John von Neumann Institute for Computing

Protein Aggregation and Unfolding Studied Using an All-Atom Model with a Simplified Energy Function

Anders Irbäck, Simon Mitternacht

published in

NIC Workshop 2006,
From Computational Biophysics to Systems Biology,
Jan Meinke, Olav Zimmermann,
Sandipan Mohanty, Ulrich H.E. Hansmann (Editors)
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **34**, ISBN-10: 3-9810843-0-6,
ISBN-13: 978-3-9810843-0-6, pp. 15-20 , 2006.

http://www.fz-juelich.de/nic-series/volume34

Protein Aggregation and Unfolding Studied Using an All-Atom Model with a Simplified Energy Function

Anders Irbäck and Simon Mitternacht

Computational Biology and Biological Physics Group,
Department of Theoretical Physics, Lund University,
Sölvegatan 14A, SE-223 62 Lund, Sweden
E-mail: {*anders, simon*}*@thep.lu.se*

We discuss two Monte Carlo studies based on a simple all-atom model for proteins. The first study explores the early aggregation steps of Aβ_{16-22}, an amyloid fibril-forming 7-residue fragment of Alzheimer's Aβ peptide. The other study deals with the mechanical and thermal unfolding of ubiquitin, a 76-residue protein that was studied in recent single-molecule constant-force experiments.

1 Introduction

Protein aggregation into amyloid fibrils is a recurrent theme in several human disorders, including Alzheimer's and Parkinson's diseases,[1] and there is evidence that amyloid structures can have a functional role, too.[2] The mechanisms of amyloid formation are currently being intensely investigated, both experimentally and by computer simulations. These studies are not limited to fibrillar aggregates; small assemblies get more and more attention, because of findings that link soluble oligomers to pathology.[3] A broad set of sequences is studied, from disease-associated proteins like the Alzheimer's Aβ peptide to designed amyloid sequences like the hexapeptide STVIIE.[4]

Here we discuss a study of small assemblies of the Aβ fragment Aβ_{16-22},[5] which was performed using an all-atom model with a simplified energy function. In addition, we discuss a study of the mechanical and thermal unfolding of ubiquitin,[6,7] based on exactly the same model. This model was developed through folding studies of a set of well characterized peptides,[8,9] including α-helical as well as β-sheet peptides. For these peptides, the model was found to give a good description of both structure and thermodynamics.[9]

Mechanical unfolding has been studied experimentally at the single-molecule level for several proteins. These studies have provided valuable insights into the elastic properties of, e.g., the muscle protein titin.[10,11] Typically, these experiments focus on the extension-versus-force behavior. Computer simulations have the potential to provide information not captured by the experiments, and thereby give a more complete picture of the unfolding process.

2 Model and Methods

The model we use contains all atoms of the protein chains, including hydrogen atoms, but no explicit water molecules. It assumes fixed bond lengths, bond angles and peptide torsion angles (180°), so that each amino acid only has the Ramachandran torsion angles ϕ, ψ and a number of side-chain torsion angles as its degrees of freedom.

The energy function

$$E = E_{\text{loc}} + E_{\text{ev}} + E_{\text{hb}} + E_{\text{hp}} \tag{1}$$

is composed of four terms. The term E_{loc} is local in sequence and represents interactions between adjacent backbone dipoles along the chain. The other three terms are non-local in sequence. The excluded volume term E_{ev} is a $1/r^{12}$ repulsion between pairs of atoms. E_{hb} represents two kinds of hydrogen bonds: backbone-backbone bonds and bonds between charged side chains and the backbone. The last term E_{hp} represents an effective hydrophobic attraction between non-polar side chains. It is a simple pairwise additive potential based on the degree of contact between two non-polar side chains. A detailed description of all the different terms can be found elsewhere.[8,9]

Despite its simplicity, this energy function is able to fold several α-helical and β-sheet peptides with about 20 amino acids.[9] One type of interaction that the model neglects is the Coulomb interaction between charged side chains. For a small peptide, these charges tend to be exposed to, and therefore screened by the solvent. To be able to study larger proteins and protein aggregates, we expect that it will be necessary to refine the energy function to take into account, e.g., the interactions between side-chain charges.

All our studies were carried out using PROFASI,[12] which is a Monte Carlo software package for simulations of this model.

Let us stress that the amino acid sequence is the only input to the model. All model parameters were thus kept the same in our different studies.

3 Aβ_{16-22} Aggregation

A characteristic "cross-β" X-ray fiber diffraction pattern reveals that the core structure of amyloid fibrils is composed of β-sheets whose strands run perpendicular to the fibril axis.[13] For Aβ_{16-22} fibrils, it has been found, by solid-state NMR, that the β-strands have an antiparallel organization.[14,15] The Aβ_{16-22} sequence (Ac-KLVFFAE-NH_2) consists of five inner residues that are hydrophobic, and two end residues that are oppositely charged. Coulomb interactions between such charges provide a possible explanation for the antiparallel β-strand organization in Aβ_{16-22} fibrils.

Computer simulations of Aβ_{16-22} aggregation have been reported by several groups[16–19]. In our calculations, we studied systems of one, three and six Aβ_{16-22} peptides. The simulations were started from random configurations, and the energy function was exactly the same as in our folding studies.[9]

The isolated Aβ_{16-22} peptide turned out to be disordered in our simulations. The three- and six-chain systems, on the other hand, self-assembled into ordered, β-strand-rich aggregates. There was no single dominating free-energy minimum, but rather a number of more or less degenerate minima. Fig. 1 shows two snapshots of such minima.

The β-strand organization in these structures is interesting since our model neglects the interactions between charged side chains, which might be responsible for the antiparallel organization seen in Aβ_{16-22} fibrils. A detailed analysis showed that mixed configurations with both parallel and antiparallel β-strand pairs were common in our simulations, but nevertheless there was a clear statistical preference for the antiparallel organization over the parallel one. Our model thus favors the antiparallel organization despite that these Coulomb interactions are neglected. This, of course, does not mean that these interactions

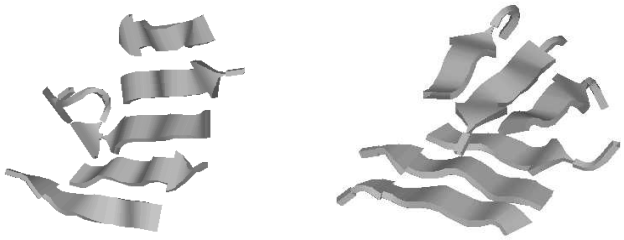

Figure 1. Two typical low-energy structures from our simulations of six Aβ_{16-22} peptides: a five-stranded β-sheet (left), and two three-stranded β-sheets "sandwiching" several of their hydrophobic side chains between them (right). Drawn with RasMol.[20]

are unimportant, but it suggests that other factors (e.g., hydrogen bond geometry) play a significant role, too. We also did some simulations with attraction/repulsion between unlike/like side-chain charges. Not unexpectedly, this led to an increased preference for the antiparallel organization.

A configuration type that did not occur at all in these simulations was closed barrel-like structures. For this sequence, it seems that six chains are not enough to permit the formation of such structures. By contrast, we have seen the formation of closed barrel-like structures in simulations for nine Aβ_{16-22} peptides.

4 Mechanical and Thermal Unfolding of Ubiquitin

Ubiquitin is a 76-residue α/β protein, whose unfolding and refolding properties have been extensively studied experimentally.[21] Its native structure contains an α-helix and a five-stranded β-sheet (see Fig. 2).

The mechanical unfolding of ubiquitin was recently investigated by the Fernandez group by single-molecule methods.[23–25] One study examined the unfolding behavior under a constant stretching force, using end-to-end linked polyubiquitin.[25] Here, the time evolution of the end-to-end distance r was followed, and a total of about 800 unfolding traces were collected. In most cases, unfolding occurred in one step, but several examples of unfolding through intermediate states were also observed. The size of the unfolding step to the typical intermediate state was consistent with what one would expect if the α-helix and the N-terminal β-hairpin (A and B in Fig. 2) remain folded in this state, whereas the rest of the molecule is unfolded. Interestingly, these two structures have been found to be the most stable ones in several different experiments at zero force.[21]

We studied the unfolding of ubiquitin under a constant stretching force using the same strengths of the applied force as in the experiments (100 pN, 140 pN and 200 pN).[25] The energy function was $E_f = E - \vec{f} \cdot \vec{r}$, where the internal energy E is the same as before (see Eq. 1), $\vec{f}$ is the applied force, and $\vec{r}$ is the end-to-end vector. Because of the existence of multiple unfolding pathways, we performed a set of 500 runs for each force.

As in the experiments, we saw both one-step unfolding and unfolding through intermediate states in our simulations. Furthermore, properties such as the size of the unfolding step, the frequency of occurrence of intermediate states, and the position of the typical

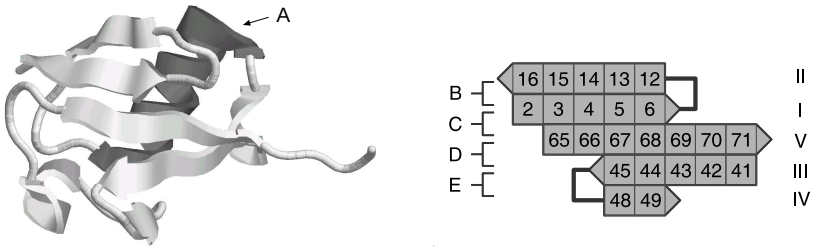

Figure 2. The native structure of ubiquitin with our labels for secondary-structure elements, A–E. Left: A 3D model (PDB code 1d3z, first model[22]) drawn with RasMol.[20] Right: The organization of the β-sheet.

intermediate state were all found to be in reasonable agreement with experimental data. Having verified these properties, we performed more detailed measurements which, in particular, aimed at characterizing the typical intermediate state.

For this purpose, we investigated the order of breaking of five secondary-structure elements, labeled A–E (see Fig. 2). The structure A is the α-helix, whereas B–E are the four pairs of adjacent strands in the five-stranded β-sheet. To determine the order of breaking, the native hydrogen bonds in these structures were monitored as a function of the end-to-end distance r (r increased essentially monotonically with time). Fig. 3 summarizes the results of this analysis, at 100 pN. From this figure it is immediately clear that the structures A–E do not break in a random order but instead in a statistically preferred order, namely CBDEA. C and B tend to break below the typical r for intermediate states, $r_{\rm int} \approx 12\,$nm, whereas D, E and A tend to break above $r_{\rm int}$. Our results thus suggest that the typical intermediate is composed of D, E and A rather than A and B.

Fig. 3 shows averages over all events, and therefore does not tell how strong the statistical preference is for the unfolding order CBDEA. To investigate this, we also did an event-by-event analysis, and found that 61 % of the events followed the unfolding pathway CBDEA. Another 23 % of the events had the order of B and D apparently interchanged, the path being CDBEA. In these events, B does unfold before D, but partially reforms after D is gone. The partial refolding of B is not a step back toward the native state, because when B reforms, D is gone, and r is larger. Hence, 84 % of the events followed the same basic pathway.

Several aspects of this calculated unfolding order can be understood in terms of native topology and pulling geometry. That C breaks first is inevitable; the other parts cannot sense the force until C is broken. The native state is mechanically resistant because C is pulled longitudinally, so that several hydrogen bonds must break at the same time. Once C is gone, nothing keeps B from unzipping, one bond at a time. Unzipping requires less force than separation by longitudinal pulling. Therefore, it seems reasonable that B breaks soon after C, as it did in our simulations.

That B breaks early implies that ubiquitin shows a different behavior in these simulations than in various experiments at zero force. In particular, there are experiments suggesting that B along with A are the thermally most stable parts of ubiquitin.[26,27] Therefore, one must ask whether one actually sees a difference between mechanical and thermal unfolding in our model. To address this question, we performed a set of 800 thermal un-

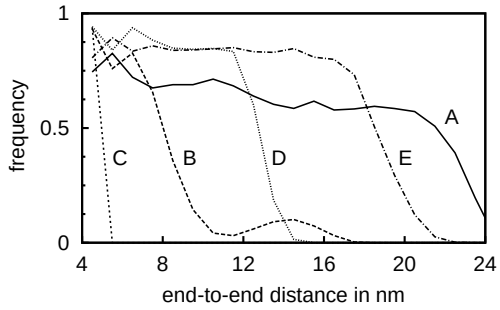

Figure 3. Frequency of occurrence of native hydrogen bonds in the structures A–E against end-to-end distance. Each curve represents an average over all the native hydrogen bonds in a given structure and over all events.

folding simulations at a fixed temperature.[7] In a majority of these events, A and B unfolded after C and D (E was not analyzed due to noisy data), which indeed is in agreement with the experiments. The agreement with experimental data in the thermal case strengthens our proposed mechanical unfolding order, which remains to be verified experimentally.

Acknowledgments

We thank Sandipan Mohanty for useful discussions. This article is based on work supported in part by the Swedish Research Council.

References

1. For a review, see C. M. Dobson, *The structural basis of protein folding and its links with human disease*, Phil. Trans. R. Soc. Lond. **B 356**, 133–145 (2001).
2. D. M. Fowler, A. V. Koulov, C. Alory-Jost, M. S. Marks, W. E. Balch, and J. W. Kelly, *Functional amyloid formation within mammalian tissue*, PloS Biol. **4**, e6 (2006).
3. S. Lesné, M. T. Koh, L. Kotilinek, R. Kayed, C. G. Glabe, A. Yang, M. Gallagher, and K. H. Ashe, *A specific amyloid-β assembly in the brain impairs memory*, Nature **440**, 352–357 (2006).
4. M. López de la Paz and L. Serrano, *Sequence determinants of amyloid fibril formation* Proc. Natl. Acad. Sci. USA **101**, 87–92 (2004).
5. G. Favrin, A. Irbäck, and S. Mohanty, *Oligomerization of amyloid $A\beta_{16-22}$ peptides using hydrogen bonds and hydrophobicity forces*, Biophys. J. **87**, 3657–3664 (2004).
6. A. Irbäck, S. Mitternacht, and S. Mohanty, *Dissecting the mechanical unfolding of ubiquitin*, Proc. Natl. Acad. Sci. USA **102**, 13427–13432 (2005).
7. A. Irbäck and S. Mitternacht, *Thermal versus mechanical unfolding of ubiquitin*, to appear in Proteins.
8. A. Irbäck, B. Samuelsson, F. Sjunnesson, and S. Wallin, *Thermodynamics of α- and β-structure formation in proteins*, Biophys. J. **85**, 1466–1473 (2003).
9. A. Irbäck and S. Mohanty, *Folding thermodynamics of peptides*, Biophys. J. **88**, 1560–1569 (2005).

10. M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, *Reversible unfolding of individual titin immunoglobulin domains by AFM*, Science **276**, 1109–1112 (1997).
11. M. S. Z. Kellermayer, S. B. Smith, H. L. Granzier, and C. Bustamente, *Folding-unfolding transitions in single titin molecules characterized by laser tweezers*, Science **276**, 1112–1116 (1997).
12. A. Irbäck and S. Mohanty, *PROFASI: a Monte Carlo simulation package for protein folding and aggregation*, to appear in J. Comp. Chem.
13. M. Sunde and C. Blake, *The structure of amyloid fibrils by electron microscopy and X-ray diffraction*, Adv. Protein Chem. **50**, 123-159 (1997).
14. 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 β-amyloid peptide, and structural characterization by solid state NMR*, Biochemistry **39**, 13748-13759 (2000).
15. D. J. Gordon, J. J. Balbach, R. Tycko, and S. C. Meredith, *Increasing the amphiphilicity of an amyloidogenic peptide changes the β-sheet structure in the fibrils from antiparallel to parallel*, Biophys. J. **86**, 428-434 (2004).
16. B. Ma and R. Nussinov, *Stabilities and conformations of Alzheimer's β-amyloid peptide oligomers ($A\beta_{16-22}$, $A\beta_{16-35}$, and $A\beta_{10-35}$): sequence effects*, Proc. Natl. Acad. Sci. USA **99**, 14126-14131 (2002).
17. D. K Klimov and D. Thirumalai, *Dissecting the assembly of $A\beta_{16-22}$ amyloid peptides into antiparallel β sheets*, Structure **11**, 295-307 (2003).
18. S. Santini, N. Mousseau, and P. Derreumaux, *In silico assembly of Alzheimer's $A\beta_{16-22}$ peptide into β-sheets*, J. Am. Chem. Soc **126**, 11509–11516 (2004).
19. S. Gnanakaran, R. Nussinov, and A.E. García, *Atomic-level description of amyloid β-dimer formation*, J. Am. Chem. Soc. **128**, 2158–2159 (2006).
20. R. A. Sayle and E. J. Milner-White, RasMol: biomolecular graphics for all, Trends Biochem. Sci. **20**, 374–376 (1995).
21. For a review, see S. E. Jackson, *Ubiquitin: a small protein folding paradigm*, Org. Biomol. Chem. **4**, 1845–1853 (2006).
22. G. Cornilescu, J. L. Marquardt, M. Ottiger, and A. Bax, *Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute crystalline phase*, J. Am. Chem. Soc. **120**, 6836–6837 (1998).
23. M. Carrion-Vazquez, H. Li H, H. Lu, P. E. Marszalek, A. F. Oberhauser, and J. M. Fernandez, *The mechanical stability of ubiquitin is linkage dependent*, Nat. Struct. Biol. **10**, 738–743 (2003).
24. J. M. Fernandez and H. Li, *Force-clamp spectroscopy monitors the folding trajectory monitors the folding trajectory of a single protein*, Science **303**, 1674–1678 (2004).
25. M. Schlierf, H. Li, and J. M. Fernandez, *The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques*, Proc. Natl. Acad. Sci. USA **101**, 7299–7304 (2004).
26. F. Cordier and S. Grzesiek, *Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR*, J. Mol. Biol. **317**, 739–752 (2002).
27. H. S. Chung, M. Khalil, A. W. Smith, Z. Ganim, and A. Tokmakoff, *Conformational changes during the nanosecond-to-millisecond unfolding of ubiquitin*, Proc. Natl. Acad. Sci. USA **102**, 612–617 (2005).