and Biochemistry, Swedish University of Agricultural Sciences, 750 07, Uppsala, Sweden. \*email: [henrik.biverstal@ki.se](mailto:henrik.biverstal@ki.se)

for structural studies of amyloid fibrils, but also for certain *in vivo* experiments, the availability of large quantities of isotope-labeled A $\beta$  is essential.

Studies of A $\beta$  aggregation *in vitro* have often been conducted with synthetically produced peptides<sup>18,19</sup>. Synthetic preparations have several drawbacks including batch-to-batch variations, intrinsic impurities and relatively high cost, especially for isotope labeling. As a consequence, several recombinant expression systems have been established. These production protocols either result in peptides with an initiating non-native methionine residue<sup>20,21</sup> or are based on solubility tags that require proteolytic cleavage to obtain the native human A $\beta$  sequence<sup>22–24</sup>. The main disadvantage of having methionine as the first residue is that it might affect processes such as posttranslational modifications, *e.g.* pyroGlu formation<sup>25–27</sup>, and metal ion binding since the metal ion-binding site is located in the N-terminus<sup>28–30</sup>. Here we describe a useful solubility tag for production of aggregation-prone proteins and peptides, and demonstrate that this tool enables very efficient production of native and isotope-labeled A $\beta$  peptides.

## Results and Discussion

**Evolutionary relationships of NT and repetitive regions of different spidroins.** A phylogenetic tree based on sequence alignment of 67 NTs found in GenBank (Supplementary Fig. S1) reveal evolutionary relationships between NT and their respective repetitive regions (Fig. 1A). The NT domains cluster according to silk type, as previously reported<sup>7</sup>. Hence, the NTs of different spidroin types, which are defined by the nature of their respective repetitive regions, have been conserved through evolution of different spider species. Structural characteristics of the repetitive regions appear to co-vary with the evolution of NTs, *e.g.* for the tubuliform (TuSp) and aciniform (AcSp) NTs, which are evolutionarily close, the corresponding repetitive regions stand out by forming globular folded domains<sup>31,32</sup> (Fig. 1A). NT from FlSp is linked to a unique repetitive region that contains several embedded spacers (each 27 residues), which are predicted to form  $\beta$ -hairpins<sup>33</sup> (Fig. 1A). NT<sub>FlSp</sub> exhibits distant evolutionary relationship (<35% sequence identity) to the previously reported NT<sub>MaSp</sub><sup>6</sup> (Fig. 1B) and MaSpS contain repetitive regions with predicted  $\alpha$ -helical and random coil structures<sup>34–38</sup>. We speculate that different NTs may have evolved to facilitate optimal solubility of their respective repeat region in the silk gland during storage conditions, where pH is neutral and NT monomeric<sup>4,5</sup>. Irrespective of any potential evolutionary co-variation between NTs and the repetitive regions, we aimed to explore whether NT<sub>FlSp</sub> could work in protein expression in an equivalent way to the previously investigated and distantly related NT<sub>MaSp</sub><sup>6</sup>.

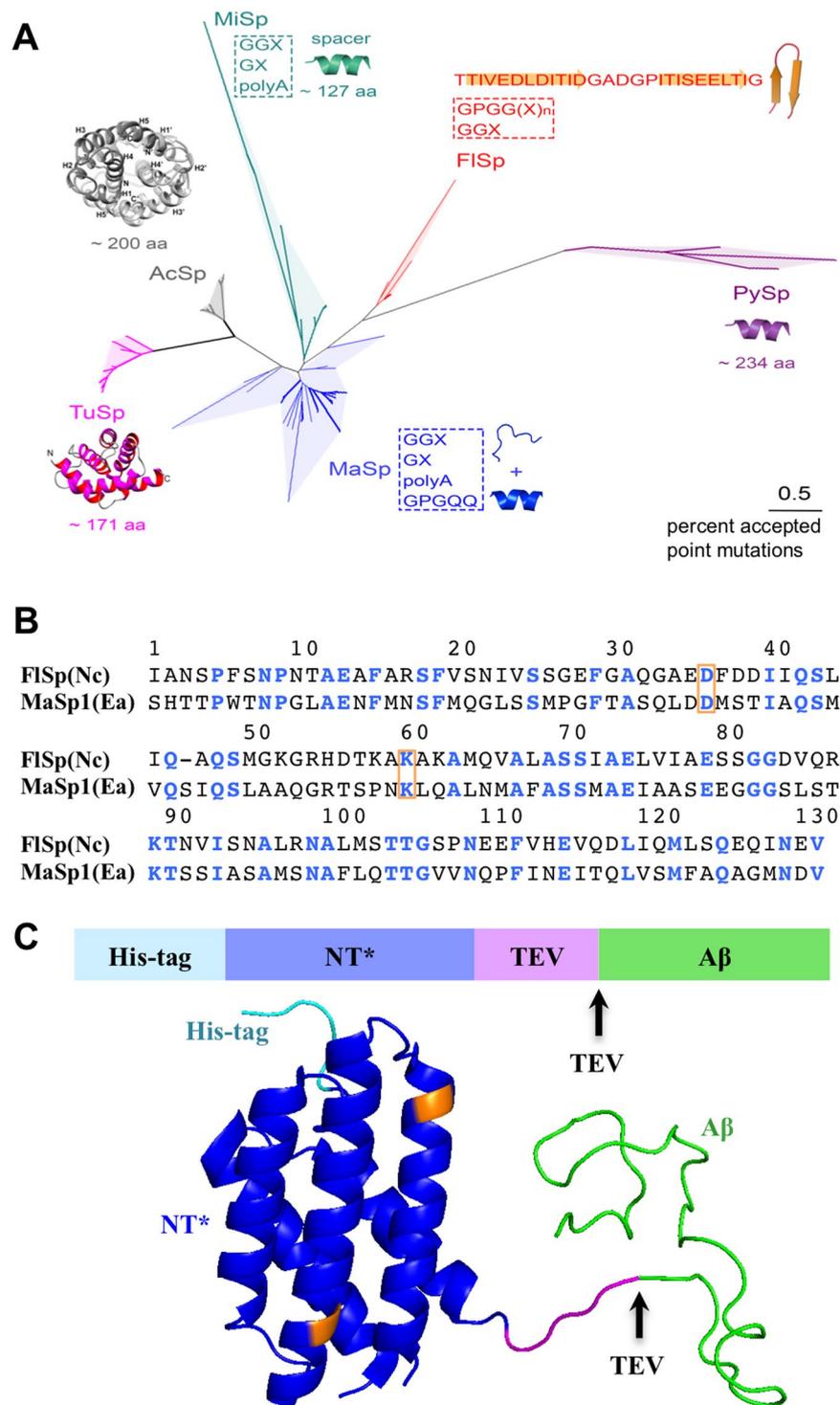
**Design of the novel solubility tag NT\*<sub>FlSp</sub>.** To prevent dimerization of NT<sub>FlSp</sub> at low pH we introduced the D40K and K65D mutations<sup>6</sup> in NT<sub>FlSp</sub> from *Nephila clavipes* (*Nc*) (Fig. 1B) (numbering as described previously<sup>6</sup>, wherefore the mutations correspond to positions 36 and 60 in Fig. 1B). NT\*<sub>FlSp</sub> has a larger number of charged residues (25 vs. 11) and stronger net charge (−7 vs. −5) compared to NT\*<sub>MaSp</sub>, which potentially enhance its solubility properties. In contrast to NT\*<sub>MaSp</sub>, NT\*<sub>FlSp</sub> has no tryptophan, whose absorbance at 280 nm would cover the intrinsic low absorbance at 280 nm of the target peptide A $\beta$ . Thus, for size exclusion chromatography (SEC), where detection relies on the protein absorbance at 280 nm, NT\*<sub>FlSp</sub> enables clearly separated intensity peaks.

**Efficient expression and purification of A $\beta$  monomers using NT\*<sub>FlSp</sub>.** We designed the fusion protein NT\*<sub>FlSp</sub>-A $\beta$  by fusing the genetic codes of the solubility tag NT\*<sub>FlSp</sub> and A $\beta$  with a TEV recognition site in-between (Fig. 1C). An overview of the expression and purification scheme is given in Fig. 2. The fusion protein was expressed in BL21(DE3) *E. coli* cells grown in rich or minimal medium, dissolved in 8 M urea after cell lysis and purified using immobilized metal ion affinity chromatography (IMAC). Urea was added as denaturant to increase binding to IMAC column. For optimal cleavage of the fusion protein by TEV, a buffer exchange was conducted, either by overnight dialysis or by column chromatography. TEV cleavage can alternatively be conducted during buffer dialysis to speed-up the purification, yet a short dialysis step to decrease the urea concentration below 2 M is recommended before the addition of TEV protease. Finally, the solution was applied to SEC with a Superdex 30 column, whereby monomeric A $\beta$  monomers were isolated.

The expression and purification protocol presented here results in highly pure A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub> monomers within 2.5–4.5 days, depending on buffer exchange and cleavage method. The yields in rich and minimal medium are listed in Table 1. The example shown in Fig. 2 represents purifications from 100 and 500 mL culture medium, yielding very similar amounts of 37 ± 7 mg of pure A $\beta$ <sub>42</sub> monomers if extrapolated to one-liter culture. Notably, the present scheme gives by far the highest yields, both in rich and minimal medium, compared to other reported protocols (Table 2).

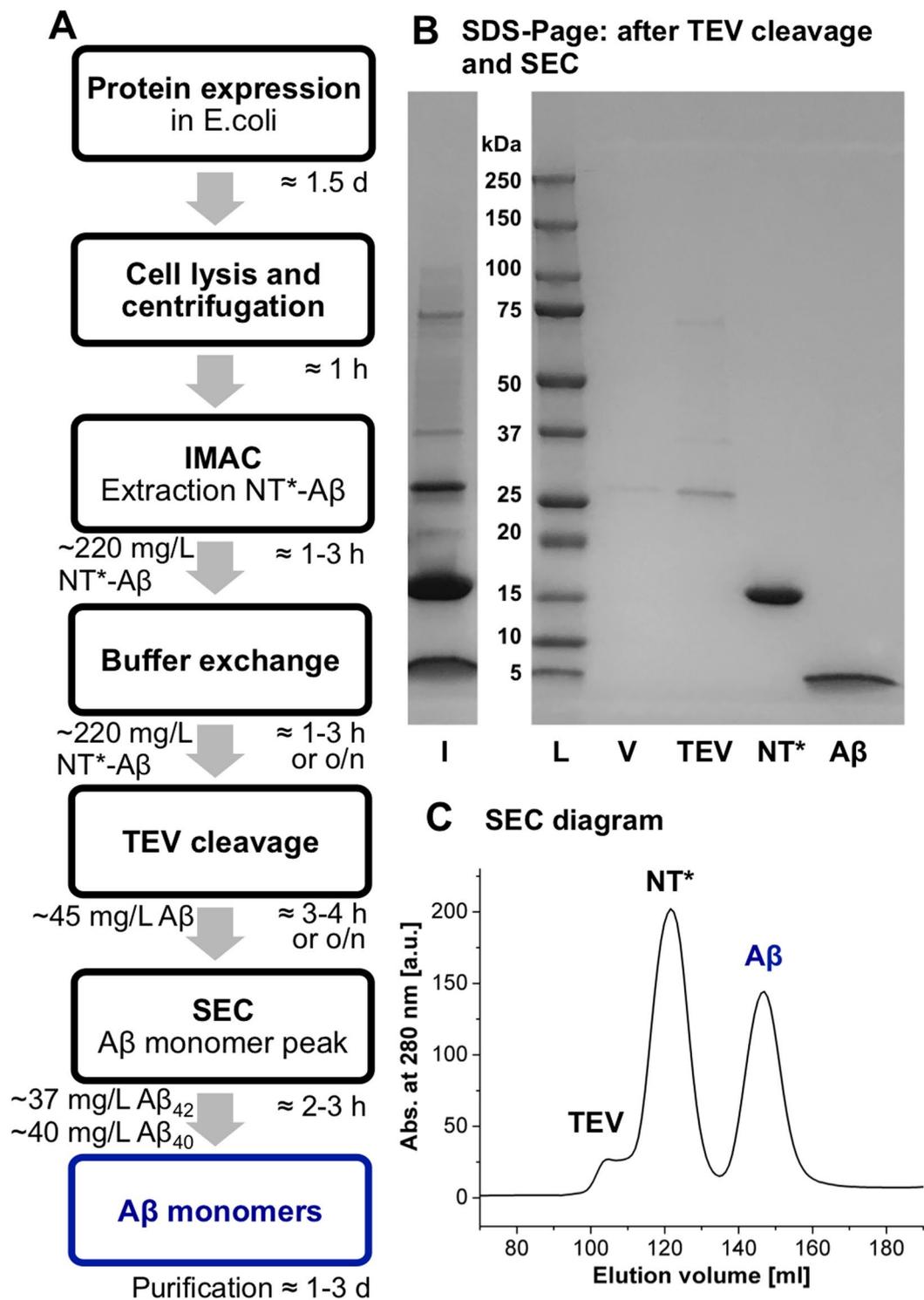
The purified peptides were investigated using mass spectrometry, confirming the expected masses for A $\beta$ , here shown for A $\beta$ <sub>40</sub> (Fig. 3A). Using <sup>13</sup>C-<sup>15</sup>N-double-labeled A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub> we performed nuclear magnetic resonance (NMR) experiments to confirm the purity and structural state of the purified peptides. We recorded <sup>1</sup>H-<sup>15</sup>N-HSQC experiments (Fig. 3B and Supplementary Fig. S3) where the chemical shifts of the cross-peaks coincide with previous assignments reported in the literature, revealing a monomeric, predominantly unstructured conformation of the purified A $\beta$  peptides<sup>39,40</sup>. To analyze the secondary structure of monomeric A $\beta$ <sub>42</sub> we applied circular dichroism (CD) spectroscopy. The initial CD spectra (Fig. 3C) indicated a predominantly unstructured conformation as previously reported<sup>20,21,41,42</sup>. Taken together, these experiments confirm that our method results in monomeric A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub> peptides.

**Production of 4-fluoro-Phe-labeled A $\beta$  peptides.** The present approach also opens new opportunities for NMR studies that require more complex isotope labeling approaches associated with reduced protein yields. For example, we have used the NT\*<sub>FlSp</sub> tag to produce monomeric A $\beta$ <sub>42</sub> incorporating 4-fluoro-Phe (4FF-A $\beta$ <sub>42</sub>) in milligram yields. The expression was performed similarly as described above, but glycosylase and DL-tyrosine



**Figure 1.** (A) Evolutionary relationships of the NTs of different spidroins. The analysis involved 67 NT amino acid sequences (Supplementary Fig. 1), revealing that spidroins from different spider species cluster according to the silk type in the phylogenetic tree. The typical repetitive regions of the respective spidroins and their known structures or main secondary structure propensities are displayed. (B) Sequence alignment of NT<sub>FISp</sub> and NT<sub>MaSp</sub> where strictly conserved residues are colored in blue. The residues marked in orange display the mutated sites in NT\*. (C) Schematic representation and structure of the NT\*-A $\beta$  fusion protein where the arrows indicate the TEV protease cleavage site. The mutated D and K residues are marked by yellow colour in the NT structure (pdb 4FBS).

was supplemented to bacterial cultures at an OD<sub>600nm</sub> value of 0.6. Further, DL-4-fluorophenylalanine was added when the OD<sub>600nm</sub> value reached 0.8 and cell expression was induced. The <sup>1</sup>H-<sup>15</sup>N-HSQC spectrum of <sup>15</sup>N-labeled 4FF-A $\beta$ <sub>42</sub> revealed again a monomeric peptide (Supplementary Fig. S3).



**Figure 2.** (A) Schematic expression and purification protocol, including typical times for performance. Yields of NT\*-A $\beta_{42/40}$  are derived from 1 L expression cultures and extrapolated from purification from 100 and 500 mL, resulting in very similar values. (B) SDS-PAGE gel, with protein ladder (L), void (V), before (I) and after SEC yielding pure A $\beta$ . An uncropped full-length gel is presented in Supplementary Fig. S2. (C) SEC diagram showing separation of TEV, NT\* and monomeric A $\beta$ .

**Aggregation kinetics of native and isotope-labeled A $\beta_{42}$ .** To ensure that the isolated peptides behave as expected, we investigated the aggregation kinetics starting from monomeric A $\beta$  peptides. Recording CD signals under continuous stirring at 37°C of 10  $\mu$ M A $\beta_{42}$ , a structural conversion from an unstructured to a  $\beta$ -structured conformation was observed (Fig. 3C), where the isodichroic point at 208 nm indicates a two-state transition. Furthermore, we used 50  $\mu$ M 4FF-A $\beta_{42}$  for real-time aggregation  $^{19}$ F-NMR studies at 25°C, revealing a

Protein/peptide	Rich medium [mg/L]	M9 medium [mg/L]
NT* <sub>FISp</sub> -A $\beta$ <sub>40</sub>	216 ± 29	74 ± 22
NT* <sub>FISp</sub> -A $\beta$ <sub>42</sub>	223 ± 42	88 ± 10
Monomeric A $\beta$ <sub>40</sub>	40 ± 5	13 ± 4
Monomeric A $\beta$ <sub>42</sub>	37 ± 7	14 ± 2

**Table 1.** Average yields of fusion proteins and monomeric A $\beta$  peptides in rich (LB) and minimal (M9) medium in mg per liter culture. Errors were estimated as standard deviations from 5 replicates by western blot analysis (see Methods).

A $\beta$ variant	Fusion partner/expression method	Purified A $\beta$ peptide yield in mg/L in rich medium	Purified A $\beta$ peptide yield in mg/L in minimal medium	Reference
A $\beta$ (1–40)	NT* <sub>FISp</sub>	40 ± 5	13 ± 4	here
	(NANP) <sub>19</sub>	22	—	22
	IFABP	4	—	24
	GST	7	1.5	23
A $\beta$ (M1–40)	Directly from inclusion bodies	10–20	—	21
	Directly from inclusion bodies	—	10–15	51
	Co-expression with Z <sub>A<math>\beta</math>33</sub>	—	4	20
A $\beta$ (1–42)	NT* <sub>FISp</sub>	37 ± 7	14 ± 2	here
	(NANP) <sub>19</sub>	19	—	22
	IFABP	3	—	24
	IFABP	—	6	52
	GST	15	—	53
	Ub	4	—	54
A $\beta$ (M1–42)	Directly from inclusion bodies	8	—	21
	Co-expression with Z <sub>A<math>\beta</math>33</sub>	—	3	20

**Table 2.** Yields of A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub> variants reported in literature and herein.

decrease of 4FF-signals over time (Fig. 3D). The signal loss can be fitted to a sigmoidal decline, with an aggregation half time of  $258 \pm 5$  min under the conditions used.

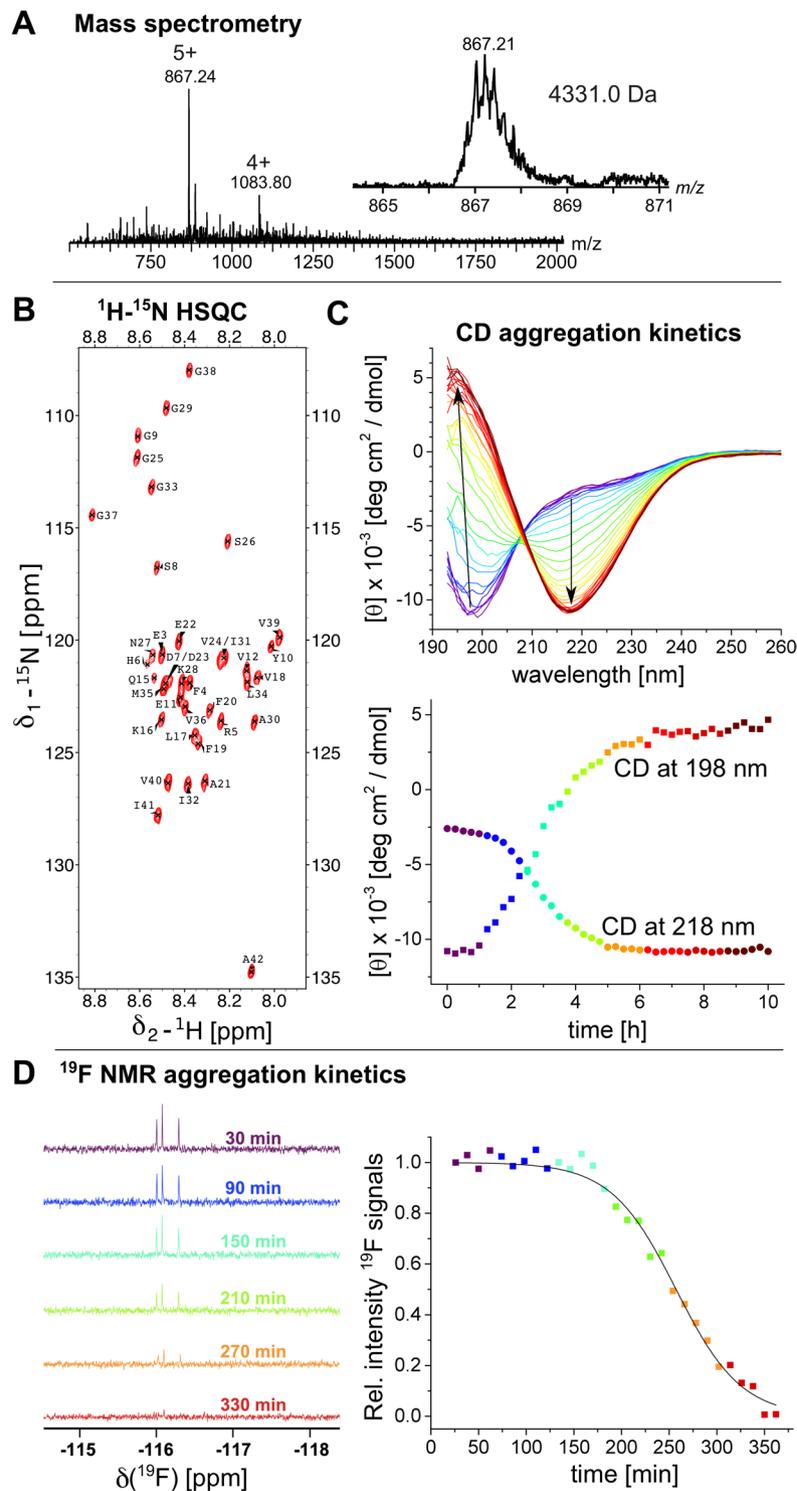
Alternatively, A $\beta$  aggregation kinetics can be monitored using the fluorescence dye thioflavin T (ThT), for a detailed elucidation of the nucleation mechanism. Here, we conducted ThT experiments on A $\beta$ <sub>42</sub> in 20 mM sodium phosphate buffer, pH 8.0, at 37 °C under quiescent conditions at different initial A $\beta$ <sub>42</sub> monomer concentrations, [A $\beta$ ] (Fig. 4). The final fluorescence intensity exhibits a linear dependence on the initial monomer concentration (Fig. 4D), suggesting that the total amount of initially monomeric peptides forms ThT-active fibril material, as previously shown for A $\beta$ <sup>12,13,42</sup>. The aggregation half times of A $\beta$ <sub>42</sub> used here exhibit a simple relation  $\tau_{1/2} \propto [A\beta]^\gamma$ , with  $\gamma = -1.0 \pm 0.1$ , corresponding to the slope in a double logarithmic plot (Fig. 4C). This value is in the same range as found for A $\beta$ <sub>M42</sub> with an initial methionine, where  $\gamma = -1.3$  was reported<sup>12</sup>. For  $\gamma = -1.0$  a multi-step secondary nucleation model describes better the observed aggregation traces compared to a single-step secondary nucleation model (Fig. 4A,B). The multi-step model additionally includes saturation of secondary nucleation and was previously shown to be applicable for the shorter A $\beta$ <sub>40</sub> and A $\beta$ <sub>M40</sub> variants<sup>13,29,42</sup> and for A $\beta$ <sub>M42</sub> at pH 7.4<sup>43</sup>, which all exhibit higher  $\gamma$ -values, but also describes well the kinetics of A $\beta$ <sub>M42</sub> at pH 8.0<sup>44</sup>. Hence, these results confirm that the native and isotope-labeled peptides obtained herein are highly pure and in a monomeric state, which is essential for accurate and reproducible aggregation kinetics experiments.

## Conclusions

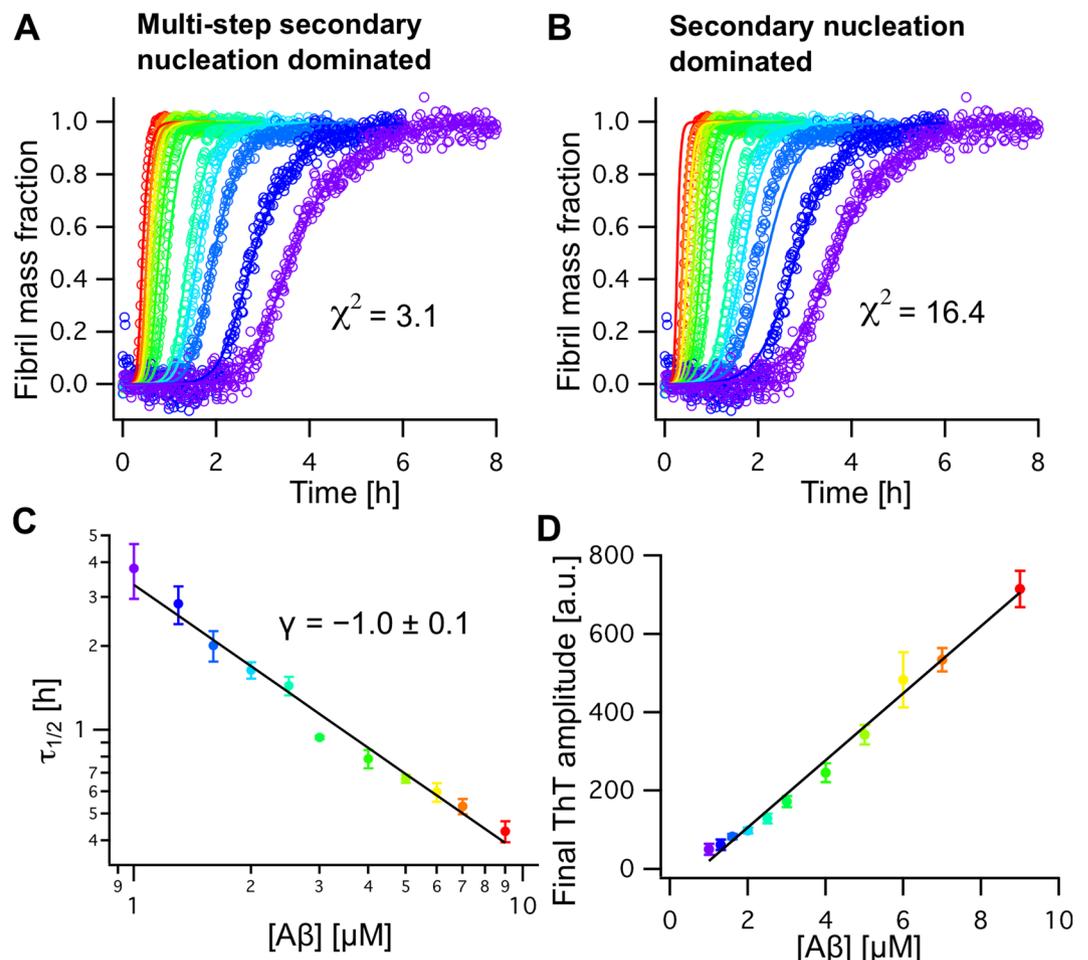
Taken together, we have developed a biomimetic tool that provides facile, fast and inexpensive production of pure and monomeric A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub> peptides. The high yield obtained also in minimal medium enables efficient generation of isotope-labeled A $\beta$  peptides. Peptides produced by our protocol recapitulate the behavior of A $\beta$  peptides obtained by other means, which indicate the applicability of using NT\*<sub>FISp</sub> for generating functional A $\beta$  peptides. The NT\*<sub>FISp</sub>-tag holds great potential, also when compared to NT\*<sub>MaSp</sub><sup>6</sup>, for efficient production of medically relevant aggregation-prone peptides and proteins. This is important since the majority of new pharmaceuticals are biologics and facile protocols for efficient production of proteins that are difficult to produce are needed.

## Methods

**Expression and purification protocol.** The synthetic gene coding for NT FISp from *Nephila clavipes* with the D40K and K65D mutations (NT\*<sub>FISp</sub>) was ordered from GenScript (GenScript Biotech, Netherlands). The NT\*<sub>FISp</sub> gene was ligated into pT7 plasmid containing TEV recognition site (TRS)-A $\beta$ <sub>40</sub>/A $\beta$ <sub>42</sub> as described



**Figure 3.** (A) Mass spectrum of  $\text{A}\beta_{40}$  showing a pure peptide with an average mass of 4331 Da. The inset shows a zoom of the 5+ charged ion. (B)  $^1\text{H}$ - $^{15}\text{N}$ -HSQC spectrum of  $15\ \mu\text{M}$   $\text{A}\beta_{42}$  recorded at  $5\ ^\circ\text{C}$ , revealing monomeric peptide. (C) Aggregation kinetics of  $10\ \mu\text{M}$   $\text{A}\beta_{42}$  at  $37\ ^\circ\text{C}$  under continuous stirring recorded by CD spectroscopy. The spectra exhibit a structural transition from a predominantly unstructured state to a  $\beta$ -structure. The lower panel shows the time dependence of the CD extremes at 198 nm (squares) and 218 nm (circles), with the same color code as used for the CD spectra. (D) Aggregation kinetics of  $50\ \mu\text{M}$  4FF- $\text{A}\beta_{42}$  at  $25\ ^\circ\text{C}$  monitored by  $^{19}\text{F}$ -NMR spectra of the signals around  $-116$  ppm, exhibiting attenuation of 4FF-signals. The color code represents the same time points in both panels.



**Figure 4.** (A,B) Aggregation kinetics of Aβ<sub>42</sub> at different concentrations from 1.0 (violet) to 9.0 μM (red) recorded by ThT fluorescence experiments fitted with a multi-step (A) and simple secondary nucleation model (B). The kinetic traces fit best to the multi-step secondary nucleation model, reflected by a lower  $\chi^2$  value. (C) Aggregation half times,  $\tau_{1/2}$ , plotted against the initial peptide concentration, [Aβ], exhibit a  $\gamma$  coefficient of  $\gamma = -1.0 \pm 0.1$ . (D) The final ThT intensity of the normalized aggregation traces in (A) exhibits a linear relation to [Aβ].

previously<sup>6</sup>. The plasmids were transformed into chemically competent *E. coli* BL21 (DE3) cells and expressed as described previously<sup>45</sup>. In short, 1 mL overnight culture was inoculated to 100 mL LB medium (1/100) or 100 mL M9 overnight culture was inoculated to 1 L M9 minimal medium (10/100) with 70 mg/l kanamycin. Cells were grown at 30 °C at 120 RPM to OD<sub>600nm</sub> around 0.8–0.9, where the temperature was lowered to 20 °C, and 0.1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added and the cells were incubated overnight. To isolate the cells from media, the bacterial culture was centrifuged at 5,000 × g for 20 minutes at 4 °C and the cell pellets resuspended in 40 mL 20 mM Tris-HCl pH 8.0, split in to two 50 mL falcon tubes and stored at –20 °C. The frozen cells were thawed and urea was added to a concentration of 8 M. The cells were sonicated to obtain a clear solution. The lysate was then loaded on 2 × 5 mL Ni-NTA column (GE Healthcare). Unbound proteins were washed away with 15 mM Imidazole in 20 mM Tris-HCl, pH 8 and 8 M urea. The fusion protein was eluted with 300 mM imidazole in 20 mM Tris-HCl, pH 8 and 8 M urea. The fractions containing the fusion protein were pooled and then dialyzed overnight against 20 mM Tris-HCl pH 8.0 at 4 °C. To remove the His6-NT\*<sub>FISP</sub> part, the fusion protein was cleaved with TEV protease (1:20–1:30, enzyme to substrate, w/w) at 4 °C overnight in 20 mM Tris-HCl pH 8, 0.5 mM EDTA and 1 mM DTT. After TEV cleavage, the sample was dissolved in 15 mL 7 M guanidine-HCl and separated on a Superdex 30 26/600PG size exclusion column (Fig. 2C). The correct size of Aβ, NT\*<sub>FISP</sub> and TEV was confirmed by SDS/PAGE in a 4–20% polyacrylamide gel, stained with Coomassie brilliant blue dye (Fig. 2B). For expression of <sup>15</sup>N- and <sup>13</sup>C- labelled NT\*<sub>FISP</sub>-Aβ, the same procedure was used except that M9 minimal medium containing <sup>15</sup>NH<sub>4</sub>Cl (1 g/L M9) and <sup>13</sup>C-glucose (4 g/L M9) was used. The plasmid pRK793 for TEV expression was obtained from addgene (addgene.org, deposited by David S. Waugh) and was expressed as described above and purified as described previously<sup>46</sup>.

**Determination of yields.** Both NT\*<sub>FISP</sub>-Aβ<sub>40</sub> and NT\*<sub>FISP</sub>-Aβ<sub>42</sub> was transformed into BL21 (DE3) *E. coli* cells and spread onto an agar plate with kanamycin. 5 starting cultures of LB and M9 were inoculated with individual colonies and incubated at 31 °C overnight. The expression was performed as described above in 100 mL LB

and M9 media. 100  $\mu\text{L}$  of each culture was taken before and after induction, lyophilized, dissolved in SDS loading buffer and boiled for 10 minutes at 96 °C. 1  $\mu\text{L}$  of each induced sample and 1  $\mu\text{L}$  uninduced sample from each condition was loaded on a 4–20% mini protean TGX gel (BioRad) and blotted on a PVDF membrane (GE healthcare). 5% w/v non-fat dry milk/PBS was used to block the membrane after blotting for 1 h, followed by incubation with 6E10 primary antibody in 5% w/v non-fat dry milk, 0.1% Tween/PBS overnight at 4 °C. The membranes were washed three times with 0.1% Tween/PBS, and ECL anti-mouse secondary antibodies in 5% w/v non-fat dry milk and 0.1% Tween/PBS were added for 1 h at room temperature. Enhanced chemiluminescence detection reagent (GE Healthcare) was added and images were acquired using an AI600 imaging system (GE healthcare). The concentration of each sample was calculated by integration of the peaks from IMAC (fusion protein) and SEC (monomeric A $\beta$ ) with an extinction coefficient  $\epsilon_{280} = 2,980 \text{ M}^{-1} \text{ cm}^{-1}$  for the fusion protein and  $1,424 \text{ M}^{-1} \text{ cm}^{-1}$  for A $\beta$ . Western blot intensities were analyzed by ImageJ software<sup>47</sup> and average and standard deviation from 5 replicates was calculated using yields from one full purification of each condition. Values are listed in Table 1.

**Expression protocol of 4FF-A $\beta$ 42.** The plasmid pT7His6NT\*<sub>F1Sp</sub>-TEV recognition site -A $\beta$ <sub>42</sub> was transformed into chemically competent *E. coli* BL21(DE3) cells. Colonies were inoculated to 10 mL LB medium with 70 mg/L kanamycin and grown at 30 °C and 200 r.p.m. to OD<sub>600nm</sub> > 1.0. 0.5 mL day culture was inoculated to 25 mL M9 medium with 70 mg/l kanamycin and grown at 30 °C and 200 r.p.m. overnight. 10 mL overnight culture was inoculated to 1 L M9 medium and cells were further grown at 30 °C. Uniform labeling with 4-fluorophenylalanine was achieved by the introduction of 1 g/L glyphosate and 75 mg/L DL-tyrosine to shaking bacterial cultures at 30 °C which had reached an OD<sub>600nm</sub> of 0.6. Once cells achieve an OD<sub>600nm</sub> of 0.8, 30 mg/L DL-4-fluorophenylalanine was added. The incubation temperature was lowered to 20 °C and expression was induced with the addition of IPTG to 0.1 mM, the cells were incubated overnight and were harvested by 7,000  $\times$  g centrifugation at 4 °C.

**Evolutionary relationships of the NT domains of different spidroins.** The evolutionary history of the NT domains from different spidroins was inferred by the Neighbor-Joining method with the Poisson correction. Evolutionary analyses were conducted in MEGA7<sup>48</sup>. The analysis involved 67 amino acid sequences. In the spider silk gland (liquid protein), the repetitive region of MaSp, consisting of GGX, polyA, GX and GPGQQ, is disordered and partially helical<sup>34–38</sup>, and MiSp and F1Sp share identical motifs<sup>33</sup>. However, there are ~127-aa spacer in MiSp, which adopt  $\alpha$ -helical conformation, whereas the 27-aa spacer in F1Sp is predicted to fold to  $\beta$ -hairpin<sup>33</sup>. The large repetitive domains of AcSp and TuSp adopt  $\alpha$ -helical conformation<sup>31,32</sup>, and the repetitive domains of PySp is also predicted to adopt  $\alpha$ -helical conformation by PSIPRED v3.3 (<http://bioinf.cs.ucl.ac.uk/psipred/>).

**Mass spectrometry.** Purified A $\beta$ <sub>40</sub> was diluted 1:10 in H<sub>2</sub>O/acetonitrile/formic acid (70:30:0.2) and directly infused into a Waters LCT Time of flight mass spectrometer (MS Vision, NL) equipped with an offline nanospray source using borosilicate capillaries (Thermo Scientific). The capillary voltage was 1.5 kV and the cone voltage was 200 V. Spectra were acquired between m/z 500 and 4000 and the mass scale was calibrated with Cesium Iodide. Data were analyzed using MassLynx 4.1 (Waters).

**Nuclear magnetic resonance (NMR).** <sup>1</sup>H-<sup>15</sup>N HSQC spectra were recorded on a 500 MHz or 700 MHz Bruker Avance spectrometer equipped with cryogenic probes. The HSQC spectrum of A $\beta$ <sub>40</sub> was recorded at 500 MHz at 8 °C using 75  $\mu\text{M}$  peptide concentration in 16 mM sodium phosphate buffer, pH 7.4, with 0.02% NaN<sub>3</sub> and 0.2 mM EDTA. For A $\beta$ <sub>42</sub> the peptide concentration was 15  $\mu\text{M}$  in 20 mM sodium phosphate buffer, pH 6.8, recorded at 5 °C and 700 MHz. The spectra were recorded using 2048  $\times$  128 complex points and 32 scans per transient. For A $\beta$ <sub>42</sub> we recorded the HSQC at 15  $\mu\text{M}$  directly after the SEC purification, ensuring the monomeric state of the peptide.

<sup>19</sup>F-NMR experiments were recorded using 50  $\mu\text{M}$  4FF-A $\beta$ <sub>42</sub> in 20 mM sodium phosphate buffer, pH 7.4 with 0.03% NaN<sub>3</sub> and 1 mM EDTA at 25 °C and 565 MHz. <sup>19</sup>F spectrum was acquired with 512 transients and 1.0 s pulse delay between each transient. Line broadening of 1.0 Hz was used to process the final spectrum. The <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of 15  $\mu\text{M}$  4FF-A $\beta$ <sub>42</sub> in 20 mM sodium phosphate buffer, pH 7.4, with 0.02% NaN<sub>3</sub> and 0.2 mM EDTA, was recorded at 4 °C on a 600 MHz Bruker Avance Neo spectrometer equipped with a cryogenic probe.

**Circular dichroism (CD).** CD measurements of 10  $\mu\text{M}$  A $\beta$ <sub>42</sub> in 20 mM sodium phosphate buffer, pH 8.0, at 37 °C were performed in a quartzglass Suprasil 10  $\times$  4 mm CD cuvette (Hellma Analytics) where the optical path length was 4 mm, using a Chirascan CD spectrometer (Applied Photophysics). A resolution of 1.0 nm and a bandwidth of 1 nm were chosen for the aggregation kinetics experiments<sup>42</sup>. During the entire measurement the sample was continuously stirred at around 1200 r.p.m and each 3 min a new CD spectrum was recorded to follow the aggregation kinetics.

**Thioflavin T (ThT) fluorescence kinetics experiments.** For ThT aggregation kinetics experiments 1 to 9  $\mu\text{M}$  monomeric A $\beta$ <sub>42</sub> was used, which was obtained after SEC purification<sup>45</sup>. ThT fluorescence was measured as described previously<sup>45</sup> using 96-well microplates, where each well contained 80  $\mu\text{l}$  sample solution with 10  $\mu\text{M}$  ThT.

**Analysis ThT aggregation kinetics.** Aggregation traces were first analyzed using a fit to a sigmoidal function, revealing the aggregation half time,  $\tau_{1/2}$ <sup>29,42,45</sup>. Subsequently, the aggregation traces were normalized and averaged over six replicates. The averaged aggregation half times are related to the initial monomer concentration, [A $\beta$ ], by  $\tau_{1/2} \propto [\text{A}\beta]^{-\gamma}$  where  $\gamma$  reflects the slope in a double-logarithmic plot (Fig. 4C). Further, we applied a

nucleation model including primary and secondary nucleation in addition to fibril-end elongation<sup>12,49,50</sup>. In order to account for saturation of secondary nucleation an equilibrium constant (Michaelis constant)  $K_M$  can be introduced, referring to a multi-step secondary nucleation model<sup>13</sup>. The kinetic equations for the time dependence of the fibril mass fractions for the two models can be found in refs. <sup>12,13,29,45</sup>. The models were applied to describe the kinetic traces using a global fit analysis<sup>12,13</sup>. The kinetic fitting parameter are listed in Supplementary Table S1.

Received: 12 October 2019; Accepted: 18 December 2019;

Published online: 14 January 2020

## References

- Rising, A. & Johansson, J. Toward spinning artificial spider silk. *Nat. Chem. Biol.* **11**, 309–315 (2015).
- Babb, P. L. *et al.* The *Nephila clavipes* genome highlights the diversity of spider silk genes and their complex expression. *Nat. Genet.* **49**, 895–903 (2017).
- Eisoldt, L., Thamm, C. & Scheibel, T. Review the role of terminal domains during storage and assembly of spider silk proteins. *Biopolymers* **97**, 355–361 (2012).
- Andersson, M. *et al.* Carbonic anhydrase generates CO<sub>2</sub> and H<sup>+</sup> that drive spider silk formation via opposite effects on the terminal domains. *PLoS Biol.* **12**, e1001921 (2014).
- Kronqvist, N. *et al.* Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation. *Nat. Commun.* **5**, 3254 (2014).
- Kronqvist, N. *et al.* Efficient protein production inspired by how spiders make silk. *Nat. Commun.* **8**, 15504 (2017).
- Rising, A., Hjalms, G., Engstrom, W. & Johansson, J. N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindrical spider silk proteins. *Biomacromolecules* **7**, 3120–3124 (2006).
- Sarr, M. *et al.* A spidroin-derived solubility tag enables controlled aggregation of a designed amyloid protein. *FEBS J.* **285**, 1873–1885 (2018).
- Haass, C. & Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. *Nat. Rev. Mol. Cell Biol.* **8**, 101–112 (2007).
- Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).
- Knowles, T. P., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 (2014).
- Cohen, S. I. A. *et al.* Proliferation of amyloid- $\beta$ 42 aggregates occurs through a secondary nucleation mechanism. *Proc. Natl Acad. Sci. USA* **110**, 9758–9763 (2013).
- Meisl, G. *et al.* Differences in nucleation behavior underlie the contrasting aggregation kinetics of the A $\beta$ 40 and A $\beta$ 42 peptides. *Proc. Natl Acad. Sci. USA* **111**, 9384–9389 (2014).
- Fändrich, M. *et al.* Amyloid fibril polymorphism: a challenge for molecular imaging and therapy. *J. Intern. Med.* **283**, 218–237 (2018).
- Masters, C. L. *et al.* Alzheimer's disease. *Nat. Rev. Dis. Prim.* **1**, 15056 (2015).
- Aleksis, R., Oleskovs, F., Jaudzems, K., Pahnke, J. & Biverstal, H. Structural studies of amyloid-beta peptides: Unlocking the mechanism of aggregation and the associated toxicity. *Biochimie* **140**, 176–192 (2017).
- Arosio, P., Cukalevski, R., Frohm, B., Knowles, T. P. & Linse, S. Quantification of the concentration of Abeta42 propagons during the lag phase by an amyloid chain reaction assay. *J. Am. Chem. Soc.* **136**, 219–225 (2014).
- Tickler, A. K., Barrow, C. J. & Wade, J. D. Improved preparation of amyloid-beta peptides using DBU as Nalpha-Fmoc deprotection reagent. *J. Pept. Sci.* **7**, 488–494 (2001).
- Zarandi, M. *et al.* Synthesis of Abeta[1–42] and its derivatives with improved efficiency. *J. Pept. Sci.* **13**, 94–99 (2007).
- Macao, B. *et al.* Recombinant amyloid beta-peptide production by coexpression with an affibody ligand. *BMC Biotechnol.* **8**, 82 (2008).
- Walsh, D. M. *et al.* A facile method for expression and purification of the Alzheimer's disease-associated amyloid beta-peptide. *FEBS J.* **276**, 1266–1281 (2009).
- Finder, V. H., Vodopivec, I., Nitsch, R. M. & Glockshuber, R. The recombinant amyloid-beta peptide Abeta1–42 aggregates faster and is more neurotoxic than synthetic Abeta1–42. *J. Mol. Biol.* **396**, 9–18 (2010).
- Long, F., Cho, W. & Ishii, Y. Expression and purification of 15N- and 13C-isotope labeled 40-residue human Alzheimer's beta-amyloid peptide for NMR-based structural analysis. *Protein Expr. Purif.* **79**, 16–24 (2011).
- Garai, K., Crick, S. L., Mustafi, S. M. & Frieden, C. Expression and purification of amyloid-beta peptides from *Escherichia coli*. *Protein Expr. Purif.* **66**, 107–112 (2009).
- Dammers, C. *et al.* Purification and Characterization of Recombinant N-Terminally Pyroglutamate-Modified Amyloid-beta Variants and Structural Analysis by Solution NMR Spectroscopy. *PLoS One* **10**, e0139710 (2015).
- Gunn, A. P., Masters, C. L. & Cherny, R. A. Pyroglutamate-Abeta: role in the natural history of Alzheimer's disease. *Int. J. Biochem. Cell Biol.* **42**, 1915–1918 (2010).
- Jawhar, S., Wirths, O. & Bayer, T. A. Pyroglutamate amyloid-beta (Abeta): a hatchet man in Alzheimer disease. *J. Biol. Chem.* **286**, 38825–38832 (2011).
- Wärmländer, S. *et al.* Biophysical studies of the amyloid beta-peptide: interactions with metal ions and small molecules. *ChemBioChem* **14**, 1692–1704 (2013).
- Abelein, A., Gräslund, A. & Danielsson, J. Zinc as chaperone-mimicking agent for retardation of amyloid beta peptide fibril formation. *Proc. Natl Acad. Sci. USA* **112**, 5407–5412 (2015).
- Danielsson, J., Pierattelli, R., Banci, L. & Gräslund, A. High-resolution NMR studies of the zinc-binding site of the Alzheimer's amyloid beta-peptide. *FEBS J.* **274**, 46–59 (2007).
- Wang, S., Huang, W. & Yang, D. Structure and function of C-terminal domain of aciniform spidroin. *Biomacromolecules* **15**, 468–477 (2014).
- Lin, Z., Huang, W., Zhang, J., Fan, J. S. & Yang, D. Solution structure of eggcase silk protein and its implications for silk fiber formation. *Proc. Natl Acad. Sci. USA* **106**, 8906–8911 (2009).
- Chen, G. *et al.* Full-length minor ampullate spidroin gene sequence. *PLoS One* **7**, e52293 (2012).
- Hijirida, D. H. *et al.* 13C NMR of *Nephila clavipes* major ampullate silk gland. *Biophys. J.* **71**, 3442–3447 (1996).
- Hronská, M., van Beek, J. D., Williamson, P. T., Vollrath, F. & Meier, B. H. NMR characterization of native liquid spider dragline silk from *Nephila edulis*. *Biomacromolecules* **5**, 834–839 (2004).
- Jenkins, J. E., Holland, G. P. & Yarger, J. L. High resolution magic angle spinning NMR investigation of silk protein structure within major ampullate glands of orb weaving spiders. *Soft Matter* **8**, 1947–1954 (2012).
- Xu, D., Yarger, J. L. & Holland, G. P. Exploring the backbone dynamics of native spider silk proteins in Black Widow silk glands with solution-state NMR spectroscopy. *Polymer* **55**, 3879–3885 (2014).
- Oktaviani, N. A. *et al.* Conformation and dynamics of soluble repetitive domain elucidates the initial beta-sheet formation of spider silk. *Nat. Commun.* **9**, 2121 (2018).

39. Roche, J., Shen, Y., Lee, J. H., Ying, J. & Bax, A. Monomeric Abeta(1-40) and Abeta(1-42) Peptides in Solution Adopt Very Similar Ramachandran Map Distributions That Closely Resemble Random Coil. *Biochemistry* **55**, 762–775 (2016).
40. Danielsson, J., Andersson, A., Jarvet, J. & Gräslund, A. 15N relaxation study of the amyloid beta-peptide: structural propensities and persistence length. *Magn. Reson. Chem.* **44**, S114–121 (2006).
41. Danielsson, J., Jarvet, J., Damberg, P. & Gräslund, A. The Alzheimer beta-peptide shows temperature-dependent transitions between left-handed 3-helix, beta-strand and random coil secondary structures. *FEBS J.* **272**, 3938–3949 (2005).
42. Abelein, A., Jarvet, J., Barth, A., Gräslund, A. & Danielsson, J. Ionic Strength Modulation of the Free Energy Landscape of Abeta40 Peptide Fibril Formation. *J. Am. Chem. Soc.* **138**, 6893–6902 (2016).
43. Meisl, G., Yang, X., Frohm, B., Knowles, T. P. & Linse, S. Quantitative analysis of intrinsic and extrinsic factors in the aggregation mechanism of Alzheimer-associated Abeta-peptide. *Sci. Rep.* **6**, 18728 (2016).
44. Yang, X. *et al.* On the role of sidechain size and charge in the aggregation of Abeta42 with familial mutations. *Proc. Natl Acad. Sci. USA* **115**, E5849–E5858 (2018).
45. Chen, G. *et al.* Bri2 BRICHOS client specificity and chaperone activity are governed by assembly state. *Nat. Commun.* **8**, 2081 (2017).
46. Kapust, R. B. *et al.* Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein Eng.* **14**, 993–1000 (2001).
47. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
48. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
49. Knowles, T. P. J. *et al.* An analytical solution to the kinetics of breakable filament assembly. *Science* **326**, 1533–1537 (2009).
50. Cohen, S. I. A., Vendruscolo, M., Dobson, C. M. & Knowles, T. P. J. From macroscopic measurements to microscopic mechanisms of protein aggregation. *J. Mol. Biol.* **421**, 160–171 (2012).
51. Bertini, I., Gonnelli, L., Luchinat, C., Mao, J. & Nesi, A. A new structural model of Abeta40 fibrils. *J. Am. Chem. Soc.* **133**, 16013–16022 (2011).
52. Sharma, S. C. *et al.* A facile method for expression and purification of (15)N isotope-labeled human Alzheimer's beta-amyloid peptides from *E. coli* for NMR-based structural analysis. *Protein Expr. Purif.* **116**, 82–89 (2015).
53. Chhetri, G., Pandey, T., Chinta, R., Kumar, A. & Tripathi, T. An improved method for high-level soluble expression and purification of recombinant amyloid-beta peptide for *in vitro* studies. *Protein Expr. Purif.* **114**, 71–76 (2015).
54. Lee, E. K., Hwang, J. H., Shin, D. Y., Kim, D. I. & Yoo, Y. J. Production of recombinant amyloid-beta peptide 42 as an ubiquitin extension. *Protein Expr. Purif.* **40**, 183–189 (2005).

## Acknowledgements

This study was supported by the Magnus Bergvall foundation (A.A., H.B.), FLPP/Latvia lzp-2018/1-0275 project support (H.B., J.P.), Swedish Research Council (A.R., J.J.), Center for Innovative Medicine (CIMED) (A.R., J.J.), FORMAS (A.R.), Vinnova (J.J.), Olle Engkvists Foundation (G.C.), Swedish Alzheimer foundation (G.C.), Åhlen Foundation (A.A., G.C., H.B.), Stiftelsen för Gamla Tjänarinnor (A.A., H.B., G.C.), P. & A. Hedlunds Stiftelse (G.C.), Instruct R&D pilot project grant APPID 272 (A.A.), Loo and Hans Osterman Foundation (A.A., G.C.), Geriatric Diseases Foundation at Karolinska Institutet (A.A., G.C.), Norsk forskningsrådet 251290, 260786 (J.P.), JPCo-fuND/EU PROP-AD 643417/Horizon 2020/EC 643417 (H.B., J.P.).

## Author contributions

A.A., J.J. and H.B. conceived and designed the project. A.A., G.C. and H.B. developed the experimental protocol for A $\beta$  production and performed and analyzed the CD, <sup>15</sup>N-NMR and ThT kinetics experiments. G.C. performed the evolutionary relationship analysis. K.K., R.A., F.O. and K.J. performed and analyzed the <sup>19</sup>F-labelling protocol and <sup>19</sup>F-NMR experiments. M.L. performed and analyzed the mass spectrometry experiments. M.S., J.P., K.N., N.K., A.R. contributed additional data and analysis tools. H.B. supervised the project. A.A., A.R., J.J. and H.B. wrote the first draft of the manuscript. All authors provided comments on the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-019-57143-x>.

**Correspondence** and requests for materials should be addressed to H.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



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

© The Author(s) 2020