Supporting Information for: Compact fibril-like structure of amyloid β-peptide (1-42) monomers†

Bogdan Barz,*[a,b] Alexander K. Buell,[c] and Soumav Nath[a,b]

1 Methods

1.1 MD simulations details

We have studied the conformational flexibility of the Aβ42 monomer in water by performing H-REMD[1] simulations using the Charmm36m force field[2] and the Charmm TIP3P water model designed for folded proteins and IDPs. The TIP3P water model was modified such that it leads to an increased protein - water interaction[2]. In Charmm36mW, the Lennard Jones well depth parameter ε of the hydrogen atoms have been modified from -0.046 kcal/mol to -0.10 kcal/mol as suggested by Huang et al.[2]. The Charmm36m force field together with the modified TIP3P water model are labeled as Charmm36mW. The simulations were performed with the Gromacs 2016.04 parallel software package[3]. Short range electrostatics and van der Waals interactions were cut at 0.1 nm, while long range electrostatic interactions were treated with the Particle Mesh Ewald method. The temperature was kept at 300 K via velocity rescaling with a stochastic term algorithm[4] and a time constant for coupling of 0.1 ps. The pressure coupling was controlled with the Parrinello-Rahman barostat[5,6] with a time constant of 1 ps. The hydrogen atoms were treated as virtual interaction sites, allowing an integration time step of 4 fs while maintaining energy conservation[7]. To simulate physiological conditions we also included 150 mM NaCl concentration.

The solution structure of Aβ42 protein monomer with PDB ID 1IYT was used as starting conformation[8]. The conformation was placed in a dodecahedral box with 1.6 nm between the protein and the box, solvated with water molecules, 43 Na and 40 Cl ions. The final systems had 47,054 atoms. This simulation box was large enough to allow free translation and rotation of the Aβ42 monomer without interacting with its periodic images that would otherwise result in simulation artifacts. After energy minimization, position restrained equilibration (with the Berendsen barostat) and a short free equilibration (with the Parrinello-Rahman barostat), the Hamiltonian replica exchange[1] simulation was performed, consisting of 12 simulations run in parallel, each simulation having a different interaction Hamiltonian where non-bonded interactions and dihedral angles are scaled with a factor λ or $\sqrt{\lambda}$. The biasing coefficients λ can be expressed as an inverse temperature (1/temperature) correspond to temperatures between 300 and 500 K and assigned to the replicas according to a geometric distribution. The average replica exchange probability was 24%. We thus performed one H-REMD simulations of 3.4 μs/replica, with 12 replicas in total, amounting to a total simulation time of 40.8 μs on the the supercomputer JURECA[9] at the Jülich Supercomputing Centre (JSC).

1.2 Analysis

Since all simulations started from the Aβ42 structure with the PDB code ID 1IYT[8], the initial state has large number of amino acids in helical conformation, see Figure S9. The highly helical conformation drops to small values below 10 residues within 1,500 ns as can be seen in Figure S9. The drop in helical conformation is accompanied by an increase in the β-sheet content. Given the random coil character of the Aβ42 monomer observed in experiments[10], we have decided to analyze only the equilibrated part the simulation, where fluctuations in the secondary structure are similar to those of random coil, i.e. without many amino acids in helical conformation. Thus, the time interval used for the analysis consist of the last 1,000 ns of the unbiased trajectory, highlighted in Figure S9.

$^3J_{HNH\alpha}$ NMR scalar couplings

$^3J_{HNH\alpha}$ NMR scalar couplings were calculated with the Karplus equation for each amino acid:

$$\langle 3J_{HNH\alpha}\rangle = \langle A\cos^2\phi + B\cos\phi + C\rangle \tag{1}$$

[a] Institut für Physikalische Biologie, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany.
[b] Institute of Biological Information Processing - Structural Biochemistry (IBI-7), Research Centre Jülich, Jülich, Germany. Fax: +49 2461 619448; Tel: +49 2461 619448; E-mail: b.barz@fz-juelich.de
[c] Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark.

with coefficients A = 7.97 Hz, B = - 1.26 Hz and C = 0.63 Hz[11]. The comparison with experimental data was done using the reduced χ^2:

$$\chi^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{J_{i,exp} - \langle J_i \rangle_{sum}}{(\Delta J_i)^2} \quad (2)$$

where $(\Delta J_i)^2 = (\Delta_{\text{block}})^2 + (\Delta_{\text{Karplus}})^2$, Δ_{block} being the simulation error calculated with block averaging and $\Delta_{\text{Karplus}} = 0.42$ Hz, the experimental error.

Secondary structure, contact maps and clustering

For snapshots of the protein structure we used the program Visual Molecular Dynamics (VMD)[12] where the secondary structure is calculated with STRIDE[13]. Contact maps were calculated using the Contact Map Explorer module implemented in Python and considering a contact between two amino acids when any two atoms from the two residues were found at a distance below a cutoff of 0.5 nm. Clustering of structures was performed with the Daura algorithm[14] implemented in Gromacs using the backbone atoms and an RMSD cutoff of 0.4 nm. In order to assess the sensitivity of the most populated cluster to the cutoff used for the RMSD, we have re-calculated the cluster distributions using additional cutoff values. Thus, we have performed in total three clustering calculations with cutoffs of 0.35 nm, 0.4 nm and 0.45 nm. The total number of clusters increases with increasing the cutoff as follows: 2435 clusters for a cutoff of 0.35 nm; 3205 clusters for a cutoff of 0.4 nm; and 4198 clusters for a cutoff of 0.45 nm. Despite the changes in the number of clusters, the population of the first cluster remained consistently at 14% of the total number of structures for each cutoff value.

In addition to cluster one, discussed in the main text, of particular interest are the clusters two and three, which contain rather extended conformations with many similarities. As can be seen from the contact maps in Figure S2, these two clusters have strong contacts with patterns that correspond to anti-parallel β-sheets between amino acids Y10-D23 and A42-S26. In the representative structures can also be identified two anti-parallel β-sheets of different lengths in agreement with the contact map. This type of structure might be relevant for the formation of the U-shape fibrils[15], where anti-parallel contacts are present between amino acids L17-G25 and V40-I32 for the side-chains.

β-sheet content, radius of gyration and hydrodynamic radius

The random coil behaviour of the Aβ42 monomer observed in NMR experiments was also confirmed by circular dichroism (CD) spectroscopy experiments where the Aβ monomers adopt between 12–25% β-sheet and much lower helical content[16,17]. Using the DSSP[18] software together with the Gromacs tool "do_dssp" we calculated an average β-sheet propensity of 19.6 $\pm$ 2.9%, in good agreement with experimental values. In order to characterize the conformational diversity of the monomeric Aβ, we have calculated the radius of gyration of all conformations in the simulation. The average radius of gyration has a value of 1.4 $\pm$ 0.3 nm, slightly smaller than values reported by other computational studies, i.e. 1.6 nm[19] and 1.59 nm[20], but within error. Using an empirically parametrized equation which relates the radius of gyration to the hydrodynamic radius for intrinsically disordered proteins[21], we have determined an averaged hydrodynamic radius of $\sim$1.7$\pm$0.1 nm, in good agreement with experimental values of both Aβ40, i.e. 1.6 nm[22], and the upper range of values for Aβ42, i.e. 0.9–1.8 nm[23–27].

RMSD of compact monomer with peptides from fibril models

Here we continue the discussion from the main text regarding the comparison of the C-terminal loop from the compact monomer with other fibril models. In the case of the double filament structure published by Colvin et al.[28], the smallest RMSD has a value of 0.19 nm. Figure S7b shows all hydrophobic amino acids pointing in the same direction, except for M35. When compared with fibrils obtained at pH of 2 and in the presence of an organic co-solvent by Gremer et al.[29], Figure S7c, the RMSD has a value of 0.17 nm. Despite the very small backbone RMSD, only L34 and M35 from the monomer point in the same direction as the residues in the fibrillar peptide. The mismatches are not surprising, given the effect of the low pH on the protonation state of charged amino acids and eventually on the overall structure of the fibril, possibly exacerbated by the organic co-solvent.

For completeness, we also compared the full length and the N-terminus loop of the monomer with the peptides from fibril models and we report the values in Table ST1 from ESI. The RMSD for the full length peptide have values between 0.82 and 1.14 nm indicating a poor structural fit. When considering only the N-terminus loop, the RMSD varies between 0.27 and 1.12 nm, with the lowest value for the fibril model with PDB ID 2MXU. This case is also displayed in Figure S8 where one can identify an extended peptide chain for both the monomer and the peptide from the fibril model.

Salt-bridge analysis

In addition to the salt-bridge discussion from the main text we include the following results regarding the salt-briges

formed by K28. From Figure S3b it is clear that E22 and A42 are close to each other and form alternating or simultaneous salt-bridges with K28. We have identified the conformations where K28 forms salt-bridges with both E22 and A41 and calculated a propensity of 29.85%. This means that within cluster one, K28 was engaged in salt-bridges with either E22 or A42 in 76.47% of the conformations.

Salt-bridge analysis details. To quantify the occurrence of the salt-bridges that stabilize the compact monomer we have calculated distances between the two oxygen atoms of the negatively charged amino acids (E22 and D23, A42) and the three hydrogen atoms of the positively charged ones (D1 and K28) involved in the salt bridge.The results for the three salt bridges obtained from all the conformations in cluster one of the simulation with Charmm36mW and NaCl are shown in Figures S4 (D23-D1), S5 (E22-K28), and S6 (A42-K28). In the case of salt-bridge D23–D1, most of the conformations have either the distance between oxygen O1 or oxygen O2 of residue D23 and the three hydrogens of residue D1 below 0.4 nm which is the threshold distance to qualify as a salt-bridge. This is also shown in the normalized distributions from Figure S4 Bottom where the largest peaks below 0.4 nm belong to distances between O2 and H1 or H2. The dynamics of this salt bridge is clear from the plots in Figure S4 Middle for both a) and b), where the three hydrogens alternate as the closest atom to the oxygen O1 or O2, with oxygen O2 being preferred for shorter distances. A similar analysis is shown in Figure S5 for salt-bridge E22–K28 and in Figure S6 for salt-bridge A42–K28. Salt-bridge E22–K28 appears in a large number of structures from cluster one, but is not as stable as D23–D1.

Notes and references

1 G. Bussi, *Mol. Phys.*, 2014, **112**, 379–384.
2 J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller and A. D. MacKerell Jr, *Nat. Methods*, 2017, **14**, 71–73.
3 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *J. Chem. Theory. Comput.*, 2008, **4**, 435–447.
4 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
5 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
6 S. Nosé and M. L. Klein, *Mol. Phys.*, 1983, **50**, 1055–1076.
7 K. A. Feenstra, B. Hess and H. J. C. Berendsen, *J. Comput. Chem.*, 1999, **20**, 786–798.
8 O. Crescenzi, S. Tomaselli, R. Guerrini, S. Salvadori, A. M. D'Ursi, P. A. Temussi and D. Picone, *Eur. J. Biochem.*, 2002, **269**, 5642–5648.
9 Jülich Supercomputing Centre, *Journal of large-scale research facilities*, 2018, **4**, year.
10 J. Roche, Y. Shen, J. H. Lee, J. Ying and A. Bax, *Biochemistry*, 2016, **55**, 762.
11 B. Vögeli, J. Ying, A. Grishaev and A. Bax, *J. Am. Chem. Soc.*, 2007, **129**, 9377–9385.
12 W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33–38.
13 D. Frishman and P. Argos, *Proteins: structure, function and genetics*, 1995, **23**, 566–579.
14 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew. Chem. Int. Ed.*, 1999, **38**, 236–240.
15 T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert and R. Riek, *Proc. Natl. Acad. Sci. U.S.A.*, 2005, **102**, 17342–17347.
16 K. H. Lim, H. H. Collver, Y. T. H. Le, P. Nagchowdhuri and J. M. Kenney, *Biochem. Biophys. Res. Commun.*, 2007, **353**, 443–449.
17 K. Ono, M. M. Condron and D. B. Teplow, *Proc. Natl. Acad. Sci.*, 2009, **106**, 14745–14750.
18 W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577–2637.
19 F. Meng, M. M. J. Bellaiche, J.-Y. Kim, G. H. Zerze, R. B. Best and H. S. Chung, *Biophys. J.*, 2018, **114**, 870–884.
20 J. Lincoff, S. Sasmal and T. Head-Gordon, *J. Chem. Phys.*, 2019, **150**, 104108.
21 M. Nygaard, B. B. Kragelund, E. Papaleo, K. Lindorff-Larsen, M. Nygaard, B. B. Kragelund, E. Papaleo and K. Lindorff-Larsen, *Biophys. J.*, 2017, **113**, 550–557.
22 D. Granata, F. Baftizadeh, J. Habchi, C. Galvagnion, A. De Simone, C. Camilloni, A. Laio and M. Vendruscolo, *Sci. Rep.*, 2015, **5**, 1–15.
23 S. Nag, B. Sarkar, A. Bandyopadhyay, B. Sahoo, V. K. A. Sreenivasan, M. Kombrabail, C. Muralidharan and S. Maiti, *J. Biol. Chem.*, 2011, **286**, 13827–13833.
24 J. J. Mittag, S. Milani, D. M. Walsh, J. O. Rädler and J. J. McManus, *Biochem. Biophys. Res. Commun.*, 2014, **448**, 195–199.
25 M. Schneider, S. Walta, C. Cadek, W. Richtering and D. Willbold, *Sci. Rep.*, 2017, **7**, 1–8.
26 S. Wennmalm, V. Chmyrov, J. Widengren and L. Tjernberg, *Anal. Chem.*, 2015, **87**, 11700–11705.
27 M. Novo, S. Freire and W. Al-Soufi, *Sci. Rep.*, 2018, **8**, 1–8.
28 M. T. Colvin, R. Silvers, Q. Z. Ni, T. V. Can, I. Sergeyev, M. Rosay, K. J. Donovan, B. Michael, J. Wall, S. Linse and R. G. Griffin, *J. Am. Chem. Soc.*, 2016, **138**, 9663–9674.
29 L. Gremer, D. Schölzel, C. Schenk, E. Reinartz, J. Labahn, R. B. G. Ravelli, M. Tusche, C. Lopez-Iglesias, W. Hoyer, H. Heise, D. Willbold and G. F. Schröder, *Science*, 2017, **358**, 116–119.
30 Y. Xiao, B. Ma, D. McElheny, S. Parthasarathy, F. Long, M. Hoshi, R. Nussinov and Y. Ishii, *Nat. Struct. Mol. Biol.*, 2015, **22**, 499–505.

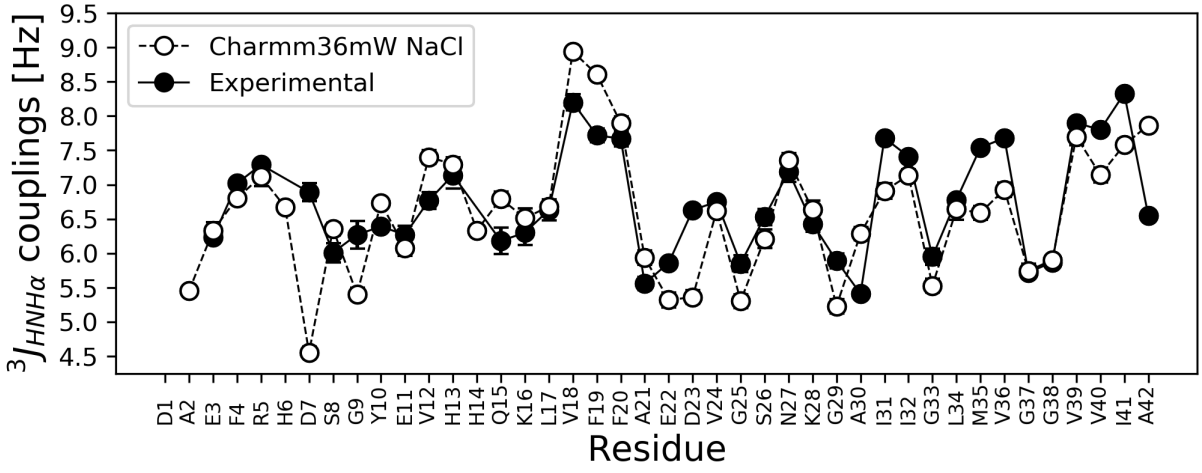

Fig. S1 $^3J_{HNH\alpha}$ NMR scalar couplings calculated for each amino acid for Charmm36mW force field with 150 mM NaCl. Experimental values are shown as black circles and those obtained from the simulations as white circles. The error bars were calculated using block averaging.

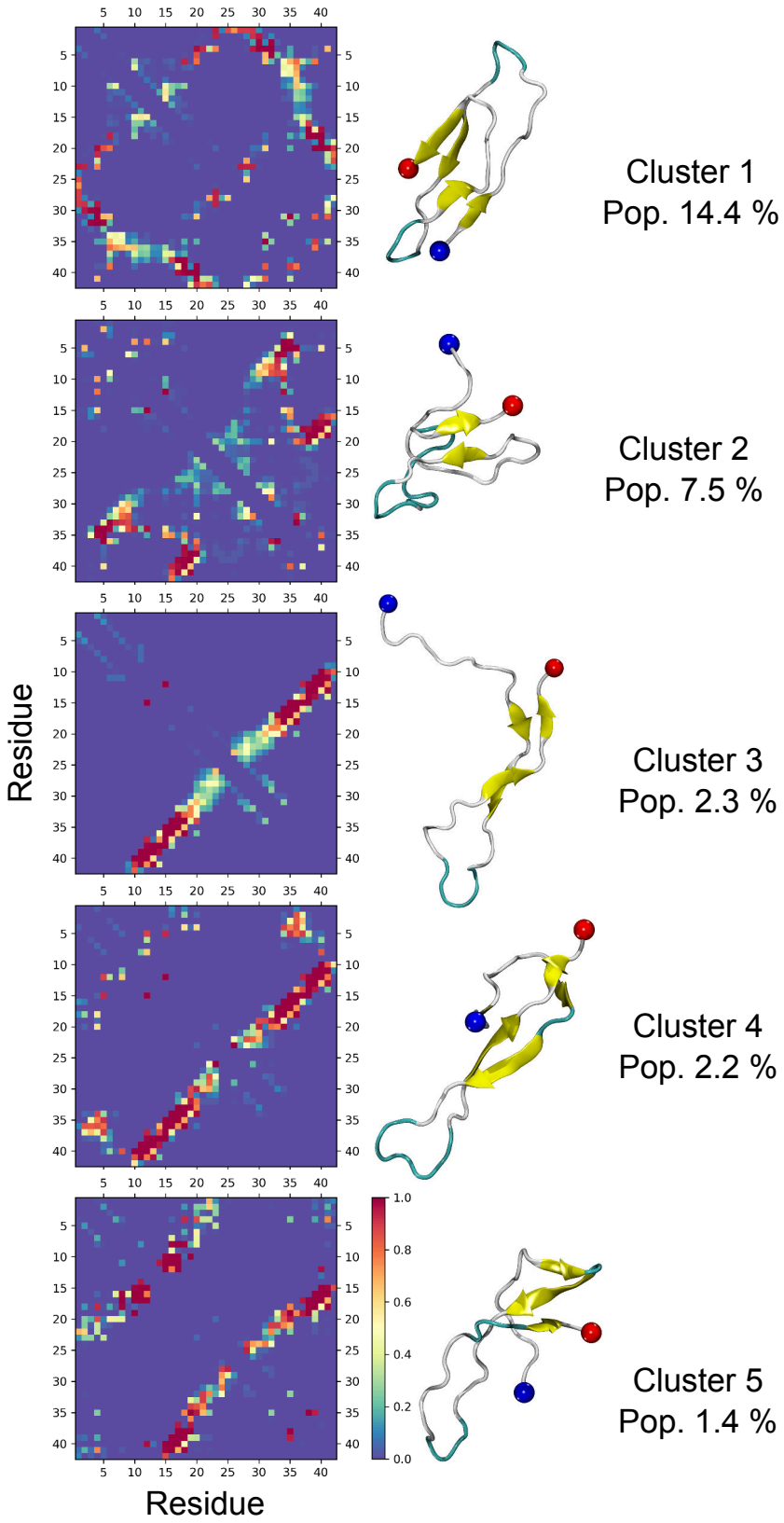

Fig. S2 Contact maps and structures for top five clusters with the highest population.

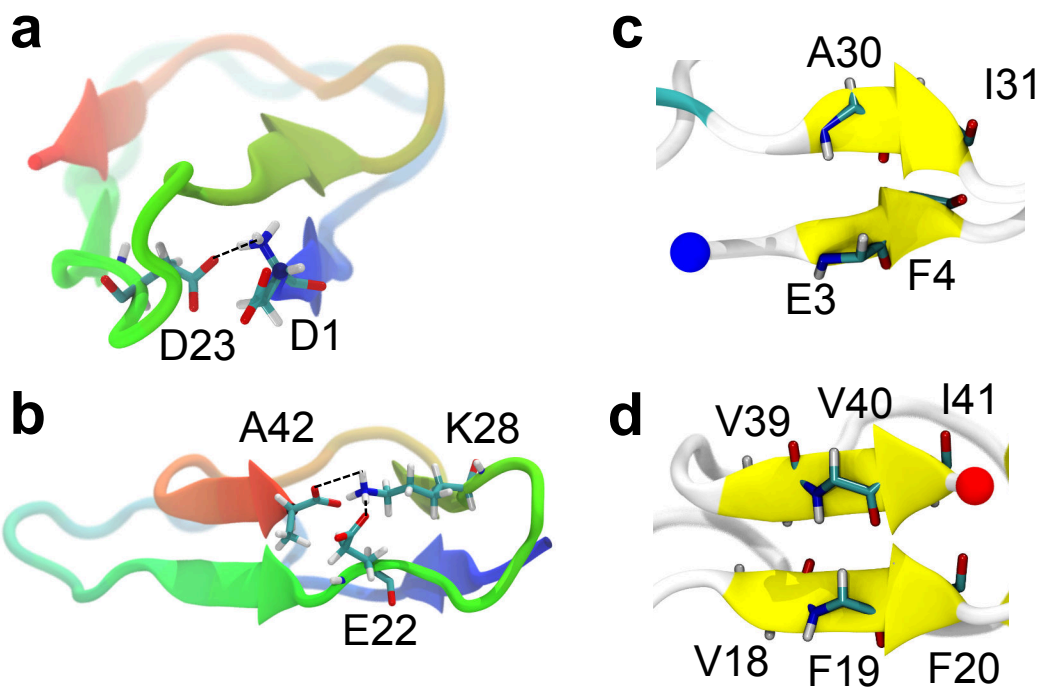

Fig. S3 Salt-bridges and β-sheets of the compact monomer. a) and b) show salt-bridges D1-D23 and K28-E22/A42, respectively. c) and d) show the two parallel β-sheets located at E3-F4 with A30-I31 and V18-F20 with V39-I41, respectively.

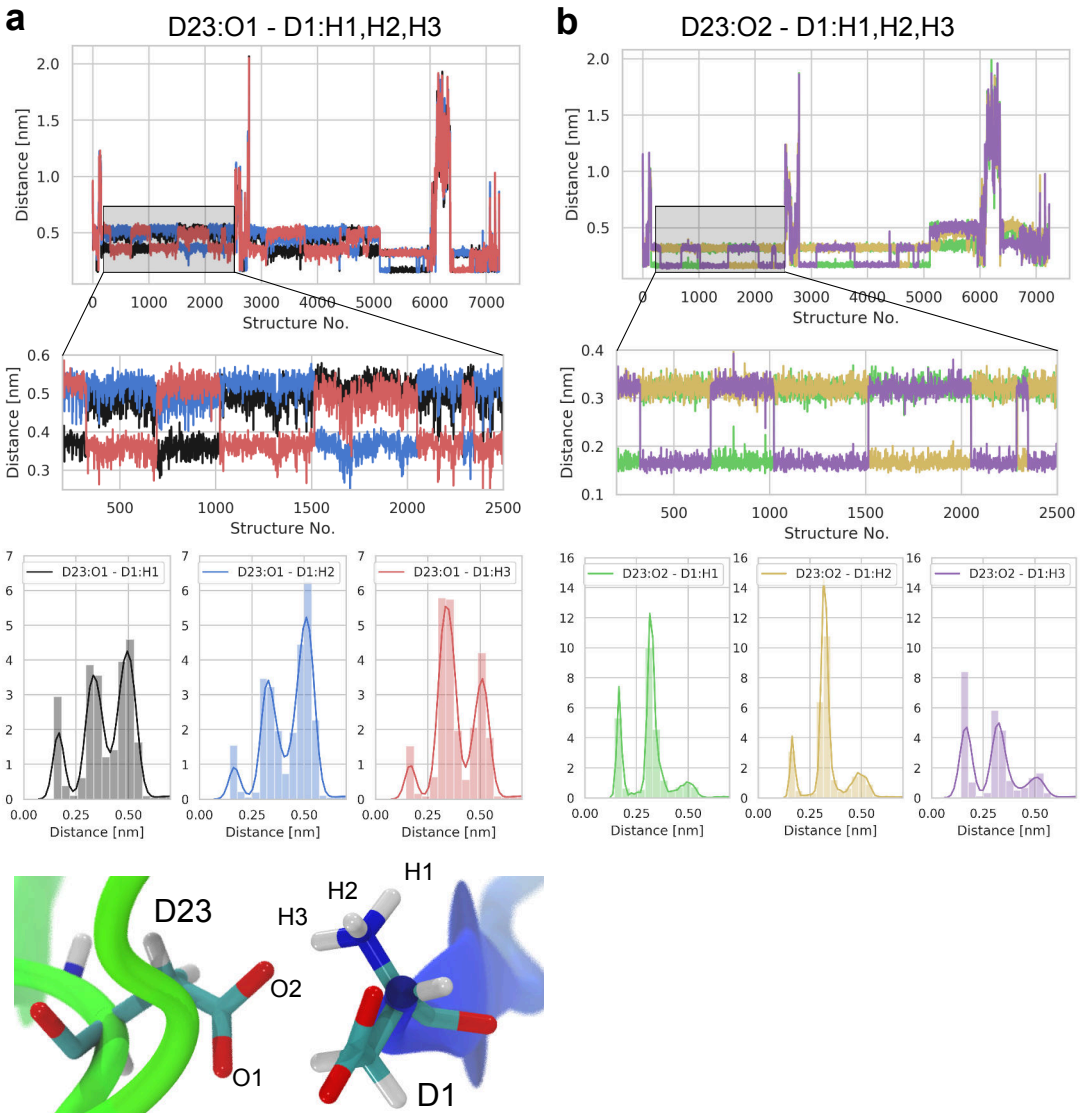

Fig. S4 Salt-bridge distances for D23–D1. a) Top - Distance between the oxygen atom O1, see Bottom snapshot, of group COO- of residue D23 and the hydrogen atoms of group NH3+ of residue D1 for all the conformations in cluster one of the simulation with Charmm36mW and NaCl. Middle - Short interval from the plot in a) which highlights the swaping of hydrogens for the shortest distance with the oxygen atom. Bottom - Normalized distributions for distances between the oxygen atom O1 and the three hydrogens. Bottom snapshot - Licorice representation of the amino acids D23 and D1 and the atoms forming the salt-bridge. b) Same as a) but for distances between oxygen O2 of group COO- of residue D23 and the hydrogen atoms of group NH3+ of residue D1.

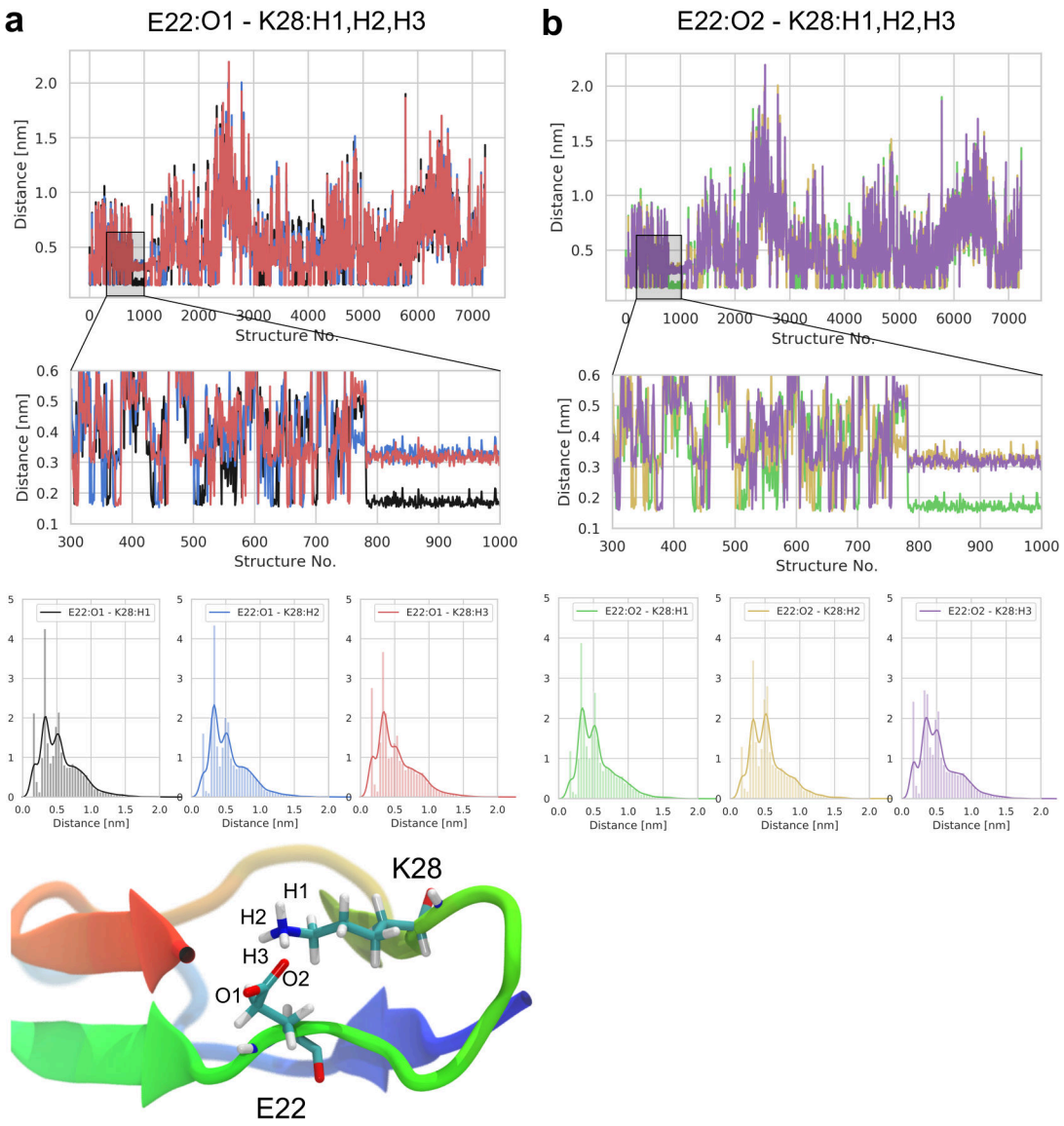

Fig. S5 Salt-bridge distances for E22–K28. a) Top - Distance between the oxygen atom O1, see Bottom snapshot, of group COO- of residue E22 and the hydrogen atoms of group NH3+ of residue K28 for all the conformations in cluster one of the simulation with Charmm36mW and NaCl. Middle - Short interval from the plot in a) which highlights the swaping of hydrogens for the shortest distance with the oxygen atom. Bottom - Normalized distributions for distances between the oxygen atom O1 and the three hydrogens. Bottom snapshot - Licorice representation of the amino acids E22 and K28 and the atoms forming the salt-bridge. b) Same as a) but for distances between oxygen O2 of group COO- of residue E22 and the hydrogen atoms of group NH3+ of residue K28.

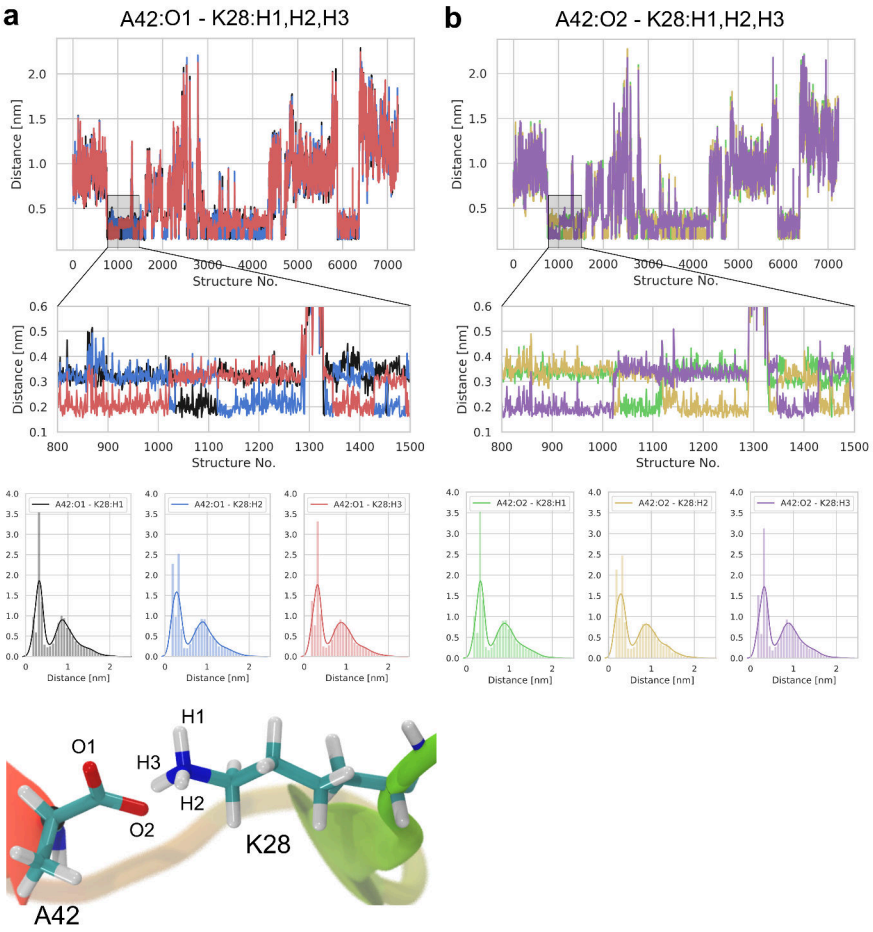

Fig. S6 Salt-bridge distances for A42–K28. a) Top - Distance between the oxygen atom O1, see Bottom snapshot, of group COO- of residue A42 and the hydrogen atoms of group NH3+ of residue K28 for all the conformations in cluster one of the simulation with Charmm36mW and NaCl. Middle - Short interval from the plot in a) which highlights the swaping of hydrogens for the shortest distance with the oxygen atom. Bottom - Normalized distributions for distances between the oxygen atom O1 and the three hydrogens. Bottom snapshot - Licorice representation of the amino acids A42 and K28 and the atoms forming the salt-bridge. b) Same as a) but for distances between oxygen O2 of group COO- of residue A42 and the hydrogen atoms of group NH3+ of residue K28.

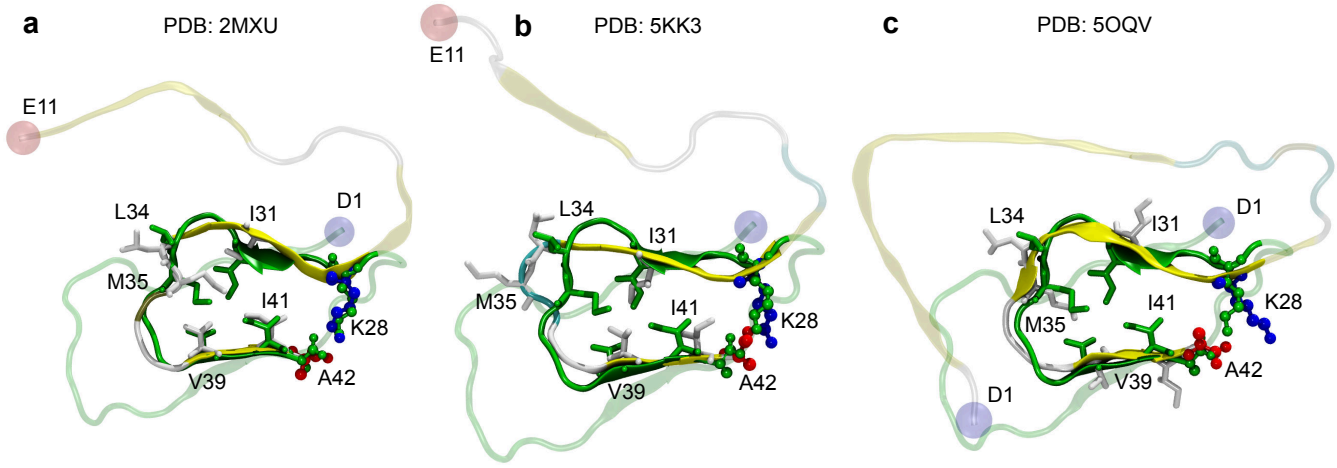

Fig. S7 Comparison of the compact monomer with S-shape fibril models. Only the C-terminus loop was used for the structural alignment. All peptides are shown in cartoon representation and selected amino acids as licorice or balls and sticks. The fibril-like monomer state is colored in green while the fibrillar peptide is colored based on its secondary structure, its hydrophobic amino acids in white, and the charged residues K28 in blue and A42 in red. Three fibril models were considered with PDB IDs: a) 2MXU[30], b) 5KK3[28], and c) 5OQV[29].

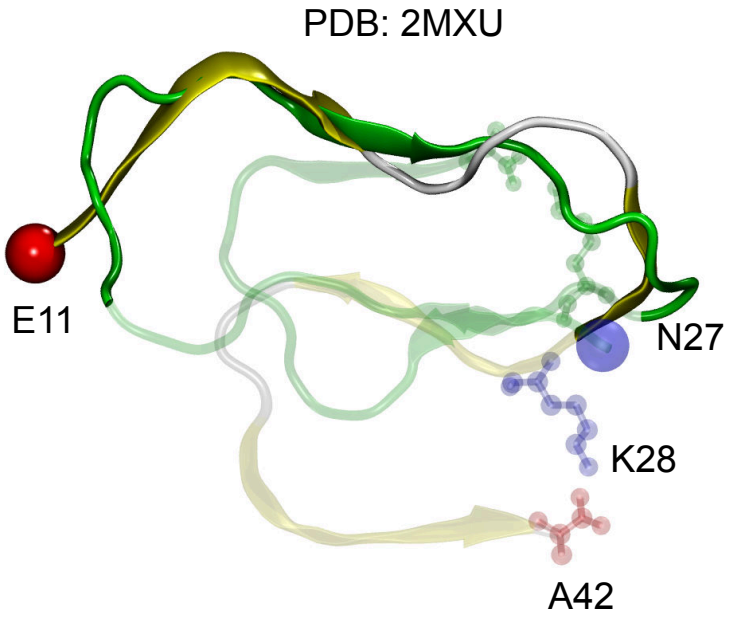

Fig. S8 Structural alignment between the N-terminus loop of the compact Aβ42 monomer and the fibril model with PDB ID 2MXU. Both proteins are shown in cartoon representation with the monomer colored in green and the fibril peptide based on the secondary structure. Region E11-N27 is highlighted, while the rest of the proteins is transparent. The backbone RMSD associated with this region is 0.27 nm.

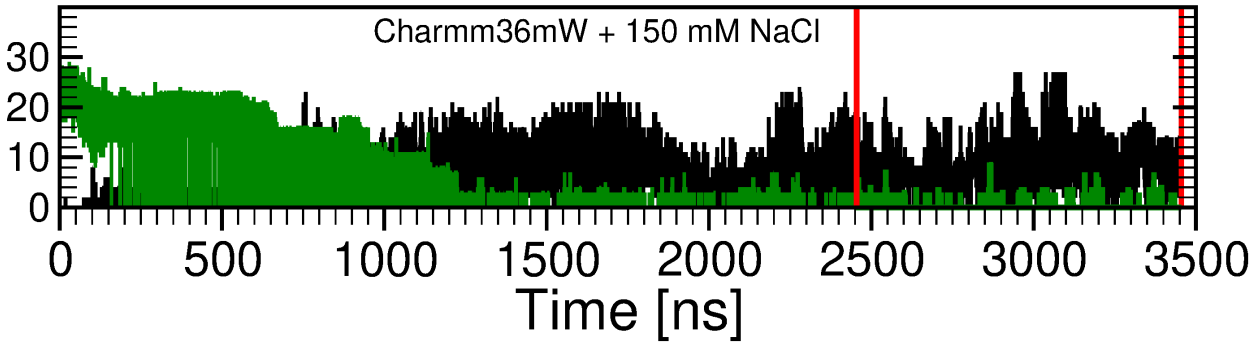

Fig. S9 Evolution of the number of residues adopting sheet (black) or helix (green) conformations during the H-REMD simulation. The vertical red lines indicate the interval considered for detailed analysis.

Table ST1 Backbone RMSD values between the compact monomer and three Aβ42 fibril models. We considered three cases where the RMSD was calculated for the full length peptide, the N-terminus loop and the C-terminus loop. Two of the fibril models (2MXU and 5KK3) lack structural information for the first 10 amino aicds, therefore the full length and the N-terminus loops are truncated accordingly when calculating the RMSD.

Sequence	RMSD [nm]		
	2MXU	5KK3	5OQV
D1(E11)-A42	0.82	0.92	1.14
D1(E11)-N27	0.27	0.42	1.12
K28-A42	0.16	0.21	0.18