



HAL
open science

Aggregation and Amyloidogenicity of the Nuclear Coactivator Binding Domain of CREB-Binding Protein

Ana Maria Garcia, Christophe Giorgiutti, Youssef El Khoury, Valentin Bauer, Coralie Spiegelhalter, Emmanuelle Leize-wagner, Petra Hellwig, Noelle Potier, Vladimir Torbeev

► To cite this version:

Ana Maria Garcia, Christophe Giorgiutti, Youssef El Khoury, Valentin Bauer, Coralie Spiegelhalter, et al.. Aggregation and Amyloidogenicity of the Nuclear Coactivator Binding Domain of CREB-Binding Protein. *Chemistry - A European Journal*, 2020, 26 (44), pp.9889-9899. 10.1002/chem.202001847 . hal-02981154

HAL Id: hal-02981154

<https://hal.science/hal-02981154>

Submitted on 24 Nov 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.

Aggregation and amyloidogenicity of nuclear coactivator binding domain of CREB-binding protein

Ana Maria Garcia Fernandez,^[a] Christophe Giorgiutti,^[b] Youssef El Khoury,^[c] Valentin Bauer,^[a] Coralie Spiegelhalter,^[d] Emmanuelle Leize-Wagner,^[b] Petra Hellwig,^[c] Noelle Potier^[b] and Vladimir Torbeev*^[a]

^[a] Dr. A. M. Garcia Fernandez, V. Bauer, Dr. V. Torbeev
ISIS (Institut de Science et d'Ingénierie Supramoléculaires) and iCFC (International Center for Frontier Research in Chemistry), University of Strasbourg and CNRS - UMR 7006
8 allée Gaspard Monge, 67083 Strasbourg (France)
Email: torbeev@unistra.fr

^[b] C. Giorgiutti, Dr. E. Leize-Wagner, Dr. N. Potier
Laboratory of Mass-Spectrometry of Interactions and Systems, University of Strasbourg and CNRS - UMR 7140
1 rue Blaise Pascal, 67070 Strasbourg (France)

^[c] Dr. Y. El Khoury, Prof. P. Hellwig
Laboratory of Bioelectrochemistry and Spectroscopy, University of Strasbourg and CNRS - UMR 7140
1 rue Blaise Pascal, 67070 Strasbourg (France)

^[d] Dr. C. Spiegelhalter
IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire)
INSERM-U964, University of Strasbourg and CNRS-UMR 7104
1 rue Laurent Fries, 67404 Illkirch (France)

Abstract: Nuclear coactivator binding domain (NCBD) of transcriptional co-regulator CREB-binding protein (CBP) is a well-known example of conformationally malleable proteins that can bind to structurally unrelated protein targets and adopt distinct folds in the respective protein complexes. Here, we show that the folding landscape of NCBD contains an alternative pathway, which results in protein aggregation and self-assembly into amyloid fibers. The initial steps of such protein misfolding are driven by intermolecular interactions of its N-terminal α -helix bringing multiple NCBD molecules in contact. Then, these oligomers undergo slow but progressive interconversion into β -sheet containing aggregates. To reveal the concealed aggregation potential of NCBD we used chemically synthesized *D*-NCBD mirror-image form. The addition of *D*-NCBD promoted self-assembly into amyloid precipitates presumably due to formation of thermodynamically more stable racemic β -sheet structures. The unexpected aggregation of NCBD needs to be taken into consideration given multitude of protein-protein interactions and resulting biological functions mediated by CBP.

Abstract: The nuclear coactivator binding domain (NCBD) of transcriptional co-regulator CREB-binding protein (CBP) is an example of conformationally malleable proteins that can bind to structurally unrelated protein targets and adopt distinct folds in the respective protein complexes. Here, we show that the folding landscape of NCBD contains an alternative pathway that results in protein aggregation and self-assembly into amyloid fibers. The initial steps of such protein misfolding are driven by intermolecular interactions of its N-terminal α -helix bringing multiple NCBD molecules into

contact. These oligomers then undergo slow but progressive interconversion into β -sheet-containing aggregates. To reveal the concealed aggregation potential of NCBD we used a chemically synthesized mirror-image D-NCBD form. The addition of D-NCBD promoted self-assembly into amyloid precipitates presumably due to formation of thermodynamically more stable racemic β -sheet structures. The unexpected aggregation of NCBD needs to be taken into consideration given the multitude of protein–protein interactions and resulting biological functions mediated by CBP.

Introduction

Protein fibrous self-assembly is involved in a variety of biological functions including intracellular molecular transport, muscle contraction, scaffolding and connection of tissues, blood clot formation and many others.^[1,2] Apart from beneficial roles, protein self-assembly can be responsible for a loss of function. For instance, an aberrant protein folding (or protein misfolding) can result in protein aggregation into insoluble amyloid fibers that are structurally rich in β -sheet content and implicated in a number of human diseases, including neurodegenerative disorders.^[3] A detailed understanding of how soluble monomeric proteins misfold and self-assemble into amyloid fibers is important for developing new strategies for the treatment of these diseases.^[4–6]

Protein misfolding is a highly complex process.^[7,8] The misfolding and aggregation into amyloids can be initiated from a fully folded protein, for example, β 2-microglobulin^[9] or trans-

thyretin,^[10] or starting from intrinsically disordered sequences such as α -synuclein^[11] or amyloid- β (A β) peptide.^[12] Typically, conditions that lead to destabilization of fully folded forms are required to trigger misfolding and aggregation.^[13,14] Various transient intermediates and oligomeric species precede the assembly into amyloid fibers.^[15,16] In neurodegenerative diseases, the soluble oligomers were shown to possess higher toxicity than mature amyloid fibers.^[16,17]

Previously, it was suggested that amyloid formation is not restricted to only a particular group of proteins but instead it can be a general feature of almost any polypeptide under the appropriate conditions.^[18,19] This is supported by a high diversity of sequences and structural properties of proteins known to form amyloids, and suggests that the formation of amyloid fibers is a self-assembly phenomenon inherently dependent on a polypeptide backbone and to lesser extent on side-chains.^[20–22] Various conditions are associated with protein misfolding and aggregation into amyloids, such as denaturing conditions (acidic pH, presence of chaotropes and surfactants), presence of transition metals (e.g., Cu²⁺) or organic solvents.^[9,23,24] These factors may induce conformations in folded or intrinsically disordered proteins that are predisposed for aggregation.

Moreover, to aid experimental work several computational approaches are available to estimate the amyloid propensity of a given protein sequence. Short amyloidogenic sequence fragments (ca. 4–6 residues long) were proposed to act as 'aggregation hotspots' defining the conversion into amyloids of an entire protein or a protein domain.^[25,26] Based on this premise, such amyloid-prone sequences can be predicted by either empirical bioinformatic algorithms evaluating the physicochemical properties of amino acids and their possible contribution to amyloid formation or with the help of structure-based methods that use atomistic description of the polypeptide chains to provide a structural interpretation of aggregation into β -sheets.^[25,26] Whereas these approaches work well for short polypeptides, their accuracy is generally not sufficient to make a precise prediction for any arbitrary protein. The mechanisms of amyloid self-assembly are at least as complex as for protein folding, and experimental studies are needed to validate the propensities of a protein to aggregate or not.^[27]

A complementary approach that may help to explore conformational space and to identify conformational states that

[a] Dr. A. M. Garcia, V. Bauer, Dr. V. Torbeev
ISIS (Institut de Science et d'Ingénierie Supramoléculaires) and
iCFC (International Center for Frontier Research in Chemistry)
University of Strasbourg and CNRS—UMR 7006
8 allée Gaspard Monge, 67083 Strasbourg (France)
E-mail: torbeev@unistra.fr

[b] C. Giorgiutti, Dr. E. Leize-Wagner, Dr. N. Potier
Laboratory of Mass-Spectrometry of Interactions and Systems
University of Strasbourg and CNRS—UMR 7140
1 rue Blaise Pascal, 67070 Strasbourg (France)

[c] Dr. Y. El Khoury, Prof. P. Hellwig
Laboratory of Bioelectrochemistry and Spectroscopy
University of Strasbourg and CNRS—UMR 7140
1 rue Blaise Pascal, 67070 Strasbourg (France)

[d] C. Spiegelhalter
Imaging Center, IGBC (Institut de Génétique et de Biologie Moléculaire et
Cellulaire)
INSERM-U964, University of Strasbourg and CNRS—UMR 7104
1 rue Laurent Fries, 67404 Illkirch (France)

[e] Prof. P. Hellwig
Institute for Advanced Study
USIAS University of Strasbourg
5 allée du Général Rouvillois, 67083 Strasbourg (France)

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/chem.202001847>.

Part of a Special Collection to commemorate young and emerging scientists. To view the complete collection, visit: Young Chemists 2020.

have distinct aggregation properties and the predisposition to self-assemble into amyloids, is to modify the chirality of polypeptide.^[28] It was shown that the site-specific incorporation of D-amino acids into Alzheimer's disease-related A β -peptide can drastically alter its aggregation properties, allowing locations in the sequence that are critical for amyloidogenicity to be identified.^[29,30] Moreover, in many examples, adding mirror-image analogues (D-enantiomers) of amyloidogenic peptides or proteins to its native L-counterpart was reported to facilitate formation of amyloid fibers.^[31–35] Probably the most striking case was reported for A β (1–42) peptide, wherein the addition of the D-enantiomer to the corresponding L-peptide accelerated the formation of amyloids and, moreover, attenuated the cytotoxicity commonly observed for A β (1–42).^[35] In this case, it was shown that racemic amyloids acted as a sink for soluble and most toxic oligomers. Presumably, the accelerated self-assembly is driven by the formation of the 'rippled' β -sheet structures in which L- and D-polypeptides adopting extended β -strand conformations alternate.^[36,37] Provided that such L/D-amyloids can be more thermodynamically stable than enantiomerically pure L- or D-amyloids,^[32,33] the experiments in the presence of the D-enantiomer may serve as an additional test for amyloidogenicity of a given peptide/protein.

In our previous work, we investigated the mechanism of complex formation between activation domain 1 (AD1) of transcriptional co-activator ACTR (also known as NCoA3 or SRC-3 from p160 family) and the nuclear coactivator binding domain (NCBD) of CREB-binding protein (CBP or p300), which were prepared by chemical protein synthesis.^[38,39] When the chemical synthesis of NCBD was performed by native chemical ligation of two peptide segments, even in the presence of 6 M guanidine hydrochloride, we frequently observed precipitate, which we suspected could be composed of amyloids (Figure 1). Given that CBP/p300 are highly important proteins for regulating gene transcription^[40] and considering the diverse roles played by amyloid structures in different biological processes,^[41] we investigated the aggregation of NCBD in detail including in the presence of mirror-image D-NCBD. Here, we present experimental data confirming the self-assembly of

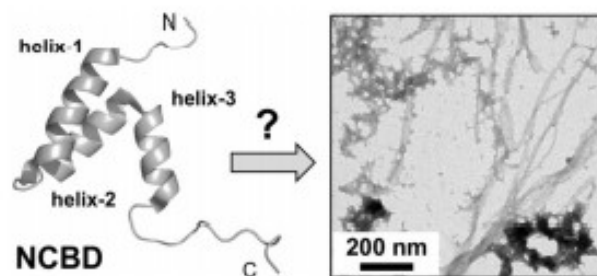


Figure 1. Largely α -helical molten globule NCBD (PDB ID: 2KKJ) aggregates and precipitates into fibrous structures that resemble amyloids (TEM image on the right). NCBD corresponds to the [2066–2112] fragment of the 270 kDa transcriptional co-regulator CREB-binding protein (CBP).^[40] CBP and its paralog p300 protein participate in multiple protein-protein interactions that fine-tune the transcription of various genes. The objective of our study is to delineate the molecular details of NCBD aggregation mechanism. Three α -helical fragments as well as the N and C termini are labeled.

NCBD into amyloid-like fibers under different experimental conditions. By using various tools, we derived a possible molecular mechanism for the earlier steps of NCBD aggregation into oligomers and their conversion into amyloids.

Results

Solubility and aggregation of NCBD

NCBD was previously shown to exhibit properties of a molten globule, possessing elements of secondary structure but lacking a well-defined tertiary structure typical for folded proteins.^[42] In its isolated form, NCBD presents a compact structure with a high content of α -helical conformation, while showing structural adaptability by forming distinct well-defined structures in complexes with other specific protein binding partners. This relatively small fragment of CBP is responsible for interactions with other relevant proteins such as transcriptional co-activators from the p160 family^[43] or interferon regulatory proteins,^[44] through the formation of stable and well-defined tertiary structures at the interaction interface.

The D-NCBD enantiomer was prepared by chemical synthesis analogously to L-NCBD,^[39] using commercially available D-amino acid building blocks (see the Experimental Section). D-NCBD folds into mirror-image molten globule, as inferred from the reciprocal circular dichroism (CD) spectra (Figure S1, see the Supporting Information). Our motivation to use D-NCBD along with L-NCBD was to bias the self-assembly of L-NCBD, so that if L-NCBD is predisposed to aggregate and self-assemble into amyloids, the addition of D-enantiomer will facilitate this process and help explore the aggregation landscape of L-NCBD at various conditions.

Samples containing L- and D-NCBD or mixtures thereof at different ratios and at two different pH values (acidic at pH 2.5 and neutral at pH 7.0) were incubated at 37 °C. The soluble fraction of NCBD in each sample was analyzed by HPLC as a function of the L/D ratio and time of incubation after ultrafiltration using centrifugal devices with membranes of 0.2 μ m porosity (Figure 2a and b). The solubility as a function of the L/D ratio suggests a distinct aggregation behavior for the two different conditions tested. At pH 2.5, analysis of L/D mixtures showed that lower amounts of soluble protein passed through

Vladimir Torbeev received his Ph.D. degree in Chemistry from the University of Chicago under direction of Prof. Stephen B. H. Kent for the work dedicated to total chemical synthesis and biophysical studies of HIV-1 protease. Then, he performed postdoctoral studies in the group of Prof. Donald Hilvert at ETH Zurich working on molecular mechanism of misfolding and aggregation of β 2-microglobulin. Since March 2014, he is a group leader at the Institut de Science et d'Ingénierie Supramoléculaires and an Assistant Professor at the University of Strasbourg.



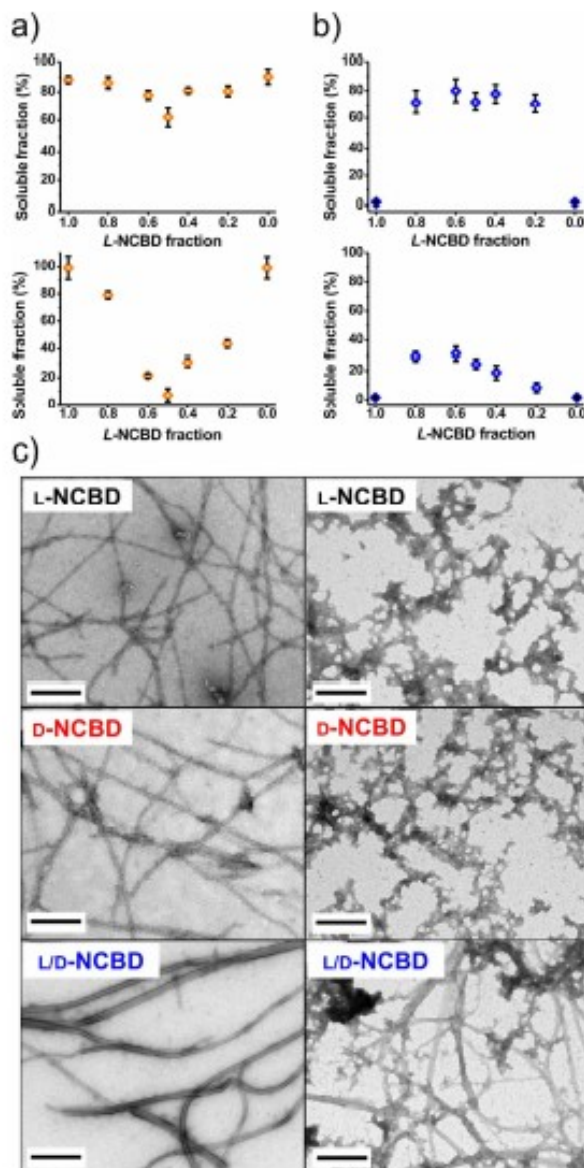


Figure 2. a,b) Solubility curves of NCBD enantiomers incubated at 37 °C at different L-NCBD/D-NCBD ratios in buffers at a) pH 2.5 and b) pH 7.0. Top panels correspond to one week of incubation and bottom panels to two weeks of incubation. The plotted data are the average of two independent experiments ($n = 2$) where samples have been analyzed in triplicate. Standard deviations are represented by black bars. c) TEM images of NCBD samples after two weeks of incubation at pH 2.5 (left column) and at pH 7.0 (right column). Scale bars = 200 nm.

0.2 μm membrane than the single enantiomers, whereas at neutral pH, single enantiomers formed aggregates that are trapped by the membrane more readily than the different L/D-NCBD mixtures. These results confirm the presence of chiral recognition effects between L- and D-NCBD. The differences in aggregation for mixtures of L- and D-NCBD can be caused by distinct conformations of polypeptides at different pH values

and, thus, distinct modes of interaction. The solubility diagrams in Figure 2a and b resemble the solubility diagrams observed for crystals formed by enantiomers of small molecules, where the situation at pH 2.5 corresponds to the solubility behavior of racemates that behave as racemic compounds and, at neutral pH, solubility curves are reminiscent of those of conglomerates.^[45–47] In the former case, heterochiral interactions are favored and result in mixed racemic crystals from the mixture of enantiomers, whereas for the latter case the preference for homochiral assemblies enables the formation of enantiopure crystals from the L/D-mixture.

A pH-dependent behavior is also reflected in the morphology of the aggregates in the samples, as shown by transmission electron microscopy (TEM) (Figure 2c and Figures S2–S9, see the Supporting Information). At acidic pH, the L- and D-NCBD single enantiomers assembled to form fibers with widths of 12.1 ± 1.3 nm, a typical width reported for other amyloids.^[48] When combinations of L- and D-enantiomers were incubated, fibers with larger widths of 25.9 ± 3.0 nm and composed of bundles of several thinner fibers were frequently observed. Aggregation at neutral pH yielded initially amorphous precipitates, but, upon two weeks of incubation, fibrous structures could also be detected. Morphologically, these fibers were thinner than the longer and thicker fibers observed at pH 2.5. Overall, fibrous self-assembly is promoted by the addition of the non-natural D-NCBD counterpart and, importantly, aggregation of NCBD generates morphologically distinct species depending on the pH of the media.

Monitoring structural changes during NCBD aggregation by FTIR spectroscopy

To follow the secondary structure adopted by NCBD aggregates over the time, attenuated total reflection Fourier transform infrared (ATR-FTIR) spectra were recorded for NCBD samples upon incubation at 37 °C at different timepoints (one, two and four weeks) (Figure 3). This spectroscopic technique is very reliable to observe the structural changes that accompany protein aggregation into β -sheets characterized by amide I stretching frequencies at around 1630 cm^{-1} .^[49] Moreover, it is suitable for the analysis of heterogeneous samples and requires small amounts of material.^[50] After one week of incubation, samples containing single NCBD enantiomers (L- or D-) showed a relatively sharp maximum in the amide I region (ca. $1650\text{--}1654\text{ cm}^{-1}$) at both pH 2.5 and 7.0, typically assigned to α -helical structures, which is not surprising given high helical content observed for NCBD molten globule.^[42] Notably, the single enantiomers at acidic pH showed a wider peak that was shifted towards the β -sheets region (ca. 1635 cm^{-1}) in comparison with those at pH 7.0. The spectra of the racemic mixtures at both pH values were dominated by the α -helix characteristic peak, with a shoulder in the β -sheet region (ca. 1626 cm^{-1}), more pronounced at neutral pH.

Over time, the spectra of samples at distinct pH values evolved differently. At pH 2.5, a band in the β -sheet region (ca. 1630 cm^{-1}) in the spectra of all three samples (L-, D- or L/D-NCBD) increased in intensity, being more intense than the α -

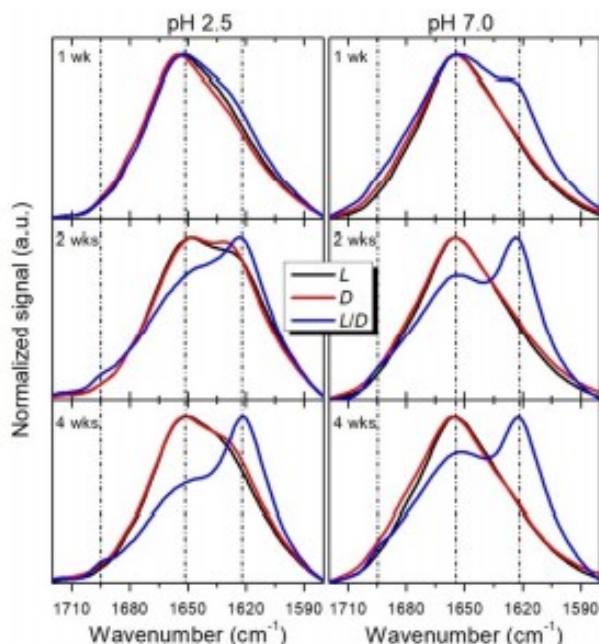


Figure 3. ATR-FTIR spectra of L-NCBD, D-NCBD, and L/D-NCBD racemic mixture in phosphate buffer at pH 2.5 (left) and pH 7.0 (right) after incubation at 37 °C for one, two, and four weeks. The vertical dotted lines highlight the frequencies of the maxima at four weeks.

helical signal for the L/D-mixture after two and four weeks of incubation, which reflects a progressive transition from α -helical to β -sheet structures. At neutral pH, the spectra of pure enantiomers were dominated by the α -helical signal (ca. 1655 cm^{-1}) even after four weeks, whereas the spectra of the racemic mixture showed a major peak centered at ca. 1623 cm^{-1} corresponding to the β -sheet structure, concomitant with the presence of the α -helical structure. These data confirm the presence of β -sheet secondary structure of fibrous aggregates typical of amyloids and again point to the chiral effects in NCBD aggregation. The D-NCBD enantiomer acts as a promoter of β -sheet self-assembly and is essential for facile formation of amyloid fibers, especially at neutral pH.

Insights into the early steps of NCBD aggregation by native mass spectrometry

To better understand the initial stages of self-association and aggregation of NCBD, we used electrospray ionization mass spectrometry (ESI-MS) under native-like conditions. This technique enables the association/dissociation of intact proteins to be monitored and provides structural insights into the corresponding protein complexes by determining their stoichiometry, topology and dynamics.^[51–53] In our case, the L- and D-NCBD enantiomers as well as the racemic mixture were separately analyzed at different concentrations (10, 150 and 300 μM). The native ESI mass spectra provide a charge-state distribution of the species within a mass-over-charge (m/z) range. By deconvolution of the charge states, the molecular

masses of the complexes were determined, and their resulting stoichiometry deduced. The results showed the formation of clusters by NCBD (monomers, dimers, trimers and even tetramers) at any concentration (Figure 4, Figure S10). Determination of stoichiometry was achieved thanks to the resolution of isotopic profiles of all ions, which led to unambiguous assignment of the charge states. All samples present the same pattern of aggregation, with a slight growth in the intensity of multimeric species upon increasing concentration but with no significant changes of coexisting stoichiometries. No multimeric interactions could be detected under denaturing conditions (50% acetonitrile, 0.1% formic acid), which strongly supports the conclusion that the spectra reflect the different populations of NCBD oligomers present in solution rather than non-specific aggregation in the gas phase.

The L- and D-NCBD were synthesized as peptide- α -carboxylic acid and peptide- α -amide, respectively, with only ca. 1 Da mass difference (see Tables S1 and S2). Given the almost identical molecular weight of L- and D-enantiomers of NCBD, native MS is not capable of clearly distinguishing between homo- and heterochiral interactions within the L/D-NCBD racemic mixture and cannot give information about the chiral composition of

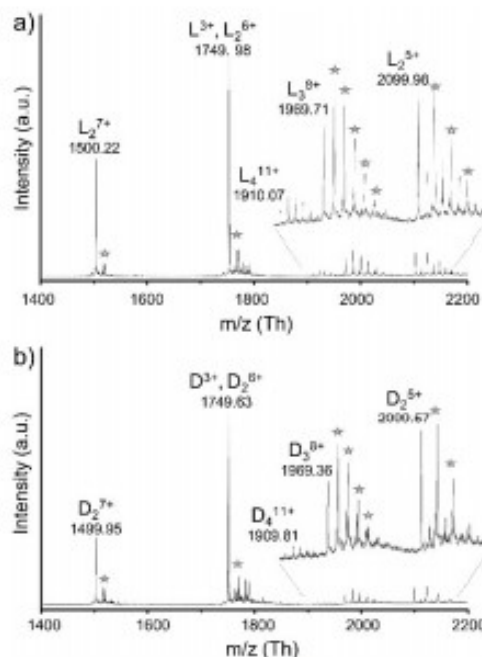


Figure 4. Native ESI mass spectra of a) L-NCBD and b) D-NCBD in 100 mM AcONH₄, pH 6.8. Resolution of the isotopic profiles allows the charge states of all multiply charged ions to be determined and thus to unambiguously assign the molecular masses of all species. Monomeric species (L and D) are detected as triply charged ions at m/z 1749.98 and 1749.63, respectively. Uneven-charged states confirm the presence of dimeric species at m/z 1500.22 and 2099.98 for L₂ and m/z 1499.95 and 2099.67 for D₂. Ions at m/z 1969.71 and 1969.36 are observed as 8⁺-charged state corresponding to trimeric species (L₃ and D₃), whereas 11⁺-charged ions at m/z 1910.07 and 1909.81 yield a molecular mass corresponding to tetrameric species (L₄ and D₄). The star symbols indicate trifluoroacetic acetate (TFA) adducts of proteins.

the aggregates. We envisaged that the addition of two more amino acids at the C terminus of the L-NCBD sequence (i.e., Gln-Pro native residues from the corresponding full-length CBP sequence) will increase the molecular weight of L-NCBD by ca. 200 Da and, therefore, can be used as a tag to efficiently distinguish L-NCBD from D-NCBD by mass. In this way, native MS can be used to identify differences in the aggregation patterns between pure enantiomers and quasi-racemic mixture. We chemically synthesized this new L-NCBD analogue (abbreviated here as L⁺-NCBD) (see the Experimental Section). By native MS analysis of an equimolar mixture of L⁺- and D-NCBD in addition to the expected L⁺- and D-NCBD dimers, a third species at 10721.9 ± 1.2 Da was observed and unambiguously confirmed the presence of a L⁺/D heterocomplex, thus demonstrating that heterochiral interactions occur in solution. No species revealing the presence of homochiral NCBD trimeric or tetrameric complexes were observed, suggesting that heterochiral oligomerization is favored over homochiral multimerization at these conditions (Figure 5). Given that each species present in solution displays a different molecular mass, all oligomeric states could be followed individually. Native MS thus appears to be a powerful technique for monitoring the concomitant formation of the distinct aggregated states of NCBD.

Identifying the aggregation-prone segment in NCBD sequence

The structure of NCBD molten globule contains three helical fragments (Figure 1).^[42] To identify which sequence fragment of NCBD is responsible for triggering aggregation, we prepared the three polypeptides corresponding to the three different helices (**helix-1**, **helix-2** and **helix-3**) observed in the structure of NCBD (PDB ID: 2KKJ).^[42] When each helix was analyzed sepa-

rately by native MS, only **helix-1** showed a pronounced aggregation, revealing clusters ranging from dimers to undecamers; **helix-2** was mostly observed as a monomer and **helix-3** demonstrated moderate aggregation propensity (Figure 6). Due to the instrument effects, the relative abundance of free peptide versus peptide clusters could not be directly extracted from the MS spectra. Indeed, all species display distinct *m/z* ratios and adopt different conformations, two parameters that are known to affect the ion intensities. Nevertheless, qualitatively, the MS spectra of the three helical segments acquired under the same ESI-MS parameters illustrate the distinct aggregation properties of these peptides. Inspection of samples of the corresponding peptides incubated at pH 2.5 and 7.0 with TEM revealed similar trend of aggregation propensities; i.e., the **helix-1** segment is particularly prone to aggregate (Figure S11).

The mixtures of full-length NCBD with every helix were also studied (Figures S12–S14). According to native MS, **helix-1** again revealed its 'sticky' behavior by forming a variety of mixed clusters in high abundance under equimolar incubation conditions with NCBD; this was not the case for **helix-2** and **helix-3**, for which the major observed species were individual NCBD and **helix-2** or **helix-3**, respectively. The association pattern of L-NCBD or D-NCBD with the three helices showed no strong chiral preferences given that the spectra for the two enantiomers and the corresponding helices were similar. The measured molecular masses allowed the stoichiometries of the various complexes coexisting in solution to be assigned and it appeared that clusters formed by the interaction of one NCBD with several units of helices (L/H_n) were favored over clusters, implying dimerization of NCBD (L₂/H_n). No higher stoichiometries of the aggregated NCBD (trimers or tetramers) were detected in the presence of **helix-1**, which clearly shows that homologous NCBD multimers were displaced in favor of heterolu-

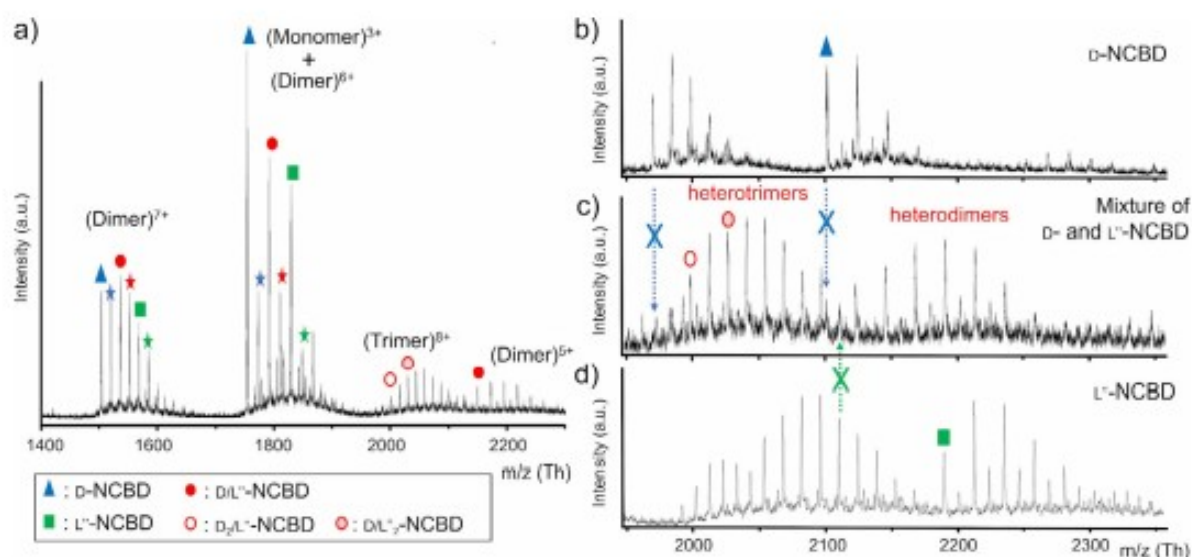


Figure 5. Evidence for self-assembly of heterochiral L/D-NCBD multimers by native ESI mass spectrometry. a) Mass spectrum of an equimolar mixture of D- and two-extra amino acids modified L-NCBD (L⁺). b–d) Expansion of the 2000–2350 *m/z* region of the spectra and a comparison between b) D-NCBD sample, c) D/L⁺-NCBD mixture, and d) L⁺-NCBD. Homochiral ions at *m/z* 1969.36 (D-NCBD homotrimer), *m/z* 2099.57 (D-NCBD homodimer) or *m/z* 2110.87 (L⁺-NCBD homotrimer) are not observed in the mass spectrum of D/L⁺-NCBD mixture (in c). The star symbols indicate TFA adducts of proteins.

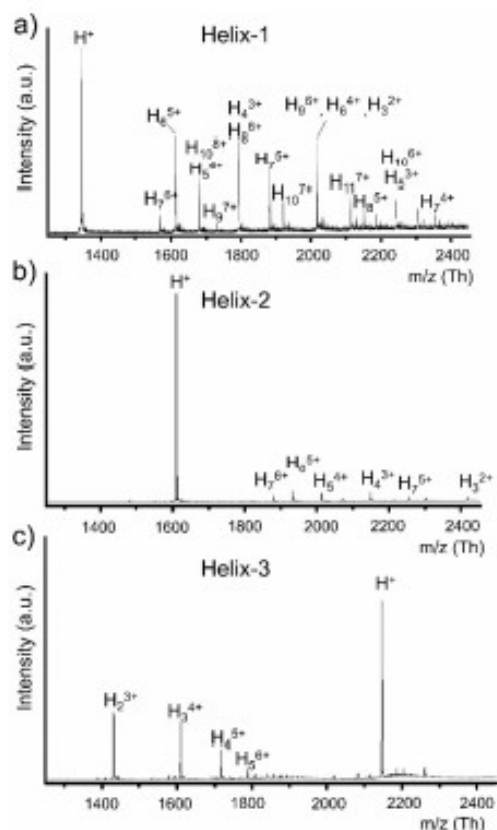


Figure 6. Characterization of the aggregation behavior of a) **helix-1**, b) **helix-2**, c) **helix-3** by native ESI mass spectrometry in 100 mM AcONH₄. Ions showing coexisting stoichiometries from monomeric (H) up to 11-mer (H₁₁) are detected for **helix-1** whereas less-pronounced aggregation is observed with **helix-2** and **helix-3**.

gous NCBD/**helix-1** multimers (Figure S12). These data suggest that **helix-1** plays a critical role in the aggregation of the NCBD sequence, promoting the self-assembly between several NCBD units.

The computational predictors point out distinct parts in the sequence of NCBD that are prone to aggregation (Figure 7). Two applied consensus methods, MetAmyl^[56] and AmylPred2,^[57] yielded consistent results showing the two probable ‘aggregation hot spots’ localized in **helix-2** and **helix-3** of the NCBD sequence (QVLN_{IL} and LMAAF_I). Indeed, inspecting samples of **helix-2** and **helix-3** after four weeks incubation revealed amyloid-like fibers, which is consistent with the computational prediction (see Figure S11). The apparent discrepancy of the aggregation-prone fragments in NCBD based on the experimental native MS measurements and the computational approaches can be explained by the fact that these methods probe different aspects of the aggregation. Native MS provides insights into the initial stages of NCBD multimerization via helix-helix interactions, whereas computational predictors suggest which sequences are likely to be found in the core of the amyloid fiber after α -helix-to- β -sheet transition. Thus, the two methods are complementary and allow a mechanistic model

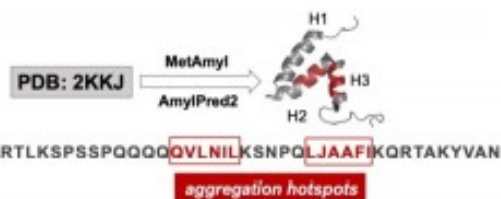


Figure 7. Results of computational prediction of amyloidogenicity of NCBD by using the consensus methods MetAmyl and AmylPred2. Two aggregation hotspots are highlighted in red and located in **helix-2** (H2) and **helix-3** (H3) fragments. Note: J – norleucine (Nle), which is a nearly isosteric replacement of methionine in the native sequence introduced to avoid oxidation of the thioether moiety.

to be proposed for sequential steps in NCBD aggregation wherein the initial aggregation is driven by promiscuous contacts of N-terminal helix of NCBD, followed by a structural conversion involving segments located in different parts of the sequence.

Discussion

As discussed previously, amyloid formation is not restricted to a particular group of proteins but can occur for highly diverse sequences under suitable conditions.^[18–22] In this article, for the first time, we characterized the misfolding and aggregation of NCBD into amyloid-like structures. We tested different pH values and used D-NCBD to modulate and explore the protein aggregation process. Indeed, the solubility and morphology of the aggregates is affected by the conditions of incubation (e.g., pH and fraction of D-NCBD in the mixture) (Figure 2, Figure S2–S9). According to FTIR data, NCBD experiences α -helix-to- β -sheet conversion over time, particularly at acidic pH (Figure 3). Notably, the L/D-racemic mixture presents a marked preference to adopt a β -sheet structure both at acidic and neutral pH. Interestingly, although it is a different self-assembly process, the formation of protein crystals from the racemic mixtures was also shown to occur more efficiently than for individual enantiomers.^[38]

Native MS data suggest that NCBD has a high propensity to form aggregates even at low concentrations (10 μ M) (Figure 4, Figure S10). The results obtained for the quasi-racemate of L^{*}- and D-NCBD as an approach to distinguish L- and D-species in the mass spectra suggest that the enantiomeric proteins behave similarly and that the initial association of homo- and heterochiral multimers occurs rapidly in solution (Figure 5). In addition, the N-terminal helix in NCBD was identified as being a trigger for the early stages of protein aggregation (Figure 6). This helix contains the leucine-rich LXXLLXXL sequence motif known to participate in various hydrophobic interactions mediating the formation of relevant biological complexes.^[59,60]

Taking into account all the results, we hypothesize that the aggregation of NCBD starts from its self-assembly into α -helical oligomers mediated through hydrophobic interactions of the ‘sticky’ N-terminal helical segment, as suggested by a high tendency of amphiphilic **helix-1** to multimerize. Modeling using CCRBuilder 2.0 software^[61] showed that the formation of

2-, 3- and 4-helix bundles by **helix-1** segment can easily be achieved (Figure S15) and this is likely an explanation of the results of native MS measurements involving **helix-1** peptide.

On the other hand, NCBD is known to adopt different structures upon binding to different binding partners. Inspection of the available structures allowed a feasible molecular mechanism of NCBD oligomerization to be deduced (Figure 8). We compared two conformational states of NCBD extracted from crystal structures of complexes with ACTR (PDB ID: 6SQC)^[39] and IRF3 (PDB ID: 1ZOQ).^[44] In these two structures the biggest difference is the orientation of N-terminal helix (**helix-1**): in 6SQC it forms tight contacts with the middle helical fragment (**helix-2**), whereas in 1ZOQ structure it is located in the groove and interacts with both middle (**helix-2**) and C-terminal (**helix-3**) helices. Apparently, the position of the middle helix in 1ZOQ also coincides with fragments of other proteins that interact with NCBD, such as helical fragment of ACTR (as in 6SQC structure) or fragment of transactivation domain of p53 (PDB ID: 2L14). Notably, the structures of the latter two fragments are very different.

NCBD being a molten globule in free state, implies conformational isomerism and heterogeneity. Thus, we envisaged that a minimal change of NCBD-6SQC conformation such as ca. 180° rotation around the axis of **helix-1** to expose its hydrophobic side to the solvent will make it predisposed to bind to another molecule of NCBD through interactions of solvent-exposed **helix-1** and a promiscuous binding groove of a second NCBD molecule formed by **helix-2-helix-3** (Figure 8). Furthermore, this conformational modification and docking can be repeated multiple times, resulting in high-order NCBD oligomers.

Over time, these α -helical oligomers undergo isomerization into more stable β -sheet structures and convert into amyloid fibers. Indeed, amyloid formation through α -helical intermediates is described to happen for many proteins, including tau

protein^[62] and α -synuclein.^[63] Thus, the results of our work can be qualitatively described by an 'association-conformational conversion' model.^[64] According to this model, amyloidogenic proteins first self-associate into soluble oligomers through hydrophobic interactions and then undergo structural conversion into β -sheet-rich amyloids. Mechanistically, the initial helical assemblies increase the local concentration necessary for the subsequent formation of intermolecular β -sheets.^[65] The importance of high local concentration of proteins can be further illustrated by "droplets" formed due to liquid-liquid phase separation of various disordered proteins.^[66] Conformational rearrangements within these compartments often lead to the formation of β -sheet "lumps" that propagate to ultimately yield more stable fibrillar aggregates with β -sheet structure.^[67,68]

Conclusions

We discovered that the NCBD sequence is aggregation-prone and amyloidogenic. The predisposition of the N-terminal helix of NCBD molten globule protein to mediate intermolecular self-assembly results in helical aggregates. They evolve towards more stable β -sheet structures that adopt fibrous amyloid-like morphologies particularly at low pH. At physiologically relevant neutral pH, this effect can be highlighted in the presence of the D-NCBD enantiomer. Thus, the use of D-NCBD helped to reveal the existence of an alternative pathway on the folding landscape of NCBD that is not obvious for enantiopure L-NCBD. These results are important given the multiple studies reported previously for this protein domain that served as a classical example of molten globule that forms distinct folded structures upon complexation with different protein binders.^[69] Further research is needed to explore the potential role of NCBD misfolding and aggregation in biological processes mediated by CBP.

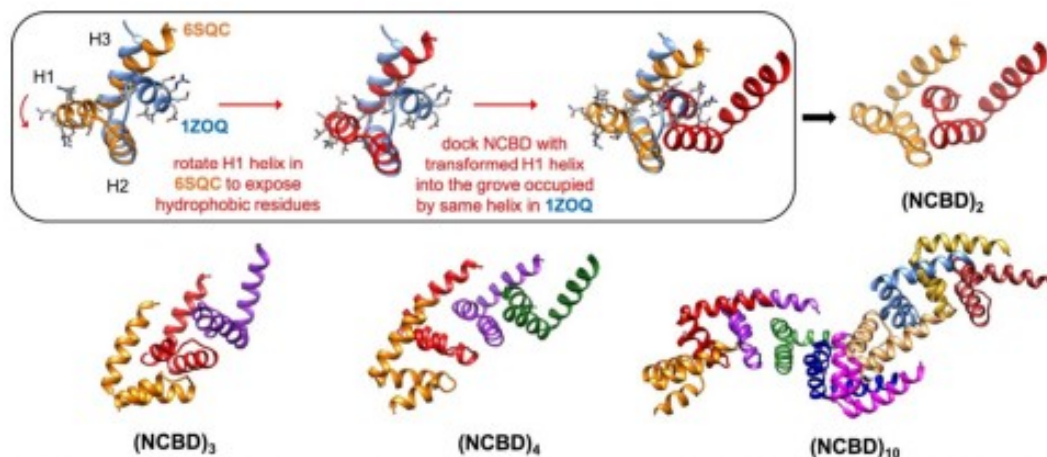


Figure 8. Model of the oligomerization of NCBD. Two distinct conformations of NCBD were extracted from the crystal structures of NCBD with ACTR (in orange, PDB ID: 6SQC) and IRF3 (in blue, PDB ID:1ZOQ). Structural alignment was performed by using region H2-H3 of the 6SQC structure as a reference with MatchMaker tool of USCF-Chimera software.^[70] Then, helix H1 was rotated to expose the hydrophobic side to the solvent making it aggregation prone. This new conformation was overlaid onto H1 helix of NCBD from 1ZOQ structure that is in the groove formed by H2 and H3 helices. Such a transformation effectively results in a (NCBD)₂ dimer. This operation can be repeated again by docking a copy of aggregation-prone NCBD conformation to the free binding groove exposed in (NCBD)₂, leading to a trimer and subsequently to a tetramer and higher order oligomers (the largest shown is a decamer).

Experimental Section

Materials: All solvents, chemicals, and reagents were purchased from commercial sources and used without further purification. Fmoc-amino acids and resins for solid-phase peptide synthesis, the coupling reagents *N,N'*-diisopropylcarbodiimide (DIC) and ethyl cyano(hydroxyimino)acetate (Oxyma pure), diisopropylethylamine, piperidine, triisopropylsilane (TIPS) and 1,4-dithiothreitol (DTT) were purchased from Iris Biotech (Marktredwitz, Germany) or Sigma Aldrich (St. Louis, Missouri, USA). Trifluoroacetic acid (Biograde) (TFA) was purchased from Halocarbon (Peachtree Corners, Georgia, USA). All other chemicals were purchased from Sigma-Aldrich. Nanosep MF centrifugal devices with Bio-Inert membrane of 0.2 μm porosity were provided by Pall Corporation (Ann Arbor, Michigan, USA).

Peptide synthesis: The L-NCBD and D-NCBD peptide fragments corresponding to the [2066–2112] fragment of CBP (CREB-binding protein, UniProtKB-P45481 (CBP MOUSE)) were synthesized by using a Liberty Blue microwave-assisted automatic synthesizer (CEM Corporation, USA), following a Fmoc/tBu-protecting group strategy, on either 0.1 or 0.25 mmol, as previously described.^[26] For the synthesis of L-NCBD and the different L-NCBD fragments (helices 1–3), the preloaded Fmoc-Asn(Trt) and Fmoc-Ser(tBu) Wang resins (0.7 mmol g⁻¹ and 0.65 mmol g⁻¹ loading, respectively) were used, while the L¹-NCBD and D-NCBD polypeptides were synthesized using Rink-Amide ChemMatrix resin (0.45 mmol g⁻¹ loading). Peptides were cleaved and fully deprotected by treatment with a mixture of TFA/TIPS/H₂O/DTT (92.5/2.5/2.5/2.5, v/v) for 2 hours (1 mL of mixture for 0.1 g of resin). Afterwards, the mixture was added onto cold diethyl ether (-20 °C) for precipitation, centrifuged, and the precipitates were dissolved in H₂O/CH₃CN (1:1, v/v) containing 0.1% TFA and lyophilized. Pure peptide fractions (> 98%) were obtained after purification by HPLC (Tables S1, S2 and Figures S16, S17).

Measurement of protein concentration: After synthesis and purification, proteins were obtained as TFA salts. To calculate concentrations of the samples used in various experiments (CD, TEM, solubility curves), the following procedure was applied: the nitrogen content of a sample of L-NCBD was determined by elemental analysis. This allowed to exactly calculate the mass portion of protein in the purified product (ca. 80%). This sample was used to calibrate the analytical HPLC system. The derived calibration was used to determine the concentrations of other NCBD samples.

Preparation of aggregated samples: L-NCBD and D-NCBD samples were dissolved at peptide concentrations of 0.5 mM (determined by analytical HPLC). The prepared samples were incubated at 37 °C, at different L-NCBD/D-NCBD ratios and under two different conditions: 200 mM sodium citrate buffer pH 2.5 (or 200 mM sodium phosphate buffer pH 2.5) and 200 mM sodium phosphate buffer pH 7.0 (with Na₂S₂O₅, 0.05%, w/v). The samples were prepared at L-NCBD/D-NCBD ratios of 100:0, 80:20, 60:40, 50:50, 40:60, 20:80, and 0:100.

Analysis of the soluble peptides by HPLC: The soluble fraction of NCBD for the samples prepared at different L-NCBD/D-NCBD ratios was analyzed by HPLC after incubation for one and two weeks. The incubated samples were centrifuged at 10,000 rpm for 10 min using Nanosep MF centrifugal devices with membrane of 0.2 μm porosity (Pall Corporation). Then, 30 μL of the samples were diluted with 120 μL of milliQ water (final volume = 150 μL) and analyzed by analytical HPLC (in triplicate). The values for the soluble fraction of NCBD for each sample were obtained by integration of the HPLC traces using a calibration curve at OD 220 nm. Incubation and analysis were repeated twice. The percentages of soluble frac-

tion for each sample were calculated based on the concentration of NCBD obtained by HPLC before incubation was started.

Computational prediction of amyloid propensity: The NCBD sequence (PDB file: 2KKJ)^[42] was screened with two different consensus methods available online, MetAmyl^[56] and AmylPred2.^[57] The first program computes the propensity of amyloid fibrillogenesis combining the following individual predictors: Pafig, SALSA, Waltz and FoldAmyloid. The score, $p(x)$, represents the probability for a fragment to form amyloid fibers, and it is calculated using a logistic regression model. AmylPred2 gives as primary output the consensus prediction that represents the overlap of at least $n/2$ of the different n methods selected by the user. The tool offers a maximum of 11 algorithms available: Aggrescan, AmyloidMutants, Amyloidogenic Pattern, Average Packing Density, Beta-strand contiguity, Hexapeptide Conformational Energy, NetCSSP, Pafig, SecStr, Tango and Waltz.

Mass spectrometry: Liquid chromatography–mass spectrometry (LC-MS) spectra of peptides were recorded with a Thermo Scientific Accela UHPLC (Hypersil II GOLD column, 50×2.1 mm², 1.9 μm) integrated with a Thermo Scientific LCQ XL linear ion trap mass spectrometer. The mass spectra were recorded in the positive ionization mode. Obtained raw data were deconvoluted using MagTran (ver. 1.03) software.

High-performance liquid chromatography (HPLC): Analytical reversed-phase HPLC of all peptides were performed with a Dionex Ultimate 3000 (Thermo Fisher) equipped with a UV detector, a column heater set to 40 °C and an autosampler. Analyzes were performed with a Reprosil-XR 120 C18 column (Dr. Maisch GmbH, particle size 3 μm , pore size 120 Å, dimensions 50×2.1 mm²) or a Kinetex XB C18 column (Phenomenex, particle size 2.6 μm , pore size 100 Å, dimensions 50×2.1 mm²) at a flow rate of 1 mL min⁻¹ and a gradient of 2–50% of eluent B (0.08% TFA in CH₃CN) in eluent A (0.1% TFA in H₂O) within 4 min. Preparative reversed-phase HPLC were performed with a Shimadzu instrument equipped with a CBM-20A communication module, a SPD-M20A UV detector, a SIL-10AP autosampler and a FRC-10A fraction collector. L-NCBD, D-NCBD and L¹-NCBD peptides were purified with a Jupiter C4 column (Phenomenex, particle size 10 μm , pore size 300 Å, dimensions 250×21.2 mm²) at a flow rate of 10 mL min⁻¹ and a gradient of 20–40% of eluent B (0.08% TFA in CH₃CN) in eluent A (0.1% TFA in H₂O) within 90 min. After purification, pure fractions were identified by analytical HPLC and LC/MS, combined and lyophilized. L-NCBD fragments (helix-1, -2, and -3) were purified with a Kinetex C4 column (Phenomenex, particle size 5 μm , pore size 100 Å, dimensions 100×21.00 mm²) at a flow rate of 10 mL min⁻¹ and a gradient of 5–30% of eluent B (0.08% TFA in CH₃CN) in eluent A (0.1% TFA in H₂O) within 40 min. After purification, pure fractions were identified by analytical HPLC and LC/MS and lyophilized.

Circular dichroism (CD): A 1 mm quartz cell was used with a Jasco J-810 spectropolarimeter. All samples were measured in quartz cuvettes with a thickness of 1 mm and a volume of 350 μL . The following parameters were set for all scans performed: scanning speed 100 nm min⁻¹ over a range of wavelengths from 185 to 280 nm, with 1 s integrations, bandwidth of 1 nm, data pitch 0.1 nm, temperature 25 °C. For measurements, protein samples were prepared at 50 μM concentration in the corresponding buffers. Freshly prepared samples of L- and D-NCBD were used to record CD data depicted in Figure S1. Every CD curve was obtained by averaging five scans and subtracting the background signal. The obtained ellipticities were transformed to mean residue molar ellipticities.

TEM analysis: TEM images were acquired with a Philips CM12 electron microscope, operating at 80 keV. TEM grids (carbon-coated 300 mesh copper EM grids) were first exposed to a UV-ozone cleaner for 5 minutes. The grids were then placed coated-side-down for 60 s onto sample drops (15 μ L). The grids were then retrieved, washed with deionized water and stained with 2% (w/v) uranyl acetate for 20 s. Grids were air-dried and kept under vacuum before analysis. Average thickness of fibers was determined by considering at least 50 individual measurements using ImageJ software.

FTIR spectroscopy: ATR-FTIR experiments were performed with a Bruker Vertex 70 FTIR spectrometer equipped with a Harrick-Diamond ATR. Residual TFA was removed from the samples by 10 mM HCl treatment and lyophilization (3 cycles). Each sample was left to dry under air. For each sample, ten spectra of 128 scans each (4000 to 700 cm^{-1} , resolution: 4 cm^{-1}) were recorded and averaged. All spectra were plotted using OriginPro 8.5 software.

Native mass spectrometry: All proteins were solubilized in 100 mM ammonium acetate (AcONH_4), pH 6.8 and analyzed at 150 μ M unless specified. MS data were acquired in the positive ionization mode with an electrospray time-of-flight mass spectrometer (MicroTOF-QII, Bruker Daltonics, Germany). Denaturing conditions measurements were conducted by injecting protein separately at a concentration of 10 μ M in $\text{H}_2\text{O}/\text{CH}_3\text{CN}$ (50:50) with 1% formic acid to check their sequence integrity and homogeneity. L- and D-NCBD synthesized as peptide- ^{13}C COOH and peptide- ^{13}C (O)NH $_2$, respectively, were detected at 5246.9 ± 0.5 Da and 5245.9 ± 0.5 Da, which is in good agreement with the corresponding theoretical monoisotopic molecular masses of the monomers (5246.93 Da and 5245.96 Da). Native mass spectrometry analyses of the noncovalent complexes were carried out in 100 mM AcONH_4 , pH 6.8. Proteins or complexes (150 μ M) were incubated for 5 min at 37 $^\circ\text{C}$ and were infused at 4 $\mu\text{L}\cdot\text{min}^{-1}$ through a KDS100 syringe pump (KD Scientific, Holliston, MA, USA). Electrospray source and transfer parameters were set to optimize the desolvation process while preventing dissociation of the noncovalent species. In particular, the pre-pulse time, which controls the delay between transfer time and TOF pulser on for injection into the TOF analyzer was set at 36 μ s. The collision energy in source (ISCID) was also carefully controlled and mass spectra were recorded as a function of the ISCID value from 0 to 100 eV. All spectra are shown at 40 eV where it was checked that no gas-phase dissociation occurred. Finally, a manual interpretation of the MS spectra was carried out to assign all charged states of the ions and infer the molecular masses of coexisting populations of species.

Acknowledgements

C.G. would like to acknowledge the "Fondation Recherche Chimie" (FRC foundation) for his PhD thesis funding. This work has been funded by the European Research Council (ERC-2016-StG, grant number 715062-1 IliChemSynPro) and the LabEx CSC (ANR-10-LABX-0026_CSC). Electron microscopy imaging was supported by the Imaging Center of IGBMC (<https://ici.igbmc.fr>).

Conflict of interest

The authors declare no conflict of interest.

Keywords: amyloid beta-peptides · conformation analysis · protein folding · self-assembly · transcription coactivators

- [1] T. Scheibel, *Curr. Opin. Biotechnol.* **2005**, *16*, 427–433.
- [2] K. Gelse, E. Pöschl, T. Aigner, *Adv. Drug Delivery Rev.* **2003**, *55*, 1531–1546.
- [3] F. Chiti, C. M. Dobson, *Annu. Rev. Biochem.* **2017**, *86*, 27–68.
- [4] D. S. T. Ong, J. W. Kelly, *Curr. Opin. Cell Biol.* **2011**, *23*, 231–238.
- [5] M. Jucker, L. C. Walker, *Nature* **2013**, *501*, 45–51.
- [6] B. Cheng, H. Gong, H. Xiao, R. B. Petersen, L. Zheng, K. Huang, *Biochim. Biophys. Acta Gen. Subj.* **2013**, *1830*, 4860–4871.
- [7] R. Gallardo, N. A. Ranson, S. E. Radford, *Curr. Opin. Struct. Biol.* **2020**, *60*, 7–16.
- [8] R. Riek, D. S. Eisenberg, *Nature* **2016**, *539*, 227–235.
- [9] T. Eichner, S. E. Radford, *FEBS J.* **2011**, *278*, 3868–3883.
- [10] S. M. Johnson, R. L. Wiseman, Y. Sekijima, N. S. Green, S. L. Adamski-Werner, J. W. Kelly, *Acc. Chem. Res.* **2005**, *38*, 911–921.
- [11] H. A. Lashuel, C. R. Overk, A. Oueslati, E. Masliah, *Nat. Rev. Neurosci.* **2013**, *14*, 38–48.
- [12] D. M. Walsh, D. B. Teplow, *Prog. Mol. Biol. Transl. Sci.* **2012**, *107*, 101–124.
- [13] T. P. J. Knowles, W. Shu, G. L. Devlin, S. Meehan, S. Auer, C. M. Dobson, M. E. Welland, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10016–10021.
- [14] E. Serebryany, J. A. King, *Prog. Biophys. Mol. Biol.* **2014**, *115*, 32–41.
- [15] P. Neudecker, P. Robustelli, A. Cavalli, P. Walsh, P. Lundström, A. Zarrine-Afsar, S. Sliemers, M. Vendruscolo, L. E. Kay, *Science* **2012**, *336*, 362–366.
- [16] P. Alam, L. Bousset, R. Melki, D. E. Otzen, *J. Neurochem.* **2019**, *150*, 522–534.
- [17] S. Li, D. J. Selkoe, *J. Neurochem.* **2020**, <https://doi.org/10.1111/jnc.15007>.
- [18] F. Chiti, P. Webster, N. Taddei, A. Clark, M. Stefani, G. Ramponi, C. M. Dobson, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3590–3594.
- [19] F. Hane, *FEBS Lett.* **2013**, *587*, 3617–3619.
- [20] M. Fändrich, C. M. Dobson, *EMBO J.* **2002**, *21*, 5682–5690.
- [21] M. T. Pastor, A. Esteras-Chopo, L. Serrano, *Prion* **2007**, *1*, 9–14.
- [22] N. S. de Groot, R. Sabate, S. Ventura, *Trends Biochem. Sci.* **2009**, *34*, 408–416.
- [23] G. Reddy, A. N. Muttathukattil, B. Mondal, *Curr. Opin. Struct. Biol.* **2020**, *60*, 101–109.
- [24] M. Calamai, F. Chiti, C. M. Dobson, *Biophys. J.* **2005**, *89*, 4201–4210.
- [25] A. B. Ahmed, A. V. Kajava, *FEBS Lett.* **2013**, *587*, 1089–1095.
- [26] G. Meric, A. S. Robinson, C. J. Roberts, *Annu. Rev. Chem. Biomol. Eng.* **2017**, *8*, 120–150.
- [27] J. S. Ebo, N. Guthertz, S. E. Radford, D. J. Brockwell, *Curr. Opin. Struct. Biol.* **2020**, *60*, 157–166.
- [28] J. A. Raskatov, D. B. Teplow, *Sci. Rep.* **2017**, *7*, 12433.
- [29] E. Y. Hayden, K. K. Hoi, J. Lopez, M. Inayathullah, M. M. Condron, D. B. Teplow, *Sci. Rep.* **2017**, *7*, 12434.
- [30] C. J. A. Warner, S. Dutta, A. R. Foley, E. Chen, D. S. Kliger, J. A. Raskatov, *Chirality* **2017**, *29*, 5–9.
- [31] J. H. Fuhrhop, M. Krull, G. Büldt, *Angew. Chem. Int. Ed. Engl.* **1987**, *26*, 699–700; *Angew. Chem.* **1987**, *99*, 707–708.
- [32] K. Nagy-Smith, P. J. Beltramo, E. Moore, R. Tycko, E. M. Furst, J. P. Schneider, *ACS Cent. Sci.* **2017**, *3*, 586–597.
- [33] J. M. Urban, J. Ho, G. Piester, R. Fu, B. L. Nilsson, *Molecules* **2019**, *24*, 1983.
- [34] K. Kar, I. Arduini, K. W. Drombosky, P. C. A. van der Wel, R. Wetzl, *J. Mol. Biol.* **2014**, *426*, 816–829.
- [35] S. Dutta, A. R. Foley, C. J. A. Warner, X. Zhang, M. Rolandi, B. Abrams, J. A. Raskatov, *Angew. Chem. Int. Ed.* **2017**, *56*, 11506–11510; *Angew. Chem.* **2017**, *129*, 11664–11668.
- [36] L. Pauling, R. B. Corey, *Proc. Natl. Acad. Sci. USA* **1953**, *39*, 253–256.
- [37] V. Torbееv, M. Grogg, J. Ruiz, R. Boehringer, A. Schirer, P. Hellwig, G. Jeschke, D. Hilvert, *J. Pept. Sci.* **2016**, *22*, 290–304.
- [38] B. Schmidtgal, O. Chaloin, V. Bauer, M. Sumyik, C. Birck, V. Torbееv, *Chem. Commun.* **2017**, *53*, 7369–7372.
- [39] V. Torbееv, B. Schmidtgal, G. Grogg, I. Dörmenc, I. Oez, Y. Nominé, C. Kretschmann, A. Cousido-Siah, A. Mitschler, N. Rochel, G. Travé, B. Kieffer, V. Torbееv, *ChemRxiv* **2019**, <https://doi.org/10.26434/chemrxiv.10113128.v1>.
- [40] N. Vo, R. H. Goodman, *J. Biol. Chem.* **2001**, *276*, 13505–13508.

- [41] D. Otzen, R. Riek, *Cold Spring Harbor Perspect. Biol.* **2019**, *11*, a033860.
- [42] M. Kjaergaard, K. Teilum, F.M. Paulsen, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 12535–12540.
- [43] S. J. Demarest, M. Martinez-Yamout, J. Chung, H. Chen, W. Xu, H. J. Dyson, R. M. Evans, P. E. Wright, *Nature* **2002**, *415*, 549–553.
- [44] B. Y. Qin, C. Liu, H. Srinath, S. S. Lam, J. J. Correia, R. Derynck, K. Lin, *Structure* **2005**, *13*, 1269–1277.
- [45] J. Jacques, A. Collet, S. H. Wilen, *Enantiomers, Racemates and Resolutions*, Krieger Publishing, Malabar, **1994**.
- [46] S. Srisanga, J. H. ter Horst, *Cryst. Growth Des.* **2010**, *10*, 1808–1812.
- [47] G. Coquerel, *J. Pharm. Pharmacol.* **2015**, *67*, 869–878.
- [48] F. T. Senguen, N. R. Lee, X. Gu, D. M. Ryan, T. M. Doran, E. A. Anderson, B. L. Nilsson, *Mol. BioSyst.* **2011**, *7*, 486–496.
- [49] A. Barth, C. Zscherp, *Q. Rev. Biophys.* **2002**, *35*, 369–430.
- [50] E. Goormaghtigh, J.-M. Ruysschaert, V. Raussens, *Biophys. J.* **2006**, *90*, 2946–2957.
- [51] A. C. Lenney, A. J. R. Heck, *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5–13.
- [52] S. K. Chaturvedi, M. K. Siddiqi, P. Alam, R. H. Khan, *Process Biochem.* **2016**, *51*, 1183–1192.
- [53] T. M. Marcinko, C. Liang, S. Savinov, J. Chen, R. W. Vachet, *J. Mol. Biol.* **2020**, *432*, 396–409.
- [54] W. Tong, G. Wang, *Methods* **2018**, *144*, 3–13.
- [55] S. Mehmood, T. M. Allison, C. V. Robinson, *Annu. Rev. Phys. Chem.* **2015**, *66*, 453–474.
- [56] M. Emily, A. Talvas, C. Delamarche, *PLoS ONE* **2013**, *8*, e79722.
- [57] A. C. Tsofis, N. C. Papandreou, V. A. Iconomidou, S. J. Hamodrakas, *PLoS ONE* **2013**, *8*, e54175.
- [58] T. O. Yeates, S. B. H. Kent, *Annu. Rev. Biophys.* **2012**, *41*, 41–61.
- [59] E. M. McInerney, D. W. Rose, S. E. Flynn, S. Westin, T. M. Mullen, A. Kronos, J. Inostroza, J. Torchia, R. T. Nolte, N. Assa-Munt, M. V. Milburn, C. K. Glass, M. G. Rosenfeld, *Genes Dev.* **1998**, *12*, 3357–3368.
- [60] J. J. Voegel, M. J. S. Heine, M. Tini, V. Vivat, P. Chambon, H. Gronemeyer, *EMBO J.* **1998**, *17*, 507–519.
- [61] C. W. Wood, D. N. Woolfson, *Protein Sci.* **2018**, *27*, 103–111.
- [62] R. Kurjithapatham, F. Y. Oliva, U. Doshi, M. Pérez, J. Avila, V. Muñoz, *Biochemistry* **2005**, *44*, 149–156.
- [63] D. Ghosh, P. K. Singh, S. Sahay, N. N. Jha, R. S. Jacob, S. Sen, A. Kumar, R. Riek, S. K. Maji, *Sci. Rep.* **2015**, *5*, 9228.
- [64] R. H. Walters, R. M. Murphy, *J. Mol. Biol.* **2011**, *412*, 505–519.
- [65] A. Abedini, D. P. Raleigh, *Protein Eng. Des. Sel.* **2009**, *22*, 453–459.
- [66] C. P. Brangwynne, P. Tompa, R. V. Pappu, *Nat. Phys.* **2015**, *11*, 899–904.
- [67] G. A. P. de Oliveira, Y. Cordeiro, J. L. Silva, T. C. R. G. Vieira, *Adv. Protein Chem. Struct. Biol.* **2019**, *118*, 289–331.
- [68] W. M. Babinchak, W. K. Surewicz, *J. Mol. Biol.* **2020**, *432*, 1910–1925.
- [69] H. J. Dyson, P. E. Wright, *J. Biol. Chem.* **2016**, *291*, 6714–6722.
- [70] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

Manuscript received: April 16, 2020

Revised manuscript received: April 30, 2020

Accepted manuscript online: May 4, 2020

Version of record online: July 14, 2020