Thermodynamic properties of amyloid fibrils in equilibrium

Tomaz Urbica,*, Sara Najemb, Cristiano L. Diasc

aFaculty of Chemistry and Chemical Technology, University of Ljubljana, Vecna pot 113, 1000, Slovenia
bNational Center for Remote Sensing, National Council for Scientific Research (CNRS), Riad al Soloh, 1107 2260 Beirut, Lebanon
cNew Jersey Institute of Technology, Physics Department, Newark,NJ 07042-1982,United States

HIGHLIGHTS
- Two-dimensional coarse-grained model to study how amyloid fibrils grow towards an equilibrium is proposed.
- Thermodynamics and structural properties of the model are studied.
- Free energies to dissociate proteins from fibrils are estimated from the residual concentration of dissolved proteins.

GRAPHICAL ABSTRACT

ABSTRACT

In this manuscript we use a two-dimensional coarse-grained model to study how amyloid fibrils grow towards an equilibrium state where they coexist with proteins dissolved in a solution. Free-energies to dissociate proteins from fibrils are estimated from the residual concentration of dissolved proteins. Consistent with experiments, the concentration of proteins in solution affects the growth rate of fibrils but not their equilibrium state. Also, studies of the temperature dependence of the equilibrium state can be used to estimate thermodynamic quantities, e.g., heat capacity and entropy.

1. Introduction

The self-assembly of proteins into amyloid-like fibrils is a ubiquitous process that arises for seemingly unrelated amino acid sequences when subjected to the right conditions [1,2]. Segments of proteins incorporated into these fibrils adopt extended conformations, i.e., $\beta$-strands, which are hydrogen bonded to neighboring proteins forming a $\beta$-sheet. Stacking of $\beta$-sheets through side chain interactions accounts for the common pattern of amyloid fibrils, also known as cross-$\beta$ structure [3,4]—see Fig. 1a. Proteins that form fibrils at physiological conditions were found to be either functional, e.g., silk fibrils that are used by spiders to capture their prey, or related to plaque formation in amyloid diseases that include Alzheimer’s and Parkinson’s [5,6]. Fibril formation proceeds through a nucleation phase followed by growth [7–9]. In the nucleation phase, some proteins form a disordered aggregate [10,11] from which the fibril nucleates while in the growth phase proteins are added to the fibril increasing its length. Important questions regarding the pathway of fibril formation [12–14] as well as the structure of the disordered aggregates remain open [11]. Also, understanding the equilibrium of mature fibrils in solution is of great interest as it could enable the development of a thermodynamic framework to study these structures. In analogy with protein folding [15], thermodynamics may provide insights into the stability of fibrils and their underlying molecular mechanisms [16].
Equilibrium thermodynamic properties of amyloid fibrils are challenging to measure using conventional experimental methods. For example, heat capacities measured as a function of temperature in calorimetric experiments were found to be irreversible for most fibrils precluding the use of analytic frameworks to compute free-energies [17,18]. Moreover, it is difficult to deconvolute contributions from molecular interactions between proteins and fibrils from these calorimetric heat capacities that depend on the overall morphology of the fibrillar network. Overcoming these limitations, recent “amyloid assembly equilibrium measurements” have shown that at least for some protein sequences mature fibrils can exist in equilibrium with dissolved proteins, i.e., monomers [19–23]. This equilibrium has been explored to measure free-energies required to dissociate a protein from a fibril, i.e., ΔG. Studies of the temperature dependence of this equilibrium can also provide insights of other thermodynamic quantities, i.e., change in enthalpy (ΔH), entropy (ΔS), and heat capacity (ΔCp). These quantities played a key role in determining the molecular driving forces of protein folding [24,25] but they remain mostly unknown for amyloid fibrils. Only few experiments have estimated ΔCp for amyloid fibrils finding both positive and negative values [18,26–28]. Currently, it is unknown how the magnitude and sign of ΔCp relate to the peptide sequence, fibril structure, and/or condition of the solvent. Notice that unfolding of proteins accounts for a large and positive (never negative) change in heat capacity [24,29–36]. Timescales to equilibrate fibrils and proteins in a solution are of the order of hours and days which is beyond reach of all-atom computer simulations [20,23,37]. Thus, models to study these equilibrium reactions require a coarse-grained approach [38–40].

Here, we develop an off-lattice coarse-grained model to study how amyloid fibrils grow towards an equilibrium state wherein free-energies ΔF to add or dissociate a protein from a fibril can be measured. The initial density of proteins in a solution affects the kinetics of aggregation and the time required for mature fibrils to form. However, we show that mature fibrils that were generated from different densities of proteins can reach the same fibril-monomer equilibrium state defined by ΔF. Moreover, ΔH, ΔS, and ΔCp can be obtained by studying the temperature dependence of this equilibrium. We discuss how the greater stability of fibrils with respect to temperature when compared to native protein structures implies a small ΔCp (close to zero). Despite its small value, we show that ΔCp can be measured in our model. We also study ΔCp for the honeycomb lattice model of amyloid fibrils which has recently been developed to study fibril solubility [41]. We find that changes in the thickness of the fibril in the honeycomb lattice accounts for an increase in the magnitude of ΔCp. This suggests that measurements of ΔCp for amyloid fibrils might provide insights into the residues that are buried in the fibril core.

2. Methodology

Our coarse grained model accounts for the two length scales that are characteristic of amyloid fibrils and which emerge from backbone hydrogen bonding and side chain interactions. Proteins are described by two-dimensional dimers, i.e., two beads separated by a fixed distance l. Beads mimic side chain and backbone atoms (see Fig. 1b) and they interact with beads from neighboring proteins through a Lennard-Jones potential defined by ϵΨΦ (binding energy) and σ (distance parameter):

\[ U_{lj}(r_{lj}) = 4\epsilon_{\Psi\Phi} \left( \frac{\sigma}{r_{lj}} \right)^{12} - 2\left( \frac{\sigma}{r_{lj}} \right)^{6} \]  

(1)

where \( r_{lj} \) is the distance between bead Φ of protein i and bead Ψ of protein j.

In addition to Lennard-Jones interactions, backbone beads can also form two hydrogen bonds (Hbonds). The directionality of Hbonds in fibrils is accounted for by two arms separated by an angle of 180° and which are perpendicular to the line joining backbone and side chain beads—see Fig. 1b. If \( \vec{r}_{ij} \) is the distance vector between centers of mass of backbone beads of proteins i and j, and \( \vec{r}_{\alpha} \) is the unit vector along arm \( \alpha \) of backbone bead i (see Fig. 1b), then the Hbond energy in the Ben-Naim’s framework is [42–44]:

\[ U_{HB}(\vec{r}_{ij}, (\vec{r}_{\alpha}), (\vec{r}_{\beta})) = \epsilon_{HB} \exp \left( -\frac{(r_{ij} - R_{HB})^2}{2\alpha_{HB}^2} \right) \times \left[ \sum_{i=1}^{2} \exp \left( \frac{-\left( r_{ij} \right)}{2\alpha_{HB}} \right) \times \left( \frac{1}{2\alpha_{HB}} \right) \right] \]

(2)

where \( \epsilon_{HB} = -1 \) and \( R_{HB} = 1 \) are the binding energy and equilibrium distance of the Hbond. The constant \( \epsilon_{HB} = 0.085\epsilon_{\Psi\Phi} \) is the attenuation parameter of the interaction. Eq. (2) favors configurations where the distance between backbone beads of molecules i and j is \( R_{HB} \), one arm of molecule i is aligned with the line joining the centers of mass of backbone beads, and the same for one arm of molecule j. Distances and energies in this work are given in units of \( R_{HB} \) and \( \epsilon_{HB} \), respectively, and the Boltzmann constant \( k_B \) is defined as one. In terms of these reduced units, the Lennard-Jones potential is defined by \( \sigma = 0.75 \) (the same distance parameter for all beads) and \( \epsilon_{\Psi\Phi} = 0.1, 0.2, \) or 0.141 if beads Φ and Ψ correspond both to side chain, backbone, or a mixture of side chain and backbone beads, respectively. This set of binding energies for the Lennard-Jones potential was chosen by trial-and-error to account for fibril formation. In particular, a same \( \epsilon_{\Psi\Phi} \) value for pairs of side chain and backbone beads resulted in circular aggregates instead of linearly elongated fibrils. The binding energy for the interaction between side chain and backbone (i.e., \( \epsilon_{\Psi\Phi} = 0.141 \)) beads was obtained from the standard Lorentz-Berthelot rule. The fixed distance between beads is \( l = 1 \).

Here, we perform Monte Carlo (MC) simulations in the canonical (NVT) ensemble [45]. Periodic boundary conditions is used to mimic an infinite system. Initial conditions are chosen randomly. At each time-step, we tried to translate, rotate and switch (180° rotation) in random order a randomly chosen protein. A MC cycle which corresponds to N time-steps is used as our unit of time. Average densities are computed from 10⁸ MC cycle simulations that are performed after an equilibration period of 10⁷ cycles. Cluster analysis is used to classify the overall bonding state of the system. It uses an energy criteria wherein proteins are considered bonded when their potential energy is less than −0.05. Small variations of this energy cutoff did not account for significant differences in the bonding state of the system. The different states of the system are: the monomeric state when more than 50% of proteins are non-bonded, the fibril state when at least 50% of the proteins are bonded to one big cluster, and the oligomeric state which comprises all other situations, i.e., when less than 50% of proteins are non-bonded and less than 50% are not part of a big cluster.

3. Results

In Fig. 1c we show the phase diagram of a system comprising \( N = 120 \) proteins. This diagram was obtained from simulations carried out at different temperatures (from 0.08 to 0.26 in steps of 0.01) and densities (from 0.01 to 0.25 in steps of 0.01). Characteristic configurations are shown in Fig. 1d–f (Supplementary material). The monomeric state occur for conditions of high temperature and
Fig. 1. (a) Structure of a fibril made from residues 16–21 of the amyloid-$\beta$ protein (PDB ID: 3OW9). Backbone and side chains of proteins are depicted using cartoon-like and bead representations, respectively. (b) Schematic representation of a protein in the model. Side chain and backbone beads are highlighted as well as H bond arms. The representation of a dimer depicts the vectors (blue arrow) used in Eq. (2) to account for H bonds. The representation of an amyloid fibril shows two $\beta$-sheets made of three peptides each stack on top of each other. The fibril axis is also shown. (c) Phase diagram showing monomeric (MR), oligomeric (OR) and fibrillar states as well as representative snapshots of the system at these states computed at the same density 0.1 and temperatures (d) 0.25, (e) 0.15, and (f) 0.1.

low density which are unfavorable to bond formation. Accordingly, proteins in Fig. 1d are mostly dispersed in the simulation box. Conversely, the fibrillar state is favored by conditions of low temperature and high density. Under these conditions, proteins are mostly bonded to each other and fibrils in Fig. 1f are characterized by $\beta$-sheet structures, i.e., extended regions of Hbonded proteins stacked on top of each others via side chain interactions. In the oligomeric state, which occurs for intermediate values of temperature and density in Fig. 1c, Hbonded proteins do not form ordered $\beta$-sheets which is required for fibril formations.

In Fig. 2, we study how systems evolve towards equilibrium. Starting from a random initial conditions, the fraction of monomeric proteins $f_o$ in the system is shown as a function of time for systems with low and high protein concentrations. Notice that in this figure equilibrium is reached in about 4000 MC cycles. For the low concentration system, the fraction of non-bonded proteins in equilibrium is greater than 50% and, therefore, the system is considered to be in the monomeric state–see Fig. 1c. Evolution towards equilibrium is a fast process which is fitted in Fig. 2 by a single exponential with a characteristic decay time of 30 MC cycles. The high concentration system reach equilibrium in the oligomeric state and a characteristic configuration is shown in Fig. 1e. The evolution of this system towards equilibrium proceeds in two steps: a very fast bonding process in which dimers and trimers are formed by local rearrangements of proteins, e.g., rotation, followed by a slower addition of proteins to $\beta$-sheets accounting for the extended but disordered Hbonded regions visible in Fig. 1e. The time dependence of this system is fitted by a double exponential with characteristic decay times of 1200 and 14 MC cycles.

Fig. 2 shows that systems evolve towards an equilibrium ondition in which bonded and monomeric proteins co-exist. Assuming that these two species interconvert, first-order rate equations predict a linear dependence of the equilibrium density of monomers $\rho_o^*$ on the overall density of proteins $\rho$.

$$\rho_o^* = \left(\frac{k_{bo}}{k_{bo} + k_{ob}}\right)\rho,$$  \hspace{1cm} (3)

where $k_{ob}$ and $k_{bo}$ are the rate constants for binding a monomer and dissolving a bonded protein, respectively. This equation is valid in equilibrium (i.e., $dN_o/dt = 0$) and the total number of proteins in the system is conserved, $N_o$ and $N_b$ are the number of monomers and bonded proteins.
the stability of mature fibrils can be quantified by measuring the capacity related to the dissociation of a protein from a fibril computed at temperature $T_0$. The temperature dependence of $\Delta F$ from our simulations is shown in Fig. 4 and the best fit of this dependence to Eq. (6) using $T_0 = 0.15$ gives $\Delta U_0 = 0.740 \pm 0.006$, $\Delta S_0 = 1.22 \pm 0.03$, and $\Delta C_0 = -4.29 \pm 0.14$. While our model does not have enough detail to account for the magnitude and sign of these quantities in real proteins, it is interesting to observe that the free-energy is dominated by $\Delta U_0$ close to $T_0$, i.e., $\Delta U_0 \gg T_0 \Delta S_0$. This behavior is expected since fibrils are highly stable, i.e., more stable than native protein conformations [50], and internal energy (or enthalpy) favors the fibrillar state while entropy opposes it [49].

In Eq. (6), $\Delta C_0$ which is related to the curvature of $\Delta F$ is expected to be small (close to zero) for amyloid fibrils [49] on the basis that the stability of these structures is not strongly dependent on temperature. This justifies the use of van’t Hoff equation:

$$\Delta F = \Delta U_0 - T \Delta S_0 + \Delta C_0 \left[ (T - T_0) - T_0 \log \left( \frac{T}{T_0} \right) \right]$$

(6)

where $\Delta U_0$, $\Delta S_0$, and $\Delta C_0$ are the change in energy, entropy, and heat capacity related to the dissociation of a protein from a fibril computed at temperature $T_0$. The temperature dependence of $\Delta F$ from our simulations is shown in Fig. 4 and the best fit of this dependence to Eq. (6) using $T_0 = 0.15$ gives $\Delta U_0 = 0.740 \pm 0.006$, $\Delta S_0 = 1.22 \pm 0.03$, and $\Delta C_0 = -4.29 \pm 0.14$. While our model does not have enough detail to account for the magnitude and sign of these quantities in real proteins, it is interesting to observe that the free-energy is dominated by $\Delta U_0$ close to $T_0$, i.e., $\Delta U_0 \gg T_0 \Delta S_0$. This behavior is expected since fibrils are highly stable, i.e., more stable than native protein conformations [50], and internal energy (or enthalpy) favors the fibrillar state while entropy opposes it [49].

In Eq. (6), $\Delta C_0$ which is related to the curvature of $\Delta F$ is expected to be small (close to zero) for amyloid fibrils [49] on the basis that the stability of these structures is not strongly dependent on temperature. This justifies the use of van’t Hoff equation:

$$\ln K_d = \frac{\Delta U_0}{k_b T} + \frac{\Delta S_0}{k_b}$$

(7)

to study amyloid fibrils as it can be derived from Eq. (6) in the limit $\Delta U_0 - T \Delta S_0 \gg \Delta C_0 \left[ (T - T_0) - T_0 \log \left( \frac{T}{T_0} \right) \right]$. Accordingly, Eq. (7) has recently been used to study the honeycomb-peptide lattice model of fibril formation [41]. In this model, peptides are represented by hexagons in a two-dimensional triangular lattice which can bind to each other along a straight line to form a $\beta$-sheet. Also, stacking of $\beta$-sheets account for fibrils with different thickness. Here, we show that despite the small value of $\Delta C_0$ for the honeycomb lattice model, this quantity can be computed for amyloid fibrils. Symbols in Fig. 5 are results from Fig. 6b of reference [41] where $i$ represents the thickness of the fibril, i.e., the number of $\beta$-sheet layers in the fibril. Results in reference [41] were obtained from Monte Carlo simulations. In particular, an initial fibril with fixed thickness and isolated peptides in solution (i.e., honeycombs) were placed in the simulation box. Allowed Monte Carlo moves were translation and rotation of peptides. Depending on the initial concentration of peptides, the
fibril elongated or shrunk until a residual concentration of peptides (i.e., $p_i$ in Fig. 6) in solution was reached. Lines in this figure are fits to Eq. (6) using $T_0 = 2$ kJ/mol and $p_i = 1$. Quantities obtained from these fits, i.e., $\Delta U_i$, $\Delta S_i$, and $\Delta C_i$, are shown in Table 1. $\Delta U_i$ and $\Delta S_i$ increase with the thickness of the fibril as more bonds need to break to dissociate peptides accounting for a greater dissociation entropy. $\Delta C_i$, which is positive for $i = 1$ (i.e., $\Delta C_i = 0.36$ kJ/mol/K) becomes increasingly more negative with increasing $i$ implying that “thicker” fibrils are more stable to both high and low temperatures. In the honeycomb model, an increase in the thickness of the fibrils (i.e., an increase in $i$) accounts for more hydrophobic interactions between side chains. Accordingly, this interaction was shown to be responsible for the dependence of $\Delta U_i$ on $i$ [41] and it can also be expected to account for changes in $\Delta C_i$ in the model.

4. Conclusion

In conclusion, we provide computational proof of principle that fibrils can exist in equilibrium with proteins dissolved in solution and this equilibrium can be used to compute free-energies $\Delta F$ to dissociate a protein [40]. We show that conditions of density accounting for different fibril growth rates can lead to the same $\Delta F$ and studies of the temperature dependence of $\Delta F$ can be used to extract thermodynamic quantities of the system. These quantities remain mostly unknown but they have the potential to provide important insights into molecular mechanisms underlying fibril stability. For example, compensation between $\Delta H_i$ and $\Delta S_i$ in protein folding provides rationalization for the marginal stability of the native state [51]. This compensation is being explored in strategies to design drugs that target protein conformations [52,53]. Thus, understanding enthalpy-entropy compensations in fibrils may be valuable to design more specific drugs for amyloid diseases. Also, $\Delta C_i$ for protein unfolding depends linearly on the number of non-polar residues in the protein core [15]. It provides, therefore, insights into the position of non-polar amino acids in the native state. Similarly, if $\Delta C_i$ for fibrils is found to be related to the position of residues, it may be used to complement other techniques to resolve the structure of fibrils. $\Delta C_i$ may also be used to predict the stability of fibrils as a function of temperature [54]. Therefore, measurements of thermodynamic quantities will be a valuable addition to our understanding of amyloid fibrils.

Acknowledgment

T. U. are grateful for the support of the NIH (GM063592) and Slovenian Research Agency (P1 0103-0201, N1-0042) and the National Research, Development and Innovation Office of Hungary (SNN 116198).

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bpc.2017.03.001.

References
