



A lowly populated, transient β -sheet structure in monomeric $A\beta^{1-42}$ identified by multinuclear NMR of chemical denaturation

Tayeb Kakeshpour¹, Venkat Ramanujam¹, C. Ashley Barnes, Yang Shen, Jinfa Ying, Ad Bax^{*}

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA

ARTICLE INFO

Keywords:

Urea denaturation
Protein folding
Paramagnetic relaxation enhancement
Triple resonance NMR
Intrinsically disordered protein
Chemical shift perturbation

ABSTRACT

Chemical denaturation is a well-established approach for probing the equilibrium between folded and unfolded states of proteins. We demonstrate applicability of this method to the detection of a small population of a transiently folded structural element in a system that is often considered to be intrinsically fully disordered. The ^1H , ^{15}N , ^{13}C , and $^{13}\text{C}'$ chemical shifts of $A\beta^{1-40}$ and $A\beta^{1-42}$ peptides and their M35-oxidized variants were monitored as a function of urea concentration and compared to analogous urea titrations of synthetic pentapeptides of homologous sequence. Fitting of the chemical shift titrations yields a $10 \pm 1\%$ population for a structured element at the C-terminus of $A\beta^{1-42}$ that folds with a cooperativity of $m = 0.06$ kcal/mol-M. The fit also yields the chemical shifts of the folded state and, using a database search, for $A\beta^{1-42}$ these shifts identified an antiparallel intramolecular β -sheet for residues I32-A42, linked by a type I' β -turn at G37 and G38. The structure is destabilized by oxidation of M35. Paramagnetic relaxation rates and two previously reported weak, medium-range NOE interactions are consistent with this transient β -sheet. Introduction of the requisite A42C mutation and tagging with MTSL resulted in a small stabilization of this β -sheet. Chemical shift analysis suggests a C-terminal β -sheet may be present in $A\beta^{1-40}$ too, but the turn type at G37 is not type I'. The approach to derive Transient Structure from chemical Denaturation by NMR (TSD-NMR), demonstrated here for $A\beta$ peptides, provides a sensitive tool for identifying the presence of lowly populated, transiently ordered elements in proteins that are considered to be intrinsically disordered, and permits extraction of structural data for such elements.

1. Introduction

Amyloid beta ($A\beta$) peptides result from proteolysis of the amyloid precursor protein and are found as amyloid deposits in the brains of patients with Alzheimer's disease (AD). $A\beta$ generally refers to short peptides of 39–43 residues in length, but the 42-residue peptide ($A\beta^{1-42}$) and its C-terminally truncated $A\beta^{1-40}$ are the forms most prevalent in AD plaques [1,2]. Thus, they are thought to play an important role in the etiology of AD pathogenesis. As is often observed for proteins implicated in amyloid diseases, in their monomeric solution state the $A\beta$ peptides lack stable secondary and tertiary structure and are commonly referred to as intrinsically disordered proteins (IDPs) [3].

While historically the extracellular plaques were identified as the hallmark of AD, they were subsequently found to correlate only weakly with the primary AD symptoms of cognitive impairment and loss of neural synapses [4]. Instead, soluble $A\beta$ oligomers are now more commonly considered to constitute the main toxic species, rather than

simply being a benign intermediate on the pathway to the observed senile plaques [5,6]. NMR studies have shown that in isotropic solution the monomeric peptides can rapidly oligomerize with rate constants that steeply depend on peptide concentration [7]. At atmospheric pressure, there exists an apparent threshold concentration of *ca* 120 μM for oligomerization of $A\beta^{1-40}$, and somewhat lower for $A\beta^{1-42}$ [8]. At elevated hydrostatic pressure, the oligomerization can be reversed and the rapid kinetics of the oligomer-monomer interconversion formed the basis for pressure-jump NMR experiments which showed that mega-Dalton oligomers can form in a few seconds [9]. The very rapid conversion from a disordered state into a highly ordered oligomer, and in particular the fast initiation of this process, is intriguing, in particular considering that there is no detectable trend for dimerization below the oligomerization threshold: ^1H - ^{15}N heteronuclear single quantum correlation (HSQC) spectra show no detectable concentration dependence of chemical shifts over a wide range of sample concentration (1–150 μM) [8] and below the oligomerization threshold, no reproducible dependence of

^{*} Corresponding author.

E-mail address: bax@nih.gov (A. Bax).

¹ T.K. and V.R. contributed equally.

transverse relaxation properties on sample concentration has been reported. Considering that NMR chemical shifts are exquisitely sensitive to structural changes associated with oligomerization, this absence of a concentration-dependence of chemical shifts puts an upper limit boundary of at most 1% for any transiently forming oligomer at a *ca* 150 μ M peptide concentration. Even highly sensitive paramagnetic relaxation enhancement studies revealed no evidence for intermolecular association [7], which appears to exclude the possibility that non-specific hydrophobic interaction precedes the formation of highly ordered oligomeric structures. Thus, the A β oligomerization process poses the same question often encountered for other intrinsically disordered systems interacting with a well-structured target protein: Does the disordered polypeptide contain amphipathic molecular recognition motifs (MoRFs) that are transiently structured and selected for binding to their interaction partner in a process often referred to as coupled folding and binding? [10] The presence of transiently ordered MoRFs would greatly decrease the requisite entropic penalty when such an element engages in formation of a stable complex with an interaction partner, while retaining the flexibility to interact with many different targets [11,12].

The higher aggregation propensity of A β^{1-42} over A β^{1-40} has been linked to its higher rate of primary nucleation, which does not involve the presence of pre-existing fibrils [1,13]. An array of sequence variations is strongly correlated with familial forms of AD, and relates to physicochemical properties such as the net peptide charge, its hydrophobicity, and its propensity to adopt secondary structure [14]. As an example of electrostatic effects, familial mutations such as E22G and D23N, which decrease the net negative charge of A β^{1-42} , accelerate its aggregation. Interestingly, this more rapid oligomerization process has been attributed to faster secondary nucleation on the surface of existing fibrils; the primary nucleation and fibril elongation rates are little impacted by these mutations [15,16]. On the other hand, the slower aggregation of the less apolar M35-oxidized A β^{1-42} form of the peptide pointed to the role of hydrophobic interactions in the aggregation process [17–19]. To date, it has remained unclear whether the differences in nucleation rates observed between A β^{1-40} and A β^{1-42} , or upon M35 oxidation, originate simply from the altered total hydrophobicity of the C-terminal region of the polypeptide, or whether their differential aggregation propensities must be attributed to differences in the conformational ensemble sampled by this region of the peptide chain. An NMR study from our laboratory confirmed prior studies [20] and concluded that the backbone chemical shifts of both A β^{1-40} and A β^{1-42} are remarkably close to random coil values, but also noted that the loss of the C-terminal two residues in A β^{1-40} impacted chemical shifts and $^3J_{\text{HNH}\alpha}$ coupling constants of residues M35-G38 to a greater extent than would be expected for a true random coil [8]. Although the reported differences were small, they are consistent with prior observations of a small increase in backbone rigidity deduced from ^{15}N NMR relaxation measurements for these residues [21], and increased short-range ^1H - ^1H NOE intensity [20].

A wide range of other NMR studies of A β or fragments thereof pointed to transiently or fully structured forms of these oligopeptides in aqueous solution [22]. Most remarkably, a folded, mostly α -helical structure was reported by Ramamoorthy and co-workers [23], and total reflectance Fourier-transform infrared spectroscopy together with solid-state NMR indicated that A β peptide oligomers can switch from a preferred β to α -helical secondary structure by simply lyophilizing such samples [24]. Single molecule Förster resonance energy transfer (FRET) measurements provide another powerful method to investigate peptide structure in solution. A recent study of both A β^{1-40} and A β^{1-42} , with fluorophores attached at opposite ends of the peptides, found large structural fluctuations on a 35-ns time scale, similar to what has been observed for other IDPs, and pointed to the absence of a significant population of transiently ordered, long-lived conformers [25]. Fluorescence studies hold the advantage over NMR that they are carried out at sub-nanomolar peptide concentrations, effectively removing any possibility of intermolecular interactions. Addition of urea showed a modest,

gradual structural expansion upon increasing the denaturant concentration, consistent with what has been reported for FRET studies of other intrinsically disordered proteins [26,27]. Combining the FRET results with replica-exchange molecular dynamics (MD) calculations, using a force field that has been carefully optimized to avoid the collapse often seen in such trajectories [28], showed a broad ensemble of rapidly interconverting conformers [25]. However, a *ca* 3% population was seen for a conformer where the C-terminus of A β^{1-42} , but not of A β^{1-40} , paired with the N-terminus to adopt an antiparallel β -sheet, thereby explaining the slightly higher FRET efficiency in A β^{1-42} [25]. A multitude of other MD calculations came to a diverse range of rather different conclusions. For example, Rosenman et al. found an array of transiently populated antiparallel β -hairpins similar to those seen in oligomer and fibril models, but with distinct differences between A β^{1-40} , A β^{1-42} , and M35-oxidized A β^{1-42} (A $\beta^{1-42-\text{Ox}}$) [29]. Head-Gordon and co-workers also found a broad range of transient local secondary structure in both A β^{1-40} and A β^{1-42} , but with these elements shifted to different locations along the two polypeptides [30], in apparent contrast to the very similar NMR chemical shifts and J-couplings of the N-terminal 33 residues of both peptides [8]. Most of the earlier MD calculations used force fields that had been parametrized using experimental data obtained for folded proteins and these can overestimate the importance of intramolecular H-bonding when used for IDPs [28]. However, with adjustment of AMBER force field parameters, Robustelli et al. [31] showed that it is possible to achieve good performance on both folded and unfolded proteins, whereas others reached this goal by modifying the CHARMM force field [32]. For A β^{1-40} , results were reported to be in very good agreement with experimental data when using a fine-tuned CHARMM22* force field [31]. This method also performed well for A β^{1-42} in a MD study that compared five of the best performing force fields to characterize the monomer structure [33].

Taking advantage of the sensitivity of NMR chemical shifts for identifying small changes in populations of a structural ensemble [34], here we pursue a somewhat different experimental approach at identifying any residual transient structure in the A β peptides. NMR holds the advantage over the FRET studies of Chung and co-workers [25] in that the backbone chemical shifts offer residue-specific reporters on any denaturant-induced changes in the equilibrium between folded and unfolded conformers, thereby offering the potential to gain residue-specific insights into possible origins of the slight chain expansion observed by FRET.

Prior NMR studies that relied on the addition of denaturants to IDPs were carried out for proteins with a significant propensity to adopt α -helical secondary structure, and this approach was quite effective at identifying such transiently structured elements [35–37]. However, considering that for the A β peptides the backbone secondary chemical shifts are very small [8], well within the range expected for a random coil, the maximum population of any transiently structured element is expected to be much lower. Using urea as a denaturant involves addition of high volume fractions of this compound to the solution, which makes it challenging to distinguish the effect of protein unfolding on chemical shifts from the effect of the changed solvent composition. In particular, weak binding of urea to the backbone amide moieties of the unfolded peptide chain has been experimentally demonstrated [38] and is considered a driving mechanism in shifting the folded-unfolded equilibrium [39]. To distinguish the two contributions to chemical shift changes, we resorted to the synthesis of peptide fragments that match triplets of the A β sequence but that are too short to adopt significant secondary structure. Comparison of the effect of urea on the chemical shifts of the full-length A β sequence and the corresponding nuclei in the short peptides then yields an exquisitely sensitive monitor of residual structure. Such analyses are shown to yield the chemical shifts of any transiently folded conformers, and thereby provide insights into the actual structure of the transiently folded elements. Using the database mining approach that underlies the CS-Rosetta method for chemical-shift-based structure determination [40], we identify a transient,

antiparallel β -sheet for residues I32-A42 in $A\beta^{1-42}$, with the two strands linked by a well-defined type I' β -turn. The turn remains intact upon oxidation of M35, but this alteration lowers the population of the structured element by about 35%, potentially shedding new light on its lower propensity to grow amyloid fibrils [17,18]. Our results are consistent with previously reported NOEs [41], and with newly recorded paramagnetic relaxation enhancement (PRE) data [42]. PRE data have been shown to offer a very sensitive probe for exploring transient long-range contacts in IDPs [43,44]. Our $A\beta^{1-42}$ PRE data point to the presence of multiple long-range chain-chain interactions, in addition to the C-terminal β -sheet, that can be attenuated by addition of denaturant.

2. Results and discussion

2.1. Extracting populations and chemical shifts from denaturant titrations

For a two-state equilibrium between folded (F) and unfolded (U) states of a protein, the free energy of unfolding is given by.

$$\Delta G_{F \rightarrow U} = -RT \ln([U]/[F]) \quad (1)$$

where R is the universal gas constant, T the absolute temperature, and $[U]$ and $[F]$ denote the concentrations of unfolded and folded states, respectively. To a very good approximation, $\Delta G_{F \rightarrow U}$ has been shown to vary linearly with denaturant concentration [45,46]:

$$\Delta G_{F \rightarrow U}(D) = \Delta G_{F \rightarrow U}(H_2O) - m[D] \quad (2)$$

where $[D]$ is the denaturant concentration, $\Delta G_{F \rightarrow U}(H_2O)$ denotes $\Delta G_{F \rightarrow U}$ in denaturant-free buffer. The coefficient m denotes the cooperativity of the unfolding transition and was shown to correlate linearly with the difference in solvent-accessible surface area between the folded and unfolded species [46,47]

$$m = (0.15 \pm 0.03 \text{ cal/mol} \cdot \text{M} \cdot \text{\AA}^{-2}) \times \Delta ASA \quad (3)$$

where ΔASA is the difference in solvent-accessible surface area between the unfolded and folded states in units of \AA^2 .

The equilibrium between the populations of the folded and unfolded states is then given by

$$[F]/[U] = A e^{-m[D]/RT} \quad (4)$$

where A is the F/U population ratio in the absence of denaturant. For IDPs like $A\beta^{1-42}$, the transiently folded local structures (if any) involve only a few hydrogen bonds (H-bonds) and hydrophobic interactions. Therefore, the sum of their strengths cannot result in a high barrier for the (un)folding process, meaning that the species in equilibrium will be in fast exchange and their peak positions are averaged in the NMR spectra. Specifically, in the ^{13}C and ^{15}N labeled $A\beta^{1-42}$ case [48], going to concentrations as high as 1 mM did not result in any detectable exchange broadening (as would have been expected for NMR in the intermediate exchange regime) nor an additional set of folded conformer peaks (as would apply for slow exchange on the NMR time scale). Therefore, any exchange between unfolded and folded conformers must be fast on the NMR time scale, and thus the experimentally observed chemical shift, $\delta_{\text{exp}}(D)$, is a weighted average of the folded and unfolded states, given by

$$\delta_{\text{exp}}(D) = \frac{A}{A + e^{m[D]/RT}} \times \delta_f + \frac{e^{m[D]/RT}}{A + e^{m[D]/RT}} \times \delta_u \quad (5)$$

where the parameters, m , δ_f , δ_u and A of this sigmoid-shaped function are obtained from the fit. In particular, the fitted values for the chemical shifts in the folded state, δ_f , which can be readily obtained for ^1H , ^{15}N , ^{13}C , and $^{13}\text{C}^\alpha$ nuclei, then can serve as input to derive an atomic resolution model for the transiently folded state using well-established protocols [49]. The effect of urea on δ_f tends to be small (*vide infra*) and does not significantly impact the fitting to Eq. (5), in particular since

the contribution of δ_f to δ_{exp} is small when $A \ll 1$, as applies for $A\beta^{1-42}$. However, there are two reasons why the above strategy has not yet been widely adopted. First, the fact that there are four adjustable parameters in the fit of Eq. (5) introduces large uncertainties into the values of the extracted parameters. Second, weak binding of the denaturant to the protein's unfolded state impacts the chemical shift, δ_u , and therefore interferes with the fitting process and the extraction of the true m , δ_f , δ_u and A values.

We have addressed these problems in the following manner. First, by measuring the effect of urea concentration on the chemical shifts of peptides that are too short to develop significant secondary structure but that match the local amino acid sequence of the target system, we account for the residue-specific effects of the denaturant concentration on the unfolded ^1H , ^{15}N , ^{13}C , and $^{13}\text{C}^\alpha$ chemical shifts. Second, by having titration curves measured simultaneously for four nuclei of each residue, which must be fitted with the same m and A values, the ratio of the number of fitted parameters relative to the number of observables is greatly reduced. Third, as a control on the reliability of the fit, it is to be expected that all residues involved in a transiently ordered structure will have the same or very similar m and A values, which in a subsequent fit can either be restrained to their averaged values or be treated as global fitting parameters to obtain highly robust and self-consistent chemical shifts for the folded state, δ_f . These δ_f values are subsequently used as input for structural modeling.

In the traditional case where denaturant is used to unfold a protein, the sigmoidal shape of the F - U equilibrium is like that of the black curve in Fig. 1.

In order to analyze the folding equilibrium curve for proteins that are largely unfolded under denaturant-free conditions (blue curve in Fig. 1) it is useful to write Eq. (5) as:

$$\delta_{\text{exp}}(D) = \delta_f \times \frac{1}{1 + A^{-1} \times e^{m[D]/RT}} + \delta_u \times \frac{A^{-1} \times e^{m[D]/RT}}{1 + A^{-1} \times e^{m[D]/RT}} \quad (6)$$

In the limit where $A \ll 1$, the denominator becomes large and $\delta_{\text{exp}}(D)$ approaches a mono-exponential function:

$$\delta_{\text{exp}}(D) \approx \delta_f (A e^{-m[D]/RT}) + \delta_u (1 - A e^{-m[D]/RT}) = \delta_u + A (\delta_f - \delta_u) e^{-m[D]/RT} \quad (7)$$

which shows that only the product $A (\delta_f - \delta_u)$ can be extracted from the

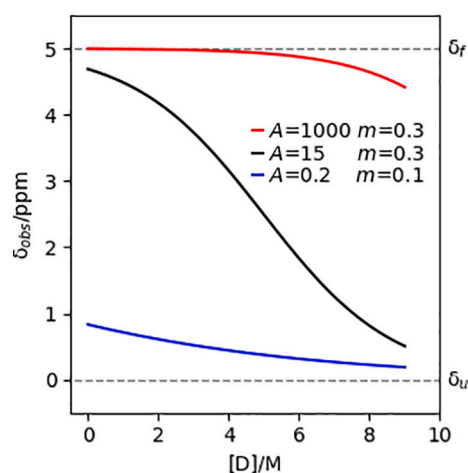


Fig. 1. Three simulated chemical denaturation curves for a typical folded protein (black), a protein that is highly resistant to chemical denaturation (red) and a transiently ordered IDP (blue). For this example, the chemical shifts of the folded and unfolded states are chosen to be 5 and 0 ppm, respectively, and are marked by dashed lines. Plots show the simulated chemical shifts under the condition of rapid exchange as a function of denaturant concentration, $[D]$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

data. Therefore, when approaching this limit, the fitting parameters A and δ_f become strongly correlated and can only be uniquely determined from the fit if the $A \ll 1$ condition does not apply. However, outside that limit, it is often possible to obtain a reasonable estimate for A by considering that, on average, the root-mean-square (rms) difference between δ_f and δ_u values is expected to be comparable to rms secondary shifts observed in folded proteins. Thus, constraining them to such values removes the degeneracy of the fit. This strategy therefore can yield δ_f values while limiting the uncertainties in the fitted thermodynamic parameters m and A . In addition, considering that there were five different peptides for which the urea titration was carried out, the δ_u values of these peptides can be treated as single variables, while treating the δ_u values for residues L34, M35, and V36 separately for the peptides with and without oxidation of M35. Also, signals from the two C-terminal residues of $A\beta^{1-40}$ and $A\beta^{1-40-Ox}$ are fitted with a common δ_u that differs from values used for $A\beta^{1-42}$ peptides. Therefore, the following

error function was used for fitting:

$$\chi^2 = \sum_D \sum_i \sum_X \left(\frac{\delta_{i,X,D}^{calc} - \delta_{i,X,D}^{exp}}{T_X} \right)^2 + \sum_i \sum_X k \left[\frac{\left(\frac{\delta_{f,i,X} - \delta_{u,i,X}}{T_X} \right)^2}{4 \times nres} - 1 \right]^2 \quad (8)$$

The first term of the error function represents the difference between calculated and measured chemical shifts and the second term is a harmonic weighting function constraining the rms secondary shift of the folded state to be close to those observed in folded protein structures. Superscripts *exp* and *calc* denote the measured and calculated chemical shifts respectively, and D , i and X subscripts denote the denaturant concentration, residue number, and nucleus type, respectively. The value of k , the weighting factor for the global harmonic function that restrains the secondary chemical shifts to rms values similar to those in folded proteins, was simply set to 1, since the solution of the fit showed

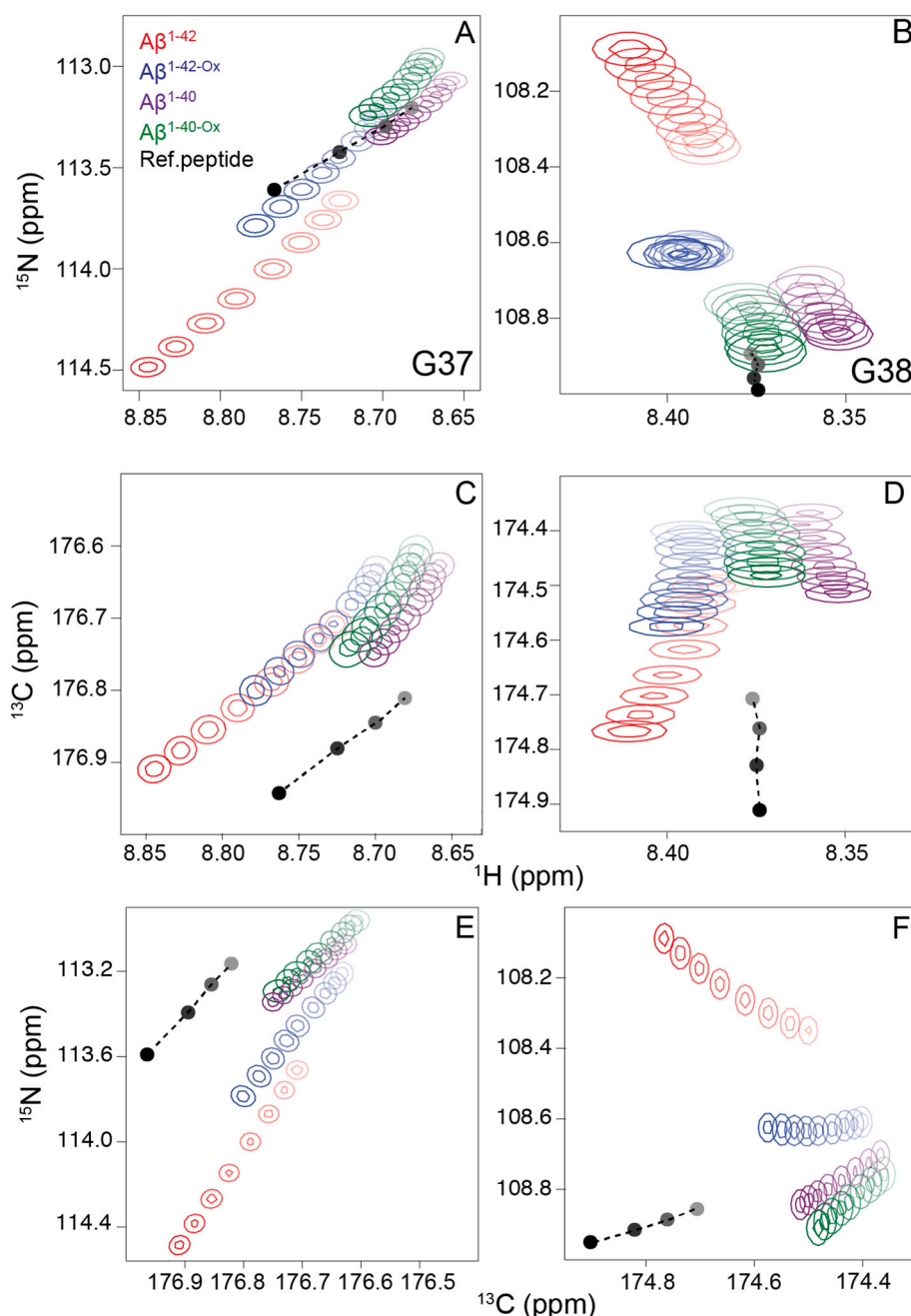


Fig. 2. Overlay of selected regions of projections of 34 3D HNCO spectra (700 MHz), recorded for four uniformly $^{15}\text{N}/^{13}\text{C}$ -enriched $A\beta$ peptides at increasing urea concentrations, marked by decreasing intensity of colour. Panels A and B correspond to the projections on the ^{15}N - ^1H plane; panels C and D are projections on the ^{13}C - ^1H plane; panel E and F are projections on the ^{15}N - ^{13}C plane for residues G37 and G38 (with the projection for panel E restricted to the 8.9–8.6 ppm region in the ^1H dimension). The black/grey dots mark the corresponding chemical shifts in the reference pentapeptide, with grey corresponding to the highest urea concentration. Urea concentrations (M) used for the titrations were $A\beta^{1-42}$: 0.00, 0.76, 1.63, 2.67, 4.04, 5.65, 7.15, 8.51; $A\beta^{1-42-Ox}$: 0.00, 0.87, 1.72, 2.62, 3.57, 4.86, 6.01, 6.98, 7.95; $A\beta^{1-40}$: 0.00, 0.81, 1.61, 2.70, 4.20, 5.78, 7.36, 8.85; $A\beta^{1-40-Ox}$: 0.00, 0.76, 1.58, 2.58, 3.45, 4.84, 5.95, 6.91, 7.78. Ref. peptide for G37 (QVGGQ): 0.00, 2.76, 5.38, 7.78; Ref. peptide for G38 (QGGVQ): 0.00, 2.72, 5.22, 7.82.

little dependence on its value in the 0.01–1.0 range. A non-zero k value is required to reach convergence, however. T_X values for nuclei $X = {}^1\text{H}^{\text{N}}$, ${}^{15}\text{N}$, ${}^{13}\text{C}'$, and ${}^{13}\text{C}^{\alpha}$ were 0.73, 4.14, 1.91, and 1.97 ppm, respectively, and represent the rms secondary shifts of these nuclei in the SPARTA+ database of assigned proteins of well-defined structure [50]. The term ' $4 \times nres$ ', in which $nres$ is the number of residues used in fitting, represents the total number of nuclei used in the fit. Although for a given nucleus type, X , T_X values have a root-mean-square spread (rmsd) over different proteins in the SPARTA+ database of ca 15%, we note that the last term in Eq. (8) sums over all four nuclei, which reduces the corresponding variation to less than 10%.

Above, the denaturant dependence of chemical shifts of the unfolded state, resulting from solvent effects and/or weak transient binding of the denaturant to polypeptide atoms was not taken into account. To a very good approximation, such effects can be eliminated by measuring the impact of denaturant on chemical shifts observed for synthetic pentapeptides of the type $\text{Ac-QX}_{i-1}\text{X}_i\text{X}_{i+1}\text{Q-NH}_2$, where the triplet of residues $\text{X}_{i-1}\text{X}_i\text{X}_{i+1}$ is chosen to match the sequence of residues in our query protein, $\text{A}\beta$. Following Kjaergaard et al. [34], flanking Gln residues were used to represent a typical "average" residue, not prone to inducing local order, and only the chemical shifts of the center residue, X_i , were used to determine the effect of denaturant on the chemical shifts of residue i in the query protein. For poorly soluble sequences, the pentapeptides were extended by three Lys residues: $\text{Ac-QX}_{i-1}\text{X}_i\text{X}_{i+1}\text{Q-K-K-K-COOH}$ to improve solubility (Supporting Information Table S1), analogous to the use of "host" peptides for water-solubilizing lipophilic peptides [51,52].

2.2. Effect of urea on $\text{A}\beta$ chemical shifts

The effect of urea on the HNCO spectra of various $\text{A}\beta$ peptides is readily observed. Superimposed projections of 34 such spectra (Fig. 2) show the differential change in chemical shift of G37 and G38 upon increasing the urea concentration from 0 to 8 M for four peptides: $\text{A}\beta^{1-40}$, $\text{A}\beta^{1-42}$ and their M35 oxidized variants $\text{A}\beta^{1-40-\text{Ox}}$ and $\text{A}\beta^{1-42-\text{Ox}}$. These spectra, all internally referenced to the IUPAC-recommended sodium-2,2-dimethyl-2-silapentane-5-sulfonate (DSS) standard [53], show a substantially different magnitude and even direction of chemical shift change upon increasing urea concentration. Remarkably, the G38 ${}^{15}\text{N}$ resonances of $\text{A}\beta^{1-40-\text{Ox}}$ and $\text{A}\beta^{1-42-\text{Ox}}$ migrate in the opposite direction from $\text{A}\beta^{1-42}$, whereas both the ${}^1\text{H}^{\text{N}}$ and ${}^{15}\text{N}$ chemical shifts of G38 in $\text{A}\beta^{1-42-\text{Ox}}$ remain nearly invariant. The latter observation results from the coincidental, approximately equal but opposite contributions of the urea "solvent" effect (black/grey symbols in Fig. 2) and the urea-induced unfolding of its residual structure on these chemical shifts. As can be seen, chemical shift differences among the four peptides progressively decrease upon addition of urea, representing the expected signature of the unfolding of transient residual structure that must be present to different extents in the four peptides.

As previously reported [8,18,19], the backbone amide chemical shift differences among the four peptides are largest in the C-terminal region, and in this study we focus only on the region extending from I32 to the C-terminus. These residues are quite insensitive to the precise pH value, thereby obviating complications that can arise from the effect of small pH changes during such titration studies. However, very small chemical shift differences are also observed more distant from the C-terminus, indicative of transient interactions between the C-terminal residues and both the center and the more polar N-terminal region of the peptide (SI Fig. S2).

To obtain a complete set of ${}^1\text{H}^{\text{N}}$, ${}^{15}\text{N}$, ${}^{13}\text{C}^{\alpha}$ and ${}^{13}\text{C}'$ chemical shifts, 3D triple resonance HNCO and HN(CO)CA spectra were recorded at all urea concentrations. Even at the modest peptide concentrations of 90 μM , such spectra yielded adequate sensitivity and high digital resolution in about 1.5 h by using non-uniform sampling of these relatively sparse data sets [54], and reconstructing the spectra with the very efficient SMILE algorithm [55]. Because both HNCO (2.8% sampled) and HN(CO)CA (4.7% sampled) spectra yield ${}^1\text{H}$ and ${}^{15}\text{N}$ chemical shifts, this

duplicate set of values was used to confirm the high reproducibility of the chemical shift measurements, typically better than 1 ppb for ${}^1\text{H}$ and better than 4 ppb for ${}^{15}\text{N}$. Recording 3D spectra also eliminated complications with partial resonance overlap during the titration series, which would reduce peak picking accuracy. For each $\text{A}\beta$ peptide, the entire set of both HNCO and HN(CO)CA spectra required less than two days of data collection, sufficiently rapid to avoid significant chemical decomposition of either urea or carbamylation of $\text{A}\beta$ amino groups.

2.3. Extracting A , m , δ_u and δ_f values

Even though, due to long range effects that restrict the conformation of an unfolded chain, the unfolded chemical shifts of an $\text{A}\beta$ residue will differ slightly from those of the center residue of the corresponding pentapeptide (black/grey dots in Fig. 2), we may safely assume that to a very good approximation the effect of urea interacting with the corresponding residues in $\text{A}\beta$ and the pentapeptide is the same. Therefore, in our analyses we remove the effect of urea on the unfolded state of $\text{A}\beta$ by assuming it is the same as that measured for the center residue of the corresponding pentapeptide. For this purpose, the urea dependencies of the chemical shifts of the short peptides were fitted to a quadratic equation, and δ_u in Eq. (5) was substituted by this nucleus- and residue-specific quadratic equation plus an offset. The offset was added for two reasons: (1) As opposed to the $\text{A}\beta$ peptides, the pentapeptides are not isotopically enriched, and hence there is some small secondary isotope effects of ${}^{13}\text{C}$ on neighboring ${}^{15}\text{N}$, and *vice versa* in $\text{A}\beta$ peptides: $\Delta\delta {}^{15}\text{N} ({}^{13}\text{C}) = 0.01$ ppm, and ${}^1\Delta\delta {}^{13}\text{C} ({}^{13}\text{C}) = {}^1\Delta\delta {}^{13}\text{C} ({}^{15}\text{N}) = 0.02$ ppm [56]. So, ${}^{13}\text{C}$ and ${}^{15}\text{N}$ labeling is expected to change the backbone ${}^{13}\text{C}'$, ${}^{13}\text{C}^{\alpha}$ and ${}^{15}\text{N}$ chemical shifts of by ca -0.04 , -0.06 and -0.02 ppm, respectively. (2) These isotope effects are in addition to small changes in ϕ/ψ distributions that result from long-range steric clashes in a fully disordered chain, which also have small impacts on the chemical shifts of disordered chains. Therefore, δ_u in Eq. (5) was replaced by the term " $a[D]^2 + b[D] + c + \text{offset}$ ", in which a , b , and c are residue- and nucleus-dependent constants obtained from short peptides, and *offset* is a fitting variable which accounts for the difference in unfolded state chemical shifts of short peptides and $\text{A}\beta$ proteins.

Furthermore, the urea dependence on the chemical shifts of the folded state was neglected because: (1) the denaturant dependence of chemical shifts in the folded state is small, presumably because the surface area accessible to urea is less compared to the unfolded state. (2) For systems such as $\text{A}\beta$ whose equilibrium is strongly shifted towards the unfolded state, the population-weighted averaged chemical shifts are little impacted by the small fraction of folded conformers. A prior study of the I-domain of bacteriophage p22 by Newcomer et al. [57] provides an example for the urea dependence of chemical shifts in the folded state. The I-domain is mostly comprised of β -sheet but also contains some disordered and α -helical segments. The rms change of ${}^1\text{H}^{\text{N}}$ and ${}^{15}\text{N}$ nuclei in β -strand residues per mole urea corresponds to 6 and 30 ppb, respectively. Assuming a similar sensitivity to urea for β -strand residues in the folded $\text{A}\beta$ conformer, the effect of 8 M urea on the ensemble-averaged chemical shifts, for a 10% folded fraction, amounts to 5 ppb and 24 ppb for ${}^1\text{H}^{\text{N}}$ and ${}^{15}\text{N}$, respectively, which is much smaller than the uncertainty in the δ_f values extracted from the fit. Therefore, after applying the δ_u corrections obtained from the pentapeptides, and neglecting the urea dependence of the folded state chemical shifts, Eq. (6) can be used directly for fitting the experimental data.

Initial, residue-specific fits of the urea dependence of the chemical shifts of $\text{A}\beta^{1-42}$ to Eq. (6) yielded values of 0.05–0.18 for A and 0.03–0.10 kcal/mol·M for m . With a $\sim 175\%$ increase in the χ^2 value of (5.8×10^{-3} vs 2.1×10^{-3}), a subsequent non-linear least-squares global fit (cf Eq. (8)) was able to also fit the chemical shift data very well but with a substantial reduction (100 to 42 for the 10-residue fragment) in the number of adjustable parameters. As can be seen in Fig. 3, this reduced set of parameters suffices to fit the experimental data very well.

The thermodynamic parameters obtained for the different peptides

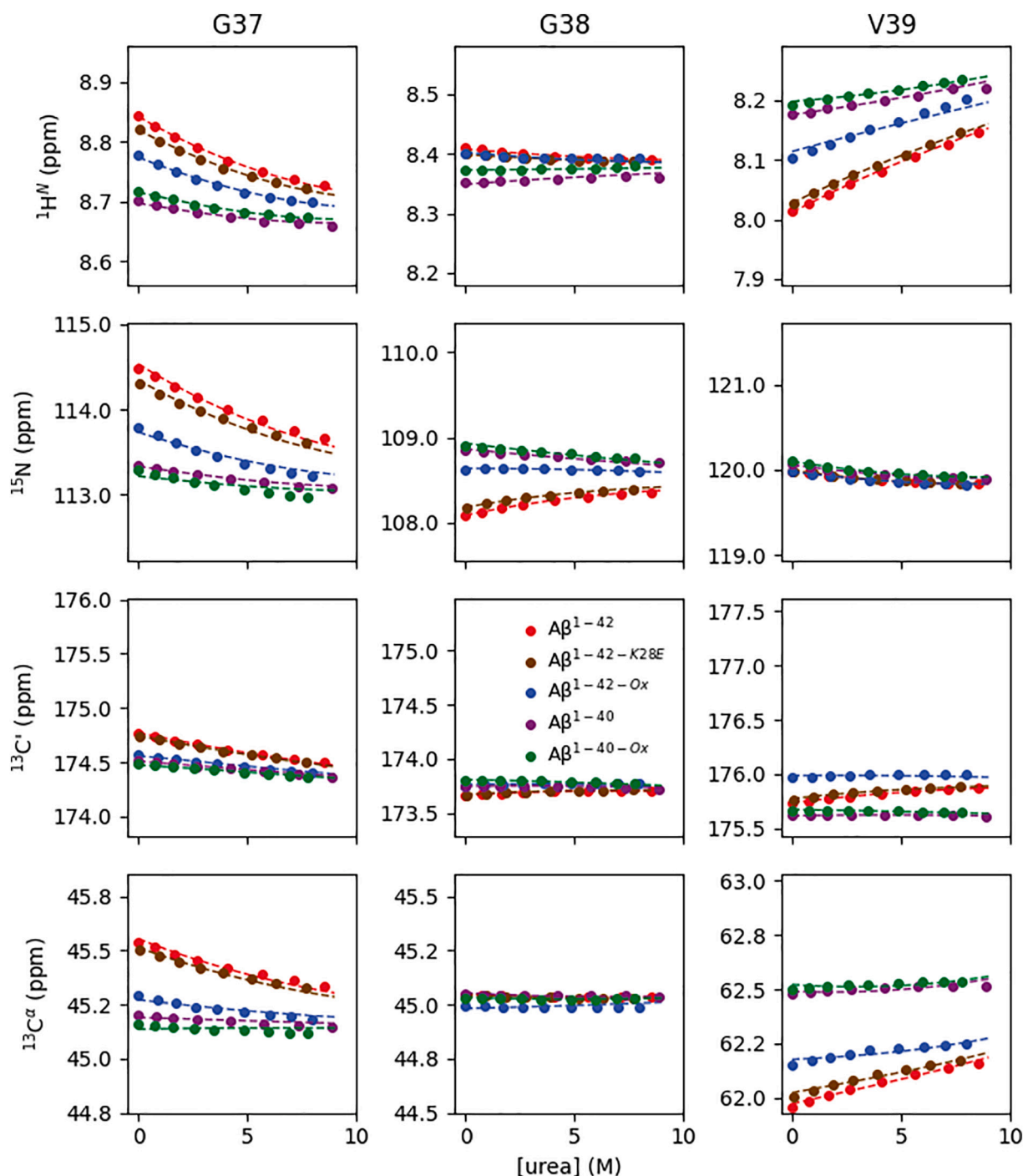


Fig. 3. Urea titrations of the backbone chemical shifts of the five A β peptides. Curves for other residues are shown in SI Fig. S4. Dashed lines correspond to fits to Eq. (6), using global m and A values across residues I32-I41 (I32-V39 for A β^{1-40}). Legend colors match those in Fig. 2.

Table 1

Fitted m and A values for various A β peptides.

Peptide	A (unitless)	m -value (kcal/mol-M)
A β^{1-40}	0.08 ± 0.01	0.06 ± 0.01
A $\beta^{1-40-Ox}$	0.08 ± 0.01	0.06 ± 0.01
A β^{1-42}	0.11 ± 0.01	0.06 ± 0.01
A $\beta^{1-42-Ox}$	0.07 ± 0.01	0.07 ± 0.01
A $\beta^{1-42-K28E}$	0.11 ± 0.01	0.06 ± 0.01

(Table 1) show populations, $A/(A + 1)$, of 7–10% for the transiently folded species, and m values in the 0.06–0.07 kcal/mol-M range, pointing to cooperativities of unfolding that are much smaller than typically encountered for folded globular proteins.

To evaluate how uniquely m and A can be extracted from the experimental data for the various peptides, we also carried out a grid search, and then for each (m , A) pair the δ_f and δ_u values were fitted to minimize the error function of Eq. (8).

For A β^{1-42} , a contour plot of the resulting χ^2/χ^2_{\min} values shows a minimum for $A = 0.11$ and $m = 0.06$ kcal/mol-M (Fig. 4). The first contour level corresponds to an increase of 100% in χ^2 , meaning that

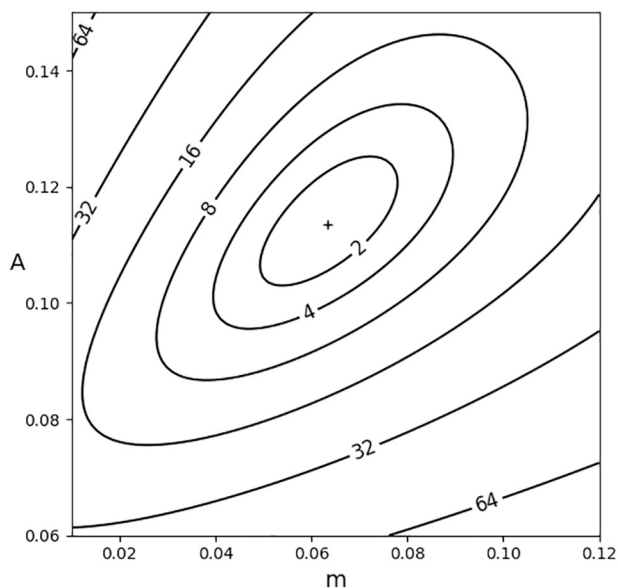


Fig. 4. Contour plot of χ^2/χ^2_{\min} , obtained for $\text{A}\beta^{1-42}$ from a 100×100 grid search of m and A , followed by best-fitting of the chemical shift parameters. The '+' marker at $A = 0.11$ and $m = 0.06$ kcal/mol-M corresponds to the global minimum, $\chi^2 = 5.8 \times 10^{-3}$. Analogous plots for the other four peptides are shown in SI Fig. S5.

small changes in m or A result in a steep increase of χ^2 , and indicating that m and A are well defined. Fits to independent data obtained for the K28E mutant, $\text{A}\beta^{1-42-\text{K28E}}$, with the site of substitution relatively distant from the C-terminus, yield m and A values that fall very close to those of the wild type peptide (Table 1). The values for δ_f (Table S3) also are in good agreement between the two peptides (rmsd = 0.1, 0.9, 0.3, 0.3 ppm for $^1\text{H}^N$, ^{15}N , $^{13}\text{C}^\alpha$, and $^{13}\text{C}'$ of residues 34–41), consistent with their nearly indistinguishable chemical shifts in the absence of denaturant.

For $\text{A}\beta^{1-42-\text{Ox}}$, the same analysis yields a *ca* 35% lower A value (Table 1), indicating that the folded state is destabilized but not eliminated by the oxidation of M35. The δ_f values obtained for this peptide differ most from those of the non-oxidized peptide for residues V36 and G37 (Table S3), suggesting that the backbone structure of the transiently folded state is impacted by the methionine oxidation and that the destabilization, presumably resulting from less favorable hydrophobic packing in the folded conformer for the oxidized M35 sidechain that contains the more polar sulfoxide moiety.

Even though the C-terminally truncated peptides, $\text{A}\beta^{1-40}$ and $\text{A}\beta^{1-40-\text{Ox}}$, yield values for A and m that are comparable to those of $\text{A}\beta^{1-42}$ peptides (Table 1), the δ_f values are quite different (Table 2 and Table S3), in particular for residues G37 and G38, pointing to a structure for the transiently folded conformer of these two peptides that is distinctly different from $\text{A}\beta^{1-42}$.

Table 2

Fitted δ_f values for the C-terminal residues of the transiently folded $\text{A}\beta^{1-42}$ conformer. Those of the other four $\text{A}\beta$ peptides are presented in SI Table S3.

Res.	δ_f ($^1\text{H}^N$) (ppm)	δ_f (^{15}N) (ppm)	δ_f ($^{13}\text{C}^\alpha$) (ppm)	δ_f ($^{13}\text{C}'$) (ppm)
I32	7.62 ± 0.16	120.3 ± 0.9	176.3 ± 0.4	61.5 ± 0.4
G33	7.75 ± 0.16	108.5 ± 0.9	171.6 ± 0.4	45.4 ± 0.4
L34	7.77 ± 0.16	122.2 ± 0.9	173.9 ± 0.4	54.3 ± 0.4
M35	8.13 ± 0.16	122.3 ± 0.9	176.8 ± 0.4	53.4 ± 0.4
V36	8.42 ± 0.16	122.2 ± 0.9	178.0 ± 0.4	60.8 ± 0.4
G37	9.46 ± 0.16	122.9 ± 0.9	176.0 ± 0.4	48.2 ± 0.4
G38	8.69 ± 0.16	102.4 ± 0.9	171.6 ± 0.4	45.5 ± 0.4
V39	6.61 ± 0.15	118.6 ± 0.9	172.7 ± 0.4	59.1 ± 0.4
V40	8.24 ± 0.16	124.4 ± 0.9	177.0 ± 0.4	63.2 ± 0.4
I41	8.53 ± 0.16	127.0 ± 0.9	175.2 ± 0.4	60.4 ± 0.4

2.4. Structure of the transiently folded $\text{A}\beta^{1-42}$ conformer

The δ_f values obtained for $\text{A}\beta^{1-42}$ show remarkable deviations from random coil values. In particular, at 122.9 and 102.3 ppm, the ^{15}N chemical shifts of G37 and G38 are quite distinct, as are the downfield shifts of the $^1\text{H}^N$ and $^{13}\text{C}'$ of G37. Whereas we lack the NOE or residual dipolar coupling (RDC) restraints required for a conventional structure determination, the unusual pattern of backbone chemical shifts suggests it should be possible to determine if these shifts correspond to known structural elements. Hence, we explored whether any previously assigned proteins of known structure contained chemical shift patterns similar to those we derived for $\text{A}\beta^{1-42}$.

Therefore, we searched the SPARTA+ database for 10-residue protein fragments that best match the backbone secondary chemical shifts and residue types for the I32-I41 segment of $\text{A}\beta^{1-42}$. The SPARTA+ database contains experimental backbone and $^{13}\text{C}^\beta$ chemical shifts as well as high-resolution X-ray structures for 580 proteins [50].

In a first pre-screening step, a sequence similarity score S_{seq} was calculated by using the BLOSUM62 amino acid similarity matrix [58]:

$$S_{\text{seq}} = \sum_i w_i \times B \left(A_{32+i}^{\text{A}\beta} A_{r_0+i}^{\text{DB}} \right) \quad (9)$$

where $i = 0-9$; w_i is the weight for residue i in the 10-residue fragment ($w_{0,9} = 0$; $w_{1-4,7,8} = 1$; $w_{5,6} = 2$) of starting residue number r_0 ; $B(a_1, a_2)$ is the BLOSUM62 similarity value for amino acids a_1 and a_2 ; and $A_{32+i}^{\text{A}\beta}$ and $A_{r_0+i}^{\text{DB}}$ are the amino acid type for residues in $\text{A}\beta$ and the database fragment, respectively. Only database fragments with a positive sequence similarity score S_{seq} are retained for subsequent chemical shift evaluation.

In the subsequent step, we derived the chemical shift matching relative to the above screened 10-residue fragments, using the following χ^2 function:

$$\chi^2 = \sum_i \sum_X w_i \times \left(\frac{\Delta\delta(X)_{32+i}^{\text{A}\beta} - \Delta\delta(X)_{r_0+i}^{\text{DB}}}{W_X} \right)^2 \quad (10)$$

where $i = 0-9$; $X = ^{13}\text{C}^\alpha, ^{13}\text{C}', ^{15}\text{N}$ and $^1\text{H}^N$; W_X is the weight for the chemical shift of different nuclei X ($W_{\text{C}\alpha} = 1.0$, $W_{\text{C}'} = 1.1$, $W_{\text{N}} = 2.5$, $W_{\text{H}^N} = 0.5$), w_i are the weights for residue type similarity defined under Eq. (9), $\Delta\delta(X)_{32+i}^{\text{A}\beta}$ is the secondary chemical shift of nucleus X of residue $32 + i$ in $\text{A}\beta$, and $\Delta\delta(X)_{r_0+i}^{\text{DB}}$ is the secondary chemical shift of X of residue $r_0 + i$ in the database fragment.

The ten top scoring fragments obtained using the above procedure are listed in Table 3. Remarkably, 8 out of 10 fragments contain a Type I' β -turn at positions that coincides with G37 and G38 in $\text{A}\beta^{1-42}$, and a single fragment shows a Type II β -turn at this position. One fragment that matches fairly well in residue type similarity, but that shows poor agreement with the ^{15}N chemical shifts of G37 and G38, lacks the reverse turn and is not considered in the analysis below. An overlay of the nine remaining segments shows a fairly well-defined antiparallel β -sheet, connected by the reverse turn (Fig. 5). Not surprisingly, considering the close similarity in δ_f values, a nearly identical set is obtained when searching for fragments that match the δ_f values of $\text{A}\beta^{1-42-\text{K28E}}$. This contrasts with the absence of a unique structural signature when searching the database for fragments that match δ_f values of $\text{A}\beta^{1-40}$ and $\text{A}\beta^{1-40-\text{Ox}}$.

The difference in solvent accessible surface area, ΔASA , derived using the DSSP program [59], between the nine 10-residue β -hairpin fragments excerpted from their respective X-ray structures, and their fully extended states ($\phi = -120^\circ$; $\psi = 120^\circ$) is $369 \pm 73 \text{ \AA}^2$ (see Table S4 for individual values). This value is well within the $400 \pm 100 \text{ \AA}^2$ range predicted by substituting the m -value of 0.06 ± 0.01 kcal/mol-M obtained from the fit in the empirical $\Delta\text{ASA}/m$ relation parameterized by Hong et al. (Eq. (3)) [47].

Table 3Protein fragments of known structure that best match the backbone secondary chemical shifts and residue types of the transiently folded segment I32-I41 in A β^{1-42} .

PDB ID	r ₀	sequence	χ^2	XGGX motif	$\delta_i^{15}\text{N}$ (G37) (ppm)	$\delta_i^{15}\text{N}$ (G38) (ppm)	$\delta_i^1\text{H}$ (G37) (ppm)	$\delta_i^1\text{H}$ (G38) (ppm)	d($^{36}\text{H}^{\text{N}}, ^{39}\text{H}^{\text{N}}$) ^a (Å)	d($^{41}\text{H}^{\text{N}}, ^{35}\text{H}^{\text{N}}$) ^a (Å)
1GRN	42	VTVMIGGEPY	38	β -Turn I'	119.96	105.34	9.385	8.606	3.12	3.35
2E7P	76	PNVFIGGKQI	41	β -Turn I'	120.59	106.28	10.514	9.556	3.26	2.72
1QBS	11	VTIKIGGQLK	46	β -Turn I'	120.11	107.27	9.590	8.730	3.19	3.57
1NOS	86	HSRTVGGYT	53	β -Turn II	119.28	107.38	9.210	8.370	3.33	3.59
1TVQ	109	ETITFGGVTL	63	β -Turn I'	116.70	105.50	8.472	8.533	3.04	3.26
1VM9	54	EGSYEGGVI	67	β -Turn I'	116.11	105.63	9.166	8.575	3.00	3.80
2NMZ	11	VTIKIGGQLK	71	β -Turn I'	118.50	106.01	9.611	8.864	2.68	3.73
2PE8	76	NGRYFGGRVV	77	β -Turn I'	117.50	106.72	8.685	7.772	2.88	3.77
1FZY	45	GTPYEGGKVF	85	None	115.75	107.32	9.360	8.300	8.99	10.93
2FA4	76	LIFYKGGKEV	85	β -Turn I'	117.09	106.83	9.292	9.472	2.95	2.54
A β^{1-42}	32	IGLMVGGVVI	–	–	122.91	102.28	9.460	8.697	–	–

^a d($^{36}\text{H}^{\text{N}}, ^{39}\text{H}^{\text{N}}$) and d($^{41}\text{H}^{\text{N}}, ^{35}\text{H}^{\text{N}}$) are the ^1H - ^1H distances in units of Å in the X-ray structure fragments for the only two medium-range NOEs previously observed in the C-terminal segment of A β^{1-42} .

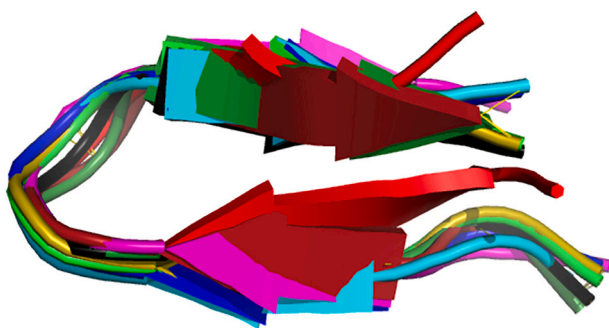


Fig. 5. Overlay of nine protein segments with secondary chemical shifts and residue types that best match the folded chemical shift (δ_i) values and residue types of I32-I41 obtained for A β^{1-42} (cf Eq. (8)). These segments resulted from a search over the SPARTA+ database that contains experimental chemical shifts for 580 proteins [50], with the best-matching segments listed in Table 3. Colors match those of the keys in Fig. 8.

2.5. Validation of the folded structure by NOE and PRE measurements

Although the above proposed structure is based on experimentally derived chemical shifts, its selection was also based on statistics and therefore only represents the highest likelihood for the structure transiently sampled by this region of A β^{1-42} . For example, the Type II β -turn element comprising residues H86-T95 of lipocalin FluA (PDB entry 1NOS) matches about as well as the other peptide sequences that contain a Type I' β -turn (Table 3). Perhaps of greater concern is the fact that one of the top ten peptides did not contain the antiparallel β -sheet but nevertheless provided a reasonable match to both residue types and chemical shifts. Therefore, it is highly desirable to have independent validation that the “consensus” represented by the eight very similar backbone structures is correct. Below, we present NOE and PRE data that are fully consistent with this β -sheet structure.

At a fractional population of $\sim 10\%$, the effective concentration of the C-terminal β -sheet in A β^{1-42} is only $\sim 9 \mu\text{M}$ for a peptide concentration of $90 \mu\text{M}$, where the peptide remains stable for prolonged durations. This low concentration combined with the relatively short rotational correlation time expected for a small structural element makes it challenging to record well-resolved 3D NOESY spectra of sufficient signal to noise. However, by resorting to high hydrostatic pressure (3 kbar), we previously were able to collect such spectra on a 1.2 mM A β^{1-42} sample while it underwent aggregation and incurred a $\sim 75\%$ signal loss over a 36 h period [48]. By using time-ordered non-uniform sampling, a 3D NOESY-HSQC spectrum of high resolution and sensitivity was obtained. That study was aimed at detecting long-range NOEs, but only two such interactions could be identified: NOEs between I41- H^{N} and M35- H^{α} and between V36- H^{N} and V39- H^{N} . As can be seen in

Table 3 (columns 10 and 11), all the corresponding ^1H - ^1H distances for the nine β -sheet fragments extracted from the PDB are consistent with the presence of these observed NOEs, whereas the single, extended conformer is not. We note that the increased concentration used for the NOE measurement at high pressure had no detectable effect on the backbone $^1\text{H}^{\text{N}}$ and ^{15}N chemical shifts, while the short-range diagonal/cross peak ratios in the 3D NOESY-HSQC spectrum stayed close to those of random coil, only increasing by the amount expected from the increase in viscosity under increased pressure. These observations indicate that increasing the A β^{1-42} concentration at high pressure did not lead to significant transferred NOE effects, and that the observed NOEs must correspond to interactions within the monomeric species.

Paramagnetic relaxation enhancement (PRE) is an excellent tool for identifying transient structures and has been widely used to identify short-lived contacts in IDPs [43,60–62]. Computational work suggested that adding the requisite chemical modification of tagging the protein with a methanethiosulfonate spin label (MTSL) at the sidechain of a Cys residue can strongly perturb the structural ensemble [63]. To investigate whether such perturbations occur upon A42C substitution of A β^{1-42} to Cys, and tagging it with MTSL, we measured their impact on the backbone amide chemical shifts. Although chemical shift perturbations in the N-terminal half of the peptide are as large as 0.15 ppm (^{15}N) and 0.03 ppm ($^1\text{H}^{\text{N}}$) (SI Fig. S2), these differences correlate fairly well with their pH dependence (SI Fig. S3) and therefore are likely to be dominated by a small change in protonation state of the three His residues between the separately prepared samples. However, somewhat larger chemical shift changes are seen for residues I32-I41 that adopt the transient β -sheet structure and which are insensitive to slight pH changes. Here, the chemical shift perturbation correlates with the differences in chemical shifts between A β^{1-42} and A $\beta^{1-42}\text{-Ox}$ (SI Fig. S6). The ~ 1.35 slope of this correlation indicates that the A42C mutation together with MTSL tagging stabilizes the β -sheet conformation by about 25% (SI Fig. S6). This analysis shows that the A42C mutant of A β^{1-42} , tagged with MTSL, provides a suitable probe for evaluating the presence of remote contacts, predicted by the β -sheet arrangement of Fig. 5.

Comparison of the ^1H - ^{15}N HSQC spectra, before and after reduction of the MTSL tag, shows that the resonances of, for example, G33 and L34 are absent in the sample with the paramagnetic MTSL tag attached to C42, whereas the intensity of the G38 amide is little impacted (Fig. 6). Indeed, resonances of all amides of I32-M35 are attenuated to below the detection limit (Fig. 7) whereas those of V36-V39 retain substantial intensity, despite being closer to the tag in terms of sequence separation. Upon addition of 8 M urea, the spectral attenuation becomes much smaller and all but the amides of the two C-terminal residues (and the two N-terminal residues, which are obliterated by rapid hydrogen exchange in both dia- and paramagnetic samples), are observed (Fig. 7). This result demonstrates that the remote PRE effects are strongly reduced, indicating that the transient structural elements must be responsible for the remote PREs. Paramagnetic relaxation scales with

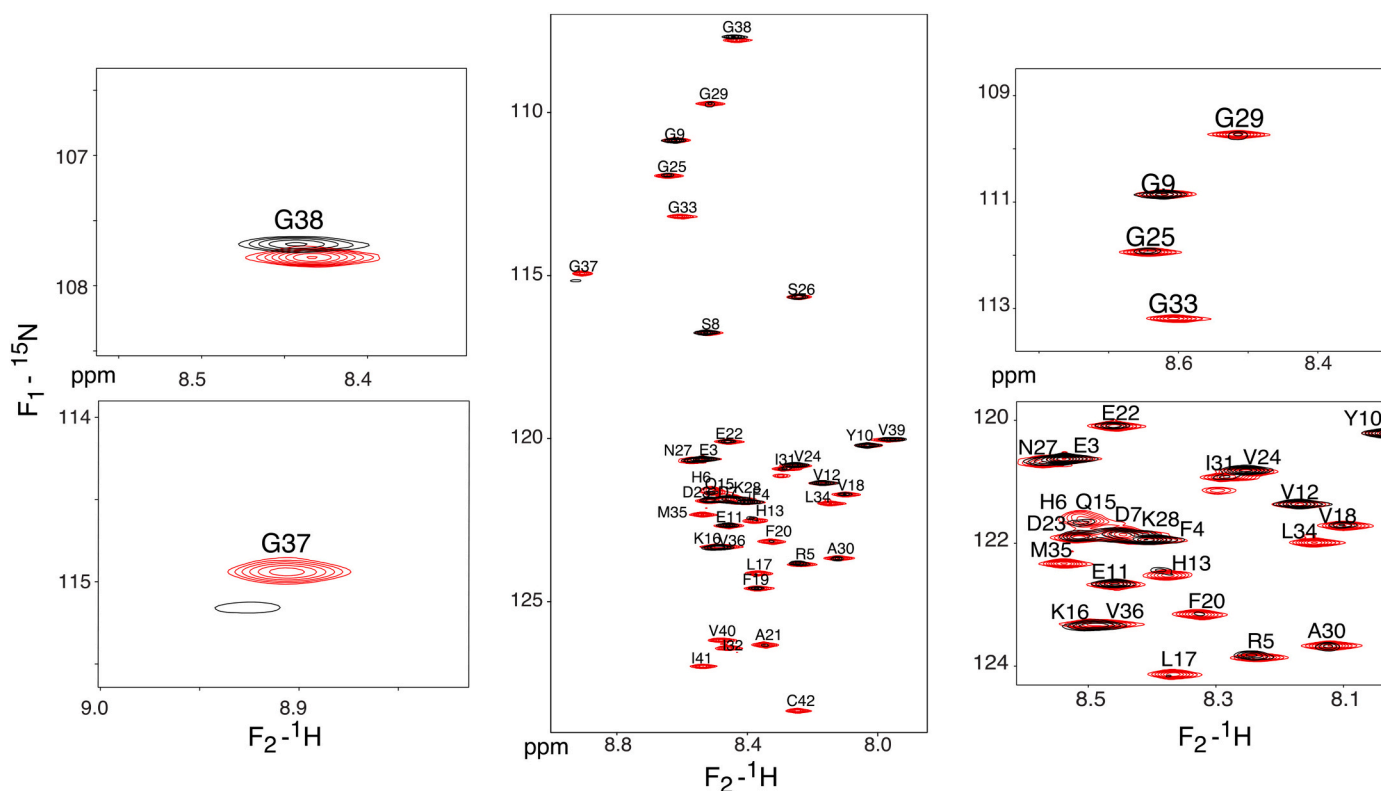


Fig. 6. Overlay of the 800-MHz ^1H - ^{15}N HSQC spectra of 75 μM $\text{A}\beta^{1-42}\text{-A42C-MTSL}$ in 44 mM Tris-HCl, 10 mM NaCl, 2 mM EDTA, pH 7.4, 278 K, recorded before (black) and after (red) reduction of the MTSL spin label by addition of a 50-fold molar excess of sodium ascorbate, pH 7.4. Both HSQC spectra were processed with 15 Hz exponential line broadening in the ^1H dimension. Smaller panels correspond to expanded regions of the full spectrum. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the inverse sixth power of the distance between the MTSL tag and the observed amide protons. Whereas it remains challenging to collect accurate quantitative PRE data on the dilute MTSL-tagged sample, which appears somewhat more prone to aggregation than wild-type $\text{A}\beta^{1-42}$, the experimental PRE attenuation pattern agrees closely with the interatomic distances observed in the nine peptides that were selected on the basis of their sequence similarity and chemical shifts (Fig. 8). This result therefore provides additional validation for the transient presence of the antiparallel β -sheet.

3. Concluding remarks

Over the past decade, the importance of transient structure in intrinsically disordered proteins has been increasingly recognized. In particular, considering that many IDPs are involved in signaling pathways, their ability to adopt transient secondary structure is believed to promote the conformational selection mechanism for forming stable protein-protein interactions [12,64,65]. Most prior experimental examples for the importance of such transient structures pertain to α -helical elements, frequently seen present at levels ranging from 30 to 70% [36,66,67]. Our current observation of a transient, well-defined β -sheet structure at the C-terminus of $\text{A}\beta^{1-42}$ extends the paradigm of pre-ordered transient structure to non- α -helical elements. Considering the low population of this transient β -sheet of only $\sim 10\%$, it is not surprising this element has remained unrecognized despite exhaustive study by a wide range of advanced molecular biophysics techniques [8,25,68].

Although the biological relevance of the transient β -sheet structure at such a low population may be questioned, there are a few observations that potentially link this element to the peptide's aggregation propensity. First, a destabilization of the structured element is seen upon oxidation of M35, lowering its population by *ca* 35%, which is consistent with its lower propensity to aggregate [18]. Qualitatively, the opposite

is seen for the A42C mutant, tagged with the hydrophobic MTSL moiety. Addition of this tag stabilizes the β -sheet structure, as judged by increased deviations for the G37 and G38 amide chemical shifts from fully unfolded values (SI Fig. S6), and also leads to increased difficulty to keep the peptide in monomeric solution, although quantitative aggregation kinetics have not yet been recorded for this peptide. Although $\text{A}\beta^{1-40}$ and its M35-oxidized variant also show propensity to form ordered, β -sheet type secondary structure at the C-terminus, the chemical shift signatures of G37 and G38 upon increasing the urea concentration distinctly differ from $\text{A}\beta^{1-42}$, and are inconsistent with the formation of the relatively rare, type I' β -turn in $\text{A}\beta^{1-40}$. Unfortunately, the δ_f values extracted for the folded conformer were insufficiently unique to identify the details of its structure other than pointing at backbone torsion angles in the β region of the Ramachandran map for L34-V36 and V39. Comparison of the δ_f values of the transiently folded conformers for the five different forms of the $\text{A}\beta$ peptides studied here (SI Fig. S7) shows that $\text{A}\beta^{1-40}$ and $\text{A}\beta^{1-42}$ peptides differ most pronouncedly for residues V36-V39 but, with the exception of L34-C', show comparable values for I31-L34. If the transient structure adopted in the C-terminal region of $\text{A}\beta^{1-40}$ also is a small β -sheet, it likely is centered at G37 as seen in some of the MD trajectories.

Our analysis utilizes the empirical knowledge that secondary chemical shifts in folded structural elements, on average, are non-zero, which forms, for example, the basis for the widely used secondary-chemical-shift-based RCI order parameter estimation [69]. To further test whether the database derived rms secondary chemical shift values are applicable for the β -hairpin structure we identified, we also calculated the corresponding rms values for the nine matching fragments from the database (Table S5). These values show that indeed these β -hairpin elements have secondary chemical shifts close to the values we used for T_X in Eq. (8).

It is important to note that the population of the folded conformer,

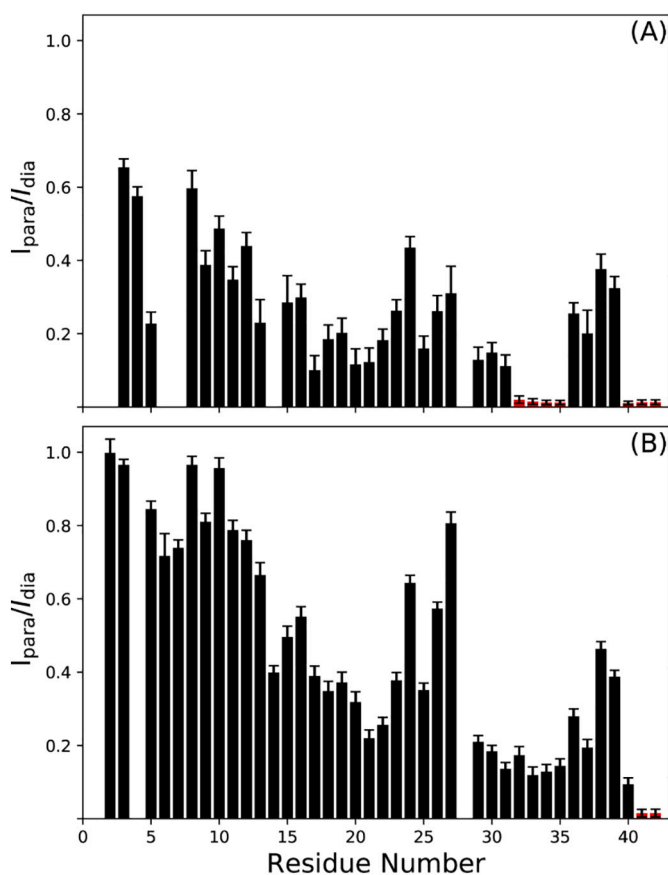


Fig. 7. Peak intensity ratios in ^1H - ^{15}N HSQC spectra of the paramagnetic and diamagnetic forms of $\text{A}\beta^{1-42}\text{-A42C-MTSL}$. (A) Spectra recorded in the absence of urea (Fig. 6). (B) Ratios obtained from spectra recorded under identical conditions except for the addition of 8 M urea. Red bars represent the upper limit for the attenuation ratios, based on the absence of detectable intensity in the paramagnetic state. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

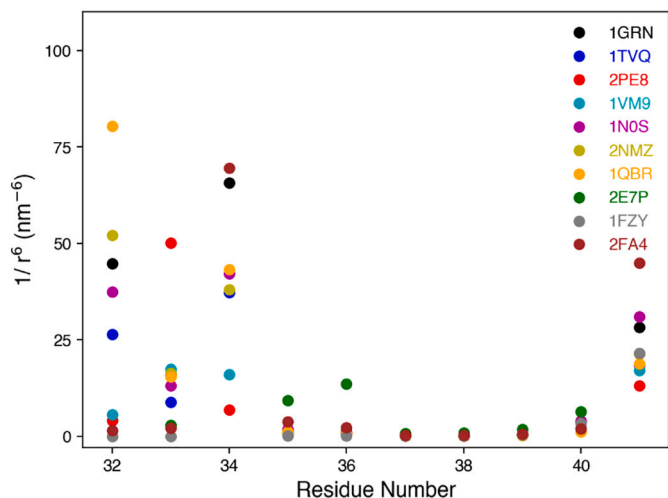


Fig. 8. Plot of the inverse sixth power of the distance between amide protons and the $\text{C}\beta$ atom of the residue that matches C42 in $\text{A}\beta^{1-42}\text{-A42C-MTSL}$. The intramolecular distances are shown for the ten best matching protein fragments listed in Table 3, nine of which contain an antiparallel β -sheet. For 2E7P, which contains a Gly residue at the position that matched C42, a pseudo- $\text{C}\beta$ atom was defined by extending the length of the $\text{C}^\alpha\text{-H}^{\alpha 3}$ bond.

extracted from the chemical shift data, is inextricably linked to the rms values, T_X , of secondary chemical shifts expected in a folded structure, because only the product $A(\delta_f - \delta_u)$ is uniquely extracted from the urea titration (cf Eq. (7)). For each nucleus, the rms spread in secondary shifts among the nine database peptides is 20–30% (Table S5). However, analogous to the RCI order parameter [69], Eq. (8) simultaneously uses the rms value of all normalized secondary chemical shift types for restraining the extracted A value, thereby reducing its uncertainty to about 15%.

There are a number of solid-state NMR (PDB entries 2MXU, 6T15-7, 5KK3, 2NAO and 2BEG) [70–74] and electron microscopy (5OQV, 6SHS and 5AEF) [75–77] structures for $\text{A}\beta^{1-42}$. Formation of multiple β -strands connected by loops or turns is a common feature of these fibrils. In all these structures, the β -strands form intermolecular H-bonds and are interacting intramolecularly via side chains. Specifically, the intramolecular interaction in the C-terminal region is through the hydrophobic side chains. Although the overall backbone conformation of the β -hairpin structure we report here for monomeric $\text{A}\beta^{1-42}$ is similar to that seen in various fibril structures, a pronounced distinguishing feature is that the H-bonds between its two short β -strands are intramolecular. Therefore, even though the backbone conformation may appear similar, breaking and reforming of a substantial number of H-bonds is required for a monomer to cross the energy barrier required for elongating a fibril structure.

Solution NMR studies in the presence of added apolar solvents (PDB entries 1IYT and 1Z0Q) [75,76] showed α -helices at various regions of the $\text{A}\beta^{1-42}$ chain and their potential importance in fibril formation remains under investigation. Interestingly, addition of dodecylphosphocholine recently was reported to induce formation of $\text{A}\beta$ tetramers (PDB entry 6RHY) and octamers with β -sheet arrangements; however these differ substantially from the β -sheet observed in the present study [77]. There also have been multiple solution NMR studies of $\text{A}\beta$ in its monomeric form in the absence of additives, but none of these studies identified the β -hairpin reported here [19,21,23,30].

Prior to the development of the newer force field optimizations, a number of MD studies already reported the presence of a β -hairpin in the general vicinity of G37 and G38 [78–80]. However, the type of turn was not discussed, and no coordinates for these structures are available, preventing a more quantitative comparison between these models and our experimental results. We consider it likely that with the newly improved force fields [28,31–33,81] it will become possible to reliably identify transient structural elements such as identified by our work.

The K28E mutation of $\text{A}\beta^{1-42}$, intended to disrupt the salt bridge interaction with D23 that is seen in solid-state NMR fibril structures [82], causes only small changes in chemical shifts for the C-terminal 10 residues. Study of an $\text{A}\beta^{1-42}$ fragment (A21-A30) confirmed that the presence of a turn in this region is stabilized by this salt bridge [83]. Indeed, we also observe small, remote chemical shift perturbations upon K28E substitution that potentially can be attributed to the transient presence of such a turn (SI Fig. S2 and Table S3). In addition, the K28E mutation causes small changes in the chemical shifts of residues A30-G37. However, urea denaturation of the transient C-terminal β -sheet in $\text{A}\beta^{1-42}$ is not affected by the K28E substitution. This latter observation suggests that in the monomeric $\text{A}\beta^{1-42}$ peptide, transient formation of the salt-bridge-stabilized turn and transient formation of the C-terminal β -sheet are independent events.

To the best of our knowledge, no prior PRE solution NMR studies have been reported for $\text{A}\beta^{1-40}$ or $\text{A}\beta^{1-42}$. Importantly, the PRE effect observed in our study does not monotonically decrease with increasing distance from the labeling site, as would be expected for random coil behavior. Instead, the rich PRE pattern seen in Fig. 7A points to the presence of multiple long-range structural features beyond the β -sheet element identified in our study. The reduction in PRE magnitude upon addition of urea suggests that it may become possible to identify these transient features in atomic detail by the TSD-NMR approach.

3.1. Experimental section

3.1.1. Pentapeptides

Pentamer peptides of the sequence Ac-Q-X_{i-1}-X_i-X_{i+1}-Q-NH₂ were chemically synthesized at natural abundance for residues I32-I41 of Aβ¹⁻⁴². A different pentamer peptide was used for *i* = 40 and 41 (Ac-Q-V₃₉-V₄₀-I₄₁-A₄₂), lacking C-terminal amidation. Likewise, for the M35 residue, the octamer peptide Ac-QL₃₄M₃₅V₃₆QKKK with charged Lys residues and lacking amidation at the C-terminus was used instead of Ac-QLMVQ-NH₂ which has very limited solubility. The methionine side chains of the short peptides corresponding to L34, M35, and V36 of M35-oxidized Aβ peptides were oxidized using hydrogen peroxide. In a typical procedure, the short peptide (5 μmol) was stirred in an aqueous solution of hydrogen peroxide (4 mL, 6%) for 4 h at ambient temperature, followed by lyophilization. All the NMR samples were in 45 mM Tris-HCl, 20 mM KCl, 2% D₂O, pH 6.8, at 10 mM peptide concentration, or at saturating concentrations for peptides with lower solubility, containing 2 mM internal DSS for chemical shift referencing. All NMR spectra were recorded at 278 K, and the molar ratios of water to urea were derived by integrating the intensities of the urea and HDO signals in a 1D ²H spectrum [84].

3.1.2. Aβ peptides

The Aβ¹⁻⁴² peptides were generated from an expression construct consisting of a 6-His tag followed by the immunoglobulin binding domain B1 of protein G (GB1), Avi-Tag (Avidity, LLC), TEV protease cleavage site, and Aβ¹⁻⁴². Details of the protocol are as described by Ying et al. [48]. Oxidation of M35 was carried out by addition of hydrogen peroxide (3 μL, 7.5% w/v) to 600 μL of 50 μM Aβ peptides in 22.5 mM Tris-HCl, 10 mM KCl, pH 6.8. The reaction was carried out at 293 K for 4 h. The resulting product was lyophilized to remove excess hydrogen peroxide and redissolved in 300 μL water containing 2% D₂O and 0.15 mM DSS. The pH was adjusted to 6.8 using a glass electrode at ambient temperature.

3.1.3. Aβ NMR sample preparation

All NMR measurements were carried out in 45 mM Tris-HCl, 20 mM KCl, pH 6.8 with 2% D₂O and 0.15 mM DSS. Briefly, lyophilized uniformly ¹³C/¹⁵N-enriched Aβ peptide (1.2 mg) was initially dissolved in 50 mM KOH to obtain a 1 mM stock solution. Next, in a separate vial on ice, 27 μL of 50 mM HCl was added to 245 μL of 50 mM Tris-HCl buffer containing 25 mM KCl, pH 6.8, which was subsequently added to 27 μL of the 1 mM peptide stock, to reach 90 μM peptide concentration. The pH was checked using a glass electrode, measured at ambient temperature, and adjusted to pH 6.8 with addition of dilute KOH or HCl solutions if deviating by more than 0.1 pH units from this target value.

3.1.4. NMR spectroscopy

All NMR data were acquired at 278 K. NMR spectra on pentapeptides were recorded on a Bruker Avance III 600-MHz spectrometer equipped with a cryogenic TCI probe, and direct ¹³C detection was used to measure the natural abundance ¹³C spectra of the short peptides. ¹H^N and ¹⁵N chemical shifts of these peptides were obtained from ¹H-¹⁵N HSQC spectra.

Aβ peptide triple resonance NMR spectra were recorded on a Bruker Avance III 700 MHz spectrometer equipped with a cryogenic TCI probe. Urea titration data were generated by stepwise addition of solid urea to a sample of uniformly ¹³C/¹⁵N-enriched Aβ peptide. The urea concentrations reported were estimated by taking the ratio of the urea and water signals, measured by ²H 1D spectra [84]. At each titration step, 3D HNCO and 3D HN(CO)CA spectra were acquired. For the 3D HNCO, the time domain matrix consisted of 85* (t₁, ¹³C) × 165* (t₂, ¹⁵N) × 1024* (t₃, ¹H) hypercomplex points with acquisition times of 96 ms (t₁), 106 ms (t₂) and 104 ms (t₃), using 2 scans per FID and 1 s interscan delay with ¹³C and ¹⁵N carriers at 175.5 and 117.5 ppm respectively. Spectral widths for ¹H, ¹⁵N and ¹³C dimensions were set to 14, 22 and 5 ppm

respectively. For the 3D HN(CO)CA, the time domain matrix consisted of 103* (t₁, ¹³C) × 165* (t₂, ¹⁵N) × 1024* (t₃, ¹H) hypercomplex points with acquisition times of 26.7 ms (t₁), 105.6 ms (t₂) and 104.4 ms (t₃), using 2 scans per FID and a 1 s interscan delay, with ¹³C and ¹⁵N carriers at 54.1 and 117.5 ppm, respectively. Spectral widths for the ¹H, ¹⁵N and ¹³C dimensions were 14, 22 and 21.8 ppm respectively. The ¹H carrier was set to 4.964 ppm for both experiments. Both the HNCO and HN(CO)CA 3D spectra were recorded in non-uniform sampling mode with 400 and 800 (t₁, t₂, t₃) hypercomplex points, respectively. Data was reconstructed using the SMILE method [55,85], to yield a final digital resolution of 2.4, 1.5, 1.7 Hz in the ¹H, ¹⁵N, and ¹³C dimensions of the HNCO spectrum respectively, and 2.4, 1.5, and 3.8 Hz for the HNCOCA spectrum. ¹H chemical shifts were referenced directly to DSS, and the ¹³C and ¹⁵N chemical shifts were defined indirectly relative to this reference [53]. All spectra were processed with the NMRPipe software [86], and analyzed using CcpNmr [87].

Data availability

Data will be made available on request.

Author statement

Tayeb Kakeshpour: NMR measurements, software development, data analysis, manuscript writing; Venkat Ramanujam: NMR measurements, data analysis, manuscript editing; Ashley Barnes: NMR measurements; data analysis; Yang Shen: database mining; Jinfa Ying: optimization of data collection; Ad Bax: directed the study, edited manuscript.

Declaration of Competing Interest

None.

Acknowledgments

We thank Galina Abdoulaeva for peptide synthesis, James L. Baber for technical support, John M. Louis for advice with protein expression, Angus Robertson for a sample of ¹³C/¹⁵N-labeled Aβ¹⁻⁴⁰ peptide, John Lloyd and the NIDDK Core Facility as well as Duck-Yeon Lee and the NHLBI Biochemistry Core Facility for mass spectrometry, and Dennis A. Torchia, G. Marius Clore, Hoi-Sung Chung, and Robert Best for valuable discussions. This work was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bpc.2020.106531>.

References

- [1] G. Meisl, X. Yang, E. Hellstrand, B. Frohm, J.B. Kirkegaard, S.I.A. Cohen, C. M. Dobson, S. Linse, T.P.J. Knowles, Differences in nucleation behavior underlie the contrasting aggregation kinetics of the A beta 40 and A beta 42 peptides, *Proc. Natl. Acad. Sci. U. S. A.* 111 (26) (2014) 9384–9389.
- [2] A.E. Roher, J.D. Lowenson, S. Clarke, A.S. Woods, R.J. Cotter, E. Gowing, M.J. Ball, Beta-amyloid-(1-42) is a major component of cerebrovascular amyloid deposits - implications for pathology of Alzheimer-disease, *Proc. Natl. Acad. Sci. U. S. A.* 90 (22) (1993) 10836–10840.
- [3] R. van der Lee, M. Buljan, B. Lang, R.J. Weatheritt, G.W. Daughdrill, A.K. Dunker, M. Fuxreiter, J. Gough, J. Gsponer, D.T. Jones, P.M. Kim, R.W. Kriwacki, C. J. Oldfield, R.V. Pappu, P. Tompa, V.N. Uversky, P.E. Wright, M.M. Babu, Classification of intrinsically disordered regions and proteins, *Chem. Rev.* 114 (13) (2014) 6589–6631.
- [4] K. Honjo, S.E. Black, N. Verhoeff, Alzheimer's disease, cerebrovascular disease, and the beta-amyloid cascade, *Can. J. Neurol. Sci.* 39 (6) (2012) 712–728.
- [5] D.M. Walsh, D.J. Selkoe, Oligomers in the brain: the emerging role of soluble protein aggregates in neurodegeneration, *Protein Pept. Lett.* 11 (3) (2004) 213–228.

- [6] T.A. Thibaut, R.T. Anderson, D.M. Smith, A common mechanism of proteasome impairment by neurodegenerative disease-associated oligomers, *Nat. Commun.* 9 (2018).
- [7] N.L. Fawzi, J. Ying, D.A. Torchia, G.M. Clore, Kinetics of amyloid beta monomer-to-oligomer exchange by NMR relaxation, *J. Am. Chem. Soc.* 132 (29) (2010) 9948–9951.
- [8] J. Roche, Y. Shen, J.H. Lee, J. Ying, A. Bax, Monomeric β (1–40) and β (1–42) peptides in solution adopt very similar Ramachandran map distributions that closely resemble random coil, *Biochemistry* 55 (5) (2016) 762–775.
- [9] C.A. Barnes, A.J. Robertson, J.M. Louis, P. Anfinsen, A. Bax, Observation of β -amyloid peptide oligomerization by pressure-jump NMR spectroscopy, *J. Am. Chem. Soc.* 141 (35) (2019) 13762–13766.
- [10] P.E. Wright, H.J. Dyson, Linking folding and binding, *Curr. Opin. Struct. Biol.* 19 (1) (2009) 31–38.
- [11] K. Teilum, J.G. Olsen, B.B. Kragelund, Functional aspects of protein flexibility, *Cell. Mol. Life Sci.* 66 (14) (2009) 2231–2247.
- [12] V. Iesmantavicius, J. Dogan, P. Jemth, K. Teilum, M. Kjaergaard, Helical propensity in an intrinsically disordered protein accelerates ligand binding, *Angew. Chem. Int. Ed.* 53 (6) (2014) 1548–1551.
- [13] F. Chiti, C.M. Dobson, Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade, *Annu. Rev. Biochem.* 86 (2017) 27–68.
- [14] F. Chiti, M. Stefani, N. Taddei, G. Ramponi, C.M. Dobson, Rationalization of the effects of mutations on peptide and protein aggregation rates, *Nature* 424 (6950) (2003) 805–808.
- [15] X. Yang, G. Meisl, B. Frohm, E. Thulin, T.P.J. Knowles, S. Linse, On the role of sidechain size and charge in the aggregation of A β 42 with familial mutations, *Proc. Natl. Acad. Sci. U. S. A.* 115 (26) (2018) E5849–E5858.
- [16] M. Torquato, T.C.T. Michaels, K. Sanagavarapu, X.T. Yang, G. Meisl, S.I.A. Cohen, T.P.J. Knowles, S. Linse, Secondary nucleation in amyloid formation, *Chem. Commun.* 54 (63) (2018) 8667–8684.
- [17] A.A. Watson, D.P. Fairlie, D.J. Craik, Solution structure of methionine-oxidized amyloid β -peptide (1–40). Does oxidation affect conformational switching? *Biochemistry* 37 (37) (1998) 12700–12706.
- [18] L.M. Hou, I. Kang, R.E. Marchant, M.G. Zagorski, Methionine 35 oxidation reduces fibril assembly of the amyloid A β (1–42) peptide of Alzheimer's disease, *J. Biol. Chem.* 277 (43) (2002) 40173–40176.
- [19] L.M. Hou, H.Y. Shao, Y.B. Zhang, H. Li, N.K. Menon, E.B. Neuhaus, J.M. Brewer, I. J.L. Byeon, D.G. Ray, M.P. Vitek, T. Iwashita, R.A. Makula, A.B. Przybyla, M. G. Zagorski, Solution NMR studies of the A β (1–40) and A β (1–42) peptides establish that the met35 oxidation state affects the mechanism of amyloid formation, *J. Am. Chem. Soc.* 126 (7) (2004) 1992–2005.
- [20] R. Riek, P. Guntert, H. Dobeli, B. Wipf, K. Wuthrich, NMR studies in aqueous solution fail to identify significant conformational differences between the monomeric forms of two Alzheimer peptides with widely different plaque-competence, A β (1–40)(ox) and A β (1–42)(ox), *Eur. J. Biochem.* 268 (22) (2001) 5930–5936.
- [21] Y. Yan, C. Wang, A β 42 is more rigid than A β 40 at the C terminus: implications for A β aggregation and toxicity, *J. Mol. Biol.* 364 (5) (2006) 853–862.
- [22] A. Abelein, J.P. Abrahams, J. Danielsson, A. Graslund, J. Jarvet, J. Luo, A. Tiiman, S.K.T.S. Warmlander, The hairpin conformation of the amyloid beta peptide is an important structural motif along the aggregation pathway, *J. Biol. Inorg. Chem.* 19 (4–5) (2014) 623–634.
- [23] S. Vivekanandan, J.R. Brender, S.Y. Lee, A. Ramamoorthy, A partially folded structure of amyloid- β (1–40) in an aqueous environment, *Biochem. Biophys. Res. Commun.* 411 (2) (2011) 312–316.
- [24] S.T. Kumar, J. Leppert, P. Bellstedt, C. Wiedemann, M. Fandrich, M. Gorch, Solvent removal induces a reversible β -to- α switch in oligomeric A β peptide, *J. Mol. Biol.* 428 (2) (2016) 268–273.
- [25] F.J. Meng, M.M.J. Bellaiche, J.Y. Kim, G.H. Zerbe, R.B. Best, H.S. Chung, Highly disordered amyloid- β monomer probed by single-molecule FRET and MD simulation, *Biophys. J.* 114 (4) (2018) 870–884.
- [26] K.A. Merchant, R.B. Best, J.M. Louis, I.V. Gopich, W.A. Eaton, Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations, *Proc. Natl. Acad. Sci. U. S. A.* 104 (5) (2007) 1528–1533.
- [27] M. Aznauryan, L. Delgado, A. Soranno, D. Nettels, J.-r. Huang, A.M. Labhardt, S. Grzesiek, B. Schuler, Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS, *Proc. Natl. Acad. Sci. U. S. A.* 113 (37) (2016) E5389–E5398.
- [28] R.B. Best, W. Zheng, J. Mittal, Balanced protein–water interactions improve properties of disordered proteins and non-specific protein association, *J. Chem. Theory Comput.* 10 (2014) 5113–5124.
- [29] D.J. Rosenman, C.R. Connors, W. Chen, C. Wang, A.E. Garcia, A β monomers transiently sample oligomer and fibril-like configurations: ensemble characterization using a combined MD/NMR approach, *J. Mol. Biol.* 425 (18) (2013) 3338–3359.
- [30] K.A. Ball, A.H. Phillips, D.E. Wemmer, T. Head-Gordon, Differences in β -strand populations of monomeric A β 40 and A β 42, *Biophys. J.* 104 (12) (2013) 2714–2724.
- [31] P. Robustelli, S. Piana, D.E. Shaw, Developing a molecular dynamics force field for both folded and disordered protein states, *Proc. Natl. Acad. Sci.* 115 (21) (2018) E4758–E4766.
- [32] J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B.L. de Groot, H. Grubmüller, A.D. MacKerell, CHARMM36m: an improved force field for folded and intrinsically disordered proteins, *Nat. Methods* 14 (1) (2017) 71–73.
- [33] M. Carballo-Pacheco, B. Strodel, Comparison of force fields for Alzheimer's A : a case study for intrinsically disordered proteins, *Protein Sci.* 26 (2) (2017) 174–185.
- [34] M. Kjaergaard, F.M. Poulsen, Disordered proteins studied by chemical shifts, *Prog. Nucl. Magn. Reson. Spectrosc.* 60 (2012) 42–51.
- [35] G.W. Daughdrill, S. Kashtanov, A. Stancik, S.E. Hill, G. Helms, M. Muschol, V. Receveur-Brechot, F.M. Ytreberg, Understanding the structural ensembles of a highly extended disordered protein, *Mol. Biosyst.* 8 (1) (2012) 308–319.
- [36] D. Kurzbach, G. Platzer, T.C. Schwarz, M.A. Henen, R. Konrat, D. Hinderberger, Cooperative unfolding of compact conformations of the intrinsically disordered protein osteopontin, *Biochemistry* 52 (31) (2013) 5167–5175.
- [37] N.Y. Haba, R. Gross, J. Novacek, H. Shaked, L. Zidek, M. Barda-Saad, J.H. Chill, NMR determines transient structure and dynamics in the disordered C-terminal domain of WASp interacting protein, *Biophys. J.* 105 (2) (2013) 481–493.
- [38] W.K. Lim, J. Rosgen, S.W. Englander, Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group, *Proc. Natl. Acad. Sci. U. S. A.* 106 (8) (2009) 2595–2600.
- [39] B.J. Bennion, V. Daggett, The molecular basis for the chemical denaturation of proteins by urea, *Proc. Natl. Acad. Sci. U. S. A.* 100 (9) (2003) 5142–5147.
- [40] Y. Shen, R. Vernon, D. Baker, A. Bax, De novo protein structure generation from incomplete chemical shift assignments, *J. Biomol. NMR* 43 (2) (2009) 63–78.
- [41] J.F. Ying, C.A. Barnes, J.M. Louis, A. Bax, Importance of time-ordered non-uniform sampling of multidimensional NMR spectra of A β (1–42) peptide under aggregating conditions, *J. Biomol. NMR* 73 (8–9) (2019) 429–441.
- [42] G.M. Clore, J. Iwahara, Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes, *Chem. Rev.* 109 (9) (2009) 4108–4139.
- [43] C.W. Bertoncini, Y.S. Jung, C.O. Fernandez, W. Hoyer, C. Griesinger, T.M. Jovin, M. Zweckstetter, Release of long-range tertiary interactions potentiates aggregation of natively unstructured α -synuclein, *Proc. Natl. Acad. Sci. U. S. A.* 102 (5) (2005) 1430–1435.
- [44] B. Mateos, R. Konrat, R. Pierattelli, I.C. Felli, NMR characterization of long-range contacts in intrinsically disordered proteins from paramagnetic relaxation enhancement in C-13 direct-detection experiments, *ChemBioChem* 20 (3) (2019) 335–339.
- [45] R.F. Greene, C.N. Pace, Urea and guanidine-hydrochloride denaturation of ribonuclease, lysozyme, α -chymotrypsin, and β -lactoglobulin, *J. Biol. Chem.* 249 (17) (1974) 5388–5393.
- [46] J.K. Myers, C.N. Pace, J.M. Scholtz, Denaturant m-values and heat-capacity changes - relation to changes in accessible surface-areas of protein unfolding, *Protein Sci.* 4 (10) (1995) 2138–2148.
- [47] J. Hong, M.W. Capp, R.M. Saecker, M.T. Record, Use of urea and glycine Betaine to quantify coupled folding and probe the burial of DNA phosphates in lac repressor–lac operator binding, *Biochemistry* 44 (51) (2005) 16896–16911.
- [48] J. Ying, C.A. Barnes, J.M. Louis, A. Bax, Importance of time-ordered non-uniform sampling of multi-dimensional NMR spectra of A β (1–42) peptide under aggregating conditions, *J. Biomol. NMR* (2019), <https://doi.org/10.1007/s10858-019-00235-7>.
- [49] Y. Shen, O. Lange, F. Delaglio, P. Rossi, J.M. Aramini, G.H. Liu, A. Eletsky, Y.B. Wu, K.K. Singarapu, A. Lemak, A. Ignatchenko, C.H. Arrowsmith, T. Szyperski, G. T. Montelione, D. Baker, A. Bax, Consistent blind protein structure generation from NMR chemical shift data, *Proc. Natl. Acad. Sci. U. S. A.* 105 (12) (2008) 4685–4690.
- [50] Y. Shen, A. Bax, SPARTA plus : a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network, *J. Biomol. NMR* 48 (1) (2010) 13–22.
- [51] X. Han, L.K. Tamm, A host-guest system to study structure-function relationships of membrane fusion peptides, *Proc. Natl. Acad. Sci. U. S. A.* 97 (24) (2000) 13097–13102.
- [52] S.C. Chiliveri, J.M. Louis, A. Bax, Concentration-dependent structural transition of the HIV-1 gp41 MPER peptide into α -helical trimers, *Angew. Chem. Int. Ed.* (2020), <https://doi.org/10.1002/anie.202008804>.
- [53] J.L. Markley, A. Bax, Y. Arata, C.W. Hilbers, R. Kaptein, B.D. Sykes, P.E. Wright, K. Wuthrich, Recommendations for the presentation of NMR structures of proteins and nucleic acids (reprinted from pure and applied chemistry, vol 70, pgs 117–142, 1998), *J. Mol. Biol.* 280 (5) (1998) 933–952.
- [54] D. Rovnyak, D.P. Frueh, M. Sastry, Z.Y.J. Sun, A.S. Stern, J.C. Hoch, G. Wagner, Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction, *J. Magn. Reson.* 170 (1) (2004) 15–21.
- [55] J. Ying, F. Delaglio, D.A. Torchia, A. Bax, Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data, *J. Biomol. NMR* 68 (2017) 101–118.
- [56] P.E. Hansen, Isotope effects on chemical shifts of proteins and peptides, *Magn. Reson. Chem.* 38 (1) (2000) 1–10.
- [57] R.L. Newcomer, L.C.R. Fraser, C.M. Teschke, A.T. Alexandrescu, Mechanism of protein denaturation: partial unfolding of the P22 coat protein I-domain by urea binding, *Biophys. J.* 109 (12) (2015) 2666–2677.
- [58] E.V. Koonin, M.Y. Galperin (Eds.), Sequence - Evolution - Function: Computational Approaches in Comparative Genomics. Vol. 4, Kluwer Academic, Boston, 2003.
- [59] W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers* 22 (1983) 2577–2637.
- [60] M.D. Mukrasch, S. Bibow, J. Korukottu, S. Jeganathan, J. Biernat, C. Griesinger, E. Mandelkow, M. Zweckstetter, Structural polymorphism of 441-residue tau at single residue resolution, *PLoS Biol.* 7 (2) (2009) 399–414.

- [61] M.R. Jensen, M. Zweckstetter, J.R. Huang, M. Backledge, Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy, *Chem. Rev.* 114 (13) (2014) 6632–6660.
- [62] J.R. Allison, P. Varnai, C.M. Dobson, M. Vendruscolo, Determination of the free energy landscape of alpha-synuclein using spin label nuclear magnetic resonance measurements, *J. Am. Chem. Soc.* 131 (51) (2009) 18314–18326.
- [63] S. Sasmal, J. Lincoff, T. Head-Gordon, Effect of a paramagnetic spin label on the intrinsically disordered peptide ensemble of amyloid-beta, *Biophys. J.* 113 (5) (2017) 1002–1011.
- [64] M. Arai, K. Sugase, H.J. Dyson, P.E. Wright, Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding, *Proc. Natl. Acad. Sci. U. S. A.* 112 (31) (2015) 9614–9619.
- [65] T. Mittag, J.D. Forman-Kay, Atomic-level characterization of disordered protein ensembles, *Curr. Opin. Struct. Biol.* 17 (1) (2007) 3–14.
- [66] R. Arai, H. Ueda, A. Kitayama, N. Kamiya, T. Nagamune, Design of the linkers which effectively separate domains of a bifunctional fusion protein, *Protein Eng.* 14 (8) (2001) 529–532.
- [67] J.M.R. Baker, R.P. Hudson, V. Kanelis, W.Y. Choy, P.H. Thibodeau, P.J. Thomas, J. D. Forman-Kay, CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices, *Nat. Struct. Mol. Biol.* 14 (8) (2007) 738–745.
- [68] B.R. Sahoo, S.J. Cox, A. Ramamoorthy, High-resolution probing of early events in amyloid-beta aggregation related to Alzheimer's disease, *Chem. Commun.* 56 (34) (2020) 4627–4639.
- [69] M. Berjanskii, D.S. Wishart, NMR: prediction of protein flexibility, *Nat. Protoc.* 1 (2) (2006) 683–688.
- [70] Y. Xiao, B. Ma, D. McElheny, S. Parthasarathy, F. Long, M. Hoshi, R. Nussinov, Y. Ishii, Aβ(1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease, *Nat. Struct. Mol. Biol.* 22 (6) (2015) 499–505.
- [71] L. Cerofolini, E. Ravera, S. Bologna, T. Wiglenda, A. Böddrich, B. Purfürst, I. Benilova, M. Korsak, G. Gallo, D. Rizzo, L. Gonnelli, M. Fragai, B. De Strooper, E. E. Wanker, C. Luchinat, Mixing Aβ(1–40) and Aβ(1–42) peptides generates unique amyloid fibrils, *Chem. Commun.* 56 (62) (2020) 8830–8833.
- [72] M.T. Colvin, R. Silvers, Q.Z. Ni, T.V. Can, I. Sergeyev, M. Rosay, K.J. Donovan, B. Michael, J. Wall, S. Linse, R.G. Griffin, Atomic resolution structure of monomeric A-beta(42) amyloid fibrils, *J. Am. Chem. Soc.* 138 (30) (2016) 9663–9674.
- [73] M.A. Wälti, F. Ravotti, H. Arai, C.G. Glabe, J.S. Wall, A. Böckmann, P. Güntert, B. H. Meier, R. Riek, Atomic-resolution structure of a disease-relevant Aβ(1–42) amyloid fibril, *Proc. Natl. Acad. Sci.* 113 (34) (2016) E4976–E4984.
- [74] T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert, R. Riek, 3D structure of Alzheimer's amyloid-β(1–42) fibrils, *Proc. Natl. Acad. Sci. U. S. A.* 102 (48) (2005) 17342–17347.
- [75] L. Gremer, D. Schölzel, C. Schenk, E. Reinartz, J. Labahn, R.B.G. Ravelli, M. Tusche, C. Lopez-Iglesias, W. Hoyer, H. Heise, D. Willbold, G.F. Schröder, Fibril structure of amyloid-β(1–42) by cryo-electron microscopy, *Science* 358 (6359) (2017) 116–119.
- [76] S. Tomaselli, V. Esposito, P. Vangone, N.A.J. van Nuland, A.M.J.J. Bonvin, R. Guerrini, T. Tancredi, P.A. Temussi, D. Picone, The α-to-β conformational transition of Alzheimer's Aβ(1–42) peptide in aqueous media is reversible: a step by step conformational analysis suggests the location of β conformation seeding, *ChemBioChem* 7 (2) (2006) 257–267.
- [77] S. Ciudad, E. Puig, T. Botzanowski, M. Meigooni, A.S. Arango, J. Do, M. Mayzel, M. Bayoumi, S. Chaignepain, G. Maglia, S. Cianferani, V. Orekhov, E. Tajkhorshid, B. Bardiaux, N. Carulla, Aβ(1–42) tetramer and octamer structures reveal edge conductivity pores as a mechanism for membrane damage, *Nat. Commun.* 11 (1) (2020) 3014.
- [78] B. Urbanc, L. Cruz, F. Ding, D. Sammond, S. Khare, S.V. Buldyrev, H.E. Stanley, N. V. Dokholyan, Molecular dynamics simulation of amyloid beta dimer formation, *Biophys. J.* 87 (4) (2004) 2310–2321.
- [79] S.J. Yun, B. Urbanc, L. Cruz, G. Bitan, D.B. Teplow, H.E. Stanley, Role of electrostatic interactions in amyloid beta-protein (A beta) oligomer formation: a discrete molecular dynamics study, *Biophys. J.* 92 (11) (2007) 4064–4077.
- [80] C. Wu, M.M. Murray, S.L. Bernstein, M.M. Condron, G. Bitan, J.-E. Shea, M. T. Bowers, The structure of Aβ42 C-terminal fragments probed by a combined experimental and theoretical study, *J. Mol. Biol.* 387 (2) (2009) 492–501.
- [81] Barz, B., A.K. Buell, and S. Nath, doi:<https://doi.org/10.1101/2020.06.23.156620v1.full>, 2020.
- [82] A.T. Petkova, W.M. Yau, R. Tycko, Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils, *Biochemistry* 45 (2) (2006) 498–512.
- [83] M.A. Grant, N.D. Lazo, A. Lomakin, M.M. Condron, H. Arai, G. Yamin, A.C. Rigby, D.B. Teplow, Familial Alzheimer's disease mutations alter the stability of the amyloid beta-protein monomer folding nucleus, *Proc. Natl. Acad. Sci. U. S. A.* 104 (42) (2007) 16522–16527.
- [84] K. Kawahara, C. Tanford, Viscosity and density of aqueous solutions of urea and guanidine hydrochloride, *J. Biol. Chem.* 241 (13) (1966) 3228–3232.
- [85] J. Ying, F. Delaglio, D.A. Torchia, A. Bax, Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data, *J. Biomol. NMR* 68 (2017) 101–118.
- [86] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, NMRpipe - a multidimensional spectral processing system based on Unix pipes, *J. Biomol. NMR* 6 (3) (1995) 277–293.
- [87] W.F. Vranken, W. Boucher, T.J. Stevens, R.H. Fogh, A. Pajon, P. Llinas, E.L. Ulrich, J.L. Markley, J. Ionides, E.D. Laue, The CCPN data model for NMR spectroscopy: development of a software pipeline, *Proteins Struct. Func. Bioinf.* 59 (4) (2005) 687–696.