Prostaglandins, Leukotrienes and Essential Fatty Acids 93 (2015) 37–43

Contents lists available at ScienceDirect

Prostaglandins, Leukotrienes and Essential Fatty Acids

journal homepage: www.elsevier.com/locate/plefa

Intervening in the β-barrel structure of lipid binding proteins: Consequences on folding, ligand-binding and aggregation propensity ☆

LM Curto, CR Angelani, JM 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

ARTICLE INFO

Keywords:
β-barrel proteins
Intestinal Fatty Acid Binding Protein (IFABP)
Folding
Ligand binding
Amyloid aggregation

ABSTRACT

Natural β-folds manage to fold up successfully. By contrast, attempts to dissect fragments or peptides from well folded β-sheet proteins have met with insurmountable difficulties. Here we briefly review selected successful cases of intervention on the well-known scaffold of intestinal fatty acid binding protein (IFABP). Lessons from these examples might set guidelines along the design of proteins belonging to this class. Impact of modifications on topology, binding and aggregation is highlighted. With the aid of abridged variants of IFABP we focus on key structural features responsible for the assembly into oligomeric forms or aggregates.

1. General structural features of lipid binding proteins (LBPs)

Lipid-binding proteins (LBPs) associate reversibly and non-covalently with lipids, greatly enhancing their aqueous solubility thus facilitating their transport between tissues and within cells. The main importance of their existence is to impart the cell with the right match of lipids, optimizing their bioavailability for particular physiological situations, thus greatly influencing those processes that depend on lipids, such as critical energy yielding or signaling steps. With few exceptions, no enzymatic activity has yet been described for most of them [1], and references cited therein. These proteins have been described in diverse species of eukaryotes in specialized cell types ranging from epithelial to neural tissue. One can attempt to understand the functional role of this hydrophobic ligand multigene family by elucidating the structure of the product proteins. This set can be classified into two groups according to the location of the members: intracellular (iLBP) and extracellular (eLBP) lipid-binding proteins [2]. Different types of LBPs include: soluble and membrane-bound proteins that can bind fatty acids, acyl-CoAs, sterols or retinoids, among other ligands. Although these proteins share a common functionality, members of diverse structural motifs are represented. For instance, unique to nematodes is the existence of helix-rich fatty acid and retinol binding proteins (FARs) and nematode polyprotein allergen/antigen (NPAs). FARs are present as various isoforms in each species and NPAs organize as large precursors comprised by several tandem repeats that may display different functions. Since these proteins have no counterparts in their plant or animal hosts, they represent potential targets for new antiparasitic drugs [3,4]. Also within the α-fold class, acyl-CoA binding protein (ACBP) is a 4-helix bundles with up-down connectivity where helices adopt a v-shaped orientation, giving rise to the ligand binding cavity [5]. On the other hand, sterol carrier protein 2 (SCP2) and phospholipid transfer protein (PLTP) belong to the αβ fold class. The former comprises a 5-strand β-sheet attached on the internal face to a α-helix-rich sub-domain [6,7], whereas the latter is a more complex array conformed by a 5 stranded β-sheet bound on one face and one edge by 9 α-helices PDB 2QGU: [8].

In spite of this wide variety, a characteristic β-barrel fold prevails as the most common binding motif. Most of the proteins belonging to the LBP family are characterized by a closely related folding pattern, although sharing as low as 20% overall sequence similarity. In fact, the distinctive feature of the majority of the LBPs is their very simple β-barrel architecture, giving rise to a large cavity for accommodating the hydrophobic ligand. The strands making up the barrel are entirely antiparallel such that the *i*th strand forms hydrogen-bonding networks with the $(i+1)$th and the $(i-1)$th strand. In the three dimensions, the first and last strands are situated so that they also form a ladder of hydrogen bonds [2]. In other words, this topology is compliant with an up-and-down barrel [9]. The

☆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. L.M.C. and J.M.D. hold teaching positions at UBA and are career researchers of CONICET.

* Corresponding author. Tel.: +54 11 4964 8290/8291x116; fax: +54 11 4962 5457.
E-mail address: delfino@qb.ffyb.uba.ar (J. Delfino).

http://dx.doi.org/10.1016/j.plefa.2014.08.001
0952-3278/

extracellular forms (eLBPs) – often also referred to as 'lipocalins'- are nearly spherical 8-strand barrels of about 175 amino acids in length. At variance with this, the intracellular members (iLBPs) are somewhat shorter, comprising about 130 amino acid residues. Because of their more flattened appearance, these 10-strand barrels are occasionally referred to as 'β-clams'.

2. Fatty acid binding proteins (FABPs)

Fatty acid binding proteins (FABPs) belong to the conserved multigene family of the iLBPs. The common function for all FABPs is their ability to bind fatty acids and/or closely related hydrophobic ligands. These proteins can be divided in four different subfamilies (named I to IV, [10]). A numerical nomenclature for the various FABPs genes has been introduced [11], but the various FABPs are still named after the tissue in which they have been first described, or are prominently expressed. It should be recognized, however, that such a classification is somewhat misleading, since most tissues express various FABP-forms [10]. The different forms of FABPs exhibit unique ligand binding profiles. Their affinity for distinct lipid species and/or the occurrence of competing ligands in the cell, including drugs, might match particular physiological situations with varying demands on the bioavailability of lipid species. For a review more focused on the tissue-specific role of ligands and FABPs on physiology we direct the reader to the review article by Storch and Thumser [12].

FABPs are monomeric antiparallel β-barrel proteins consisting of two five-strand β-sheets (named βA–βE and βF–βJ). These sheets are arranged in a nearly orthogonal orientation enclosing the ligand binding cavity. The FABP barrel is flattened and is not continuously hydrogen bonded, presenting a wide discontinuity between βD and βE. All β-strands are connected by β-turns with the exception of βA and βB, where an intervening helix-turn-helix motif appears. The latter is formed by two short but well-defined α-helices, which pack onto the top of the fatty acid-binding site. At variance with most globular proteins, the interior of FABPs is occupied by a large solvent-filled cavity, whereas the hydrophobic core is small and displaced from the protein center.

From a structural standpoint, one of the most studied members of this family is rat intestinal fatty acid binding protein (IFABP, Fig. 1). This is a 131-amino acid residue protein suitable for bacterial expression and straightforward purification. It contains neither Pro nor Cys residues, thus simplifying the folding mechanism, therefore avoiding the formation of inclusion bodies [13]. Its structure has been solved at high resolution by X-ray crystallography in the apo (ligand free) and holo forms (with oleate or palmitate bound) [14–16]. Additionally, the structure in solution of these forms has been elucidated by NMR [17–19]. In the wild-type complex, the methyl end of the fatty acid interacts with side chains in the α-helical region and the carboxylate of the bound fatty acid is buried deep inside, where it interacts primarily with Arg106, forming an ion pair/hydrogen bonded network.

A hierarchical folding mechanism has been proposed for IFABP [20]. According to this scheme, the unfolded polypeptide first collapses into a semi-compact structure surrounding a hydrophobic core consisting of Phe47, Phe62, Leu64, Phe68, Trp82, Met84 and Leu89. Next, strands βB to βG propagate outwardly from the hydrophobic core in a manner analogous to interlacing zippers. Although these are not tightly matched due to side-chain interactions, they will suffice to establish the native topology. Finally, this scaffold serves as a template to fold the remaining three β-strands (βH–βJ), completing the native hydrogen bonding network. As such, the formation of the native structure relies on the coordinated interaction between both halves of the protein. It has been argued that peptides of IFABP composed of either half of the clamshell would not form by themselves a β-sheet structure (I. J. Ropson, unpublished results, as cited in [20]). In this scenario, the closure of the β-barrel is a critical step to thoroughly consolidate the β-barrel fold. More specifically, this region comprises a hydrogen bonded network involving residues of the distal part of strands βA and βJ, each of them belonging to different β-sheets. Twisting of βA is a necessary prerequisite to achieve this goal. This feat allows βA to contact its neighbors: the N-terminal end makes contact with βB, whereas the C-terminal end interacts with βJ. Overall, this represents a sensible strategy used by β-sheet proteins to avoid edge-to-edge aggregation. It should be pointed out that this is not the only strategy to carry out this task [21].

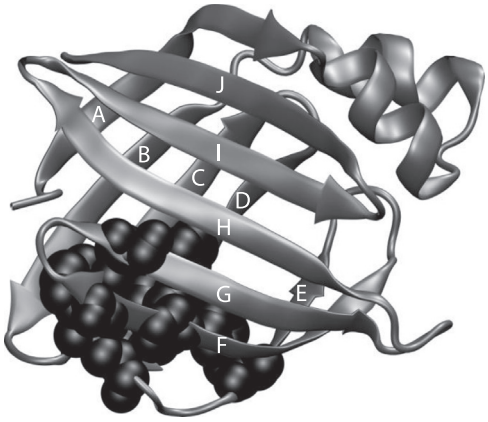

Fig. 1. Ribbon structure of IFABP (PDB 2IFB). Side chains of amino acid residues belonging to the hydrophobic core are depicted in space-filling representation.

Aggregation and ensuing insolubility are common occurrences in the design of β-sheet proteins. Nevertheless, natural β-folds are produced in the cell and manage to fold up successfully. By contrast, attempts to dissect fragments or peptides from well folded β-sheet proteins have met with insurmountable difficulties of insolubility. On the other hand, the build-up of β-structure dictates the process of amyloid genesis, leading even to speculate that the default most stable structure of many polypeptide chains is a β-sheet fiber [22]. Particularly, because of these problems, the *de novo* design of β-proteins has resulted in an extremely difficult endeavor [23–25]. In general, *de novo* β-designs are seldom soluble and monomeric, and have often required many rounds of redesign to achieve a successful construct. Many designs originally intended as globular β-sheet proteins have turned out instead to be valuable as models for amyloid-fiber formation. Retrospectively, the obvious feature that imparts regular β-sheets or β-sandwiches the ability to be aggregation prone is the solvent exposure of free edges where donor and acceptor groups are already set up to interact with any other β-strands in the vicinity. One way to rationalize this conundrum is to consider this a problem in 'negative design', i.e. in addition to those features necessary to reach the desired folded form, other traits should be embedded in the protein sequence to prevent the occurrence of a different alternative structure [21]. As a consequence of these difficulties, there are not many examples of β-barrel intervention. In this context, the accumulated knowledge available on IFABP makes it attractive as a target for structural intervention (Fig. 2). Judicious choice of appropriate modification of this scaffold might serve one to evaluate its impact on topology, binding and aggregation.

3. Interventions into the β-barrel of intestinal FABP (IFABP)

With the aim of populating a non-native state of the protein, a C-terminal truncated form of IFABP ($^{W6F}IFABP_{1\text{-}128}$) was designed by rational manipulation of the amino acid sequence [26]. By prior

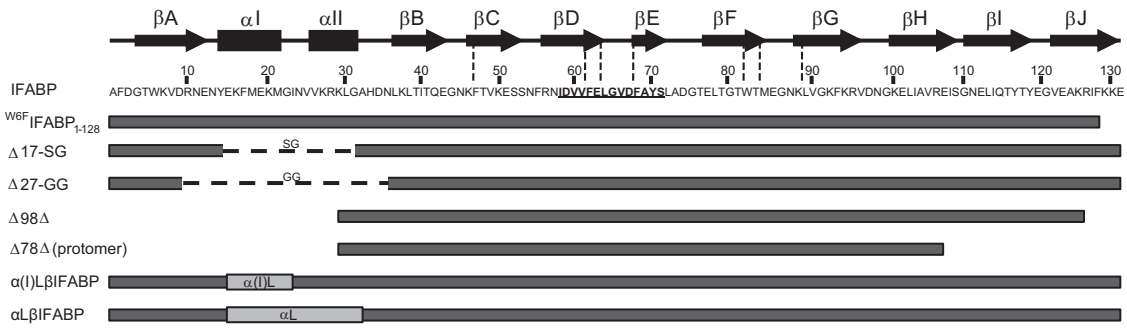

Fig. 2. Amino acid sequence and schematic representation of the secondary structure elements of IFABP. Amino acids belonging to the hydrophobic core are marked with dotted lines. The putative aggregation-prone peptide 58–71 is underlined. Linear representations of several IFABP variants are shown below.

inspection of the crystallographic model the authors reasoned that removing residues 129–131 would open the barrel, setting the two sheets free to slide against each other, and allowing free rotation of side-chains at the internal interface. The underlying assumption was that since the last three residues have charged side-chains, their removal would likely not interfere with the hydrophobic collapse of the molecule. However, the removal of a few specific interactions would hamper the formation of the native state. In fact, the controlled truncation produced a monomeric and less compact non-native state. All the spectroscopic evidence suggests a significant disruption with altered secondary and tertiary structure. Evidence from UV absorption and fluorescence spectroscopies reveals partial exposure of internal aromatic residues. However, Trp82 remains in a solvent-shielded environment more hydrophobic than that in the native state. Consistently, ^{W6F}IFABP$_{1-128}$ shows high susceptibility to proteolytic attack and an overall lower thermodynamic stability. Interestingly, this construct binds ANS with affinity typical of molten globules. In hindsight, according to the hierarchical folding scheme [20] the crucial role for folding of this last C-terminal segment will be related to the very late folding event whereby the β-barrel attains full closure.

The X-ray crystal structure of apo-IFABP [15,27] does not differ significantly from that of the holo protein [27–28]. Indeed, the mobility appears to be limited only to subtle conformational differences involving the helices, the turn between βE and βF, and the side chain of F55. These differences have led to the hypothesis that the helical region may be able to form an entry portal for the ligand [28]. However, neither structure reveals an opening large enough to accommodate the entry or exit of the ligand. Later on, NMR evidence revealed that more extensive backbone dynamics takes place at this region [17–19]. The disorder at this dynamic portal region is believed to play an essential role for regulating the transit of the fatty acid. Once the fatty acid binds, through a series of long range cooperative interactions the ligand imparts the backbone less flexibility.

The high degree of topological conservation of the helical domain in the iLBP protein family implies that it might serve an important structural and/or functional role. Much insight into this matter was gained after the intervention on the helix-turn-helix motif of the protein. In this regard, two helix-less variants were engineered to interrogate the system: Δ17-SG and Δ27-GG. The first construct deletes 17 amino acid residues comprising this subdomain. According to the available NMR and CD evidence, this protein is still able to fold into an essentially all β-sheet structure [29–31], although its stability is compromised by comparison with the wild-type protein [31]. This was conclusive evidence that the helix-turn-helix motif represents a dispensable moiety for the folding process of the β-clam, but it adds to the overall stability of the protein. The α-helical domain does not serve as a nucleation site for protein folding because its absence appears to bear little effect on the kinetics of refolding [31]. From a structural viewpoint, the lower stability might be rationalized in terms of lost hydrogen bonds and van der Waals contacts between α-helical and β-sheet subdomains. The removal of this motif discloses a large opening to the interior cavity, uncovering additional hydrophobic surface as a consequence of the loss of the relatively polar helical domain.

The ability of this helix-less variant to bind fatty acid ligands was probed by kinetic and NMR experiments [29]. The role of the α-helical domain appears to regulate the affinity of fatty acid binding by selectively altering the dissociation rate constant. In addition, HSQC and NOESY data served to elucidate the details of the binding mode of palmitate to this variant. The carboxyl end of the fatty acid binds to Δ17-SG in a manner similar to that observed in wild-type IFABP, forming an ion pair/hydrogen bonding network. By contrast, few or no nearest neighbor contacts are found between the fatty acid methyl protons and the protein. Because of the large opening to the interior cavity, the transfer of fatty acid is unimpeded. Thus, in full length IFABP the rate-limiting step would be the transition leading to a somewhat more open state. At variance, in Δ17-SG a protein conformational transition is not a prerequisite for binding. Removal of the proposed structural barrier to fatty acid entry -represented by the portal region- correlates with the loss of a kinetic barrier for binding.

Another turn of the screw along this line of research is exemplified by the design of Δ27-GG [32]. This is a second-generation helix-less variant of IFABP engineered by deleting 27 amino acid residues spanning the distal half of the first β-strand along with those that define the entire helical domain. This deletion includes the 10 residues that constitute an ill-defined loop in Δ17-SG. The optimization of loop length (with respect to Δ17-SG) led to a more stable and compact all-β-sheet protein. Interestingly, in this new construct the changes begin to compromise structural elements belonging to the β-barrel. More specifically, the pairwise interactions between βA and βJ are absent. Comparison of the free energies of unfolding obtained using chemical denaturation experiments revealed that the stability of Δ27-GG falls between that of the wild-type and that of Δ17-SG. However, as predicted by the authors, Δ27-GG bearing a short reverse turn between the first two strands is more stable than Δ17-SG, which is endowed with a longer unstructured loop. As regards the binding ability of Δ27-GG, the authors claim that this construct will be able to bind more than one fatty acid molecule, although the relative binding affinity could not be determined accurately [32]. This difference can hardly be supported by structural alterations between the two variants.

A different helix-less variant was also engineered based on the scaffold of the ileal lipid binding protein (ILBP) also known as bile acid binding protein (BABP) [33–34]. Here a stretch of 27 amino acids including most of the α-helical motif was replaced by a short flexible linker (Gly-Gly-Ser-Gly). The resulting construct was referred to as

Δα-ILBP. This protein is intrinsically disordered under physiological conditions, but folds in a ligand-assisted fashion with a binding affinity that is comparable to that of the wild type protein.

IFABP and liver FABP (LFABP) employ markedly different mechanisms of fatty acid transfer to acceptor model membranes. Transfer from IFABP occurs during protein-membrane collisional interactions, while for LFABP the transfer needs diffusion through the aqueous phase [35]. Chimeric proteins were engineered with the 'body' (ligand binding domain) of IFABP and either the αI-helix or the whole α-helical domain of LFABP exchanged, named α(I)LβIFABP [36] or αLβIFABP [37], respectively. All evidence shows that the fatty acid transfer properties of the resulting chimeric FABPs become mainly determined by the properties of the grafted moieties. Control experiments suggest that no major alterations occur in the conformation and binding site properties of the chimeric proteins relative to their parent wild-type protein IFABP. Key to the choice of mechanism is the charge distribution of the αI helix, which is involved in the primary electrostatic interaction step with the target membrane. In addition, subsequently the αII-helix participates in the consolidation of a protein-membrane 'collisional complex'.

A more radical approach to advance in the knowledge on key structural and functional determinants is to systematically abbreviate the parent protein insofar as stable and well behaved constructs are obtained. Rather than applying a 'rational' design procedure to achieve new forms of the protein, we took advantage of the 'selection' imposed by controlled proteolysis. This approach was based on the grounds established by the description of conformational sub-states of IFABP with defined properties that proved to be differently susceptible to enzymatic attack [38]. With this tool in hand, it was possible to deconstruct the β-barrel, thus dissecting the well known motif [39–41]. An equilibrium exists between at least two major conformations with different flexibility at the helical domain, namely, an open state and a closed more rigid state. The latter form corresponds to holo-IFABP showing the highest resistance to proteolysis. This crucial piece of evidence gave birth to Δ98Δ, the limiting fragment obtained after digestion with clostripain (ArgC).

Δ98Δ is a stable, monomeric and functional truncated form of IFABP, including only 98 amino acids corresponding to sequence 29–126 of IFABP. In comparison with the full-length protein, this protein is devoid of βA, most of the helical domain and the last five amino acids belonging to βJ. Most significantly, this truncation leads to the loss of both stretches involved in the closure of the β-barrel. Despite this fact, this fragment retains substantial β-sheet content and native tertiary interactions. The reasons for this remarkable behavior would lie on the ancillary role played by the segments deleted and on the conservation of all the critical residues of the hydrophobic core, that is, those involved in the nucleation event leading to the folded state. Δ98Δ might fold through a similar hierarchical scheme as IFABP. This process might lead to a native-like topology with a compact core but endowed with an expanded periphery. By its very nature, Δ98Δ would constitute itself into a minimalist model potentially capable of populating discrete intermediates sparsely represented in the conformational ensemble of the full-length protein. Moreover, despite the extensive truncation, Δ98Δ retains the ability to bind fatty acids with affinities similar to those of the helix-less variants and IFABP. The binding event would result crucial for the stabilization of side-chain contacts leading to an ultimate readjustment of the tertiary structure. It is noteworthy that Δ98Δ lacks any tendency to aggregate under the conditions assayed. This happens despite the likely occurrence of free edges in the construct.

Similarly, in a second round of proteolysis holo-Δ98Δ also gives rise to Δ78Δ. This new limiting fragment lacks both the 1–28 and the 107–131 stretches. Here again, all the residues belonging to the hydrophobic core of IFABP are preserved. Diverse spectroscopic techniques support the fact that Δ78Δ adopts a stable well-folded state. The most distinctive feature of Δ78Δ is that it folds into a dimer preserving the ability to bind fatty acids. In fact, the interaction with the ligand brings about a gain in structure. In principle, the extensive stretches deleted in Δ78Δ could determine the appearance of free edges prompting dimerization, thus avoiding the formation of higher order aggregates.

Binding of trans-parinaric acid uncovers the enhanced flexibility of both abridged forms when compared with IFABP. On the other hand, all three proteins are able to bind the fluorescent probe ANS. Although both variants bind the probe less tightly than the parent protein, they still exhibit an affinity well above that shown by a typical molten globule state. However, they differ in their displacement behavior with oleic acid. Indeed, in the abridged variants the fatty acid effectively competes with ANS to cause the removal of a major part of the bound probe. However, the fact that a sizeable amount remains bound, even in the presence of a large excess of fatty acid, discloses the existence of a binding site in the constructs that is absent in the full-length protein.

In hindsight, one can rationalize that preservation of the hydrophobic core of the protein arises as an essential demand to attain a proper fold and function. As has already been proven in subsequent work, perturbation around this region is conducive to aggregation. Accordingly, consensus prediction of aggregation-prone regions fit well with this view because the putative segment is embedded in the most hydrophobic section of the protein (underlined stretch shown in Fig. 2) [42–43]. In this scenario, Δ98Δ and Δ78Δ emerge as useful models to explore critical determinants leading to β-aggregates, a hot topic of study nowadays for its relevance in the formation of amyloid fibrils.

4. Peering into the mechanism of aggregation of the β-barrel of IFABP: insights derived from the kinetics of the phenomenon

Achieving full understanding of the mechanism of protein aggregation will represent a breakthrough in the context of both physiological and pathological phenomena occurring in nature. In regard to this general aim, the protein IFABP provides an interesting scaffold amenable to molecular intervention. In particular, the structure of IFABP and those of the excised forms Δ98Δ and Δ78Δ described above allowed us to interrogate the system in terms of key structural features responsible for assembly into oligomeric forms or aggregates. Characteristic signatures of the aggregation phenomenon become apparent from the examination of the time-course of aggregation reaction [42–43]. In this section we highlight key facts arising from this analysis that permit to establish a rigorous link between macroscopic observations and the underlying microscopic events. In regard to the methods used, developments in this direction have already been reported by Ferrone [44], Flyvbjerg et al. [45] and Cohen et al. [46] and summarized by Morris et al. [47].

Among the biophysical techniques used to monitor the aggregation process [47], turbidity (apparent absorbance at 350 nm) [42,43,45,48] or fluorescent signals [49–51] have been most frequently used to monitor the process in real time. As a general observation, all kinetics of aggregation share a consistent sigmoid character. In a first approximation, the empirical treatment by Wang and Kurganov [48] was useful to extract descriptive kinetic parameters. Importantly, this formalism only deals with the elongation phase of fibril formation and proposes that kinetics should obey the relationship $A = A_{\lim}(1-\exp(-k(t-t_0)))$. In this fashion, the calculation of parameters such as lag time (t_0), overall pseudo-first order rate constant (k) and limiting turbidity (A_{lim}) was accomplished [42,43]. A_{lim} values may vary from protein to protein, but are always directly proportional to the total protein concentration (P_0), thus validating the use of the measurement of turbidity for evaluating the

magnitude of aggregation. Moreover, for IFABP and its variants all protein undergoes a transition to fibrillar aggregates, leaving in the end scant or none soluble form. Starting from $A/A_{lim} \sim 0.1$, all kinetics are satisfactorily described by pseudo-first-order processes. As will be later explained, the duration of the lag phase (t_0) holds a hyperbolic dependence on protein concentration, i.e. it is inversely proportional. On the other hand, a linear dependence of the empirical pseudo-first order constant k with protein concentration is also observed. Consequently, the aggregation rate, i.e. the product kA_{lim} represented by the slope of the kinetics at the onset of the elongation phase, results a linear function of the square of protein concentration. As such, the order of aggregation with respect to protein concentration is equal to 2. In this regard, Δ78Δ shows the highest rate, followed by full-length IFABP and Δ98Δ [42,43].

This empirical treatment simply ignores the early evolution of turbidity, equating this period to a flat line. This approximation leaves behind the initial region that, although demanding in terms of data acquisition due to the low value of the signal, provides further insight into the process of aggregation. However, a deeper comprehension of the fundamental process can be grasped after accomplishing a global analysis of all time series recorded at different protein concentrations. Moreover, this action can later be extended to different proteins to evaluate if they comply or not with a single mechanism of aggregation. One cannot underestimate the power of the so-called phenomenological scaling, as depicted in Fig. 3A, where several time series for IFABP, Δ98Δ and Δ78Δ run in the concentration range 3–60 μM were found to precisely superimpose after scaling both A_{lim} and time (for the latter, any characteristic time (t_c) of each trace, such as t_0 or $t_{1/2}$ can be used). The meaning of this scalability condition $A/A_{\mathrm{lim}} = f[t/t_c]$ is that all kinetics can be described by a single function f [40]. In other words, fulfilling scalability means sharing a common kinetic mechanism. In mathematical terms, after scaling the variables a single set of differential equations should represent all data (see below).

In fact, early and late time dependences of reaction time provide information on the dominant process leading to aggregation [46]. In the case of our trio of IFABP variants, the aggregation signature can be readily explained by a common primary nucleation event followed by elongation fed by bulk protein. This means that at early stages, a polynomial time course is verified, whereas at late stages an exponential dependence takes place. In this fashion, we could safely rule out secondary pathways to aggregation, such as fibril fragmentation giving rise to additional nuclei or surface catalyzed nucleation. Relevant for this is the fact that avoidance of mechanical stirring was the rule for all samples.

The phenomenological data analysis imposes constraints and reveals demands on any model that attempts to describe the experimental data. At this point, our rationale was to consider that kinetic scheme representing the minimal number of assumptions able to explain the body of experimental data. In this fashion we proposed the following simple aggregation scheme, comprising two subsequent irreversible steps. The first describes the genesis of the active nucleus and the second depicts the growth of aggregates:

$$nP \xrightarrow{f_0} P_n$$
$$P + P_n \xrightarrow{f_1} F$$

The following set of differential equations fully describes this system:

$$dP/dt = -f_0 P^n - f_1 P P_n$$
$$dP_n/dt = f_0 P^n - f_1 P P_n$$
$$dF/dt = f_1 P P_n$$

Please note that the measured signal A is proportional to F (the bulk concentration of fibrillar aggregates expressed in terms of protein concentration). Therefore, A/A_{lim} results equal to F/F_∞. In this context, the scalability condition means that after due scaling of turbidity $\hat{F} = F/F_\infty$, protein $\hat{P} = P/P_0$ and nucleus $\hat{P}_n = P_n/\mu$ concentrations, and time $\hat{t} = t/t_c$ the system reduces to the following single set of equations:

$$d\hat{P}/d\hat{t} = -\hat{P}^n - \hat{P}\hat{P}_n$$
$$d\hat{P}_n/d\hat{t} = \hat{P}^n - \hat{P}\hat{P}_n$$
$$d\hat{F}/d\hat{t} = \hat{P}\hat{P}_n$$

provided that the following conditions are met: $P_0 = F_\infty$ and $t_c = (1/f_0 f_1)^{1/2}/P^{n/2}$.

In this fashion, the system becomes independent from f_0 and f_1, so that all time traces should overlap, regardless of protein concentration. Moreover, different proteins formally complying with the same mechanistic scheme will also give rise to superimposable time traces. Interestingly, this was the case of our trio of β-barrel proteins when the trigger for aggregation is 25% v/v TFE [42,43].

Moreover, this approach permits to (i) elucidate the number of steps required to build an active nucleus, and (ii) define the exact

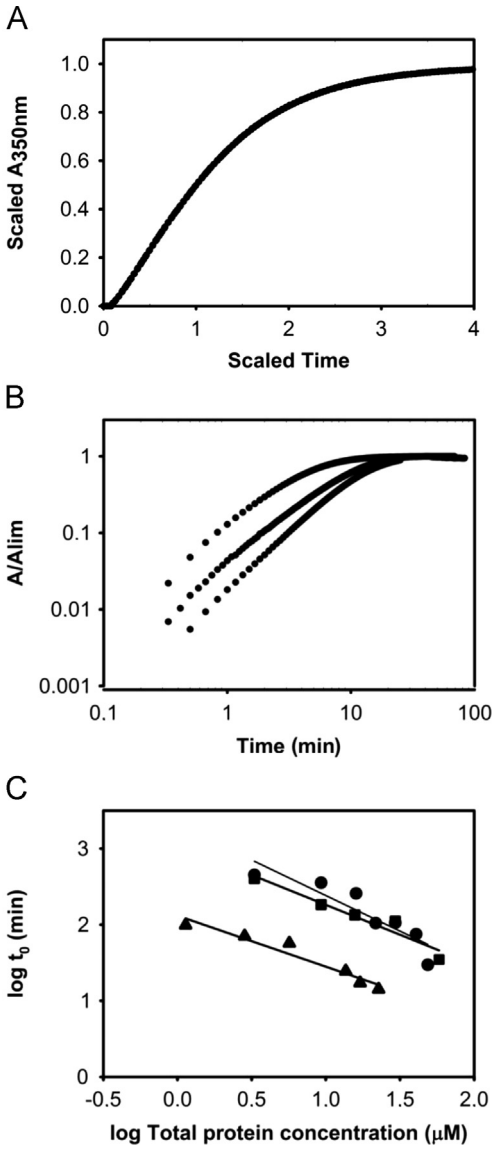

Fig. 3. Challenging the scaffolds of IFABP (■), and those of the extensively abbreviated forms Δ98Δ (●) and Δ78Δ (▲) with 25% v/v trifluoroethanol (TFE) readily triggers aggregation. (a) Scalability condition, double-log plots of (b) the evolution of the signal (A/A_{lim}) during the early time course, (c) the dependence of t_0 on total protein concentration.

oligomeric state (stoichiometry) of this nucleus. To delve into this matter, one should also recognize other phenomenological restraints and their consequences on the model. Along the initial (lag) phase where $P \sim P_0, P_n \sim 0, F \sim 0$, a defined relationship was found (Fig. 3B) $F/F_\infty \sim f_0 f_1 P_0^n t^2/2$, giving rise to: $\log(F/F_\infty) \sim 2 \log t + \text{constant}$.

The meaning of the slope represents the number $k+2$, where k is the number of intermediate assembly stages of the nucleus P_n. As such, a value of ~2 (meaning that $k \sim 0$ for our three proteins, see Figs. 3B and 4) indicates that nucleation occurs in one step (the first step of the scheme), i.e. no intermediate stages exist from protein to the stable nucleus [45].

A second defined relationship also holds true between the duration of t_0 and A_{lim}: $\log t_0 \approx slope \log A_{\lim} + \text{constant}'$. The slope of this plot (Figs. 3C and 4) contains information on the stoichiometry n of the nucleus P_n. The measured values are ~ -1 for IFABP and Δ98Δ, and $\sim -1/2$ for Δ78Δ. As regards the meaning of this relationship, from the proposed model it readily appears that before A reaches 10% of its maximum value A_{lim} (i.e. $A/A_{lim} \sim 0.1$) the following holds true: $0.1 \sim f_0 f_1 P_0^n t_0^2/2$. Knowing the equivalence $P_0 = F_\infty$, the parameter t_0 follows the equation $t_0 = (0.2/f_0 f_1)^{1/2}/P_0^{n/2}$ (note that t_0 complies with the form for t_c, see above), giving rise to the following double log plot: $\log t_0 \sim (-n/2)\log F_\infty + C$ from which stoichiometry n can be inferred. Interestingly, prior to the onset of aggregation the monomeric forms IFABP and Δ98Δ will require to self-associate into dimeric entities (in both cases, $n=2$ transitions are verified, Figs. 3C and 4). At variance with this, the naturally occurring dimeric form Δ78Δ might only need to undergo a unimolecular transition to the activated dimeric nucleus (a $n=1$ transition, Fig. 3C), not excluding a conceivable dissociation/reassociation occurrence [38]. Most relevantly, the behavior observed for the latter holds formal resemblance to the incidence of monomeric nuclei (another $n=1$ transition), as has already been reported for the protein CRABP I [50], and for the polyE peptides similar to the N-terminal extension of huntingtin [52,53]. Additionally, one should note that by itself or in the presence of Δ98Δ, the major excised peptide 1–28 does not form aggregates [42].

In summary, in our case the evolution of turbidity measurements demonstrates a common scaling behavior of the kinetic parameters, pointing to a primary nucleation-elongation mechanism, whereby the stabilization of dimeric nuclei precedes the association of protein to the growing amyloid-like aggregates (Fig. 4). In a more general context, from the systematic kinetic examination carried out on bulk experimental data on protein aggregation we were able to derive detailed and quantitative insights into the underlying microscopic mechanisms. With this knowledge available, one could attempt to understand the effect of small or large deletions or insertions, transpositions, point mutations or chemical modifications of the target protein in terms of defined kinetic parameters on key steps of the process of aggregation. Such information will likely be of great use to shed light on normal processes or -when the outcome goes astray- on the origin of pathologies. Undoubtedly, this new comprehension will be of fundamental value in establishing modes of intervention on aggregation diseases with effector molecules, hopefully leading to the development of new therapies.

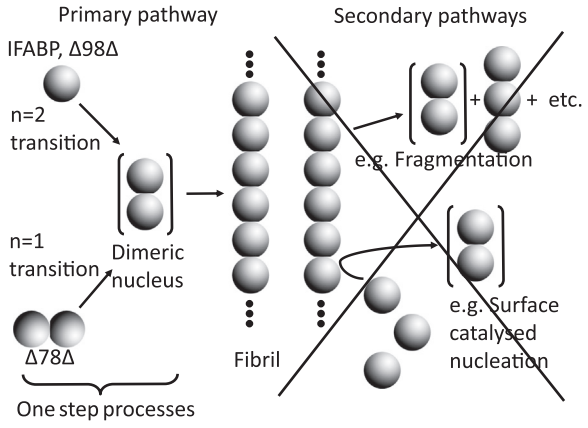

Fig. 4. Pictorial conception on how the β-barrel structures of IFABP and abridged variants will proceed to form fibrillar aggregates.

Acknowledgments

We thank Dr. Betina Corsico and the members of her group for sharing their enthusiasm. The authors thank Dr. Mariano González Lebrero for his help with rendering Fig. 1.

References

[1] J.F.C. Glatz, Lipids and lipid binding proteins: A perfect match, Prostaglandins, Leukotrienes and Essential Fatty Acids (2014), http://dx.doi.org/10.1016/j.plefa.2014.07.011.
[2] L. Banaszak, N. Winter, Z. Xu, D.A. Bernlohr, S. Cowan, T.A. Jones, Lipid-binding proteins: a family of fatty acid and retinoid transport proteins, Adv. Protein Chem. 45 (1994) 89–151.
[3] M. Kennedy., The nematode polyprotein allergens/antigens, Parasitol. Today 16 (2000) 373–380.
[4] A.S. Solovyova, N. Meenan, L. McDermott, A. Garofalo, J.E. Bradley, M.W. Kennedy, O. Byron, The polyprotein and FAR lipid binding proteins of nematodes: shape and monomer/dimer states in ligand-free and bound forms, Eur. Biophys. J. 32 (2003) 465–476.
[5] B.B. Kragelund, J. Knudsen, F.M. Poulsen, Acyl-coenzyme a binding protein (ACBP), Biochim. Biophys. Acta 1441 (1999) 150–161.
[6] F.P. De Berti, S. Capaldi, R. Ferreyra, N. Burgardt, J.P. Acierno, S. Klinke, H.L. Monaco, M.R. Ermácora., The crystal structure of sterol carrier protein 2 from Yarrowia lipolytica and the evolutionary conservation of a large, non-specific lipid-binding cavity, J. Struct. Funct. Genomics 14 (2013) 145–153.
[7] N.J. Stolowicha, A.D. Petrescub, H. Huang b, G.G. Martinb, A.I. Scott c, F. Schroeder, Sterol carrier protein-2: structure reveals function, Cell. Mol. Life Sci. 59 (2002) 193–212.
[8] A.P. Kuzin, Y. Chen, S. Jayaraman, C.X. Chen, Y. Fang, K. Cunningham, L.C. Ma, R. Xiao, J. Liu, M.C..Baran, T.B. Acton, B..Rost, G.T..Montelione, J.F..Hunt and L.. Tong, Three-dimensional structure of the phospholipid-binding protein from Ralstonia solanacearum Q8XV73_RALSQ in complex with a phospholipid at the resolution 1.53 A, Unpublished results.
[9] C. Branden, J. Tooze, Introduction to Protein Structure, Garland Publishing, New York (1999) 68.
[10] N.H. Haunerland, F. Spener, Fatty acid-binding proteins – insights from genetic manipulations, Prog. Lipid Res. 43 (2004) 328–349.
[11] A. Vogel Hertzel, D.A. Bernlohr, The mammalian fatty acid-binding protein multigene family: molecular andgenetic insights into function, Trends Endocrinol. Metab. 11 (2000) 175–180.
[12] J. Storch, A.E. Thumser, Tissue-specific functions in the fatty acid-binding protein family, J. Biol. Chem. 285 (2010) 32679–32683.
[13] J.C. Sacchettini, L.J. Banaszak, J.I. Gordon, Expression of rat intestinal fatty acid binding protein in E. coli and its subsequent structural analysis: a model system for studying the molecular details of fatty acid-protein interaction, Mol. Cell. Biochem. 98 (1990) 81–93.
[14] J.C. Sacchettini, J.I. Gordon, L.J. Banaszak, Crystal structure of rat intestinal fatty-acid-binding protein. Refinement and analysis of the Escherichia coli-derived protein bound with palmitate, J. Mol. Biol. 208 (1989) 327–339.
[15] G. Scapin, J.I. Gordon, J.C. Sacchettini, Refinement of the structure of recombinant rat intestinal fatty acid-binding apoprotein at 1.2-Å resolution, J. Biol. Chem. 267 (1992) 4253–4269.
[16] G. Scapin, A.C. Young, A. Kromminga, J.H. Veerkamp, J.I. Gordon, J.C. Sacchettini, High resolution X-ray studies of mammalian intestinal and muscle fatty acid-binding proteins provide an opportunity for defining the chemical nature of fatty acid: protein interactions, Mol. Cell. Biochem. 123 (1993) 3–13.
[17] M.E. Hodsdon, J.W. Ponder, D.P. Cistola, The NMR solution structure of intestinal fatty acid-binding protein complexed with palmitate: application of a novel distance geometry algorithm, J. Mol. Biol. 264 (1996) 585–602.
[18] M.E. Hodsdon, D.P. Cistola, 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 (1997) 1450–1460.
[19] M.E. Hodsdon, D.P. Cistola, Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by 15N NMR relaxation and 1H exchange, Biochemistry 36 (1997) 2278–2290.
[20] S. Yeh, I.J. Ropson, D.L. Rousseau, Hierarchical folding of intestinal fatty acid-binding protein, Biochemistry 40 (2001) 4205–4210.

[21] J.S. Richardson, D.C. Richardson, Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation, Proc. Natl. Acad. Sci. USA 99 (2002) 52754–52759.
[22] C.M. Dobson, Experimental investigation of protein folding and misfolding, Methods 34 (2004) 4–14.
[23] A. Pessi, E. Bianchi, A. Crameri, S. Venturini, A. Tramontano, M. Sollazzo, A designed metal-binding protein with a novel fold, Nature 362 (1993) 367–369.
[24] T.P. Quinn, N.B. Tweedy, R.W. Williams, J.S. Richardson, D.C. Richardson, Betadoublet: de novo design, synthesis, and characterization of a beta-sandwich protein, Proc. Natl. Acad. Sci. USA 91 (1994) 8747–8751.
[25] M.W. West, W. Wang, J. Patterson, J.D. Mancias, J.R. Beasley, M.H. Hecht, De novo amyloid proteins from designed combinatorial libraries, Proc. Natl. Acad. Sci. USA 96 (1999) 11211–11216.
[26] E.M. Clérico, S.G. Peisajovich, M. Ceolín, P.D. Ghiringhelli, M.R. Ermácora, Engineering of a compact non-native state of intestinal fatty acid binding protein, Biochim. Biophys. Acta 36047 (1999) 1–16.
[27] J.C. Sacchettini, J.I. Gordon, L. Banaszak, Refined apoprotein structure of rat intestinal fatty acid binding protein produced in Escherichia coli, Proc. Natl. Acad. Sci. USA 86 (1989) 7736–7740.
[28] J.C. Sacchettini, G. Scapin, D. Gopaul, J.I. Gordon, 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 (1992) 23534–23545.
[29] D.P. Cistola, K. Kim, H. Rogl, C. Frieden, Fatty acid interactions with a helix-less variant of intestinal fatty acid-binding protein, Biochemistry 35 (1996) 7559–7565.
[30] R.A. Steel, D.A. Emmert, J. Kao, M.E. Hodson, C. Frieden, D.P. Cistola, The three-dimensional structure of a helix-less variant of intestinal fatty acid-binding protein, Protein Sci. 7 (1998) 1332–1339.
[31] K. Kim, D.P. Cistola, C. Frieden, Intestinal fatty acid-binding protein: the structure and stability of a helix-less variant, Biochemistry 35 (1996) 7553–7558.
[32] B. Ogbay, G.T. Dekoster, D.P. Cistola, The NMR structure of a stable compact all-β-sheet variant of intestinal fatty acid-binding protein, Protein Sci. 13 (2003) 1227–1237.
[33] N. Kouvatsos, J.K. Meldrum, M.S. Searle, N.R. Thomas, Coupling ligand recognition to protein folding in an engineered variant of rabbit ileal lipid binding protein, Chem. Commun. (2006) 4623–4625.
[34] N. Kouvatsos, V. Thurston, K. Ball, N.J. Oldham, N.R. Thomas, M.S. Searle, Bile acid interactions with rabbit ileal lipid binding protein and an engineered helixless variant reveal novel ligand binding properties of a versatile β-clam shell protein scaffold, J. Mol. Biol. 371 (2007) 1365–1377.
[35] K.T. Hsu, J. Storch, Fatty acid transfer from liver and intestinal fatty acidbinding proteins to membranes occurs by different mechanisms, J. Biol. Chem. 271 (1996) 13317–13323.
[36] G.R. Franchini, J. Storch, B. Corsico., The integrity of the α-helical domain of intestinal fatty acid binding protein is essential for the collision-mediated transfer of fatty acids to phospholipid membranes, Biochim. et Biophys. Acta 1781 (2008) 192–199.
[37] B. Córsico, H.L. Liou, J. Storch, The α-helical domain of liver fatty acid binding protein is responsible for the diffusion mediated transfer of fatty acids to phospholipid membranes, Biochemistry 43 (2004) 3600–3607.
[38] C.N. Arighi, J.P.F.C. Rossi, J.M. Delfino, Temperature-induced conformational switch in intestinal fatty acid binding protein (IFABP) revealing an alternative mode for ligand binding, Biochemistry 42 (2003) 7539–7551.
[39] L.M. Curto, J.J. Caramelo, J.M. Delfino, Δ98Δ, a functional abridged form of intestinal fatty acid binding protein, Biochemistry 44 (2005) 13847–13857.
[40] L.M. Curto, J.J. Caramelo, G.R. Franchini, J.M. Delfino, Δ98Δ, a minimalist model of antiparallel β-sheet proteins based on intestinal fatty acid binding protein, Protein Sci. 18 (2009) 735–746.
[41] G.R. Franchini, L.M. Curto, J.J. Caramelo, J.M. Delfino, Dissection of a β-barrel motif leading to a functional dimer: the case of intestinal fatty acid binding protein, Protein Sci. 18 (2009) 2592–2602.
[42] L.M. Curto, C.R. Angelani, J.J. Caramelo, J.M. Delfino, Truncation of a β-barrel scaffold dissociates intrinsic stability from its propensity to aggregation, Biophys. J. 103 (2012) 1929–1939.
[43] C.R. Angelani, L.M. Curto, I.S. Cabanas, J.J. Caramelo, V. Uversky, J.M. Delfino, Toward a common aggregation mechanism for a β-barrel protein family: insights derived from a stable dimeric species, Biochim. Biophys. Acta 2014 (1844) 1599–1607.
[44] F. Ferrone, Analysis of protein aggregation kinetics, Methods Enzymol. 309 (1999) 256–274.
[45] H. Flyvbjerg, E. Jobs, S. Leibler, Kinetics of self-assembling microtubules: an inverse problem in biochemistry, Proc. Natl. Acad. Sci. USA 93 (1996) 5975–5979.
[46] S.I.A. Cohen, M. Vendruscolo, C.M. Dobson, T.P.J. Knowles, From macroscopic measurements to microscopic mechanisms of protein aggregation, J. Mol. Biol. 421 (2012) 160–171.
[47] A.M. Morris, M.A. Watzky, R.G. Finke, Protein aggregation kinetics, mechanism, and curve-fitting: a review of the literatura, Biochim. Biophys. Acta 2009 (1794) 375–397.
[48] K. Wang, B.I. Kurganov, Kinetics of heat- and acidification-induced aggregation of firefly luciferase, Biophys. Chem. 106 (2003) 97–109.
[49] M.R. Nilsson, Techniques to study amyloid fibril formation in vitro, Methods 34 (2004) 151–160.
[50] Z. Ignatova, L.M. Gierasch, Aggregation of a slow-folding mutant of a beta-clam protein proceeds through a monomeric nucleus, Biochemistry (2005) 7266–7274.
[51] D.A. Yushchenko, J.A. Fauerbach, S. Thirunavukkuarasu, E.A. Jares-Erijman, T.M. Jovin, Fluorescent ratiometric MFC probe sensitive to early stages of alpha-synuclein aggregation, J. Am. Chem. Soc. 132 (2010) 7860–7861.
[52] S. Chen, F.A. Ferrone, R. Wetzel, Huntington's disease age-of-onset linked to polyglutamine aggregation nucleation, Proc. Natl. Acad. Sci. 99 (2002) 11884–11889.
[53] A.K. Thakur, M. Jayaraman, R. Mishra, M. Thakur, V.M. Chellgren, I.J. Byeon, D.H. Anjum, R. Kodali, T.P. Creamer, J.F. Conway, A.M. Gronenborn, R. Wetzel, Polyglutamine disruption of the huntingtin exon 11N terminus triggers a complex aggregation mechanism, Nat. Struct. Mol. Biol. 16 (2009) 380–389.