IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Aggregation of peptides in the tube model with correlated sidechain orientations

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2015 J. Phys.: Conf. Ser. 627 012028

(http://iopscience.iop.org/1742-6596/627/1/012028)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 14.177.143.108
This content was downloaded on 16/07/2015 at 05:30

Please note that terms and conditions apply.

Aggregation of peptides in the tube model with correlated sidechain orientations

Nguyen Ba Hung[1,2] **and Trinh Xuan Hoang**[1]

[1] Institute of Physics, Vietnam Academy of Science and Technology, 10 Dao Tan, Ba Dinh, Ha Noi, Viet Nam
[2] Vietnam Military Medical University, 160 Phung Hung, Ha Dong, Ha Noi, Viet Nam

E-mail: `hungnb80@yahoo.com.vn`

Abstract. The ability of proteins and peptides to aggregate and form toxic amyloid fibrils is associated with a range of diseases including BSE (or mad cow), Alzheimer's and Parkinson's Diseases. In this study, we investigate the the role of amino acid sequence in the aggregation propensity by using a modified tube model with a new procedure for hydrophobic interaction. In this model, the amino acid sidechains are not considered explicitly, but their orientations are taken into account in the formation of hydrophobic contact. Extensive Monte Carlo simulations for systems of short peptides are carried out with the use of parallel tempering technique. Our results show that the propensity to form and the structures of the aggregates strongly depend on the amino acid sequence and the number of peptides. Some sequences may not aggregate at all at a presumable physiological temperature while other can easily form fibril-like β-sheet struture. Our study provides an insight into the principles of how the formation of amyloid can be governed by amino acid sequence.

1. Introduction

Amyloid fibrils are stable aggregates of proteins implicated in a range of diseases including Alzheimer's, Parkinson's and spongiform encephalopathies [1]. These fibrils have been found for a variety of proteins and synthetic peptides including those that do not involve in any disease [2, 3]. From these experimental evidences and simulations [4, 5, 6, 8, 9] it is generally accepted that that the tendency to form amyloid fibrils is a generic property of polypeptide chain. Yet, it is also believed that the amino acid sequence is a dominant factor in determining the propensity to form fibrils (see e.g. Ref. [10]) eventhough the role of sequence is much less known. Studies of the tube model of protein have shown that even homopolymeric peptides can aggregate and form fibril-like structures [5, 8, 9]. A general framework provided by tube model [4, 5, 7] also suggests that amyloid-like aggregates belong to a few minima in a presculpted free energy landscape of protein that is governed by symmetry of a tube-like polymer and geometrical constraints of hydrogen bonds. The role of amino acid sequence is to select or to avoid these minima.

In this study we investigate the role of sequence in the formation of amyloid by taking the advantages of the tube model in capturing the key features of polypeptide chain and of being strongly coarse-grained so that a full equilibrium characteristics of a pool of peptides could be studied. Furthermore, in the present study we incorporate sidechain orientations to the tube model in order to define better hydrophobic contacts. We will consider peptides of

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

different hydrophobic-polar sequences but with the same hydrophobic fraction for making useful comparison.

2. Models and methods

We consider the tube model of protein [4, 11], in which each amino acid is considered as a single residue whose position is specified by the center of its C_α atom. The distance between two consecutive residues along the chain is 3.8 $\AA$. The residues are placed along the axis of a self-avoiding tube of cross-sectional radius $\Delta = 2.5\AA$. The finite thickness of the tube was imposed by requiring the radius of circle drawn through any three C_α atoms must be larger than Δ [12, 11]. Also, sterics requires that two non-consecutive C_α's cannot be closer than 4 $\AA$ from each other. The bond angles are restricted to the range of $82^o - 148^o$, and a bending stiffness is introduced by applying a local bending penalty energy of $e_R = 0.3\epsilon > 0$ if the local radius of curvature at a given bead is less than 3.2 $\AA$, where ϵ is an energy unit equal to the energy of a local hydrogen bond. Hydrogen bonds between amino acids are required to satisfy a set of distance and angular constraints [4] on the C_α's as found by a statistical analysis of PDB's native protein structures [5]. A local hydrogen bond is formed between residues that are separated by three peptide bonds along the chain, and is given an energy $-\epsilon$. A non-local hydrogen bond is given an energy of -0.7ϵ. To avoid spurious effects of the chain termini, hydrogen bonds (local and non-local) involving a terminal residue are given a reduced energy of -0.5ϵ. Additionally, a cooperative energy of -0.3ϵ is given for each pair of hydrogen bonds that are formed by pairs of consecutive amino acids in the sequence.

Hydrophobic interactions are based on pairwise contacts between amino acids. We assume that there are only two kinds of amino acids: hydrophobic (H) and polar (P). A contact between two non-consecutive beads is formed if they are found within a distance of 7.5 $\AA$. Only contacts between hydrophobic residues are favorable and are assigned an energy of $e_{HH} = -0.5\epsilon$ per contact. Contacts involving polar residues are given zero energy.

In the present study, we consider a modified tube model with a new procedure for hydrophobic interaction. The sidechains are not considered explicitly, but their orientations as given by anti-normal vectors are taken into account in hydrophobic contacts (Fig. 1). Apart from the distance constraint of 7.5 $\AA$ as described above, our new procedure requires that two residues i and j make a hydrophobic contact if $\mathbf{n}_i \cdot \mathbf{c}_{ij} < 0.5$ and $-\mathbf{n}_i \cdot \mathbf{c}_{ij} < 0.5$ where $\mathbf{n}_i$ is the normal vector associated with bead i, and $\mathbf{c}_{ij}$ is an unit vector pointing from bead i to bead j. The normal and connecting vectors are given by:

$$\mathbf{n}_i = \frac{\mathbf{r}_{i-1} + \mathbf{r}_{i+1} - 2\mathbf{r}_i}{|\mathbf{r}_{i-1} + \mathbf{r}_{i+1} - 2\mathbf{r}_i|} , \tag{1}$$

and

$$\mathbf{c}_{ij} = \frac{\mathbf{r}_j - \mathbf{r}_i}{|\mathbf{r}_j - \mathbf{r}_i|} , \tag{2}$$

respectively, where $\mathbf{r}_i$ is the position vector of bead i.

Monte Carlo (MC) simulations are carried out with periodic boundary condition. A parallel tempering [13] Monte Carlo scheme is employed for obtaining the ground state as well as other equilibrium characteristics of the system. For each system, 16 to 24 replicas are considered, each evolving at its own selected temperature T_i. For each replica, the simulation is carried out with standard pivot and crankshaft move sets and the Metropolis algorithm for move acceptance. An attempt to exchange replicas is made every 100 MC steps. The exchange of replicas i and j is accepted with a probability equal to $p_{ij} = \min\{1, \exp[k_B^{-1}(T_i^{-1} - T_j^{-1})(E_i - E_j)]\}$, where k_B is the Boltzmann constant, and E_i and E_j are the energies of the replicas at the time of the exchange.

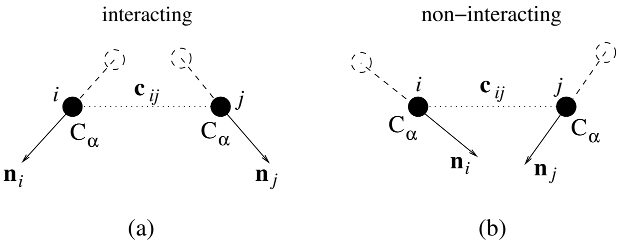

Figure 1. Contact interaction with sidechain orientation. The sidechain is approximately placed in the anti-normal direction from the C_α atom. Two residues are in contact if their sidechains are oriented not too apart from the other residue.

Table 1. HP sequences of amino acids of peptides considered in present study (H – hydrophobic, P – polar). The parameter s denotes sequence separation between two H amino acids.

Sequence name	Sequence	s
S1	P P P H H P P P	1
S2	P P H P H P P P	2
S3	P P H P P H P P	3
S4	P H P P P H P P	4
S5	P H P P P P H P	5
S6	H P P P P P H P	6
S7	H P P P P P P H	7

The temperature range in parallel tempering simulations are chosen such that it covers the transition from a dilute gas of separated peptides at high temperature to an aggregation phase at low temperature. The parallel tempering temperatures T_i are chosen such that acceptance rate of replica exchanges is significant, especially near the peak of the specific heat where energy fluctuation is large. Typically we need to change the set of temperatures several times in such a way that there are more temperatures for the replicas near the specific heat's peak. In each simulation, the number of Monte Carlo attempted moves for each temperature is of the order of 10^9.

The weighted multiple-histogram technique [14] is used to compute the specific heat of the system, which is given by:

$$C = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2} \ , \tag{3}$$

where E is the energy of the system and $\langle \cdot \rangle$ denotes thermodynamic average.

3. Results and discussion

We consider systems of $M = 10$ identical peptides of length $N = 8$ residues. The amino acid sequences of the peptides denoted as S1–S7 are given in Table 1. We have selected the sequences in a way that they have the same fraction of 25% of hydrophobic amino acids, i.e. each sequence has two H residues. The sequences differ in the separation between the two H residues which is given by the parameter s (see Table 1).

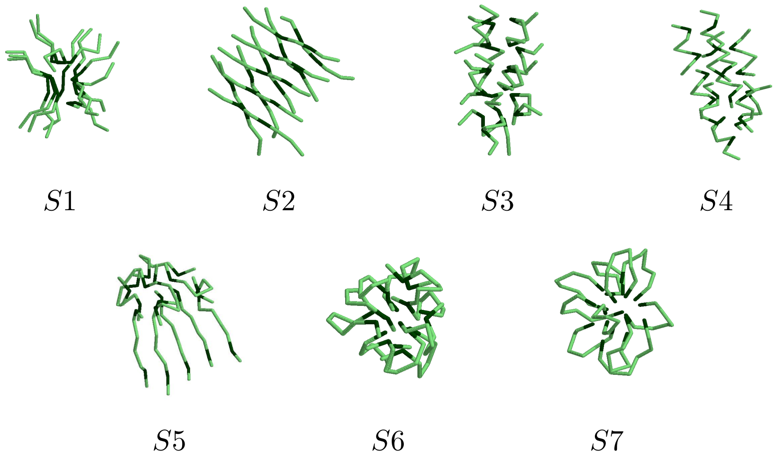

Figure 2. RASMOL visualization of aggregate structures for systems of 10 chains of the sequences (S1-S7) given in Table 1. H and P amino acids are shown in green and blue, respectively. The structures shown are lowest energy conformations obtained in parallel tempering Monte Carlo simulations.

The systems are simulated in a cubic box of box size $L = 255.15\mathring{A}$ with periodic boundary conditions. The box size was chosen such that the peptide concentration is equal to 1 mM. Figure 2 shows the lowest energy conformations obtained from parallel tempering simulations which are supposed to be the ground states of the systems. It is shown that the forms of the aggregates are remarkably different for different sequences. Only the sequence S2 system forms a double layer β-sheet structure which is similar to cross-β structures found in amyloid. α-helix bundles are seen for the S3 and S4 systems. Other sequences have disordered aggregates with the β-sheet appeared more clearly for sequence S5. Our results shows that the propensity to form amyloid-like structure strongly depends on the sequence.

Remarkable differences between sequences are also seen in the temperature dependence of the specific heats for the systems considered (Figure 3). Each sequence has its own specific heat profile with the temperature of the main peak varies between 0.17 and 0.26 ϵ/k_B and the heights of the peaks differ up to about one order of magnitude (Figure 4). The main peak in the specific heat signifies an aggregation transition and its height is a measure of the cooperativity of the transition. A sharp peak means that the transition is first-order like and the aggregate is an ordered structure. Note that the highest peak in the specific heat belongs to sequence S2 which forms the double layer β-sheet structure. This result suggests that the propensity to form amyloid may be linked to the cooperativity of the aggregation transition. The least cooperative aggregation transition belongs to sequence S1 which has two main peaks in the specific heat. For this sequence aggregation happens at the temperature of the higher peak but this transition just corresponds to a strong hydrophobic collapse and the aggregated structure remains disordered if the temperature is higher than the temperature of the lower peak.

The differences in the transition temperatures T_{max} (corresponding to the maximum of the specific heat) among sequences suggest a very interesting consequence. Suppose that we are interested in the behavior of systems at the physiological temperature, T^*. A rough estimate of T^* in our model is about 0.2 ϵ/k_B (which corresponds to an energy of a local hydrogen bond of 5 k_BT with T is room temperature). With this estimate from Figure 4B one can see that out of 7, three systems corresponding to sequences S3, S4 and S5 have the transition temperature T_{max} lower than the physiological temperature T^*. Thus at $T = T^*$ these three systems do not

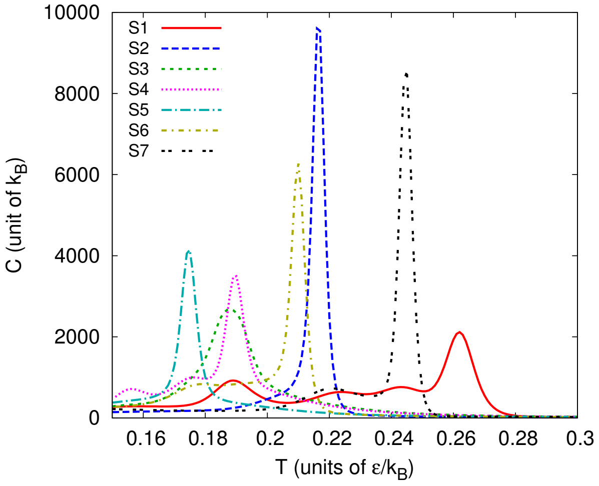

Figure 3. Temperature dependence of the specific heat C for systems of 10 peptides of HP sequences as given in Table 1 (S1–S7).

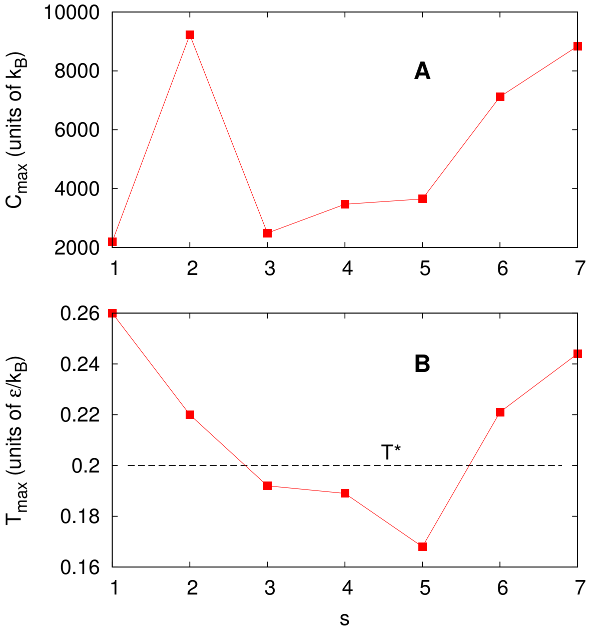

Figure 4. Dependence of the maximum of the specific heat C_{max} (A) and its temperature T_{max} (B) on the sequence separation s between two H residues in the HP sequence for the 7 systems considered. The horizontal dashed line in (B) indicate a presumed physiological temperature T^*.

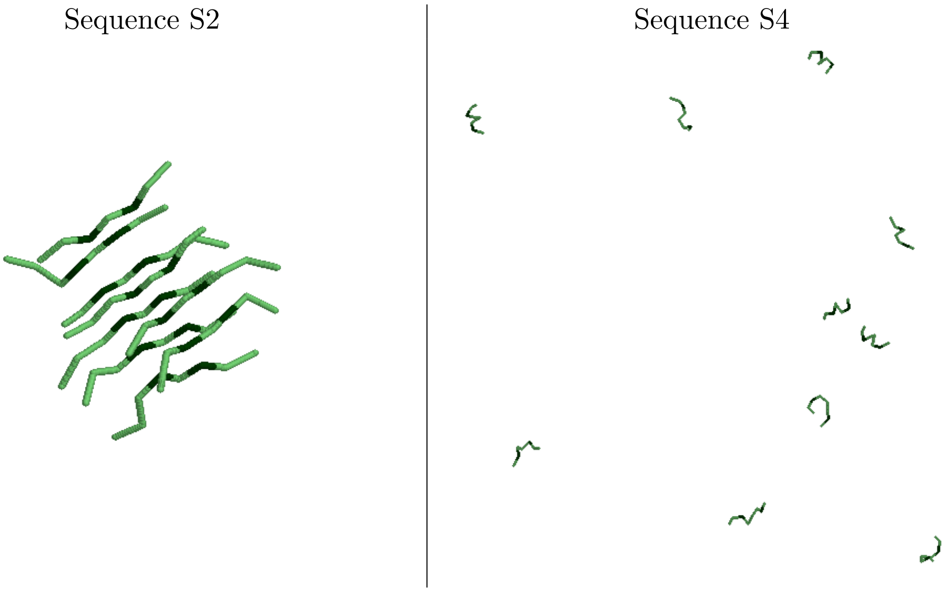

Figure 5. Snapshots of conformations at $T = 0.2\epsilon/k_B$ for peptides of sequence S2 (left) and sequence S4 (right). Note that at the given temperature the S2 peptides aggregate and form an amyloid-like structure whereas the S4 ones do not aggregate.

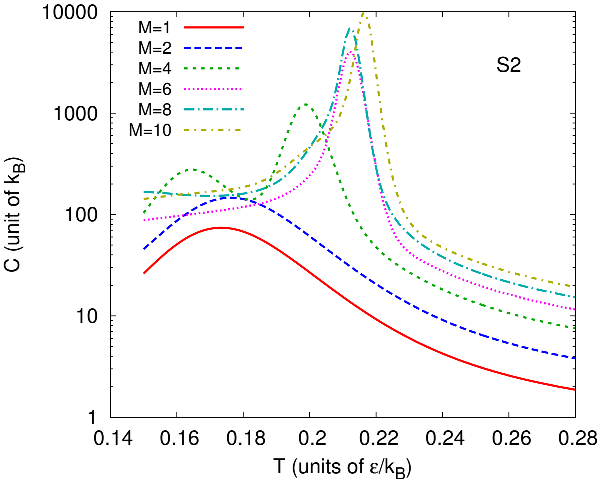

Figure 6. Temperature dependence of the specific heat for sequence S2 with various numbers of chains, M, as indicated.

aggregate while the other systems including that of sequence S2 do. This situation is illustrated in Figure 5 which shows snapshots obtained at $T = 0.2\epsilon/k_B$ for S2 and S4 systems. One can see that the S2 system has already formed a amyloid-like two-layer β-sheet structure while the S4 system has not aggregated. About 50% of chains in the S4 system form stand-alone α-helices.

Figure 6 shows the temperature dependence of the specific heat for increasing number of chains for sequence S2. It is shown that the height and the position of the peak increases as M increases which suggests that the propensity of peptides to aggregate grows with the density of peptides. For single chain the ground state of this sequence is an α-helix. The amyloid-like β-sheet structure starts to form when $M > 4$. This result agrees with the nucleation and growth mechanism of amyloid formation [3, 8, 9]. For the case of $M = 4$, the specific heat has two peaks that correspond to a collapse transition at the higher temperature and a disorder-oder transition at the lower temperature.

In a recent study [15] Dobson and coworkers considered a lattice model with side chain directions for amino acids to study amyloid formation of short peptides. Their designed sequence that was shown to form fibrillar structure is remarkably similar to sequence S2 in our model. Our present study confirms the important role of sequence in amyloid formation in a more realistic model. Two features of the sequence that are important for amyloid formation are the parallel pairing of hydrophobic sidechains within a cross β-sheet and the anti-parallel facing of hydrophobic sidechains between β-sheets in a fibirilar structure.

4. Conclusion

We have studied aggregation of peptides by performing Monte Carlo simulations using a modified tube model with correlated sidechain orientations. We have shown that the structure of the aggregates strongly depends on the HP sequence of peptides, and so does the aggregation transition as indicated by the specific heat. In particular, for a sequence with two H amino acid separated by one P, the aggregate is a two-layer β-sheet structure similar to that found in amyloid fibrils. The aggregation transition for this sequence is also the most cooperative and first-oder like among all sequences. Interestingly, our simulations indicate that some sequences have a very low propensity to form amyloid and they may not aggregate at all at a presumable physiological temperature. We conclude that the hydrophobic-polar patterns in the amino acid sequence play an important role in protein aggregation and they could be a dominant factor in determining the propensity of a polypeptide chain to form amyloid fibrils.

Acknowledgments

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number “103.01-2013.16”.

References

[1] Dobson C M 1999 *TIBS* **24** 329
[2] Chiti F, Dobson C M 2006 *Annu Rev Biochem.* **75** 333
[3] Hellstrand E, Boland B, Walsh D M and Linse S 2010 *ACS Chem. Neurosci.* **1** 13
[4] Hoang T X, Trovato A, Seno F, Banavar J R and Maritan A 2004 *Proc. Natl. Acad. Sci. USA.* **101** 7960
[5] Banavar J R, Hoang T X, Maritan A, Seno F and Trovato A 2004 *Phys. Rev.* E **70** 041905
[6] Hoang T X, Marsella L, Trovato A, Seno F, Banavar J R and Maritan A 2006 *Proc. Natl. Acad. Sci. USA.* **103** 6883
[7] Banavar J R and Maritan A 2007 *Annu. Rev. Biophys. Biomol. Struct.* **36** 261
[8] Auer S, Dobson C M and Vendruscolo M 2007 *HFSP Journal.* **1** 137
[9] Auer S, Dobson C M, Vendruscolo M and Maritan A 2008 *Phys. Rev. Lett.* **101** 258101
[10] Li M S, Co N T, Reddy G, Hu C K, Straub J E and Thirumalai D 2010 *Phys. Rev. Lett.* **105** 218101
[11] Maritan A, Micheletti C, Trovato A and Banavar J R 2000 *Nature.* **406** 287
[12] Gonzalez O, Maddocks J H 1999 *Proc. Natl. Acad. Sci. USA.* **96** 4769
[13] Swendsen R H, Wang J S 1986 *Phys. Rev. Lett.* **57** 2607
[14] Ferrenberg A M, Swendsen R H 1989 *Phys. Rev. Lett.* **63** 1195
[15] Abeln S, Vendruscolo M, Dobson C M and Frenkel D 2014 *PLOS One* **9** e85185