Simulating Alzheimer's on the XD1

Jan H. Meinke[1,*] and Ulrich H.E. Hansmann[1,2,†]

[1]*John-von-Neumann Institute for Computing, Forschungszentrum Jülich, D-52425 Jülich, Germany*
[2]*Department of Physics, Michigan Technological University, Houghton, MI 49931, USA*

ABSTRACT: We study the folding and aggregation of 6 chains of the β-amyloid fragment 16-22 using parallel tempering, which scales nearly linearly up to 72 processors on a Cray XD1 . While the isolated fragment prefers a helical form at room temperature, in the system of 6 interacting fragments one observes both parallel and anti-parallel β-sheets below a cross over temperature $T_x \approx 420$K. The anti-parallel sheets have lower energy and are therefore more stable. Above the nucleation temperature the aggregate quickly dissolves into widely separated, weakly interacting chains.

1. INTRODUCTION

Isolated proteins fold into unique 3D structures that are determined by their amino-acid sequence. In the cell, however, proteins are never isolated. Already during folding proteins interact with the ribosomes and with each other. Chaperones, e.g., help the protein fold correctly. But the crowded environment can also lead to unwanted consequences. At least for some sequences it depends on the local environment whether they form α-helices or β-strands. In some cases, the presence of a β-strand can catalyze formation of a sheet even in sequences that otherwise form helices [16, 17]. The newly-folded sheet is often at the exterior making such misfolded structures prone to aggregate and, as the catalytic process repeats, successively form fibrils. Such fibrils seem to be involved as a general mechanism in a number of diseases such as Alzheimer's, Huntington's or spongiform encephalopathies (prion-mediated) [20]. The most common of these diseases is Alzheimer's. Associated with its neuropathology are amyloid deposits composed mainly of the β-amyloid peptide (Aβ). It is found in body fluids in a soluble form that has partial α-helical structure. In Alzheimer's disease, Aβ undergoes a conformational change toward a β-sheet structure in which it is insoluble and assembles in fibrils 60-90Å in diameter. Fibrillar amyloids form lesions 10-200μm in diameter known as senile plaques. These plaques are surrounded by degenerating and swollen nerve terminals, and found in extracellular space of the brain. The neurotoxity of the Aβ-peptide is related to the degree of β-aggregation. Hence, an understanding of the aggregation mechanism could contribute to a developing understanding of the biogenesis of this devastating neurological disorders [20] and may lead to more targeted treatments.

Simulations provide an important tool to research the mechanism of aggregation in the amyloid peptide and other proteins. Even using all-atom models such numerical investigations are possible because experimental evidence indicates that already several different fragments of Alzheimer's β-amyloid can form fibrils. For instance, some authors find evidence for the formation of parallel sheets in Aβ_{10-35} [2] and Aβ_{1-40} [19]. Others have reported anti-parallel strands, for example, in Aβ_{34-42} [11], Aβ_{11-25} [18], and Aβ_{16-22} [1, 7]. Recently Lührs *et al.* looked at the aggregation of the whole Aβ_{1-42} protein. They determined through NMR studies [12] that the residues 18–26 form a β-sheet that is in registry with a β-sheet formed by residues 31-42 of a second β-amyloid molecule. Almost all of these studies contain the fragment 16-22, which has been identified as a key element already in 1996 by Tjernberg *et al.* [21]. Hence, a number of different computational studies have focused on this fragment. Using MD simulations, Klimov *et al.* [10] find that the formation of the aggregate requires an intermediate helical phase. They emphasize the distinct roles of the hydrophobic interactions, which provide the driving force for the initial collapse, and the electrostatic interaction, which result in the formation and stabilization of anti-parallel β sheets. In fact they conclude that "the ordered (anti-parallel) orientation is only obtained upon the formation of salt bridges." However, Favrin *et al.* [5] find anti-parallel β-sheets as the free-energy minimum of their MC simulation as well despite the fact that they don't have charged end groups in their model.

In order to resolve this discrepancy and to test whether the results of Ref. 5 are artifacts of the specific protein model or universal we have reproduced these simulations using the ECEPP/3 force field with an additional implicit solvent term and different end groups. We use parallel tempering with Monte-Carlo updates on the level of the individual replicas. The parallel tempering approach scales almost linearly up to 72 processors on a Cray XD1. Our simulation results show that the isolated fragment Aβ_{16-22} has rather a propensity to form α-helices. However, in a crowded environment, it becomes energetically favorable to form strands, and both parallel and anti-parallel β-sheets are observed by us, with the anti-parallel form being more stable.

2. A PROTEIN PRIMER

Proteins are the workhorses of the cell. They catalyze biochemical reactions, transport oxygen and nutrients, and fight diseases. Naturally occurring proteins are long chain molecules synthesized out of 20 different amino acids by forming peptide bonds. Their sequence — or primary structure — is encoded in the genome. The sequence of a protein is commonly written down using 1- or 3-letter abbreviations for the amino acids. The amino-acid sequence determines the 3D

*Electronic address: j.meinke@fz-juelich.de
†Electronic address: u.hansmann@fz-juelich.de

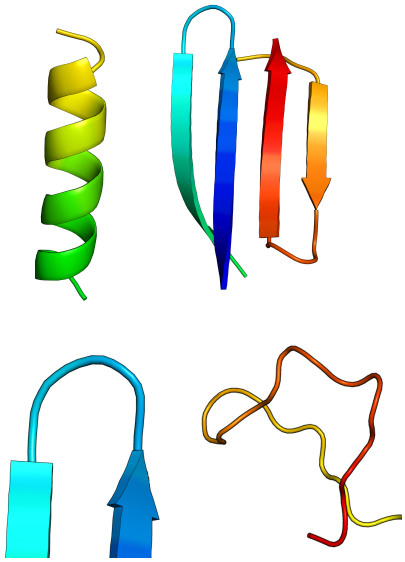


Figure 1: Secondary structure elements in proteins. Commonly observed structure elements are α-helices (top left), β-sheets (top right), and turns (bottom left). An unstructured segment is called a random coil (bottom right).

structure of a protein and the 3D structure in turn determines the functionality of the protein. How a protein folds from a linear chain to a compact, well-defined structure is know as the protein-folding problem.

Some motifs are frequently observed in proteins (see Fig. 1). Sections of proteins can form helices, which are stabilized by internal hydrogen bonding, others form nearly straight β-strands. Several β-strands can combine to form β-sheets, stabilized by hydrogen bonding between different strands. Two strands of a β-sheet are often connected by a turn. Helices, β-strands and sheets, and turns make up the secondary structure of a protein. The segments that don't fall into any of the other categories are called random coils. Often proteins are displayed using secondary structure elements to hide atomic details.

The arrangement of the secondary structure elements within a domain is called the tertiary structure (see Fig. 2 for an example). It determines the functionality of the protein. Most random amino-acid sequences do not fold into a defined 3D structure. The proteins used in biology have been selected to quickly fold into a stable tertiary structure. They cannot do this by random sampling alone. Even small proteins would need the life time of the universe to sample their entire configurational space. But proteins fold within fractions of a second. This is the Levinthal paradox. The Levinthal paradox can be resolved if we assume that as the protein gets to lower and lower energies, the number of possible configurations reduces until it is left with a single native state at the free energy minimum.

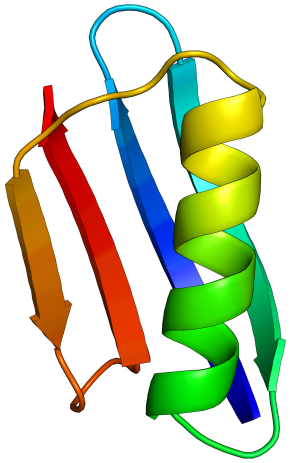


Figure 2: Tertiary structure of Protein G. This protein consists of an α-helix lying in front of 4 β-strands linked by hydrogen bonds to a β-sheet.

3. METHODS

In order to study aggregation of the 7-residue fragment β-amyloid$_{16-22}$ we performed a parallel tempering simulation of a system of 6 interacting molecules in a 50 Å box with hard walls. In contrast to Ref. 5 we used as N-terminal end group NH_2 (instead of a N-acetyl group) and as C-terminal end group $COOH$ (instead of NH_2). The intramolecular interactions are described by the ECEPP/3 force field [14] and given as the sum of electrostatic energy E_C, Lennard-Jones term E_{LJ}, hydrogen-bond energy E_{hb} and torsion energy E_{tor}:

$$\begin{aligned} E_{\text{ECEPP/3}} &= E_C + E_{LJ} + E_{hb} + E_{tor} \\ &= \sum_{(i,j)} \frac{332 q_i q_j}{\epsilon r_{ij}} \\ &\quad + \sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \\ &\quad + \sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \\ &\quad + \sum_{l} U_l (1 \pm \cos(n_l \xi_l)) \ , \end{aligned} \tag{1}$$

where r_{ij} is the distance between the atoms i and j, ξ_l is the l-th torsion angle, and energies are measured in kcal/mol. The intermolecular interaction includes hydrogen bonding E_{hb}, van-der-Waals interactions E_{LJ}, and coulomb interactions E_C. These energies are defined in the same way as for intramolecular terms. The protein-solvent interactions are approximated by a solvent accessible surface term

$$E_{solv} = \sum_i \sigma_i A_i \ . \tag{2}$$

Here A_i is the solvent accessible surface area of the ith atom in a given configuration, and σ_i a solvation parameter for the

T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}
250	267	284	300	325	345	266	380	393	400	408	411
T_{13}	T_{14}	T_{15}	T_{16}	T_{17}	T_{18}	T_{19}	T_{20}	T_{21}	T_{22}	T_{23}	T_{24}
415	418	421	423	425	430	435	437	440	442	446	450
T_{25}	T_{26}	T_{27}	T_{28}	T_{29}	T_{30}	T_{31}	T_{32}				
456	460	466	475	500	550	600	700				

Table I: Temperatures used for parallel tempering. The calculation for each replica was performed on a separate processor.

atom i. Our simulation relies on the implementation of these interactions in version 3 of the program package SMMP [3, 4], the first version to support multi-molecule simulations and described in detail in [13].

The above atomistic model of our system of 6 interacting peptides contains both attractive and repulsive interactions. As a consequence, the resulting energy landscape is characterized by a multitude of local minima separated by high energy barriers. Hence, sampling of low-energy conformations is a hard computational task. Physical quantities cannot be calculated accurately from simple low-temperature molecular dynamics or Monte Carlo simulations. One way to overcome this so-called multiple-minima problem is parallel tempering [6, 9] first used in protein science in Ref. 8. In this method, several replicas of the system are simulated in parallel at different temperatures. Configurations are exchanged between replicas every n_{ex} steps with probability

$$P(i \leftrightarrow j) = \min(1, \exp((\beta_i - \beta_j)(E_i - E_j))) \; . \quad (3)$$

For the present simulation we used 32 replicas with temperatures from $T_{\min} = 250$ to $T_{\max} = 700$ as listed in Table I. An exchange was attempted every 10 Monte Carlo sweeps. The simulation ran over 100,000 sweeps, where each sweep consists not only of a series of Metropolis updates of the internal angles in all chains but also of the rigid body coordinates of each molecule. The maximum step size along each axis is 0.5Å and the angles are changed randomly between $-\pi$ and π. The starting configurations were obtained from an initial equilibration run that also served to adjust the temperatures of the replicas. For comparison, we also simulate with the same protocol and energy function a single 7-residue fragment β-amyloid$_{16-22}$.

3.1. Parallel scaling of parallel tempering

We run each replica on a different processor. The exchange of configurations in parallel tempering is most probable between replicas that are adjacent in temperature. The exchange can be done in two ways. We can either exchange temperatures between replicas or configurations. The current implementation exchanges temperatures to minimize communication. Only a single double precision number needs to be exchanged. But it makes tracking of observables more difficult. Observables such as the energy or the compactness of a configuration are statistically linked to the temperature and not the configuration. Tracking extrema, and creating time series

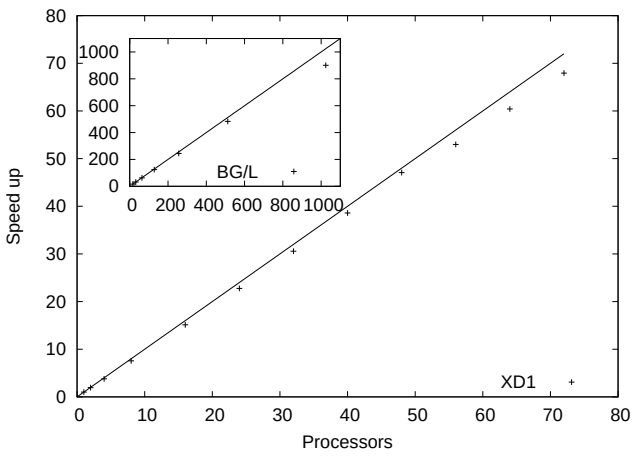


Figure 3: Parallel scaling of parallel tempering. The algorithm scales nearly linearly up to the 72 processors we have available on our Cray XD1. The inset shows the scaling on JUBL, NIC's IBM BlueGene/L, for comparison. The scaling remains nearly linear for up to 1024 processors. To scale to an even larger number of processors requires a direct neighbor-to-neighbor approach to the replica exchange. For the scaling tests, we used Protein G (Protein Data Bank ID 1PGB) as an examples.

at fixed temperature is much easier if we exchange configurations.

The replica exchange is performed on a master node. Energies are gathered through collective MPI calls. After the replica exchange the temperature distribution together with extremas is distributed to the nodes using collective calls as well. The master-node approach scales nearly optimally for the 72 processors, we have available on our XD1. We were also able to test the scaling on the new IBM BlueGene/L at the John von Neumann Institute for Computing in Jülich for up to 1024 processors. The scaling remains nearly linear (see Fig. 3).

4. SIMULATION RESULTS

We start by presenting our results for an isolated fragment β-amyloid$_{16-22}$. Fig. 4 displays the secondary structure content as a function of temperature. Here we observe two distinct temperature regions. At high temperatures, the peptide is a random coil with low sheet and helical content while at low temperatures helical configurations dominate. Here and in the following we define a residue as helical if its dihedral angle pair (ϕ, ψ) takes values $(-70^\circ \pm 30^\circ, -37^\circ \pm 30^\circ)$. Similarly, we define a residue as sheet-like if this pair takes values $(-150^\circ \pm 30^\circ, 150^\circ \pm 30^\circ)$. The cross-over between the two temperature regions occurs around $T \approx 325$K where the specific heat has a corresponding peak (see the inset of Fig. 4). Hence, in our model at room temperature the isolated fragment has an intrinsic tendency to form helices. This result is in contrast to Ref. 5, where they report only an insignificant increase in the helix propensity at low temperatures, but not unreasonable. For various peptides (see, for instance, Ref. 15)

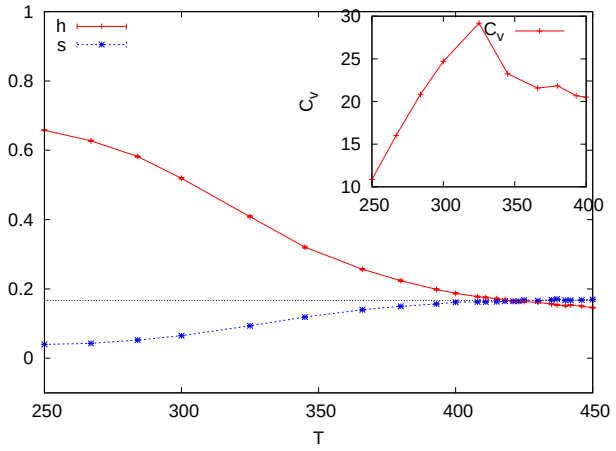
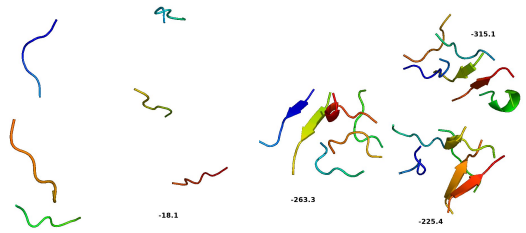


Figure 4: [Color online] Secondary structure content for a single β-amyloid$_{16-22}$ fragment. At temperatures below 325K, the helical content h (solid line) of the fragment dominates. It falls off to the value expected for a random coil at high temperatures. The sheet contents s is suppressed below the random value at low temperatures. The propensity values h and s for a random are indicated by the straight dotted line. The inset shows the specific heat of the random coil-helix cross over of a single fragment. The significant change in the helix propensity coincides with the specific-heat peak at $T \approx 325$K .

it has been observed that such short sequences prefer to adapt a helical configuration when isolated while they may take different shapes when interacting with other molecules.

Hence, the question arises how the situation changes when we no longer have a single isolated molecule but a system of six interacting β-amyloid 16-22 fragments. In Fig. 6 we display the specific heat of this compound system. Again, one observes a peak in the specific heat, shifted now to a temperature $T \approx 420$K, that separates a high temperature phase from one at low temperatures. The corresponding plot of the radius of gyration $R_{\rm gyr}$ (calculated over all atoms in the system) shows that this peak marks a sharp cross over from compact, aggregated structures at low temperatures to extended and isolated configurations at high temperatures. Note that the apparent saturation of $R_{\rm gyr}$ (and the shoulder in the spec. heat) is an artifact of our bounding box. This compactification happens despite an *increase* in the solvation energy with decreasing temperature (see inset of Fig. 7). It is driven by a decrease in the interaction energy that is almost an order of magnitude larger so that the total energy also decreases sharply with decreasing T.

Typical snapshots of the system in these two phases are displayed in Fig. 5(a) (which shows a typical high temperature configuration of our system) and Fig. 5(b) (showing dominant low temperature configurations). Consistent with the snapshots, the corresponding plot of the average percentage of sheetness in Fig. 8 indicates that the high temperature phase has little secondary content, while in the transition region the propensity for β-strands increases significantly up to a value of 23% just below the cross over. Note that the frequency of configuration with residues in helical configurations increase



(a) Weakly interacting chains at $T > T_\times$ (b) Aggregates at $T < T_\times$

Figure 5: [Color online] Sample configurations above and below the cross over.

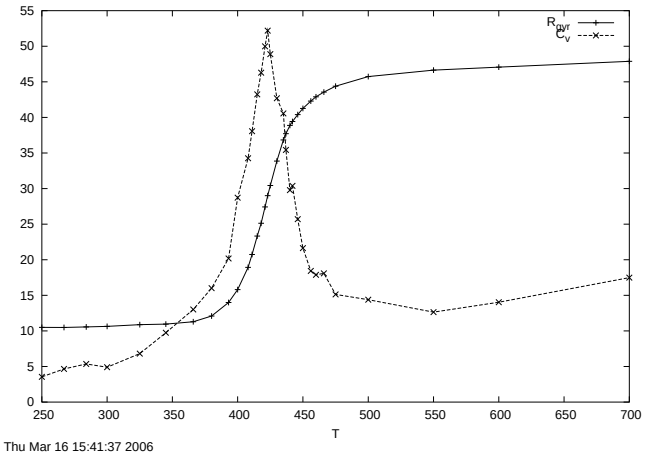


Figure 6: Radius of gyration vs. temperature. Radius of gyration $R_{\rm gyr}$ vs. temperature T (top). The error bars are one standard deviation. A sharp cross over is evident around 410K. The apparent saturation of $R_{\rm gyr}$ is an artifact of our fixed size bounding box. The crossover also shows up as a peak in the specific heat (bottom). The specific heat has been scaled down by a factor of 15 to emphasize the correspondence.

again below $T \approx 375$K, however, these configurations are at room temperature with 15% still much less frequently observed. With decreasing temperature, the importance of β-strands becomes more and more pronounced. The substantial propensity of helical structures is consistent with the results by Klimov *et al.* [10] that also observe an intermediate helical phase. The decrease in sheet propensity below $T \approx 375$K. and the magnitude of the increase in the helix propensity may in part be due insufficient sampling. Lower temperature require longer runs to equilibrate and sample configuration space sufficiently. Detailed results of these longer runs will be published elsewhere.

As one can see from Fig. 5(b), β-amyloid$_{16-22}$ forms stable aggregates quickly in agreement with previous simulations. Note that in agreement with previous work by Favrin *et al.* anti-parallel β-sheets occur at lower energies than parallel ones.

These authors found an increasing fraction of anti-parallel sheets for larger systems and lower temperatures although

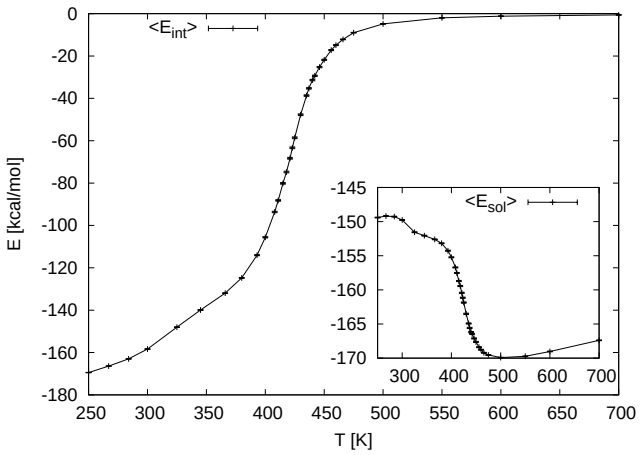


Figure 7: Partial energies vs. T. The decrease in the intermolecular interactions energy with decreasing temperature drives the aggregation despite an increase in the solvent energy term $E_{\rm sol}$ (inset). The intramolecular energy contributions don't show the cross over.

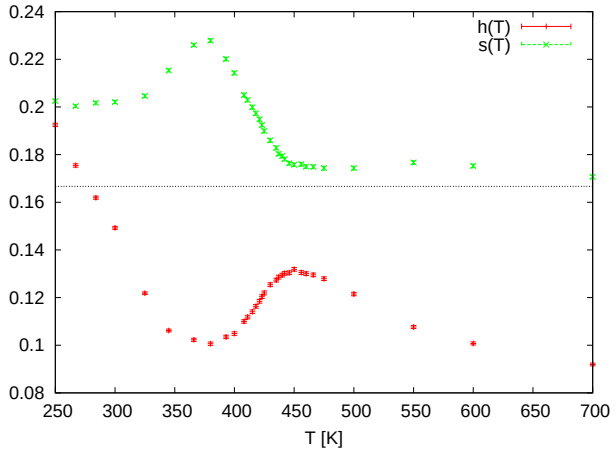


Figure 8: [Color online] Secondary structure content for a 6-chain system of β-amyloid$_{16-22}$. The sheet propensity increases with decreasing temperature around the cross-over regions up to about 23% while the helix contents clearly decreases to as little as 10%. Below $T \approx 375$K, this trend starts to reverse probably due to an insufficiently long equilibration time. The equilibration time increases with decreasing temperature.

their model does not include any Coulomb interaction. They argue that other effects, e.g., the arrangement of hydrophobic side chains or simple steric constraints may prefer anti-parallel alignment of sheets. This hypothesis is in contradiction to Klimov *et al.* [10] who emphasize that the anti-parallel sheets are favored because of salt-bridge formation.

5. CONCLUSIONS

We have studied the folding and aggregation of 6 chains of the β -amyloid fragment 16-22. While the isolated fragment prefers a helical form at room temperature, in the system of the 6 interacting fragments aggregates of parallel and anti-parallel β-sheets dominate below a cross over temperature $T_x \approx 420$K.The anti-parallel sheets have considerably lower energy and therefore are more stable. Even using efficient sampling algorithms such as parallel tempering, low temperatures require long equilibration and sampling runs. Parallel tempering is an excellent algorithm for parallel environments. It scales nearly linearly on a Cray XD1 up to 72 processors. Even using 1024 processors on an IBM BlueGene/L the speed up is only 6% below optimal.

Acknowledgments

U.H. acknowledges support by a research grant (CHE-9981874) of the National Science Foundation (USA).

About the Authors

Jan Meinke is a Postdoc in the NIC Research Group Computational Biology and Biophysics, E-mail: j.meinke@fz-juelich.de. Ulrich Hansmann is the head of the NIC Research Group Computational Biology and Biophysics and Professor of Physics at Michigan Technological University, E-mail: u.hansmann@fz-juelich.de. Both authors can be reached at NIC, FZ Jülich, 52425 Jülich, Germany.

[1] J.J. Balbach, Y. Ishii, O.N. Antzutkin, R.D. Leapman, N.W. Rizzo, F. Dyda, J. Reed, and R. Tycko. Amyloid fibril formation by A beta(16-22), a seven-residue fragment of the Alzheimer's beta-amyloid peptide, and structural characterization by solid state NMR. *BIOCHEMISTRY*, 39(45):13748 – 13759, NOV 14 2000.

[2] T. S. Burkoth, T. L. S. Benzinger, V. Urban, D. M. Morgan, D. M. Gregory, P. Thiyagarajan, R. E. Botto, S. C. Meredith, and D. G. Lynn. Structure of the b-amyloid(10-35) fibril. *J. Am. Chem. Soc.*, 122:7883 – 7889, 2000.

[3] F. Eisenmenger, U. H. E. Hansmann, S. Hayryan, and C.-K. Hu. [SMMP] A modern package for simulation of proteins. *Computer Physics Communications*, 138:192–212, August 2001.

[4] Frank. Eisenmenger, Ulrich. H. E. Hansmann, Shura Hayryan, and Chin-Kun Hu. An enhanced version of SMMP—an open-source software package for simulation of proteins. *Computer Physics Communications*, 174:422–429, 2006.

[5] G. Favrin, A. Irbäck, and S. Mohanty. Oligomerization of Amyloid Aβ16-22 Peptides Using Hydrogen Bonds and Hydrophobicity Forces. *Biophys. J .*, 87(6):3657 – 3664, 2004.

[6] C. J. Geyer and E. A. Thompson. Annealing Markov chain Monte Carlo with applications to ancestral inference. *J. Am. Stat. Assn.*, 90:909, 1995.

[7] D.J. Gordon, J.J. Balbach, R. Tycko, and S.C. Meredith.

Increasing the amphiphilicity of an amyloidogenic peptide changes the beta-sheet structure in the fibrils from antiparallel to parallel. *BIOPHYSICAL JOURNAL*, 86(1):428 – 434, JAN 2004.

[8] Ulrich H. E. Hansmann. Parallel tempering algorithm for conformational studies of biological molecules. *J. Chem. Phys. Lett.*, 281(1–3):140–150, 12 1997.

[9] Koji Hukushima and Koji Nemoto. Exchange Monte Carlo method and application to spin glass simulations. *Phys. Soc. (Jap)*, 65:1604 – 1608, 1996.

[10] D. K. Klimov and D. Thirumalai. Dissecting the assembly of abeta16-22 amyloid peptides into antiparallel beta sheets. *Structure*, 11(3):295 – 307, 2003.

[11] P. T. Lansbury, P. R. Costa, J. M. Griffiths, E. J. Simon, M. Auger, K. J. Halverson, D. A. Kocisko, Z. S. Hendsch, T. T. Ashburn, R. G. S. Spencer, B. Tidor, and R. G. Griffin. Structural model for the b-amyloid fibril based on interstrand alignment of an antiparallelsheet comprising a c-terminal peptide. *Nat. Struct. Biol.*, 2:990 – 998, 1995.

[12] T. Luehrs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Doebeli, D. Schubert, and R. Riek. 3D structure of Alzheimer's amyloid-β(1-42) fibrils. *Proc. Natl. Acad. Sci. USA*, 102(48):17342–17347, 2005.

[13] Jan H. Meinke and Ulrich H. E. Hansmann. SMMP v. 3 — A modern software package for simulating proteins and protein interactions (unpublished). You can download SMMP from our web page http://www.fz-juelich.de/nic/cbb.

[14] G. Nemethy, K. D. Gibson, K. A. Palmer, C. N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, and H. A. Scheraga. Energy parameters in polypeptides. 10. improved geometrical parameters and nonbonded interactions for use in the ecepp/3 algorithm, with application to proline-containing peptides. *Journal of Physical Chemistry*, 96(15):6472 – 6484, 1992.

[15] Y. Peng and U. H. E. Hansmann. Helix versus sheet formation in a small peptide. *Phys. Rev. E*, 68:041911, 2003.

[16] Yong Peng and Ulrich H. E. Hansmann. Solvation model dependency of helix-coil transition in polyalanine. *Biophys. J.*, 82(6):3269–3276, 2002.

[17] Yong Peng, Ulrich H. E. Hansmann, and Nelson A. Alves. Solution effects and the order of the helix coil transition in polyalanine. *J. Chem. Phys.*, 118(5):2374–2380, 2003.

[18] A. T. Petkova, G. Buntkowsky, F. Dyda, R. D. Leapman, W.-M. Yau, and R. Tycko. Solid state nmr reveals a ph-dependent antiparallel beta-sheet registry in fibrils formed by a beta-amyloid peptide. *J. Mol. Biol.*, 335(1):247 – 260, 2004.

[19] A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, and R. Tycko. A structural model for Alzheimer's b-amyloid fibrils based on experimental constraints from solid state nmr. *Proc. Natl. Acad. Sci. USA*, 99:16742 – 16747, 2002.

[20] J. C. Rochet and P. T. Lansbury Jr. Amyloid fibrillogenesis: themes and variations. *Curr Opin Struct Biol*, 10(1):60 – 68, 2000.

[21] L. O. Tjernberg, J. Naslund, F. Lindqvist, J. Johansson, A. R. Karlstrom, J. Thyberg, L. Terenius, and C. Nordstedt. Arrest of beta-amyloid fibril formation by a pentapeptide ligand. *J. Biol. Chem.*, 271(15):8545 – 8548, 1996.