Available online at www.sciencedirect.com

ScienceDirect

Physics Procedia 53 (2014) 90 – 95

Physics Procedia

Thermodynamics of Protein Aggregation

Kenneth L. Osborne[1], Bogdan Barz[1], Michael Bachmann[2], and Birgit Strodel[1,3]*

[1] Institute of Complex Systems: Structural Biochemistry, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
[2] Center for Simulational Physics, The University of Georgia, Athens, Georgia 30602, USA
[3] Institute of Theoretical and Computational Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

Abstract

Amyloid protein aggregation characterizes many neurodegenerative disorders, including Alzheimer's, Parkinson's, and Creutzfeldt-Jakob disease. Evidence suggests that amyloid aggregates may share similar aggregation pathways, implying simulation of full-length amyloid proteins is not necessary for understanding amyloid formation. In this study we simulate GNNQQNY, the N-terminal prion-determining domain of the yeast protein Sup35 to investigate the thermodynamics of structural transitions during aggregation. We use a coarse-grained model with replica-exchange molecular dynamics to investigate the association of 3-, 6-, and 12-chain GNNQQNY systems and we determine the aggregation pathway by studying aggregation states of GNNQQNY. We find that the aggregation of the hydrophilic GNNQQNY sequence is mainly driven by H-bond formation, leading to the formation of β-sheets from the very beginning of the assembly process. Condensation (aggregation) and ordering take place simultaneously, which is underpinned by the occurrence of a single heat capacity peak only.

Peer-review under responsibility of The Organizing Committee of CSP 2013 conference

Keywords: amyloid aggregation; coarse-grained simulations; replica exchange molecular dynamics; statistical analysis; orientational order parameters; aggregation states

1. Introduction

Diseases such as Alzheimer's disease, Parkinson's disease, type II diabetes, and others share a general characteristic: each of them is associated with the misfolding and subsequent aggregation of soluble peptides and proteins into soluble oligomeric assemblies, and later, into insoluble amyloid fibrils [1–3]. During the past decade compelling evidence has emerged that soluble, low molecular weight oligomers are more toxic than fully formed fibrils [4–7]. The study of protein fragments that retain the essential amyloid characteristics of the full length sequences is attractive because the short peptide length allows for a systematic computational investigation of aggregation kinetics and thermodynamics. One such example fragment is GNNQQNY, a polar heptapeptide from the N-terminal region of the yeast prion protein Sup35 that exhibits amyloidogenic properties similar to the prion-determining domain of Sup35 [8].

The knowledge of the atomic structure of GNNQQNY fibrils [8–10], together with its small size, has led GNNQQNY to become a model amyloid test system for experimental and theoretical studies. One of the first computational studies of GNNQQNY examined the behavior of the trimer using molecular dynamics (MD) simulations at a constant temperature of 330 K and revealed that the in-register parallel β-sheet observed in experiments is stabilized by side-chain hydrogen bonding and π-stacking of the aromatic tyrosine residues [12]. This study was followed by various other simulation studies, which characterized the structures and free energies of small GNNQQNY aggregates ranging from dimers to 20-mers starting from disordered states, or studied the stability of pre-formed GNNQQNY assemblies with cross-β or annular morphologies [13–22].

Given the large time and length scales involved in aggregation, coarse-grained (CG) models provide the possibility of extracting general characteristics of the thermodynamics and kinetics of aggregation. CG models utilized for studying amyloid aggregation include the model by Caflisch and coworkers, where each peptide consists of four spherical backbone beads and six spherical side-chain beads of hydrophobic and hydrophilic nature [24–26], the mid-resolution Shea model with two backbone and one side chain bead per residue [27, 28], discontinuous CG models used in connection with discrete molecular dynamics

*Email address: b.strodel@fz-juelich.de (Birgit Strodel)

Peer-review under responsibility of The Organizing Committee of CSP 2013 conference
doi:10.1016/j.phpro.2014.06.032

[29–33], and the OPEP model by Derreumaux and coworkers that uses a detailed representation of all backbone atoms and reduces each side-chain to a single bead [34–37].

In this paper, we perform REMD simulations using the CG peptide model developed by Bereau and Deserno [38] to investigate the thermodynamics of structural transitions during the aggregation of the 3-, 6-, and 12-chain GNNQQNY systems.[22] Their intermediate resolution level CG force field was shown to be able to fold proteins exhibiting both helical and β conformations with tertiary structures and amino acid sequences different from the one used for parameter tuning [38–40]. The aim of the current study is analyze the structures and thermodynamics during aggregation of the GNNQQNY peptide. The resulting aggregation dynamics in temperature space is characterized in terms of orientational order, secondary structure and aggregation states.

2. Model and Methods

The primary motivation for using a CG model was to reach the greater than microsecond time scales associated with peptide aggregation. The C_β-type CG force field we use here was built to sample a balanced proportion of α-helical and β-extended configurations, with the aim of avoiding a bias toward any particular secondary structure [38]. The backbone is represented by three beads per residue and one bead per side chain with the latter located in the position of the C_β atom. The force field was parametrized to reproduce local, secondary, and tertiary conformations. The nearly atomistic resolution of the backbone beads allows the force field to model physically relevant secondary structures, such as β-sheets and α-helices, without imposing a given secondary structure prior to simulation [38].

To sample the conformational space, we employed the REMD method [23] as implemented in the ESPRESSO simulation package [41]. Each REMD simulation was started from 3, 6, or 12 chains of GNNQQNY peptides placed at random starting positions and in random conformations in a simulation box at a concentration of 80 mM. Each REMD simulation was performed for 500 million time steps per replica, where the initial 100 million steps were used for system equilibration. This results in about 500 ns simulation time per replica. Studies of the 3-chain system used 10 temperature replicas, while the studies of the 6- and 12-chain systems used 16 replicas. Within each temperature thread of our REMD simulations resides a canonical (NVT) simulation. The weighted histogram analysis method (WHAM) combines the data from all replicas to extract the multicanonical ensemble averages [42]. It was used to calculate thermodynamic quantities, such as the heat capacity (C_v), and the average behavior of impact parameters ($\langle P \rangle$) defined as a function of T.

For the characterization of the aggregation behavior, we computed the liquid crystal order parameters P_1 and P_2 to determine the (anti)parallel structural order of the system [13]. This is useful because amyloidogenic sequences are expected to form β-sheets, and thus be aligned in characteristic forms. For the classification of the GNNQQNY oligomers in terms of their

State Name	Oligomeric Species	Population
0	Monomer	3
1	Monomer	1
	Dimer	1
2	Trimer	1

Table I: Definition of the unique aggregation states of a system containing three GNNQQNY peptides.

aggregation state, we analyzed the aggregate size for each structure sampled in our REMD simulations. Here, a single chain is said to belong to a given oligomer cluster if it shares one or more hydrogen bonds with another chain in that oligomer. We determined all possible combinations of oligomer sizes for the 3- and 6-chain system. As an example we enumerated aggregation states for 3-chain system in Table I. Lower numbered aggregation states correspond to a less aggregated system, with 0 being the completely monomeric state. Conversely, the higher numbered aggregation states correspond to aggregated systems that include larger fragments. A similar casification was done for the 6-chain system while the aggregation behavior for the 12-chain system was characterized by tracking only the largest β-sheet in the system [22].

3. Results

We simulated 3-, 6-, and 12-chain GNNQQNY systems. We will first present the results of the 3- and 6-chain systems, as we have performed 3 and 12 runs for these systems, respectively, and can thus provide an in depth statistical analysis of the results. Conversely, because the computational cost of simulating the 12-chain system was high, we only performed one run for this system. WHAM [42] was used to calculate various quantities important for understanding the behavior of the system as a function of temperature T. In our study, we examined the temperature dependence of the heat capacity and the order parameters $\langle P_1 \rangle$ and $\langle P_2 \rangle$ as shown in Figure 1 for 3- and 6-chain systems. Fluctuations in energy (A and C) appear maximal at the transition point T_{trans}. The peaks in C_v for the 3-chain system in (A) occurs at $T_{\text{trans}} \approx 0.86$ RTU and for the 6-chain system in (C) at $T_{trans} \approx 0.89$ RTU. From (B) and (D) we notice that $\langle P_1 \rangle$ and $\langle P_2 \rangle$ decrease over the entire temperature range with the sharpest decrease at the transition point. This link between change in energy and change in system order suggests aggregation into ordered structures at T_{trans}.

In order to understand how aggregation proceeds we discuss characteristic features of representative structures in the different phases (i.e. at the lowest temperature T_{low}, T_{trans}, and at the highest temperature T_{high}). Figure 2 displays typical structures of