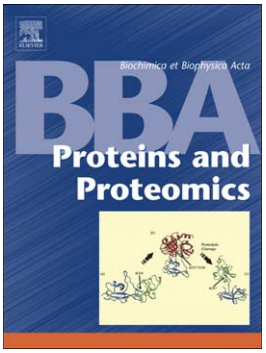


Toward a common aggregation mechanism for a β-barrel protein family: Insights derived from a stable dimeric species

Carla R. Angelani, Lucrecia M. Curto, Inés S. Cabanas, Julio J. Caramelo, Vladimir N. Uversky, José M. Delfino

PII: S1570-9639(14)00151-4
DOI: doi: 10.1016/j.bbapap.2014.06.002
Reference: BBAPAP 39365

To appear in: *BBA - Proteins and Proteomics*

Received date: 1 April 2014
Revised date: 30 May 2014
Accepted date: 2 June 2014

Please cite this article as: Carla R. Angelani, Lucrecia M. Curto, Inés S. Cabanas, Julio J. Caramelo, Vladimir N. Uversky, José M. Delfino, Toward a common aggregation mechanism for a β-barrel protein family: Insights derived from a stable dimeric species, *BBA - Proteins and Proteomics* (2014), doi: 10.1016/j.bbapap.2014.06.002

Toward a common aggregation mechanism for a β-barrel protein family: Insights derived from a stable dimeric species.

Carla R. Angelani†, Lucrecia M. Curto†, Inés S. Cabanas†, Julio J. Caramelo‡, Vladimir N. Uversky§,¶, and José M. Delfino†*

† Department of Biological Chemistry and Institute of Biochemistry and Biophysics (IQUIFIB), School of Pharmacy and Biochemistry, University of Buenos Aires, Junín 956, C1113AAD Buenos Aires, Argentina.
‡ Instituto de Investigaciones Bioquímicas de Buenos Aires-CONICET and Laboratory of Structural Cell Biology, Leloir Institute Foundation, Av. Patricias Argentinas 435, C1405BWE, and Department of Biological Chemistry-School of Sciences-University of Buenos Aires, Buenos Aires, Argentina.
§ Department of Molecular Medicine, USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, Florida 33612, USA;
¶ Institute for Biological Instrumentation, Russian Academy of Sciences, 142292 Pushchino, Moscow Region, Russia

Carla R. Angelani and Lucrecia M. Curto contributed equally to this work.

* Corresponding author: José María Delfino, Department of Biological Chemistry and Institute of Biochemistry and Biophysics (IQUIFIB), School of Pharmacy and Biochemistry, University of Buenos Aires, Junín 956, C1113AAD, Buenos Aires, Argentina; e-mail: delfino@qb.ffyb.uba.ar; tel: +54 11 4964-8289; fax: +54 11 4962-5457.

ABSTRACT

Δ78Δ is a second generation functional all-β sheet variant of IFABP (intestinal fatty acid binding protein) corresponding to the fragment 29-106 of the parent protein. This protein and its predecessor, Δ98Δ (segment 29-126 of IFABP), were initially uncovered by controlled proteolysis. Remarkably, although IFABP and Δ98Δ are monomers in solution, Δ78Δ adopts a stable dimeric structure. With the aim of identifying key structural features that modulate the aggregation of β-proteins, we evaluate here the structure and aggregation propensity of Δ78Δ. The 2,2,2-trifluoroethanol (TFE) induced aggregation of this protein shows a primary nucleation-elongation mechanism, characterized by the stabilization of a dimeric nucleus. Its rate of production from the co-solvent induced aggregation prone state governs the kinetics of polymerization. In this context, the value of Δ78Δ lies in the fact that -being a stable dimeric species- it reduces an otherwise bimolecular reaction to a unimolecular one. Interestingly, even though Δ78Δ and IFABP display similar conformational stability, the abrogated form of IFABP shows an enhanced aggregation rate, revealing the ancillary role played on this process by the free energy of the native proteins. Δ78Δ share with IFABP and Δ98Δ a common putative aggregation-prone central peptide. Differences in the exposure/accessibility of this segment dictated by the environment around this region might underlie the observed variations in the speed of aggregation. Lessons learnt from this natural dimeric protein might shed light on the early conformational events leading to β-conversion from barrels to amyloid aggregates.

KEYWORDS

- β-barrel protein
- intestinal fatty acid binding protein
- trifluoroethanol
- amyloid-like aggregation
- protein conformation
- abridged proteins

ABBREVIATIONS

IFABP, intestinal fatty acid binding protein; Δ98Δ, a truncated variant of IFABP corresponding to the fragment 29-126 of the parent protein; Δ78Δ, a truncated variant of IFABP corresponding to the fragment 29-106 of the parent protein; TFE, 2,2,2-trifluoroethanol; Th-T, Thioflavin-T; CR, Congo Red; λ_{max}, position of the maximum wavelength of fluorescence emission; TEM, transmission electron microscopy.

β-Sheet proteins are continuously produced and fold up successfully in the cell, with only a few isolated cases reported of misfolding, aggregation, and insolubility. The understanding of the stability and folding mechanism of this class of proteins is becoming increasingly relevant because many 'conformational diseases', with severe consequences on animal and human health, are based on the generation of β-sheet structures [1]. Natural β-sheet structures in proteins present different mechanisms to avoid edge-to-edge mediated aggregation. Particularly, β-barrel motifs escape this situation, because they tend not to expose free edges. To accomplish this, a continuous β-hydrogen bonding network organizes all around the barrel [2]. There are few proteins of the β-class useful as model systems for protein engineering, mainly due to their conspicuous tendency to aggregate. For this reason, the intestinal fatty acid binding protein (IFABP) family arises as a very helpful target that allows the exploration of structural determinants underlying conformational change, folding, misfolding and aggregation. Although fatty acid binding proteins (FABPs) display variable sequence identity, they share a common three dimensional structure consisting of a β-barrel formed by two five-stranded β-sheets arranged in a nearly orthogonal orientation (Fig. 1). This structure differs from most globular proteins since its interior is occupied by a large solvent-filled cavity that binds non-polar ligands, while the hydrophobic core is small and displaced from the geometrical centre of the protein.

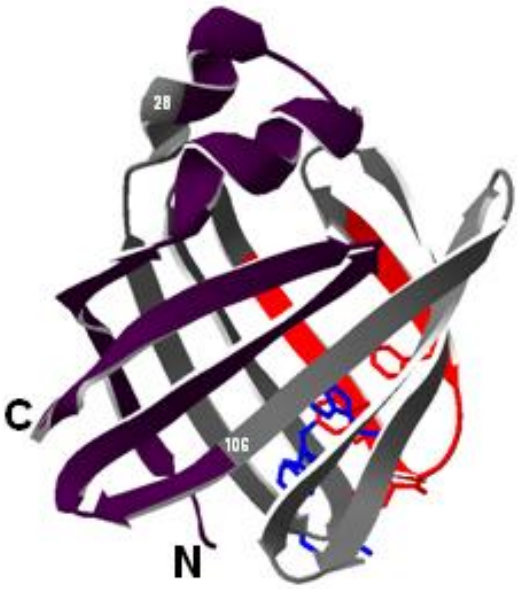


FIGURE 1 Ribbon structure of IFABP (PDB 2IFB). The Δ78Δ adopts a dimeric structure and this picture depicts the protomer painted in gray and the excised N- and C-termini stretches (1-28 and 107-131, respectively) in purple. The red segment corresponds to the predicted aggregation prone region. Residues belonging to the hydrophobic core are depicted with their side-chains in polytube representation.

By using proteolysis with the enzyme Arg-C, a natural tool sensitive to local flexibility, we generated two all-β sheet abridged variants of IFABP: Δ98Δ [3, 4] and Δ78Δ [5]. Δ98Δ comprises a stretch of 98 amino acids corresponding to fragment 29-126 of IFABP. Interestingly, this truncation leads to a loss of residues involved in the closure of the β-barrel. Although lacking one quarter of the sequence of the parent protein, Δ98Δ retains significant β-sheet content and tertiary interactions, being a stable and functional form of IFABP [3]. This variant was submitted to a new round of controlled proteolysis, yielding a second generation functional abridged variant of IFABP named Δ78Δ [5]. This protein lacks one third of the sequence of the parent protein, comprising only 78 amino acids corresponding to

sequence 29-106 of IFABP. Δ78Δ is devoid of the first β-strand, most of the helical domain and the last 25 amino acids belonging to the C-terminal β-sheet (Fig. 1). It should be stressed that both truncated forms preserve all the critical residues of the hydrophobic core, i.e. those involved in the nucleation step leading to the folded state [6]. Interestingly, while Δ98Δ displays a monomeric state, Δ78Δ adopts a highly stable dimeric form [5]. It can be speculated that the dimerization of Δ78Δ would help the protein to avoid potential problems arising from the increased hydrophobic surface exposure of the protomer. In principle, the extensive stretches deleted in Δ78Δ (53 out of 131 amino acids) could determine the exposure of hydrophobic residues in the protomer and the appearance of free edges prompting dimerization. The association of each protomer with the other might implicate an interface involving primarily a new set of backbone-backbone hydrogen bonds that make their own contributions to the overall stability of the dimeric structure [5].

Due to their remarkable stability, both abridged forms were challenged with physical and chemical agents to evaluate their impact on structure. In this general context, conformational change in solution along with the occurrence of aggregation was observed. Finally, the co-solvent 2,2,2-trifluoroethanol (TFE) was found to be the best choice to study these phenomena. It is well known that this chemical agent induces helical structures in peptides. However, this co-solvent has also the ability to alter the native structure of proteins [7]. In this sense, the consequences of adding TFE upon secondary structural elements within a protein is distinct from that observed with a standalone discrete peptide [8]. TFE does not operate like classical denaturants such as urea, but it shows a dual concentration-dependent effect. At low concentrations TFE interacts with carbonyl oxygen atoms and hydrophobic groups on the surface of proteins and increases internal protein interactions. However, at relatively high concentrations, TFE also penetrates the hydrophobic core of the proteins disrupting the internal core stability thus initiating unfolding. This is in direct contrast to the observed effects on peptidic structures whereby TFE interacts weakly with non-polar side chain groups and does not disrupt hydrophobic interactions between peptide side chains [8]. At around 30 % v/v TFE, alcohol molecules associate so as to minimize their contact with water, resulting in the formation of micelle-like clusters with the hydrophobic groups located inside. Such clusters provide a highly hydrophobic local environment where polarity decreases. Upon binding to these clusters, proteins undergo a conformational transition, leading to non-native states [7]. Due to its ability to promote the conversion of proteins into native-like aggregation-prone species, TFE has been widely used to trigger protein aggregation or amyloid formation [9 - 13]. In our previous work [14] we presented a comparative study of the conformation and aggregation propensity of IFABP and Δ98Δ upon addition of 25% v/v TFE. Since Δ98Δ might expose free edges and displays a less stable native-like structure than IFABP, an increased rate of aggregation for this variant would have been expected. On the contrary, Δ98Δ displays a similar (or even lower) tendency to aggregate. Once we uncovered that the critical nucleus for protein aggregation constitutes into a dimer, it became all the more interesting to take full advantage of the Δ78Δ dimer as probe to study this phenomenon.

In this work we elucidate mechanistic aspects of the aggregation process with emphasis on the relationship between amino acid sequence, conformation and aggregation propensity. With the critical evidence in hand presented in this paper we can now postulate that a common aggregation mechanism indeed can be established for all three proteins. Lessons learnt from this structure might shed light on the early conformational events leading to amyloid aggregation.

MATERIALS AND METHODS

Materials

Recombinant Δ78Δ was cloned using primers 5´-GGAATTCCATATGAAGCTTGGAGCTCATG-3´and 5´-CGCGGATCCTCATCGGACAGCAATCAGC-3´. The PCR product was digested with both NdeI and BamHI and cloned in the pET-22b(+) vector. The protein was purified as described previously (5).
Protein concentration was estimated by ultraviolet absorption: ε_{280nm}= 6970 M^{-1} cm^{-1}.
For end-point protein aggregation measurements, samples were incubated at the indicated TFE concentration for at least 2 h, a time sufficient to attain a stable turbidity reading. To avoid fragmentation of the fibrils, no stirring was applied [15]. Unless otherwise stated, the final concentration of protein and TFE are 0.24 mg mL^{-1} and 25 % v/v, respectively.

Dye binding assays
Th-T and CR binding assays were performed as originally described by Naiki [16] and Klunk [17], respectively and according to the technique reported by Nilsson [18]. Briefly, for the Th-T assay, protein samples (20 µL, 0.24 mg mL^{-1}) incubated for 2 h at the concentrations of TFE shown were diluted in Th-T solution (2 mL, 50 µM). The fluorescence intensity of this solution (λ_{max} 482 nm) is subtracted from that of a protein sample at 0 % v/v TFE (which in turn is equal to a sample without protein). For the CR assay, the wet protein pellets obtained after centrifugation of samples incubated for 2 h at the concentrations of TFE shown (5 µL) were diluted in CR solution (1 mL). The absorbance at 540 nm of a sample without protein is subtracted.

Transmission electron microscopy
Samples containing Δ78Δ (0.24 mg mL^{1}) in 25 % v/v TFE were incubated overnight. The following day, pellets were resuspended in Tris-HCl buffer (20 mM, pH 8.0). A sample (5 µL) was placed on a carbon-coated copper grid, and allowed to stand for 2 min. The grids were then washed with distilled water and stained with uranyl acetate (2 % w/v) for 2 min. Images were taken with a Philips EM 301 transmission electron microscope located at the Center for Advanced Microscopies (School of Sciences, University of Buenos Aires).

Aggregation kinetics
The extent of aggregation was followed by the evolution of turbidity at 350 nm (A_{350nm}) on a Jasco J-550 spectrophotometer. Samples containing Δ78Δ (0.02-0.50 mg mL^{-1}, equivalent to 2.3-56.8 µM) in 25 % v/v TFE were measured at 25 ºC. A fixed amount of neat TFE (213 µL) was added to the protein solution (637 µL) and the content was mixed immediately by inverting the cuvette three times prior to beginning the collection of turbidity readings. The whole procedure does not represent a delay longer than 10 s, an acceptable dead-time for kinetic measurements in a time-scale of minutes. Here again, no stirring was applied to samples, due to its known effect on fibril fragmentation [15].
All kinetic data describing the aggregation process share a similar sigmoid shape (Fig. 3). As such, they were found to comply with the phenomenological scaling [19], meaning that a *single* function *f* is able to describe all time courses. In mathematical terms, this can be stated as follows: $A / A_{\text{lim}} = f\left[t / t_{char}\right]$, where A_{lim} is the limiting value of the turbidity as time tends to infinity whereas t_{char} is a characteristic time for the process. The implication of satisfying this scaling property is that all kinetics share *the same* assembly pathway.
Independently, an empirical model [20] was fitted to the measured turbidity to account for the main evolution of the aggregation kinetics. Here, the process is considered an irreversible reaction starting with the association of *n* molecules of non-aggregated protein P:

$$nP \xrightarrow{k_n} P_{agg}$$

where P_{agg} is the aggregated form and k_n is the n^{th} order rate constant. The rate of aggregation may be written as the time-dependent decay of the non-aggregated protein concentration ($v_{agg} = -d[P]/dt$). Since the turbidity of the protein solution is assumed to be proportional to the amount of aggregated protein, i.e. $A \propto [P_{agg}]$, the evolution of turbidity with time follows the expression:

$$dA/dt = k(A_{\lim} - A)^n \quad (1)$$

where A is the turbidity at time t, A_{lim} is the limiting value of the turbidity as time tends to infinity and $k = nk_n [P]_0^{n-1} / A_{\lim}^{n-1}$, in which $[P]_0$ is the initial (total) protein concentration. When $n=1$ (first order process of aggregation with respect to protein) Eq. 1 becomes the following:

$$dA/dt = k_1(A_{\lim} - A) \quad (2)$$

where k_1 is the first order rate constant. Integration of Eq. 2 gives the expression describing the expected dependence of A on t:

$$A = A_{\lim}\left\{1 - e^{\left[-k_1(t-t_0)\right]}\right\} \quad (3)$$

where t_0 is a lag time during which $A \simeq 0$. The model was fitted to the data by non-linear regression analysis using the Microsoft Excel solver tool. The agreement between the experimental data and the calculated values was estimated by the R^2 correlation coefficient [20].

Circular dichroism

Spectra were recorded on a Jasco J-810 spectropolarimeter. Data in the near UV (250-320 nm) or in the far UV (200-250 nm) regions were collected at 25 ºC using a 10 or a 1 mm path length cuvettes, respectively. A scan speed of 20 nm min^{-1} with a time constant of 1 s was used for both proteins. Each spectrum was measured at least three times and the data was averaged to minimize noise. Molar ellipticity was calculated as described elsewhere [21], using a mean residue weight value of 111.5.

Fluorescence measurements

Fluorescence measurements were performed at 25 ºC in a Jasco FP-6500 spectrofluorimeter equipped with a thermostated cell. A 3 mm path cuvette sealed with a Teflon cap was used. The intrinsic emission spectra were obtained in the absence or presence of appropriate concentrations of TFE. In this case, excitation wavelength was 295 nm and emission was collected in the range 310-410 nm. The excitation and emission monochromator slit widths were both set at 5 nm. For each spectrum, the total integrated intensity and the maximum wavelength of fluorescence emission (λ_{max}) were the parameters used for further analysis.

Prediction of aggregation propensity and intrinsic disorder

The AMYLPRED 2 tool [22] was used to evaluate the existence of aggregation prone peptides along the sequence of IFABP. Independently, the propensity of the protein for intrinsic disorder was evaluated by several computational tools from the PONDR family, such as PONDR® VLXT [23], PONDR® VSL2 [24, 25], and PONDR-FIT [26]. Although PONDR® VLXT is not the most accurate predictor, it is known to be very sensitive to the local compositional biases and provides useful information on the local peculiarities of intrinsic disorder [23], a feature which often correlates with the presence of disorder-based potential molecular interaction motifs [27, 28]. PONDR® VSL2 is one of the most accurate stand-alone disorder predictors, whereas PONDR-FIT represents a metapredictor which is moderately more accurate than each of the component predictors [26].

RESULTS

TFE-induced aggregation

End-point turbidity at 350 nm of Δ78Δ (incubated for 2 h) at various TFE concentrations was measured. A representative set of data is shown in Fig. 2. In the absence of TFE, there is no evidence of aggregation, not even after incubation at 50 ºC for 6 h (data not shown). Up to 10 % v/v TFE, no significant change was observed. Starting at 15 % v/v there is a sharp increment of turbidity, consistently displaying a maximum value at ~ 25 % v/v TFE. The turbidiy decreases slightly at 30 % v/v TFE and remains fairly constant up to 60 % v/v TFE. At the highest concentration of cosolvent assayed, the turbidity lowers but it is not abolished. In view of these facts, the time evolution of this parameter at 25 % v/v TFE -the condition of maximum turbidity- was recorded at various protein concentrations (Fig. 3 *A*). The sigmoidal shape of all kinetic curves reflects the occurrence of a concentration-dependent lag period before the onset of aggregation (Fig. 3 *B*).

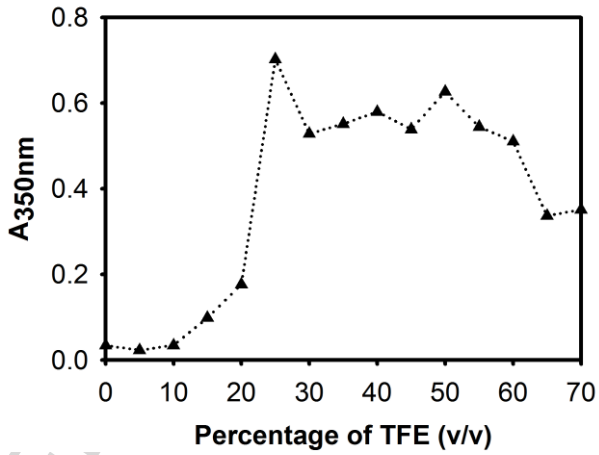


FIGURE 2 TFE-induced aggregation. End-point turbidity measured at 350 nm of a 0.24 mg mL^{-1} solution of Δ78Δ was measured at increasing concentrations of TFE. For the sake of clarity, a dotted line linking the experimental points was added.

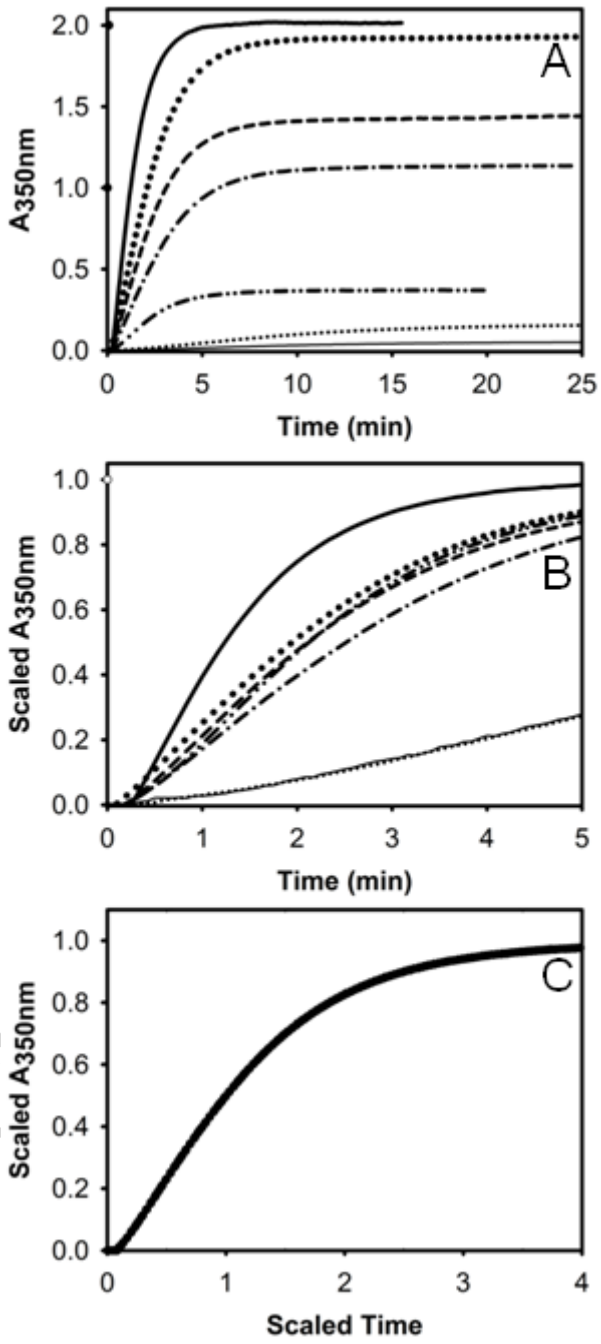


FIGURE 3 Kinetics of aggregation of Δ78Δ at different protein concentrations: 2.3 (solid thin line), 5.7 (small dotted line), 11.4 (dash-dot-dot-dash line), 27.2 (dash-dot-dash line), 34.1(dashed line), 45.4 (large dotted line) and 56.8 μM (solid thick line). *(A)* Effect of concentration on the kinetics of aggregation at 25 % v/v TFE, as followed by measuring turbidity at 350 nm. *(B)* Zoom in on the early time-points along the kinetics shown in *A* after scaling turbidity (A/A_{lim}), and *(C)* plot of the data shown in *A* after scaling turbidity (A/A_{lim}) and time ($t/t_{1/2}$).

As a first step, the aggregation kinetics was analyzed by plotting the *scaled turbidity* (A/A_{lim}, where A_{lim} is the asymptotic value of turbidity measured at a given protein concentration) against the *scaled time* ($t/t_{1/2}$, where $t_{1/2}$ is the characteristic time needed for the turbidity to reach half the value of A_{lim}). Interestingly, the scaled curves measured for all protein

concentrations fall into a single curve (Fig. 3 *C*), suggesting a common underlying aggregation mechanism [19]. To further investigate the process at 25 % v/v TFE, the Wang and Kurganov model [20] was fitted to the data. The aggregation curves from $A/A_{lim} \sim 0.1$ to higher values are satisfactorily described by a pseudo-first order kinetics (Eqs. 2 and 3, $R^2 > 0.9999$). This is inferred from the linear dependence of the rate constant k_1 with protein concentration (Fig. 4 *A*). The parameters A_{lim} (Fig. 4 *B*) and t_0 (lag time, Fig. 4 *C*) -as derived from Eq. 3- can be determined by this procedure as well. A_{lim} values are proportional to total protein concentration, suggesting that the turbidity value faithfully reflects the extent of protein aggregation. One should note that most protein originally in solution undergoes the transition to aggregates (> 95 %, data not shown). Considering that, at $t \geq t_0$ aggregation follows a first order kinetics, the overall rate of this process may be estimated by the slope of a tangent to the curve at $t \simeq t_0$ (mathematically equivalent to the product $A_{lim}\, k_1$). Since both A_{lim} and k_1 were found to depend linearly on protein concentration, this product results proportional to the square of protein concentration (Fig. 4 *D*), a fact indicative of a second order kinetics. Under the assumption that the concentration of nuclei remains constant during the elongation phase, the second order association reaction can be satisfactorily described by an apparent first order kinetics. This is consistent with the sequential addition of protein to the preformed nuclei. In summary, the TFE-induced aggregation kinetics of Δ78Δ shows: (i) a lag period showing an inverse relationship on protein concentration, and (ii) a second-order dependence of this process with respect to protein concentration.

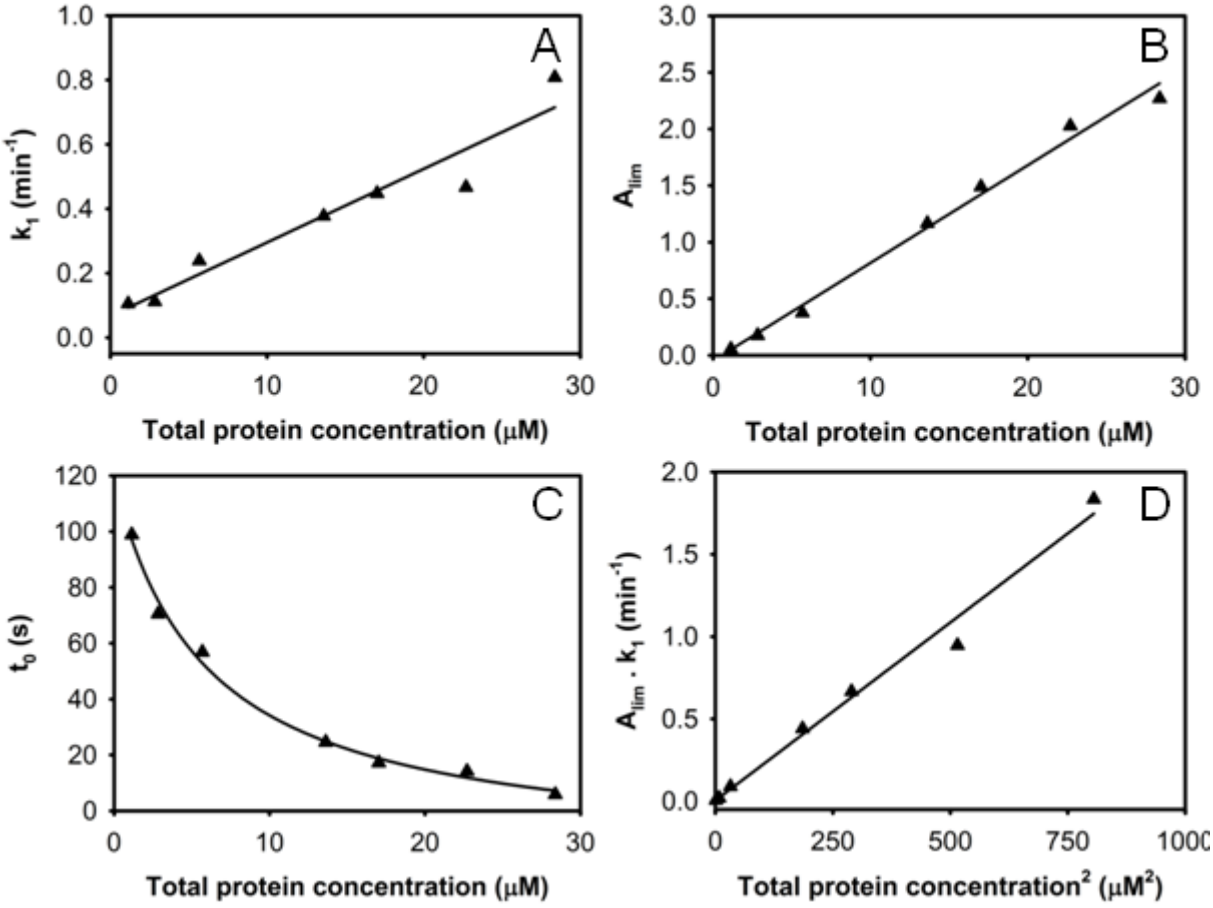


FIGURE 4 Dependence on protein concentration of the aggregation kinetic parameters. The dependence on total protein concentration of the rate constant k_1 *(A)*, plateau value A_{lim} *(B)*

and the lag time t_0 *(C)* are shown. Panel *D* shows the rate of aggregation at $t=t_0$ (equal to the product $A_{lim}\ k_1$) as a function of the squared of total protein concentration.

Given that we deliberatedly avoided stirring of the samples, minimal effects due to fragmentation of the fibrils are expected [15]. One particularly revealing scaling law reflecting primary nucleation relates $t_{1/2}$ to A_{lim} as a power law [29]. One should be reminded that A_{lim} depends linearly on total protein concentration (Fig. 4 *B*).

$$t_{1/2} \propto Alim^{\gamma} \qquad (4)$$

where γ is the scaling exponent. The latter equation becomes a linear relationship with slope equal to γ when plotted as a log-log plot, as seen for Δ78Δ in Fig. 5. As for systems dominated by primary nucleation $\gamma = -n/2$, the reaction order of the nucleation process (*n*) can be simply extracted. In this case $\gamma \simeq 0.5$, therefore implying *n*=1. Consequently, this imposes a restriction on the kinetic model by which the active nucleus is indeed *a single Δ78Δ molecule*, which is a homodimer in aqueous solution.

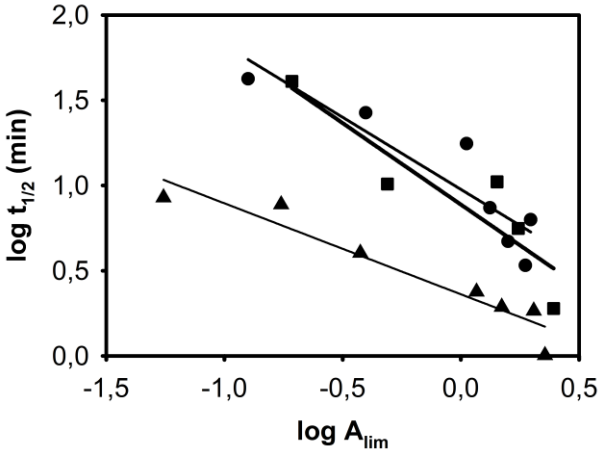


FIGURE 5 Double log plot of the characteristic time $t_{1/2}$ versus A_{lim}. For the sake of clarity, data of IFABP (■) and Δ98Δ (●) are plotted alongside those of Δ78Δ (▲). The slopes of the fitted curves are 0.96 ± 0.26, 0.84 ± 0.17, and 0.53 ± 0.07 for IFABP, Δ98Δ and Δ78Δ, respectively.

The possibility that the aggregate formed at 25 % v/v TFE displayed amyloid-like properties was examined through its ability to bind amyloid-sensitive dyes and by inspecting its structural characteristics using electron microscopy (Fig. 6). Both Th-T and CR interact with the cross-β-sheet structure common to a variety of amyloid structures, leading to an increase in the fluorescence at 482 nm or to an enhanced differential absorbance at 540 nm, for Th-T and CR respectively [18, 30]. For both parameters, no significant changes were detected at low TFE concentrations up to 10 % v/v TFE. By contrast, a sharp increase was observed at 25 % v/v TFE (Fig. 6 *A*). Beyond this point, the Th-T fluorescence emission intensity and CR absorbance reach a plateau. In all respects, the dye binding behavior of this abridged variant is similar to that observed for IFABP and Δ98Δ [14]. For all proteins assayed, at the highest concentrations of TFE (60-70 % v/v), the CR absorbance of each sample was considerably higher than that measured for sub-aggregating conditions (e.g. up to 15 % v/v).
The transmission electron micrograph of aggregates is shown in Fig. 6 *B*. Although mature fibril morphology is not immediately apparent, annular and several rod-like structures exhibiting different lengths and curvatures are present, features indicative of bundles of fibrillar material.

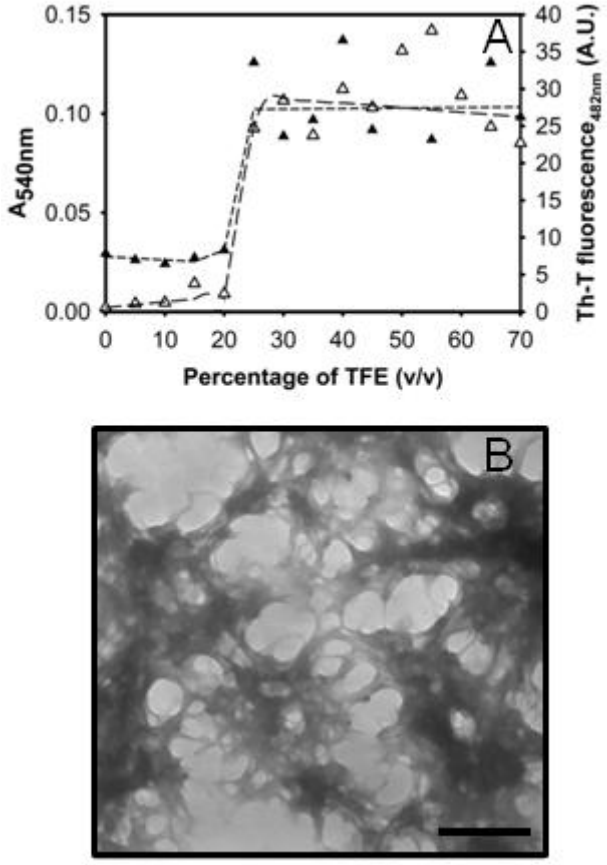


FIGURE 6 Binding of amyloid dyes to the proteins. *(A)* The absorbance of CR (▲, short-dashed line) and the fluorescence emission of Th-T (△, long-dashed line) were measured at different TFE concentrations. *(B)* Transmission electron microscopy of protein aggregates. Micrographs of Δ78Δ aggregates produced after incubation at room temperature for one day in the presence of 25 % v/v TFE. The scale bar represents 200 nm.

Effect of TFE on protein conformation

TFE-induced structural changes of Δ78Δ were monitored by far and near UV circular dichroism spectroscopy. The shape and magnitude of the spectral changes observed for Δ78Δ are very similar to those observed for Δ98Δ, but of larger magnitude than those observed for IFABP [14]. The far UV CD spectrum of native Δ78Δ in buffer displays a minimum at ~213 nm, a distinctive feature of β-sheet structure (Fig. 7 *A*, see also ref. 5). Interestingly, in the presence of a sub-aggregating concentration of TFE the shape and the intensity of the spectrum are substantially modified. At 15 % v/v TFE the spectrum of Δ78Δ becomes very similar to that of Δ98Δ and IFABP measured under identical conditions and to the spectrum of native IFABP in the absence of TFE [14]. At 25 and 45 % v/v TFE the shape of the spectra remains β-like, but a ~ 10 nm red shift of the main negative band occurs. In addition, at 45 % v/v TFE a 2-fold enhancement in the intensity of this dichroic signal is observed. The position of this new minimum is typical of largely extended β-sheet conformation, a structure prevalent in amyloid fibrillar aggregates [13, 31]. Conversely, at 65 % v/v TFE the spectrum shows a change in its shape toward one with two negative bands: one at ~222 nm and an incipient band at ~208 nm, a feature typical of α-helical structure. All in all, the new abridged variant likely exhibits a higher conformational flexibility as compared to the full-length protein.

In the near UV CD range, minor spectral changes take place by the addition of up to 15 % v/v TFE, indicating full conservation of the fine structure. Conversely, at 25 % v/v TFE and higher co-solvent concentrations there are large deviations from the spectrum measured in the absence of TFE (Fig. 7*B*), essentially erasing all spectral features. Here again a similar trend is observed for IFABP and Δ98Δ [14].

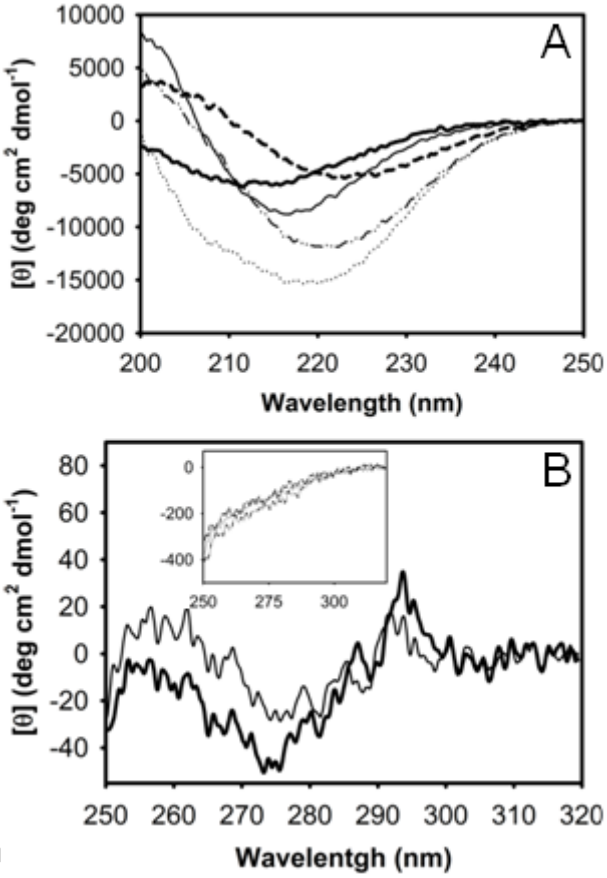


FIGURE 7 Circular dichroism spectroscopy of Δ78Δ. Far *(A)* and near *(B)* UV CD spectra at increasing TFE concentrations: 0 (solid thick line), 15 (solid thin line), 25 (dashed line), 45 (dash-dot-dot-dash line) and 65 (dotted line) % v/v. For the sake of clarity, near UV CD spectra corresponding to 25, 45 and 65 % v/v TFE are shown as an inset.

IFABP contains two Trp residues: Trp82, which is buried within the hydrophobic core at the bottom of the ligand binding pocket (Fig. 1), and Trp6, placed at the N-terminal β-strand. Both residues are almost completely occluded from the aqueous solvent [32]. In the full-length protein, Trp82 (the only remaining Trp in Δ78Δ) is known to contribute approximately 75 % of the fluorescence emission and there is almost no cross-talk between both Trp residues [33]. Previous work from our laboratory showed that this residue in Δ78Δ is placed in a milieu akin to that found in IFABP [5], thus becoming a useful spectroscopic probe to evaluate the integrity of the hydrophobic core. In this sense, Fig. 8 shows the evolution of the fluorescence intensity and the position of the maximum wavelength of emission (λ_{max}) as a function of TFE concentration. From 0 to 15 % v/v TFE, λ_{max} remains almost invariant (333 nm). By contrast, a 4 nm red shift occurs between 15 and 30 % v/v TFE and, from that point up to the highest concentration assayed (70 % v/v), an additional 5 nm red shift is observed. It should be noted the 1.3-fold enhancement in fluorescence intensity occurring upon addition of as little as 5 % v/v TFE, a feature also displayed by Δ98Δ [14]. This is at variance with IFABP, which shows a monotonous descending trend of fluorescence intensity. On the other

hand, while at 70 % v/v TFE half of the fluorescence intensity of native Δ78Δ is conserved, the emission intensity for the other two proteins is slightly lower (20-30 % of the initial value). It should be noted that the expected effect of TFE on Trp fluorescence emission (as revealed by N-acetyl-L-tryptophanamide) is a blue shift of its emission peak (from 353 to 350 nm), and a more than 4-fold decrease in quantum yield [34]. In agreement with results derived from CD measurements, co-solvent-induced fluorescence quenching of Trp residues does not suffice to account for the above observations, revealing the occurrence of conformational changes as well. In this regard, as well as for Δ98Δ and IFABP, at 65 % v/v TFE, the build-up of a helix-rich form concomitant with a loss of the rigid aromatic environment are evidence of a transition to a molten globule state of Δ78Δ. In this last condition, the λ_{max} of emission is 10 nm blue-shifted with respect to unfolded Δ78Δ in 3 M guanidinium hydrochloride (352 nm, result not shown). This goes in line with the existence of a flexible and somewhat solvent occluded environment for the only remaining Trp residue in the abridged molecule.

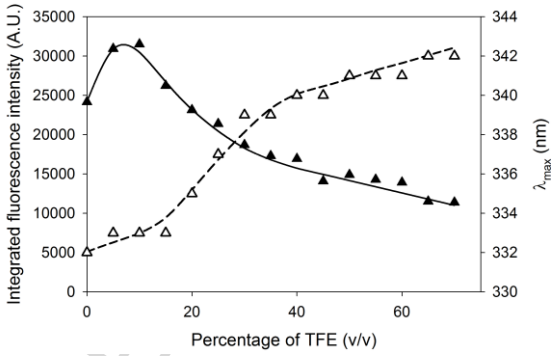


FIGURE 8 Fluorescence spectroscopy of Δ78Δ. Integrated intensity (*solid symbols*) and position of the maximum wavelength of emission (λ_{max}, *open symbols*) at the indicated concentrations of TFE are shown.

Prediction tools for amyloidogenic segments

AMYLPRED2 is a consensus method for the prediction of aggregation prone peptides in globular proteins [22]. It combines 11 of the existing tools and defines a consensus when a hit coincides for at least half of the selected methods. The result obtained for the full length protein is shown in Fig. 9. The three regions considered to be aggregation prone are: 23-26, 58-71 and 114-118. Additionally, FoldAmyloid [35] also predicts regions 59-64 and 114-118 of IFABP to be amyloidogenic. In summary, most of the algorithms used highlight two regions as putative aggregation prone peptides: a short C-terminal segment (114-118, which is absent in Δ78Δ) and a longer central stretch encompassing residues 58-64/71. The short N-terminal peptide predicted only by AMYLPRED2 is absent in both abridged variants. Moreover, the aggregation propensity of the whole excised N-terminal region (segment 1-28 of IFABP) has been experimentally assessed and no aggregation at 25 % v/v TFE occurs [14].

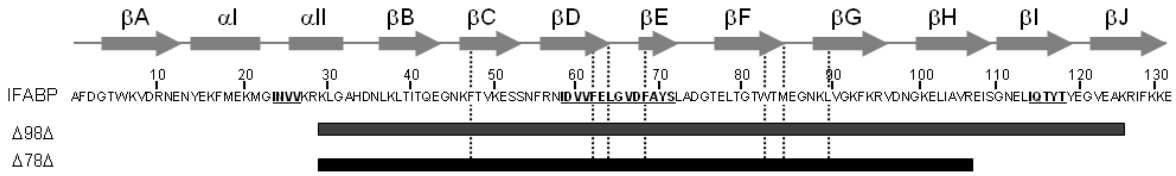


FIGURE 9 Consensus prediction (AMYLPRED 2) for aggregation-prone peptides. Segments sharing positive hits are underlined. Amino acid residues belonging to the hydrophobic core are marked with dotted lines.

Analysis of intrinsic disorder propensity

With the aim at uncovering structural determinants relevant for the generation of aggregation prone species, we undertook a bioinformatic analysis that might potentially reveal the influence on conformation of non-native environmental effects, be those of the protein sequence itself or those arising from the solvent milieu.

Despite the fact that both the crystallographic and NMR aqueous solution structures [36 - 39] do not reveal any region of IFABP to be highly disordered, we analyzed the protein (UniProt IDF: P02693) with intrinsic disorder propensity predictors. Surprisingly, according to this analysis, sizeable N- and C-terminal regions (~15 first residues and ~30 last residues, respectively) of IFABP are predicted to be the least ordered ones (Fig. 10). In the PONDR VLXT plot, the C-terminal predicted disordered region of IFABP includes an insertion of a fragment with increased order propensity, which is manifested as a typical dip centered at residue 120. Also, the PONDR VSL2 plot contains a similar dip in the N-terminal part of the protein (centered at residue 9). Typically, such dips in disorder plots correspond to potential binding regions [27, 28]. The presence of such sites was further evaluated by the MoRFpred tool [40], which is the leading predictor of molecular recognition features (MoRF). This analysis revealed the existence of MoRFs at residues 8-13 and 118-131. Another MoRF was predicted in the 60-68 region, which coincides with the main central amyloidogenic segment. Remarkably, a major part of the predicted disordered C-terminal tail is preserved in Δ98Δ, whereas both of them are eliminated in the Δ78Δ construct.

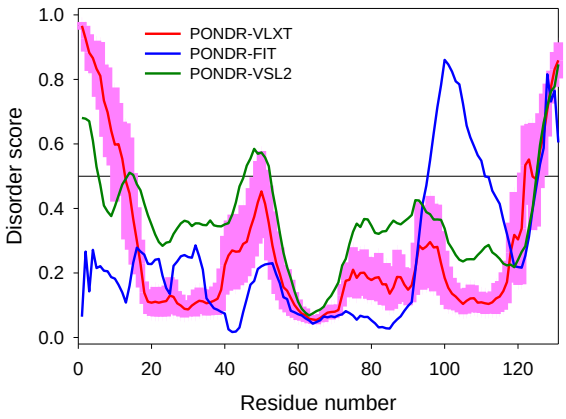


FIGURE 10 Intrinsic disorder propensity of IFABP. Disorder propensity was evaluated with PONDR® VLXT (red curve), PONDR® VSL2 (green curve), and PONDR-FIT (blue curve). The light pink shadow around the corresponding PONDR-FIT curve represents the distribution of statistical error.

DISCUSSION

The main consequences of TFE addition on the conformation of Δ78Δ can be analyzed from the information presented in Figs. 7 and 8. The abridged variants Δ78Δ and Δ98Δ exhibit features indicating higher conformational plasticity than the parent protein IFABP [14]. Increasing TFE concentration from 0 to 15 % v/v causes a marked enhancement in Δ78Δ´s fluorescence emission and an intensification of all spectral features in both far and near UV CD. This phenomenon holds resemblance to that observed for Δ98Δ and is in sharp contrast with the behavior of the full length protein [14]. On the other hand, the onset of turbidity (Fig. 2), the overall descending trend of the fluorescence intensity (Fig. 8) and the evolution of CD signals (Fig. 7) reveal that further conformational changes are associated to the aggregation

phenomenon as well. Particularly at 25 % v/v TFE, the structural rearrangements observed by CD spectroscopy occur concurrently with a step-down in fluorescence intensity and a red shift of λ_{max}. These results, combined with those obtained from the binding of dyes (Th-T and CR), and from TEM (Fig. 6), sustain the amyloid-like character of the aggregated protein.

It has been previously reported that IFABP and Δ98Δ share a common aggregation mechanism at 25 % v/v TFE, in which the first step is a dimerization event giving rise to the nucleus, followed by the subsequent addition of protein molecules to the growing fibril [14]. Noteworthy, the doubly scaled curves shown in Figure 3*B* are indistinguishable from those obtained for IFABP and Δ98Δ (Fig. 3*D* in [14]). The implication of satisfying this scaling property is that all three protein kinetics share *the same* assembly pathway [19], i.e. a primary nucleation route, meaning that the nucleus arises solely from the initiating protein and not from fibril fragmentation and/or from surface catalysis by the fibril [29]. In this sense, the aggregation phenomenon can be explained by the following nucleation-elongation mechanism (Scheme 1).

$$P \rightleftarrows P^* \xrightarrow{f_0} P_n^* \xrightarrow[P^* \nearrow]{f_1} F$$

Scheme 1

The first event preceding this irreversible process is a fast equilibrium step, where the co-solvent TFE induces a conformational rearrangement of the native protein (P) giving rise to an aggregation-prone state (P^*). Subsequently, the formation of the nucleus (P_n^*) with a defined oligomeric state *n* takes place. Finally, the growth of the fibrillar aggregates (F) occurs by the irreversible apposition of more P^* to the primordial nucleus. Interestingly, despite differing in their oligomerization state under standard buffer conditions, all three proteins share a bimolecular nucleus upon challenging them with 25 % v/v TFE. For IFABP and Δ98Δ -which are monomeric proteins in solution- the *n* value is 2 [14]. At variance with this, the value measured for the dimeric variant Δ78Δ is 1, meaning that remains the same as in the originating protein P. It can be speculated that genesis of the nucleus might imply a co-solvent induced rearrangement of the pre-existing dimeric form. However, a dissociation-reassociation event cannot be excluded. For the Δ78Δ variant, the lag period of the kinetics – which is intimately related to the nucleation stage- is considerably lower than for the monomeric proteins Δ98Δ and IFABP. Consistently, the initial rate of aggregation of Δ78Δ (i.e. the $A_{lim}\ k_1$ value) is one half to one order of magnitude higher than that of IFABP and Δ98Δ, respectively. Both kinetic parameters measured for Δ78Δ point to a protein form particularly susceptible to aggregation.

There is a common tenet in the field that the perturbation of the native compact fold would inescapably populate aggregation-prone species [41]. Our evidence confirms that there is not a straightforward correlation between the decrease in the stability of P and the enhancement of the aggregation propensity. Proof of this is the fact that IFABP and Δ78Δ display similar stability but the latter exhibits a much higher rate of aggregation. Conversely, even though Δ98Δ presents the least stable native conformation under standard conditions, it shows the lowest aggregation rate. Thirumalai and col. [42] postulated that the kinetics of polymerization is determined by the rate of production of P^*, which in turn is controlled by barriers separating P and P^*. In this scenario, considering the stability of P in isolation may not be a good indicator of its rate of fibrillation. Although the generality of this observation has not yet been established, our variants might be examples of this complex behavior. In this regard, it should be noted that the free energy of unfolding $\Delta G^{\circ}_{H_2O}$ refers to the difference between native P and the unfolded state. The modification of the energy landscape for the

protein due to the effect exerted by chemical perturbants such as TFE should certainly be taken into consideration to adequately predict the transition of the native state to fibrillar forms.

Several factors might modulate the aggregation conduct of this protein family. It can be speculated that the central peptide (58-64/71), common to all three proteins, might be the region mainly responsible for aggregation. This piece of evidence highlights the fact that the aggregation propensity of a given construct is not only a function of the local amino acid sequence, but that long-range interactions between different regions might also play an important role in this process. In this sense, variations in the speed of aggregation might be modulated by differences in the conformation of the protein, which will certainly dictate the exposure/accessibility of this stretch. Remarkably, as shown in red in Figure 1, the putative central amyloidogenic peptide found here maps to the region where a gap occurs between the fourth and fifth β-strands of IFABP, a spot with few inter-strand backbone hydrogen bonds [38]. Accordingly, a significant overlap exists between the regions of the sequence found to be solvent exposed and those identified to be critical as rate-determining steps of aggregation, indicating that a considerable extent of solvent exposure is a feature of the regions that initiate the process of aggregation [43]. Indeed, a Kyte-Doolittle hydropathy plot (using the GRAVY score and a 9 amino acids window size) reveals the common segment 58-64/71 to be the one with the highest hydrophobicity (data not shown). It can be speculated that the tight dimerization event characteristic of Δ78Δ renders a stable soluble form preventing aggregation by hindering critical hydrophobic segments [5].

Suggestively, one of the three MoRF predicted coincides with the main central amyloidogenic segment. The other two MoRFs lay on each of the putative disordered tails of IFABP. Remarkably, a major part of the predicted disordered C-terminal tail is preserved in Δ98Δ, whereas both of them are eliminated in the Δ78Δ construct. Biologically active proteins or protein regions that do not assume unique three-dimensional structures by themselves are very common in nature [44 - 46] and are abundantly involved in numerous biological processes, where they are found to play different roles in regulation of the function of their binding partners and in promotion of the assembly of supra-molecular complexes [47 - 57]. Disorder is unevenly distributed within the protein sequences, and residues located in the protein termini are on the average more disordered than residues in the middle of the protein chain [58]. Among various functions ascribed to such disordered tails are protection of binding sites, stabilization of a protein molecule, and multifarious interactions with various binding partners [58]. A function potentially related to this study is their ability to serve as entropic bristles, which by random movements about their point of attachment, would sweep out a significant region in space and entropically exclude large particles without excluding small molecules, therefore acting as entropic bristles or entropic bristle domains [59]. In this sense, their ability to prevent proteins from interactions was used to design a set of specific protein solubilizers that extend away from the partner and sweep out large molecules, therefore allowing the target protein to fold free from interference [60]. It is likely that potentially disordered (or at least more flexible than the compact hydrophobic core) tails of IFABP can have at least two different functions. First, they are involved in the closure of the β-barrel and second, they might serve as entropic bristle domains that prevent IFABP from oligomerization as the hierarchical folding [6] proceeds. More precisely, NMR data [38, 39] shows that strand βA interacts with βJ closing the β-barrel thus avoiding the exposure of free edges. For both abridged variants, the removal of such interacting tails opens the entry to the β-barrel. It can be speculated that, for Δ98Δ, the remaining part of the predicted C-terminal region might act as an entropic bristle helping this variant to attain a proper fold and to sterically avoid aggregation. By contrast, because of its absence in Δ78Δ, the oligomerization of this shortest construct is no longer prevented.

On the other hand, it can be speculated that TFE might alter the native structure of Δ78Δ, thus interfering with interfacial protomer contacts and leading to rearrangement of this area, perhaps even implying dissociation and reassociation of protomers. The considerable hydrophobic surface involved in this molecular recognition event might underlie the enhanced aggregation propensity of Δ78Δ. Even though the native state of many proteins is indeed dimeric, dimer formation is also often the earliest event leading to aggregation for many amyloidogenic proteins [61]. The latter phenomenon is usually under kinetic control, with the dimer having a higher energy than the monomeric state. Importantly, dimerization can lock misfolded conformations, shifting the equilibrium away from the native state. In addition, dimerization can also catalyze β-strand conversion, and in this scenario, a constrained β-strand dimer would enhance the rate of oligomer formation [61].

CONCLUSIONS

In summary, IFABP and its truncated variants Δ98Δ and Δ78Δ have become useful models for studying the molecular determinants related to aggregation of the β-clam type of β-barrel proteins. A straightforward correlation between the decrease in the stability of *P* and the enhancement of the aggregation propensity cannot be established. This fact reveals the ancillary role played in these cases by considering solely the free energy of unfolding of the native form P. Additionally, as the three proteins discussed here share the same mechanism of aggregation, and a common aggregation prone fragment, variations in the speed of the process might be rationalized in terms of differences in the conformation of the protein, which in turn will determine the exposure/accessibility of this critical stretch. Noteworthy, for all three proteins the nucleus is bimolecular. The value of Δ78Δ lies in the fact that -because it is naturally dimeric- it reduces an otherwise bimolecular reaction to a unimolecular one. It is now clear that in order to properly understand those factors governing protein aggregation, it is mandatory to investigate the early events leading to the genesis of the critical nucleus.

This research has been supported by grants to J.M.D. and L.M.C. from the University of Buenos Aires (UBACyT B-901 and IJ-069), the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET PIP 1936) and the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT PICT 2011-0861 and 2010-0460). C.R.A. has been awarded a graduate student fellowship from CONICET and I.S.C., a fellowship from Consejo Interuniversitario Nacional (CIN). L.M.C., J.J.C. and J.M.D. hold teaching positions at UBA and are career researchers of CONICET.

REFERENCES

1. Carrell, R. W., and D. A. Lomas. 1997. Conformational diseases. *Lancet* 350:134-138.

2. Richardson J. S., Richardson D. C. (2002) Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. Proc Natl Acad Sci U. S. A 99:2754-2759

3. Curto, L. M., J. J. Caramelo, and J. M. Delfino. 2005. Δ98Δ, a functional abridged form of intestinal fatty acid binding protein. *Biochemistry.* 44:13847-13857.

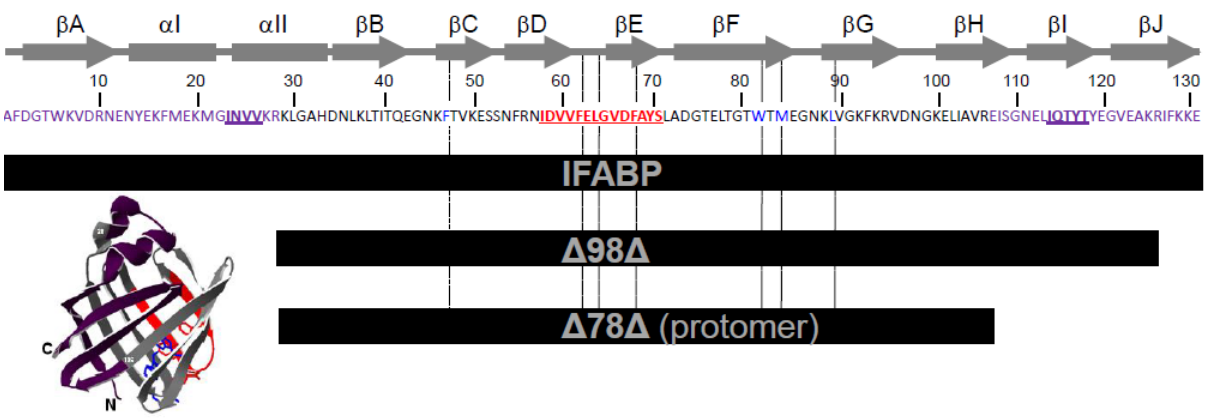
4. Curto, L. M., J. J. Caramelo, G. R. Franchini, and J. M. Delfino. 2009. Δ98Δ, a minimalist model of antiparallel β-sheet proteins based on intestinal fatty acid binding protein. *Protein Sci.* 18:735-746.

5. Franchini, G. R., L. M. Curto, J. J. Caramelo, and J. M. Delfino. 2009. Dissection of a β-barrel motif leading to a functional dimer: the case of intestinal fatty acid binding protein. *Protein Sci.* 18:2592-2602.

6. Yeh, S., I. J. Ropson, and D. L. Rousseau. 2001. Hierarchical folding of intestinal fatty acid-binding protein. *Biochemistry*. 40:4205-4210.
7. Hong, D., Hoshino, M., Kuboi, R. and Y. Goto. 1999. Clustering of Fluorine-Substituted Alcohols as a Factor Responsible for Their Marked Effects on Proteins and Peptides. J. Am. Chem. Soc. 121:8427–8433.
8. Povey, J. F., C. M. Smales, S. J. Hassard, and M. J. Howard. 2007. Comparison of the effects of 2,2,2-trifluoroethanol on peptide and protein structure and function. *J. Struct. Biol.* 157:329-338.
9. Chiti, F. and C. M. Dobson. 2009. Amyloid formation by globular proteins under native conditions. Nat. Chem. Biol. 5:15-22;
10. Calamai, M., Tartaglia, G. G., Vendruscolo, M., Chiti, F. and C. M. Dobson. 2009. Mutational analysis of the aggregation-prone and disaggregation-prone regions of acylphosphatase. J. Mol. Biol. 387:965-974;
11. Chiti, F., Taddei, N., Bucciantini, M., White, P., Ramponi, G. and C. M. Dobson. 2000. Mutational analysis of the propensity for amyloid formation by a globular protein. EMBO J. 19:1441-1449;
12. Gast, K., Zirwer, D., Müller-Frohne1, M. and G. Damaschun. 1999. Trifluoroethanol-induced conformational transitions of proteins: Insights gained from the differences between α-lactalbumin and ribonuclease A. Protein Sci. 8: 625–634;
13. Rezaei-Ghaleh, N., Ebrahim-Habibi, A., Moosavi-Movahedi, A. A. and M. Nemat-Gorgani. 2007. Role of electrostatic interactions in 2,2,2-trifluoroethanol-induced structural changes and aggregation of alpha-chymotrypsin. Arch. Biochem. Biophys. 457:160-169)
14. Curto, L. M., C. R. Angelani, J. J. Caramelo, and J. M. Delfino. 2012. Truncation of a β-barrel scaffold dissociates intrinsic stability from its propensity to aggregation. *Biophysical Journal.* 103:1929-1939.
15. Knowles, T. P., C. A. Waudby, G. L. Devlin, S. I. Cohen, A. Aguzzi, M. Vendruscolo, E. M. Terentjev, M. E. Welland, and C. M. Dobson. 2009. An analytical solution to the kinetics of breakable filament assembly. *Science*. 326:1533-1537.
16. Naiki, H., K. Higuchi, M. Hosokawa, and T. Takeda. 1989. Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavin T. *Anal Biochem.* 177:244-249.
17. Klunk, W. E, J. W. Pettegrew and D. J. Abraham. 1989. Quantitative evaluation of congo red binding to amyloid-like proteins with a beta-pleated sheet conformation. *J Histochem Cytochem.* 37:1273-1281.
18. Nilsson, M. R. 2004. Techniques to study amyloid fibril formation in vitro. *Methods.* 34:151-160.
19. Flyvbjerg, H., E. Jobs, and S. Leibler. 1996. Kinetics of self-assembling microtubules: An "inverse problem" in biochemistry. *Proc. Natl. Acad. Sci. U S A.* 93:5975–5979.
20. Wang, K., and B. I. Kurganov. 2003. Kinetics of heat- and acidification-induced aggregation of firefly luciferase. *Biophys. Chem.* 106:97–109.
21. Schmid, F. 1989. Spectral methods of characterizing protein conformation and conformational changes. In Protein Structure: a practical approach (Creighton, T.E., ed.), IRL, New York. 251.
22. Tsolis, A. C., N. C. Papandreou, V. A. Iconomidou, and S. Hamodrakas. 2013. A consensus method for the prediction of 'aggregation-prone' peptides in globular proteins. *PLoS One*. 8:e54175J.
23. Romero, P., Z. Obradovic, X. Li, E. C. Garner, C. J. Brown, and A. K Dunker. 2001. Sequence complexity of disordered protein. *Proteins.* 42:38-48.
24. Obradovic, Z., K. Peng, S. Vucetic, P. Radivojac, and A. K. Dunker.2005. Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins.* 61 Suppl 7:176-182.

25. Peng, K., S. Vucetic, P. Radivojac, C. J. Brown, A. K. Dunker, and Z. Obradovic. 2005. Optimizing long intrinsic disorder predictors with protein evolutionary information. *J. Bioinform. Comput. Biol.* 3:35-60.
26. Xue, B., R. L. Dunbrack, R. W. Williams, A. K. Dunker, and V. N. Uversky. 2010. PONDR-FIT: A meta-predictor of intrinsically disordered amino acids. *BBA-Proteins Proteom.* 1804:996-1010.
27. Oldfield, C. J., Y. Cheng, M. S. Cortese, P. Romero, V. N. Uversky, and A. K. Dunker. 2005. Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry.* 44:12454-12470.
28. Cheng, Y., C. J. Oldfield, J. Meng, P. Romero, V. N. Uversky, and A. K. Dunker. 2007. Mining alpha-helix-forming molecular recognition features with cross species sequence alignments. *Biochemistry*. 46:13468-13477.
29. Cohen, S. I. A., M. Vendruscolo, C. M. Dobson, and T. P. J. Knowles. 2012. From macroscopic measurements to microscopic mechanisms of protein aggregation. *J. Mol. Biol*. 421:160–171.
30. Le Vine 3rd, H. 1993. Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: detection of amyloid aggregation in solution. Protein Sci. 2:404–410.
31. Pallarès, I., J. Vendrell, F. X. Avilés, and S. Ventura. 2004. Amyloid fibril formation by a partially structured intermediate Sstate of α-chymotrypsin. *J. Mol. Biol.* 342: 321–331.
32. Arighi, C. N., J. P. F. C. Rossi, and J. M. Delfino. 1998. Temperature-induced conformational transition of intestinal fatty acid binding protein enhancing ligand binding: a functional, spectroscopic, and molecular modeling study. *Biochemistry.* 37:16802-16814.
33. Dalessio, P. M., S. E. Fromholt, and I. J Ropson. 2005. The role of Trp-82 in the folding of intestinal fatty acid binding protein. *Proteins.* 61:176-183.
34. Mukhopadhyay, K., and S. Basak. 1998. Conformation induction in melanotropic peptides by trifluoroethanol: fluorescence and circular dichroism study. *Biophys. Chem.* 74:175-186.
35. Garbuzynskiy, S. O., M. Y. Lobanov, and O. V. Galzitskaya. 2010. FoldAmyloid: a method of prediction of amyloidogenic regions from protein sequence. *Bioinformatics.* 26:326-332.
36. Scapin, G., J. I. Gordon, and J. C. Sacchettini. 1992. Refinement of the structure of recombinant rat intestinal fatty acid-binding apoprotein at 1.2-Å resolution. *J. Biol. Chem.* 267:4253-4269.
37. Sacchettini, J. C., G. Scapin, D. Gopaul, and J. I. Gordon. 1992. Refinement of the structure of *Escherichia coli*-derived rat intestinal fatty acid binding protein with bound oleate to 1.75 Å resolution. Correlation with the structures of the apoprotein and the protein with bound palmitate. *J. Biol. Chem.* 267:23534-23545.
38. Hodsdon, M. E., J. W. Ponder, and D. P. Cistola. 1996. The NMR solution structure of intestinal fatty acid-binding protein complexed with palmitate: application of a novel distance geometry algorithm. *J. Mol. Biol.* 264: 585-602.
39. Hodsdon, M. E., and D. P. Cistola. 1997. Discrete backbone disorder in the nuclear magnetic resonance structure of apo-intestinal fatty acid-binding protein: implications for the mechanism of ligand entry. *Biochemistry.* 36:1450-1460.
40. Disfani, F. M., W. L. Hsu, M. J. Mizianty, C. J. Oldfield, B. Xue, A. K. Dunker, V. N. Uversky, and L. Kurgan. 2012. MoRFpred, a computational tool for sequence-based prediction and characterization of disorder-to-order transitioning binding sites in proteins. *Bioinformatics.* 28:i75-83.
41. Chiti. F., and C. M. Dobson. 2006. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* 75:333-366.
42. Thirumalai, D., D. K. Klimov, and R. I. Dima. 2003. Emerging ideas on the molecular basis of protein and peptide aggregation. *Curr. Opin. Struct. Biol.* 13:146–159.

43. Monti, M., B. Lelj Garolla di Bard, G. Calloni, F. Chiti, A. Amoresano, G. Ramponi, and P. Pucci. 2004. The Regions of the Sequence Most Exposed to the Solvent Within the Amyloidogenic State of a Protein Initiate the Aggregation Process. *J. Mol. Biol.* 336:253–262.
44. Dunker, A. K., Z. Obradovic, P. Romero, E. C. Garner, and C. J. Brown. 2000. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* 11, 161-171.
45. Ward, J. J., J. S. Sodhi, L. J. McGuffin, B. F. Buxton, and D. T. Jones. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* 337:635-645.
46. Xue, B., A. K. Dunker, and V. N. Uversky. 2012. Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* 30:137-149.
47. Wright, P. E., and H. J. Dyson. 1999. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* 293: 321-31.
48. Uversky, V. N., J. R. Gillespie, and A. L. Fink. 2000. Why are "natively unfolded" proteins unstructured under physiologic conditions?. *Proteins.* 41: 415-427.
49. Dunker, A. K., C. J.Brown, J. D. Lawson, L. M. Iakoucheva, and Z. Obradovic. 2002. Intrinsic disorder and protein function. *Biochemistry.* 41:6573-6582.
50. Uversky, V. N. 2003. Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding cross-roads: which way to go? *Cell. Mol. Life Sci*. 60:1852-1871.
51. Uversky, V. N., C. J. Oldfield, and A. K. Dunker. 2005. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit.* 18:343-384.
52. Uversky, V. N. 2011. Multitude of binding modes attainable by intrinsically disordered proteins: a portrait gallery of disorder-based complexes. Chem. Soc. Rev. 40:1623-1634.
53. Tompa, P. 2012. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* 37:509-516.
54. Uversky, V. N. 2013. A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci*. 22:693-724.
55. Uversky, V. N. 2013. Unusual biophysics of intrinsically disordered proteins. *Biochim. Biophys. Acta.* 1834:932-951.
56. Uversky, V. N. 2013. Intrinsic disorder-based protein interactions and their modulators. Curr. Pharm. Des. 42:4191-4213.
57. Uversky V. N. 2013. Under-folded proteins: Conformational ensembles and their roles in protein folding, function and pathogenesis. *Biopolymers*. 99:870-997.
58. Uversky V. N. 2013. The most important thing is the tail: Multitudinous functionalities of intrinsically disordered protein termini. *FEBS Letters*. 587:1891-1901.
59. Hoh, J. H. 1998. Functional protein domains from the thermally driven motion of polypeptide chains: a proposal. *Proteins*. 32: 223-228.
60. Santner, A. A., C. H. Croy, F. H. Vasanwala, V. N. Uversky, Y. Y. Van, and A. K. Dunker. 2012. Sweeping away protein aggregation with entropic bristles: intrinsically disordered protein fusions enhance soluble expression. *Biochemistry.* 51:7250-7262.
61. Ma, B., and R. Nussinov. 2006. Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. *Curr. Opin. Chem. Biol.* 10:445–452.



Graphical abstract

HIGHLIGHTS

- IFABP, Δ98Δ and Δ78Δ share a primary nucleation-elongation mechanism of aggregation
- Irrespective of their original oligomerization state, the nucleus is bimolecular
- Dimeric Δ78Δ reduces an otherwise bimolecular reaction to a unimolecular one
- Accessibility of a common aggregation prone fragment dictates the speed of the process