Supplementary Material

A Molecular Dynamics Approach to the Structural Characterization of Amyloid Aggregation

M. Cecchini, R. Curcio, M. Pappalardo[†], R. Melki[*], and A. Caflisch[‡]

Department of Biochemistry,

University of Zurich,

Winterthurerstrasse 190,

CH-8057 Zurich, Switzerland

tel: +41 44 635 55 21,

fax: +41 44 635 68 62,

[‡]*:Corresponding author,*

e-mail: caflisch@bioc.unizh.ch

[†]*Dipartimento di Scienze Chimiche,*

Università di Catania,

Viale Andrea Doria 6,

95125 Catania, Italy

[*]*Laboratoire d'Enzymologie et Biochimie Structurales, CNRS,*

Avenue de la Terrasse,

91198 Gif-sur-Yvette, France

(Dated: March 13, 2006)

1. Single-peptide cluster analysis

A single-peptide cluster analysis was performed to investigate the conformations visited by each peptide segment in the perturbation field of its oligomeric system and to determine the statistical relevance of the conformations sampled. The simulation trajectory of the oligomeric system was first split into n single-peptide trajectories, where n is the number of peptide copies in the simulation box. For each snapshot, the Cartesian coordinates of single peptides were extracted and collected according to their identification number, i.e., coordinates of copy 1 were gathered in the first trajectory, coordinates of copy 2 in the second and so on. Single-peptide trajectories were then merged one after the other. The resulting trajectory, which is made of a number of snapshots n times larger than the original trajectory, underwent a cluster analysis based on structural similarity[1]. Single-peptide conformations were grouped according to the C_α–C_β root-mean-square deviation (RMSD) for *each* pairs of structures after optimal superposition[2]. A 1.5 Å RMSD cutoff was used to determine the cluster centers. The clustering algorithm proceeded iteratively until each structure was assigned to a cluster center.

2. Role of condensation in amyloid peptide aggregation

To further investigate the role of the condensation equilibrium in amyloid peptide aggregation, the simulation trajectories of two non aggregation-prone stretches ($A\beta_{25-31}$ and $A\beta_{29-35}$) were analyzed in detail. Unlike the free-energy profiles along $\overline{P_2}$ and R_g, a single-peptide cluster analysis (see above) revealed that the two oligomeric systems have intrinsically different β-aggregation propensities. For both systems, the most populated clusters sampled along the whole trajectory are shown in Fig. S1 (top). For $A\beta_{29-35}$, already the third most populated cluster (4.1%) is the extended conformation which favors β-sheet formation. Moreover, if one considers only the fraction of condensed conformations sampled along the trajectory ($R_g < {R_g}^C$) the statistical weight of the cluster increases substantially and the extended conformation becomes the most populated cluster (6.3%). Hence, $A\beta_{29-35}$ does have a certain β-sheet propensity, which due to the condensation equilibrium remains hidden at 330 K and cannot be detected by the order parameter analysis. For $A\beta_{25-31}$, the situation is different: the cluster producing the extended conformation ranks 19^{th} (statistical

weight of only 0.7%). However, even considering the condensed fraction, its ranking does not improve. Irrespective of the degree of condensation $A\beta_{25-31}$ does not show any β-sheet propensity. Thus, the condensation propensity, though not sufficient to describe amyloidogenicity, is a necessary condition to promote the growth of ordered amyloid nuclei. It is not sufficient because amyloidogenic sequences must also contain β-sheet propensity.

3. Effect of the number of peptides on β-aggregation

To investigate the effect of the dimensionality of the system on the $A\beta_{42}$ β-aggregation profile, the self-assembly process of hexamers was considered. Eighteen MD simulations of six peptides were performed (see Table I) and each trajectory was analyzed by means of the order parameters. At 330 K, β-aggregation propensities ranged between 0.28 and 0.71, thus confirming a strong heterogeneity in aggregation tendencies. A direct comparison with β-aggregation propensities measured from trimeric simulations is however not possible. Although by definition the order parameters do not depend on the number of molecules (n) in the simulation box, when this number is rather small a n-dependent "background" order[3] is detected[4]. β-aggregation propensities obtained from hexameric and trimeric trajectories are therefore not directly comparable. Nonetheless, if one considers only triplets of peptides from each hexameric snapshot, the order information becomes homogeneous and different oligomeric systems can be compared. To guarantee a fair comparison, only triplets of non-isolated peptides, i.e., triplets where each chain forms at least a C_α contact ($C_\alpha - C_\alpha < 5.5$ Å) with another chain, were considered. Hexameric and trimeric amyloidogenicity profiles are shown in Fig. S2 (blue and red circles, respectively). Remarkably, the correlation between the two is quantitative and the grey bars, which show β-aggregation differences ($\Delta\beta$) along the profile, are rather small (below 0.05) almost everywhere. For both systems three *hot-spots* separated by non aggregation-prone regions have been identified. Interestingly, the C-terminal *hot-spot* in the trimeric profile (red spots) was not originally observed (see Fig. 1). Here, the third aggregation-prone region originates from the restriction of the order analysis to non-isolated peptide triplets and gives further evidence of the role played by the condensation equilibrium in amyloid aggregation (see previous section). Free energy profiles along R_g, used to monitor the condensation propensity of the hexameric system, were also measured (data not shown); in agreement with previous findings, the C-terminal stretches

(from $A\beta_{27-33}$ to $A\beta_{35-41}$) showed little condensation propensity.

The structural properties of the aggregates sampled by the MD simulations were also analyzed. By cluster analysis, the trajectories of the stretches located in the *hot-spots* revealed the existence of fully ordered states. The majority of these low-energy states were structurally diverse six-stranded β-sheets characterized by different register alignment, twist and polarity content. Although a statistically relevant preference for in-register parallel arrangements was observed, no hexameric system exhibited a single dominating free energy minimum. This finding is consistent with the very recent discovery that the well-known amyloid-fibril polymorphism results from significant variations in the molecular structure at the protofilament level[5]. The observed "fibril-like" aggregates are fully consistent with the ordered arrangements sampled along the trimeric trajectories and can be thought as their logic extension. The emergence of multiple and competitive low free-energy states suggests, on one hand, that the critical nucleus size has not been reached yet and, on the other hand, that kinetically trapped states might be involved in the structural evolution of ordered oligomers[6].

4. Effect of the segment length on β-aggregation

The robustness of the simulation results with respect to the length of the segments that are identified along the polypeptide sequence was also investigated. The $A\beta_{42}$ primary structure was decomposed into overlapping hendecamers (Table S2) and the trimeric systems of individual segments were studied by implicit solvent MD simulations. The increasing complexity arising from the presence of longer peptide chains resulted in insufficient sampling of the conformational space by constant temperature MD. In fact, a few preliminary runs at 330 K produced trajectories where the system remained trapped in a single free-energy minimum for the whole length of the simulation ($\sim 2\mu$s, data not shown). Thus, replica exchange MD (REMD)[4] was preferred to investigate the aggregation process of 11-residue stretches. Profiles of β-aggregation propensity as a function of temperature are shown in Fig. S3. At elevated temperature, the profiles converge to 0.5, which indicates that no orientational order is present and different stretches are indistinguishable. Close to the physiological temperature, a strong heterogeneity becomes apparent: (i) the stretches centered at residues 14 and 16 show a steep sigmoidal profile which means large β-aggregation propensity over a

wide range of temperatures; (ii) stretches 6, 12, 18 and 20 display a maximum indicating the existence of an equilibrium that disfavors ordered states at low temperature; (iii) stretches 8 and 22 show monotonically growing profiles which resemble the behavior of highly amyloidogenic sequences, though shifted to lower temperatures; and (iv) the remaining eight stretches (10, 24, 26, 28, 30, 32, 34 and 36) do not show any β-aggregation propensity over the range investigated. Whereas both monotonically growing and non-amyloidogenic profiles have been already observed for 7-residue peptides[4], trends characterized by a maximum are detected here for the first time. Secondary structure analysis (data not shown) revealed that the competition between self (intramolecular) and cross (intermolecular) interactions is responsible for such behavior. By lowering the temperature the self interactions (e.g., α-helical and loop conformations) of certain segments are favored at the expenses of the extended conformations and the β-aggregation propensity drops.

The values of β-aggregation propensity from implicit solvent REMD trajectories were used to draw the $A\beta_{42}$ amyloidodgenicity profile at 330 K. However, to compare the results obtained from different segment decompositions (i.e., 7- and 11-residue peptide segments) the profile was re-derived by considering 7-residue substretches on the 11-residue segments. For each hendecapeptide three substretches were identified (e.g., $\mathtt{D}_1\mathtt{AEFRHD}_7$, $\mathtt{E}_3\mathtt{FRHDSG}_9$ and $\mathtt{R}_5\mathtt{HDSGYE}_{11}$ from $\mathtt{D}_1\mathtt{AEFRHDSGYE}_{11}$; $\mathtt{E}_3\mathtt{FRHDSG}_9$, $\mathtt{R}_5\mathtt{HDSGYE}_{11}$ and $\mathtt{D}_7\mathtt{SGYEVH}_{13}$ from $\mathtt{E}_3\mathtt{FRHDSGYEVH}_{13}$; and so on) and the time series of the nematic order parameter was computed for each of them along the trajectory. β-Aggregation profiles obtained from 7- and 11-residue segment simulations are comparable (Fig. S4, top); the trends are very similar and display a maximum in the region 10-22 and two minima at $\mathtt{S}_8\mathtt{G}_9$ and the C-terminus. Due to the errors inherent to the implicit solvation model, the N-terminus is more aggregation-prone than the C-terminus in both cases. Similar conclusions can be drawn from the comparison of the β-sheet structure profiles (Fig. S4, bottom). Again, maxima and minima are similarly located along the $A\beta_{42}$ sequence. The analysis shows that the segment length does not affect the simulation results.

5. Structural interpretation of the $A\beta_{42}$ aggregation profile

The $A\beta_{42}$ β-aggregation profile from simulations of three 11-residue peptides (Table S2) shows a high propensity region encompassing residues 14-22 (Fig. 2, top). Interestingly,

β-aggregation propensities indicate the following ranking $14 \approx 16 \gg 18 > 20 \gg 22$. To structurally characterize the amyloidogenic trend in this central *hot-spot*, a secondary structure analysis of the conformations saved along the trajectories was performed. Secondary structure histograms plotted along the sequence are shown in Fig. S5. A α-helical/β-strand equilibrium, consistent with NMR studies that highlighted a trend to helical structures for the segment 16-24 in aqueous solution[7], accounts for the shape of the profile. On the left edge of the central *hot-spot* (stretch 12), α-helical conformations are dominating and very low β-aggregation propensity is observed. In stretches 14 and 16 the β-strand content increases dramatically, the α-helical propensity vanishes and the sequences are highly amyloidogenic. In stretches 18 and 20 the α/β equilibrium restores and although β-strand conformations are still preferred at this stage, α-helical structures become more stable. Finally, in stretch 22 a turn-like motif ($\mathtt{G}_{25}\mathtt{S}_{26}$) shows up and a second large drop in the profile occurs. Hence, β-aggregation propensity correlates with β-strand content, anticorrelates with α-helical content and seems very sensitive to β-turn or bend propensity. The simulation results indicate that both α-helical/β-strand and β-turn/β-strand equilibria modulate the aggregation properties of this region.

6. Structure stability from explicit water runs

For a subset of four $A\beta_{42}$ 7-residue stretches ($A\beta_{1-7}$, $A\beta_{5-11}$, $A\beta_{13-19}$, and $A\beta_{29-35}$) explicit solvent MD simulations were started from in-register parallel β-sheet conformations observed in implicit solvent runs to evaluate their structural stability. For each peptide system, 50-ns explicit water production runs were performed and the time series of $\overline{P_2}$ were computed along the trajectories. Average values taken at time intervals of increasing length (0.002, 1, 2, 5, 10, 20, 30, and 50 ns) are shown in Fig. S6. The quick drop of the average $\overline{P_2}$ in the case of the $A\beta_{1-7}$ and $A\beta_{5-11}$ segments indicates marginal structural stability of the in-register parallel β-sheet arrangements. On the contrary, segments $A\beta_{13-19}$ and $A\beta_{29-35}$ show rather stable β-sheet conformations in agreement with experimental observations[8–10].

---

[1] P. Ferrara, A. Caflisch, Native topology or specific interactions: What is more important for peptide folding?, J. Mol. Biol. 306 (2001) 837–850.

[2] A. Cavalli, U. Haberthür, E. Paci, A. Caflisch, Fast protein folding on downhill energy landscape, Protein Science 12 (2003) 1801–1803.

[3] T. P. Doerr, D. Herman, H. Mathur, P. L. Taylor, Randomness in nanoscopic liquid-crystal droplets - how small is small?, Europhys. Lett. 59 (3) (2002) 398–402.

[4] M. Cecchini, F. Rao, M. Seeber, A. Caflisch, Replica exchange molecular dynamics simulations of amyloid peptide aggregation, J. Chem. Phys. 121 (21) (2004) 10748–10756.

[5] A. T. Petkova, R. D. Leapman, Z. Guo, Y. W., M. P. Mattson, R. Tycko, Self-propagating, molecular level polymorphism in Alzheimer's β-amyloid fibrils, Science 307 (2005) 262–265.

[6] W. Hwang, Z. Shuguang, R. D. Kamm, M. Karplus, Kinetic control of dimer structure formation in amyloid fibrillogenesis, Proc. Natl. Acad. Sci. USA. 101 (35) (2004) 12916–12921.

[7] R. Riek, P. Güntert, H. Döbeli, B. Wipf, K. Wüthrich, NMR studies in aqueous solution fail to identify significant conformational differences between the monomeric forms of two Alzheimer peptides with widely different plaque-competence, Aβ(1-40)^{ox} and Aβ(1-42)^{ox}, Eur.J.Biochem. 268 (2001) 5930–5936.

[8] L. O. Tjernberg, J. Näslund, F. Lindqvist, J. Johansson, A. R. Karlstrom, J. Thyberg, L. Terenius, C. Nordstedt, Arrest of β-amyloid fibril formation by a pentapeptide ligand., J. Biol. Chem. 271 (1996) 8545–8548.

[9] A. T. Petkova, Y. Ishii, J. J. Balbach, O. N. Antzutkin, R. D. Leapman, F. Delaglio, R. Tycko, A structural model for Alzheimer's β-amyloid fibrils based on experimental constraints from solid state NMR, Proc. Natl. Acad. Sci. USA. 99 (26) (2002) 16742–16747.

[10] R. Liu, C. McAllister, Y. Lyubchenko, M. R. Sierks, Residues 17-20 and 30-35 of β-amyloid play crucial roles in aggregation, J. Neurosci. Res. 75 (2003) 162–171.

TABLE S1: $A\beta_{42}$: 7-Residue stretches simulations of four disease-related and one non-pathological variants

Segment	Variant	Peptide Sequence	Central Residue	3 peptides (μs) 310 K	330 K
$A\beta_{15-21}$	A21G	DAEFRHDSGYEVHH**QKLVFFG**EDVGSNKGAIIGLMVGGVVIA	18	3×1.0	
$A\beta_{17-23}$	A21G	DAEFRHDSGYEVHHQK**LVFFGED**VGSNKGAIIGLMVGGVVIA	20	3×1.0	
$A\beta_{19-25}$	A21G	DAEFRHDSGYEVHHQKLV**FFGEDVG**SNKGAIIGLMVGGVVIA	22	3×1.0	
$A\beta_{21-27}$	A21G	DAEFRHDSGYEVHHQKLVFF**GEDVGSN**KGAIIGLMVGGVVIA	24	3×1.0	
$A\beta_{17-23}$	E22G	DAEFRHDSGYEVHHQK**LVFFAGD**VGSNKGAIIGLMVGGVVIA	20	3×1.0	
$A\beta_{19-25}$	E22G	DAEFRHDSGYEVHHQKLV**FFAGDVG**SNKGAIIGLMVGGVVIA	22	3×1.0	
$A\beta_{21-27}$	E22G	DAEFRHDSGYEVHHQKLVFF**AGDVGSN**KGAIIGLMVGGVVIA	24	3×1.0	
$A\beta_{17-23}$	E22K	DAEFRHDSGYEVHHQK**LVFFAKD**VGSNKGAIIGLMVGGVVIA	20	3×1.0	
$A\beta_{19-25}$	E22K	DAEFRHDSGYEVHHQKLV**FFAKDVG**SNKGAIIGLMVGGVVIA	22	3×1.0	
$A\beta_{21-27}$	E22K	DAEFRHDSGYEVHHQKLVFF**AKDVGSN**KGAIIGLMVGGVVIA	24	3×1.0	
$A\beta_{17-23}$	E22Q	DAEFRHDSGYEVHHQK**LVFFAQD**VGSNKGAIIGLMVGGVVIA	20	3×1.0	
$A\beta_{19-25}$	E22Q	DAEFRHDSGYEVHHQKLV**FFAQDVG**SNKGAIIGLMVGGVVIA	22	3×1.0	
$A\beta_{21-27}$	E22Q	DAEFRHDSGYEVHHQKLVFF**AQDVGSN**KGAIIGLMVGGVVIA	24	3×1.0	
$A\beta_{13-19}$	F19S	DAEFRHDSGYEV**HHQKLVS**FAEDVGSNKGAIIGLMVGGVVIA	16		1×1.9
$A\beta_{15-21}$	F19S	DAEFRHDSGYEVHH**QKLVSFA**EDVGSNKGAIIGLMVGGVVIA	18		1×1.7
$A\beta_{17-23}$	F19S	DAEFRHDSGYEVHHQK**LVSFAED**VGSNKGAIIGLMVGGVVIA	20		1×1.3
$A\beta_{19-25}$	F19S	DAEFRHDSGYEVHHQKLV**SFAEDVG**SNKGAIIGLMVGGVVIA	22		1×1.3

TABLE S2: $A\beta_{42}$: 11-Residue stretches simulations

Segment	Peptide Sequence	Central Residue	3 peptides (μs) CTMD 330 K	3 peptides (μs) REMD 294-413 K
$A\beta_{1-11}$	**DAEFRHDSGYE**VHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA	6	1×1.5	10×1.0
$A\beta_{3-13}$	DA**EFRHDSGYEVH**HQKLVFFAEDVGSNKGAIIGLMVGGVVIA	8	1×1.4	10×1.0
$A\beta_{5-15}$	DAEF**RHDSGYEVHHQ**KLVFFAEDVGSNKGAIIGLMVGGVVIA	10	1×1.3	10×1.0
$A\beta_{7-17}$	DAEFRH**DSGYEVHHQKL**VFFAEDVGSNKGAIIGLMVGGVVIA	12	1×1.5	10×1.0
$A\beta_{9-19}$	DAEFRHDS**GYEVHHQKLVF**FAEDVGSNKGAIIGLMVGGVVIA	14	1×1.5	10×1.0
$A\beta_{11-21}$	DAEFRHDSGY**EVHHQKLVFFA**EDVGSNKGAIIGLMVGGVVIA	16	1×1.5	10×1.0
$A\beta_{13-23}$	DAEFRHDSGYEV**HHQKLVFFAED**VGSNKGAIIGLMVGGVVIA	18	1×1.5	10×1.0
$A\beta_{15-25}$	DAEFRHDSGYEVHH**QKLVFFAEDVG**SNKGAIIGLMVGGVVIA	20	1×1.7	10×1.0
$A\beta_{17-27}$	DAEFRHDSGYEVHHQK**LVFFAEDVGSN**KGAIIGLMVGGVVIA	22	1×1.8	10×1.0
$A\beta_{19-29}$	DAEFRHDSGYEVHHQKLV**FFAEDVGSNKG**AIIGLMVGGVVIA	24	1×1.7	10×1.0
$A\beta_{21-31}$	DAEFRHDSGYEVHHQKLVFF**AEDVGSNKGAI**IGLMVGGVVIA	26	1×1.9	10×1.0
$A\beta_{23-33}$	DAEFRHDSGYEVHHQKLVFFAE**DVGSNKGAIIG**LMVGGVVIA	28	1×2.0	10×1.0
$A\beta_{25-35}$	DAEFRHDSGYEVHHQKLVFFAEDV**GSNKGAIIGLM**VGGVVIA	30	1×2.0	10×1.0
$A\beta_{27-37}$	DAEFRHDSGYEVHHQKLVFFAEDVGS**NKGAIIGLMVG**GVVIA	32	1×2.1	10×1.0
$A\beta_{29-39}$	DAEFRHDSGYEVHHQKLVFFAEDVGSNK**GAIIGLMVGGV**VIA	34	1×3.0	10×1.0
$A\beta_{31-41}$	DAEFRHDSGYEVHHQKLVFFAEDVGSNKGA**IIGLMVGGVVI**A	36	1×2.3	10×1.0

TABLE S3: hIAPP_{1-37}: 7-Residue stretches simulations

Segment	Peptide Sequence	Central Residue	3 peptides (μs) 330 K
hIAPP_{1-7}	**KCNTATC**ATQRLANFLVHSSNNFGAILSSTNVGSNTY	4	1×1.1
hIAPP_{3-9}	KC**NTATCAT**QRLANFLVHSSNNFGAILSSTNVGSNTY	6	1×1.1
hIAPP_{5-11}	KCNT**ATCATQR**LANFLVHSSNNFGAILSSTNVGSNTY	8	1×1.1
hIAPP_{7-13}	KCNTAT**CATQRLA**NFLVHSSNNFGAILSSTNVGSNTY	10	1×1.1
hIAPP_{9-15}	KCNTATCA**TQRLANF**LVHSSNNFGAILSSTNVGSNTY	12	1×1.1
hIAPP_{11-17}	KCNTATCATQ**RLANFLV**HSSNNFGAILSSTNVGSNTY	14	1×1.1
hIAPP_{13-19}	KCNTATCATQRL**ANFLVHS**SNNFGAILSSTNVGSNTY	16	1×1.1
hIAPP_{15-21}	KCNTATCATQRLAN**FLVHSSN**NFGAILSSTNVGSNTY	18	1×1.1
hIAPP_{17-23}	KCNTATCATQRLANFL**VHSSNNF**GAILSSTNVGSNTY	20	1×1.1
hIAPP_{19-25}	KCNTATCATQRLANFLVH**SSNNFGA**ILSSTNVGSNTY	22	1×1.1
hIAPP_{21-27}	KCNTATCATQRLANFLVHSS**NNFGAIL**SSTNVGSNTY	24	1×1.1
hIAPP_{23-29}	KCNTATCATQRLANFLVHSSNN**FGAILSS**TNVGSNTY	26	1×1.1
hIAPP_{25-31}	KCNTATCATQRLANFLVHSSNNFG**AILSSTN**VGSNTY	28	1×1.1
hIAPP_{27-33}	KCNTATCATQRLANFLVHSSNNFGAI**LSSTNVG**SNTY	30	1×1.1
hIAPP_{29-35}	KCNTATCATQRLANFLVHSSNNFGAILS**STNVGSN**TY	32	1×1.1
hIAPP_{31-37}	KCNTATCATQRLANFLVHSSNNFGAILSST**NVGSNTY**	34	1×1.1

TABLE S4: Ure2p$_{20-70}$: 7-Residue stretches simulations

Segment	Peptide Sequence	Central Residue	3 peptides (μs) 330 K
Ure2p$_{20-26}$	**NIGNRNS**NTTTDQSNINFEFSTGVNNNNNNNSSSNNNNVQNNNSGRNGSQN	23	1×1.8
Ure2p$_{22-28}$	NI**GNRNSNT**TTDQSNINFEFSTGVNNNNNNNSSSNNNNVQNNNSGRNGSQN	25	1×1.8
Ure2p$_{24-30}$	NIGN**RNSNTTT**DQSNINFEFSTGVNNNNNNNSSSNNNNVQNNNSGRNGSQN	27	1×1.7
Ure2p$_{26-32}$	NIGNRN**SNTTTDQ**SNINFEFSTGVNNNNNNNSSSNNNNVQNNNSGRNGSQN	29	1×2.0
Ure2p$_{28-34}$	NIGNRNSN**TTTDQSN**INFEFSTGVNNNNNNNSSSNNNNVQNNNSGRNGSQN	31	1×2.1
Ure2p$_{30-36}$	NIGNRNSNTT**TDQSNIN**FEFSTGVNNNNNNNSSSNNNNVQNNNSGRNGSQN	33	1×1.8
Ure2p$_{32-38}$	NIGNRNSNTTTD**QSNINFE**FSTGVNNNNNNNSSSNNNNVQNNNSGRNGSQN	35	1×1.7
Ure2p$_{34-40}$	NIGNRNSNTTTDQS**NINFEFS**TGVNNNNNNNSSSNNNNVQNNNSGRNGSQN	37	1×1.6
Ure2p$_{36-42}$	NIGNRNSNTTTDQSNI**NFEFSTG**VNNNNNNNSSSNNNNVQNNNSGRNGSQN	39	1×1.8
Ure2p$_{38-44}$	NIGNRNSNTTTDQSNINF**EFSTGVN**NNNNNNSSSNNNNVQNNNSGRNGSQN	41	1×2.0
Ure2p$_{40-46}$	NIGNRNSNTTTDQSNINFEF**STGVNNN**NNNNSSSNNNNVQNNNSGRNGSQN	43	1×2.3
Ure2p$_{42-48}$	NIGNRNSNTTTDQSNINFEFST**GVNNNNN**NNSSSNNNNVQNNNSGRNGSQN	45	1×1.8
Ure2p$_{44-50}$	NIGNRNSNTTTDQSNINFEFSTGV**NNNNNNN**SSSNNNNVQNNNSGRNGSQN	47	1×1.6
Ure2p$_{46-52}$	NIGNRNSNTTTDQSNINFEFSTGVNN**NNNNNSS**NNNNVQNNNSGRNGSQN	49	1×1.8
Ure2p$_{48-54}$	NIGNRNSNTTTDQSNINFEFSTGVNNNN**NNNSSSN**NNNVQNNNSGRNGSQN	51	1×2.1
Ure2p$_{50-56}$	NIGNRNSNTTTDQSNINFEFSTGVNNNNNN**NSSSNNN**NVQNNNSGRNGSQN	53	1×2.1
Ure2p$_{52-58}$	NIGNRNSNTTTDQSNINFEFSTGVNNNNNNNS**SSNNNNV**QNNNSGRNGSQN	55	1×1.9
Ure2p$_{54-60}$	NIGNRNSNTTTDQSNINFEFSTGVNNNNNNNSSS**NNNNVQN**NNSGRNGSQN	57	1×1.7
Ure2p$_{56-62}$	NIGNRNSNTTTDQSNINFEFSTGVNNNNNNNSSSNN**NNVQNNN**SGRNGSQN	59	1×1.7
Ure2p$_{58-64}$	NIGNRNSNTTTDQSNINFEFSTGVNNNNNNNSSSNNNN**VQNNNSG**RNGSQN	61	1×1.7
Ure2p$_{60-66}$	NIGNRNSNTTTDQSNINFEFSTGVNNNNNNNSSSNNNNVQ**NNNSGRN**GSQN	63	1×1.1
Ure2p$_{62-68}$	NIGNRNSNTTTDQSNINFEFSTGVNNNNNNNSSSNNNNVQNN**NSGRNGS**QN	65	1×2.1
Ure2p$_{64-70}$	NIGNRNSNTTTDQSNINFEFSTGVNNNNNNNSSSNNNNVQNNNS**GRNGSQN**	67	1×2.0

TABLE S5: Ure2p$_{20-70}$: 7-Residue stretches simulations of single- and double-point N-to-S substitutions at positions 47, 48 and 49.

Segment	Variant	Peptide Sequence	Central Residue	3 peptides (μs) 330 K
Ure2p$_{44-50}$	N47S	NIGNRNSNTTTDQSNINFEFSTGV**NNNSNNN**SSSNNNNVQNNNSGRNGSQN	47	1×1.5
Ure2p$_{44-50}$	N48S	NIGNRNSNTTTDQSNINFEFSTGV**NNNNSNN**SSSNNNNVQNNNSGRNGSQN	47	1×1.5
Ure2p$_{44-50}$	N49S	NIGNRNSNTTTDQSNINFEFSTGV**NNNNNSN**SSSNNNNVQNNNSGRNGSQN	47	1×0.9
Ure2p$_{44-50}$	N4748S	NIGNRNSNTTTDQSNINFEFSTGV**NNNSSNN**SSSNNNNVQNNNSGRNGSQN	47	1×1.0
Ure2p$_{44-50}$	N4749S	NIGNRNSNTTTDQSNINFEFSTGV**NNNSNSN**SSSNNNNVQNNNSGRNGSQN	47	1×1.0
Ure2p$_{44-50}$	N4849S	NIGNRNSNTTTDQSNINFEFSTGV**NNNNSSN**SSSNNNNVQNNNSGRNGSQN	47	1×1.0
Ure2p$_{42-48}$	N4748S	NIGNRNSNTTTDQSNINFEFST**GVNNNSS**NNSSSNNNNVQNNNSGRNGSQN	45	1×1.4
Ure2p$_{44-50}$	N4748S	NIGNRNSNTTTDQSNINFEFSTGV**NNNSSNN**SSSNNNNVQNNNSGRNGSQN	47	1×1.0
Ure2p$_{46-52}$	N4748S	NIGNRNSNTTTDQSNINFEFSTGVNN**NSSNNSS**NNNNVQNNNSGRNGSQN	49	1×1.4
Ure2p$_{48-54}$	N4748S	NIGNRNSNTTTDQSNINFEFSTGVNNNN**SNNSSSN**NNNVQNNNSGRNGSQN	51	1×1.5

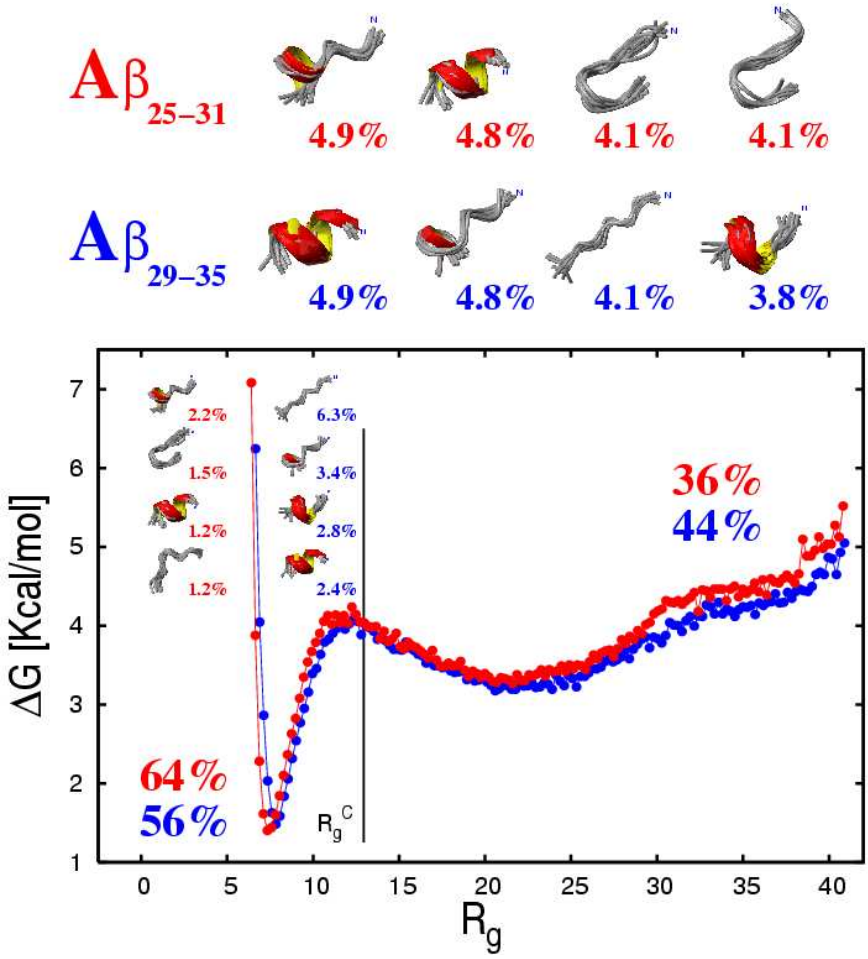

FIG. S1: Analysis of the non aggregation-prone 7-residue stretches $A\beta_{25-31}$ (red) and $A\beta_{29-35}$ (blue). (Top) Single-peptide cluster analysis on the whole trajectory. (Bottom, inset) Cluster analysis for fraction of conformations with $R_g < {R_g}^C$. ${R_g}^C$ corresponds to the lowest radius of gyration detected for conformations where all inter-peptide atomic distances are larger than the long-range interactions cutoffs (7.5 Å in this case). For both fractions (the whole and the condensed), the most populated clusters are represented with their statistical weight. (Bottom, main plot), free-energy projection along the radius of gyration of the trimeric system. Unlike the free-energy profiles on R_g, the single-peptide cluster analysis shows that the two trimeric systems are distinguishable and reveals the role played by the condensation equilibrium in β-aggregation.

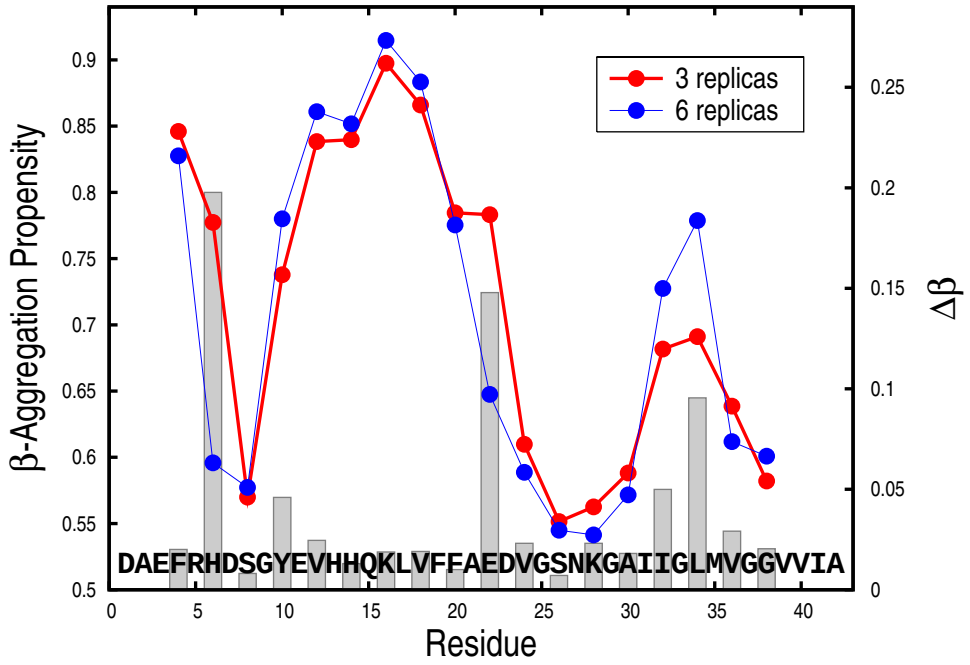

FIG. S2: Analysis of dependence on the size of the simulation system. Values of the β-aggregation propensity from 330 K constant temperature MD simulations of trimeric (red) and hexameric (blue) 7-residue peptide systems. To compare the different oligomeric systems, only triplets where any chain forms at least one C_α contact (C_α – C_α < 5.5 Å) with another chain were considered. β-Aggregation propensity differences ($\Delta\beta = \|\beta^{3rep} - \beta^{6rep}\|$) shown as gray bars with y-axis legend on the right highlight the largest discrepancies.

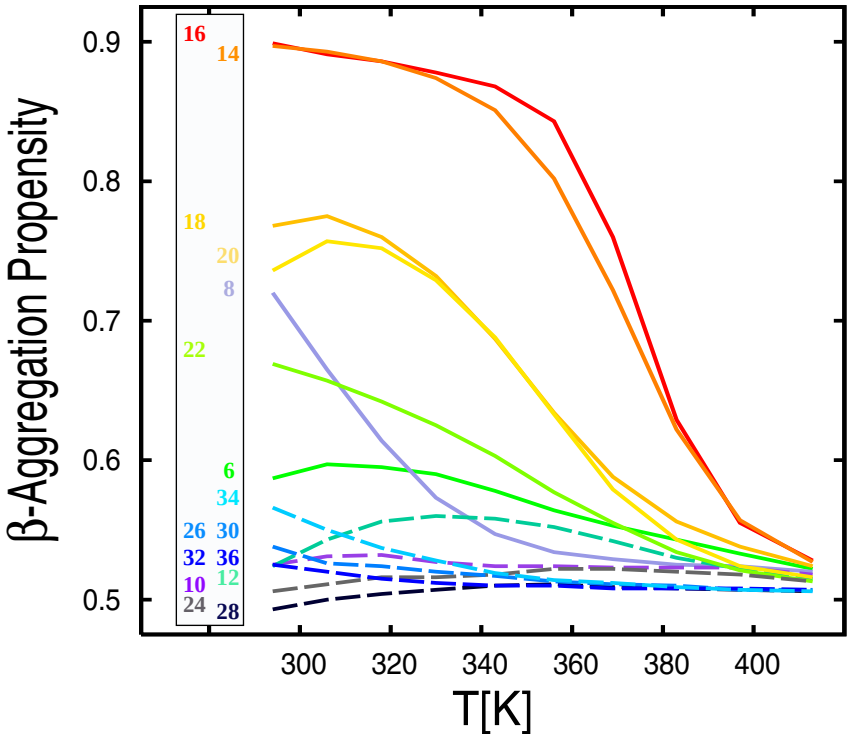

FIG. S3: $A\beta_{42}$: β-Aggregation propensity profiles as a function of temperature from REMD simulations of trimeric systems (11-residue peptide segments, Table S2 in Supplementary Material). Stretches with high β-aggregation propensity at physiological temperature values are highlighted by solid lines. The segment identification numbers, which are reported in the vertical panel on the left, correspond to the position of the central residue in the $A\beta_{42}$ full-length sequence.

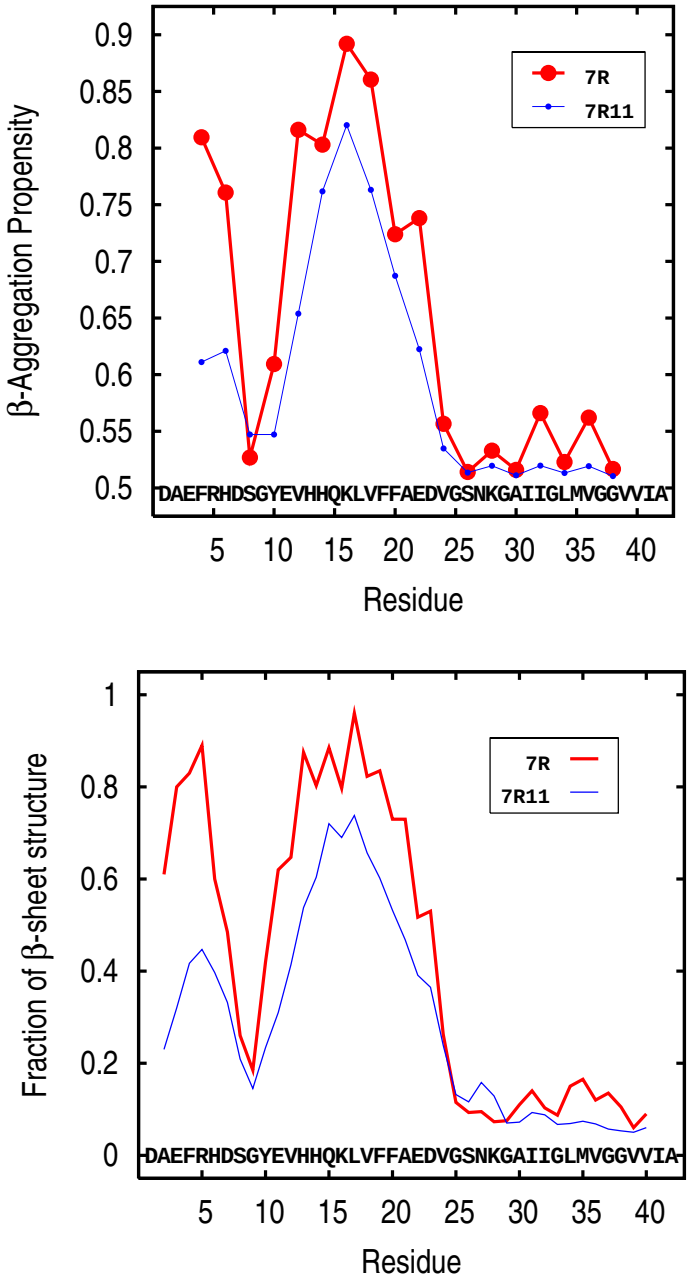

FIG. S4: Analysis of dependence on the length of the segments considered along the $A\beta_{42}$ sequence. β-Aggregation (top) and β-sheet structure (bottom) propensities have been extracted from 330 K MD trajectories of trimeric 7-residue (red) and 11-residue (blue) peptide systems. To allow for a direct comparison, β-aggregation and β-sheet structure propensities extracted from the 11-residue segment simulations have been computed by considering only 7-residue substretches (e.g., $\mathtt{D}_1\mathtt{AEFRHD}_7$, $\mathtt{E}_3\mathtt{FRHDSG}_9$ and $\mathtt{R}_5\mathtt{HDSGYE}_{11}$ from $\mathtt{D}_1\mathtt{AEFRHDSGYE}_{11}$).

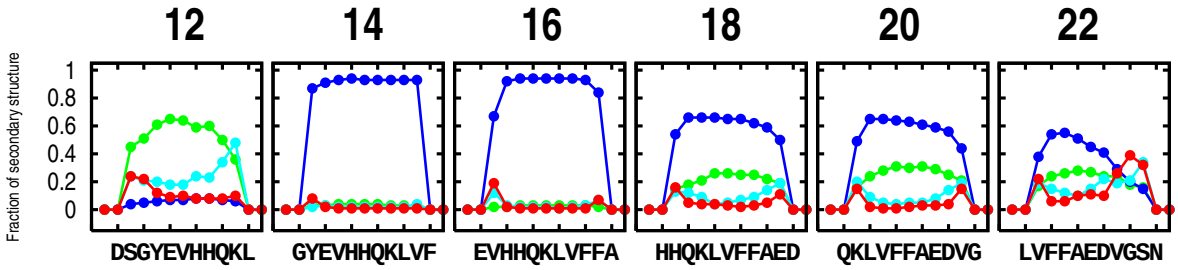

FIG. S5: Secondary structure histograms over REMD trajectory segments at 306 K for six $A\beta_{42}$ 11-residue central stretches. Green, blue, red and cyan dots correspond to α-helical, β-strand, β-turn or bend and random coil content, respectively. The histograms highlight an α-helical/β-strand competition in the central region of the $A\beta_{42}$ sequence.

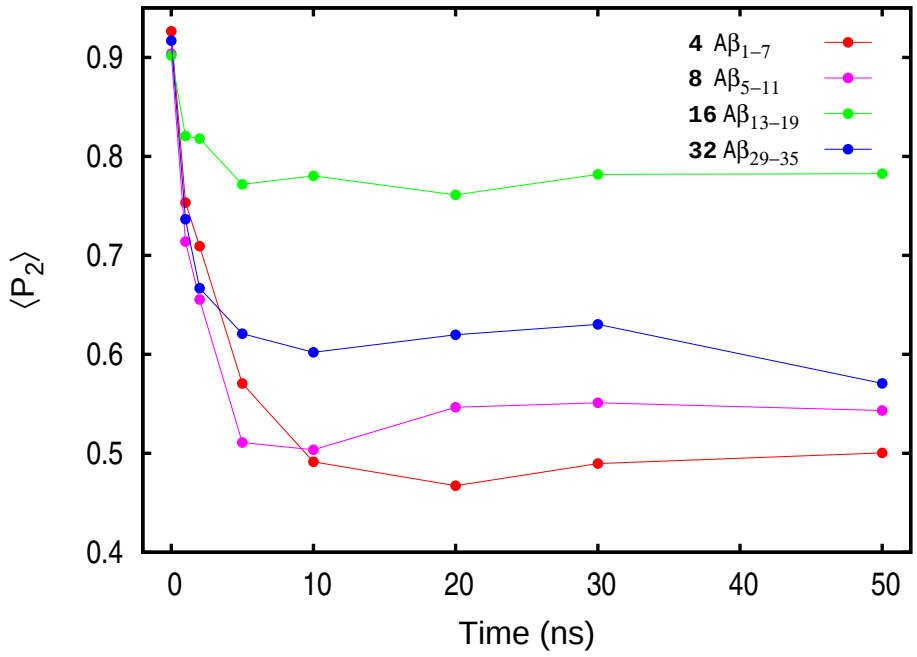

FIG. S6: Average $\overline{P_2}$ values computed from explicit solvent trajectories started from parallel β-sheet conformations observed in implicit solvent runs for four $A\beta_{42}$ 7-residue stretches: $A\beta_{1-7}$ (red), $A\beta_{5-11}$ (magenta), $A\beta_{13-19}$ (green), and $A\beta_{29-35}$ (blue). From left to right, the data points represent average values taken at time intervals of growing length: 0.002, 1, 2, 5, 10, 20, 30, and 50 ns.