

Resolving the atomistic modes of anle138b inhibitory action on peptide oligomer formation.

Dirk Matthes,[†] Vytautas Gapsys,[†] Christian Griesinger,[‡] and Bert L. de Groot^{*,†}

Computational Biomolecular Dynamics Group, Department of Theoretical and Computational Biophysics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany, and Department of Structural Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

E-mail: bgroot@gwdg.de

Phone: ++49-551-2012308. Fax: ++49-551-2012302

*To whom correspondence should be addressed

[†]Computational Biomolecular Dynamics Group

[‡]Department of Structural Biology

Supporting Methods

Simulation setup.

Simulation system and aggregate structure models. Simulations of ten residue long segments from multiple highly amyloidogenic peptide and protein sequences were carried out with anle138b, anle234b and without any small molecule as an internal control (Tables S1 and S2). The N- and C-termini of the peptides were capped with acetyl and N-methyl groups, respectively. For the simulations of 1, 2, 3, 4 and 8 peptide molecules a cubic box with 5.5, 6.94, 7.95, 8.75 nm and 11.0 nm edge length was used, respectively. All systems were solvated with explicit water molecules resulting in a system size of 16k to 132k particles. Counterions (Na^+ , Cl^-) were added to yield an ionic strength of 150 mM and to neutralize the net system charge. The resulting peptide concentration was 10 mM in all cases. Aspartate, glutamate and lysine residues were simulated in their charged states corresponding to the protonation state in the physiological pH range around 7.4. For simulations with the anle138b or anle234b compound, the molecule was placed randomly in the solvent at least 1 nm away from any protein heavy atom. The simulations were initiated from either fully randomized monomeric peptide conformations or pre-assembled β -sheet arrangements to ensure proper sampling of the aggregated state (Figure S2). The crystal structures from the Protein Data Bank (PDB) used to construct the sets of β -sheet models are listed in Table S2.

Small molecule parameterization. Every anle138b and anle234b conformation was geometry optimized at the HF/6-31G* level using the computational chemistry software Gaussian09 (1). Subsequently, each conformer was parameterized for the MD simulations by using the Generalized Amber Force Field (GAFF) (2) to assign the bonded parameters. The restrained electrostatic potential (RESP) fit was performed to ascribe partial charges to the atoms (3). The conversion into the file format compatible with the MD simulation package GROMACS was performed by means of the ACPYPE (4) program. To account for σ -holes, i.e. regions of positive electrostatic potential in compounds containing halogen atoms (Cl, Br, I) (5), an additional virtual particle was constructed on the bromine atom. According to the procedure described in (6), the virtual particle was placed along an extension of the bond between carbon and bromine at the 0.222 nm distance from the bromine atom. The RESP fit was performed on the molecule including the constructed virtual σ -hole, resulting in a small positive partial charge ascribed to the virtual particle. Hydrogen atoms for the compounds were replaced with virtual sites to remove the fastest degrees of freedom, allowing an increased integration timestep to 4 fs.

Analysis protocol.

Decomposing higher-order oligomeric structures into dimeric substructure units.

To compare the substructure of trimeric and tetrameric oligomers with dimers, the following approach of decomposition into pairs of peptide chains was used (7): For each oligomeric structure, all peptide chain pairs that share at least one heavy atom contact within a distance of 0.45 nm were stored. By removing identical and permuted combinations, i.e. molecule pairs AB and BA were treated as equivalent and one of them was discarded, only unique molecule pair combinations were considered. The obtained conformational ensemble of dimeric structures was subjected to an iterative relabeling scheme to assign the same chain identifiers to structurally similar molecule pairs. This relabeling enables a consistent superposition to a reference structure using the variance minimization method (8). Main-chain atoms were used for fitting and for the root-mean-square deviation (RMSD) calculation.

Collective variable definition. Drawing from the body of work by Laio *et al.* (9), the fraction of peptide chains in extended β -strand packing was expressed as a collective variable to comprehensively describe the conformational state of each peptide aggregate structure. A smoothly switched cutoff criterion of $\text{RMSD}_0 = 0.2$ nm was used to define whether or not each individual chain c_i of an aggregate with the total number of N chains was adopting an extended β -strand conformation (1) or not (0). The RMSD_i for chain i was calculated with respect to existing crystallographic data of the fibrillar state for the following seven residue long sequence stretches: hTau40_{305–314}: SVQIVYK; A β _{14–23}: KLVFFAE; hIAPP_{20–29}: NFGAILS; Sub35N_{7–16}: GNNQQNY.

$$\text{CV}_{\beta\text{-strand}} = \frac{\sum_{i=1}^N c_i}{N}; \quad \text{with} \quad c_i = \frac{1 - (\text{RMSD}_i/\text{RMSD}_0)^8}{1 - (\text{RMSD}_i/\text{RMSD}_0)^{12}}.$$

Anle138b contact probability and binding mode classification. To quantify the frequency of short-range interatomic interactions between the anle138b small molecule and the peptide aggregates, the contact search algorithm provided by the g_contacts program (10) with a cutoff distance between heavy atoms of 0.45 nm was used.

Supporting Figures

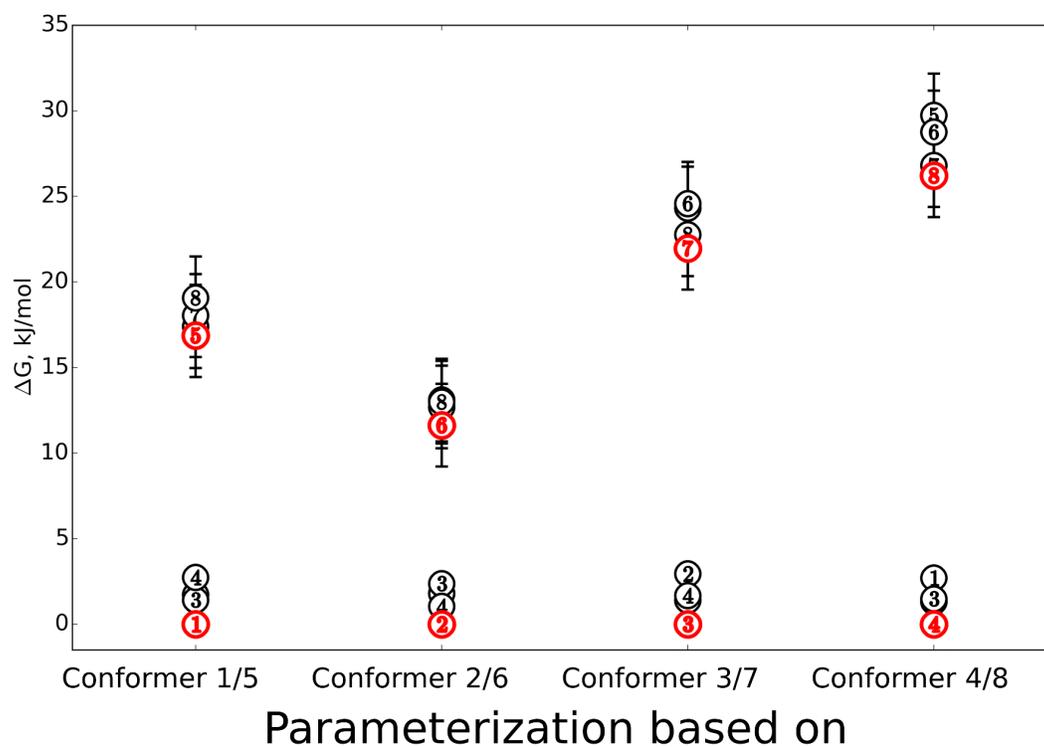


Figure S1: Free energy differences between anle138b conformers and dependence on structure used for parameterization. Free energy differences between the anle138b conformers for each of the two possible protonation patterns obtained by replica exchange simulations.

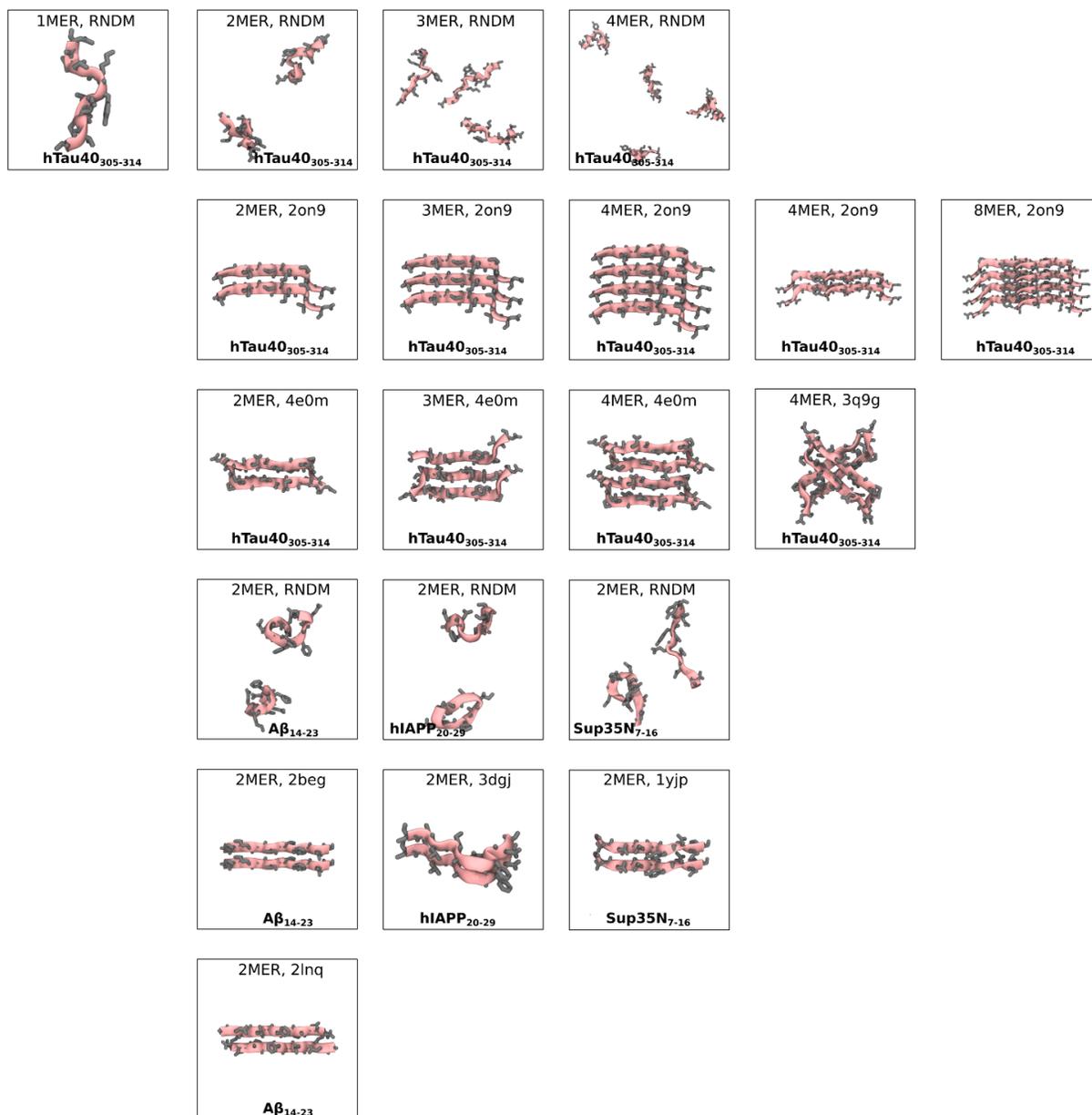


Figure S2: Summary of the starting conformations used for the simulation setup. Overview of the individual starting structures used for the different sets of simulations: Both setups with either random, dispersed peptide conformations or β -sheet model structures are shown for all investigated sequences and system sizes.

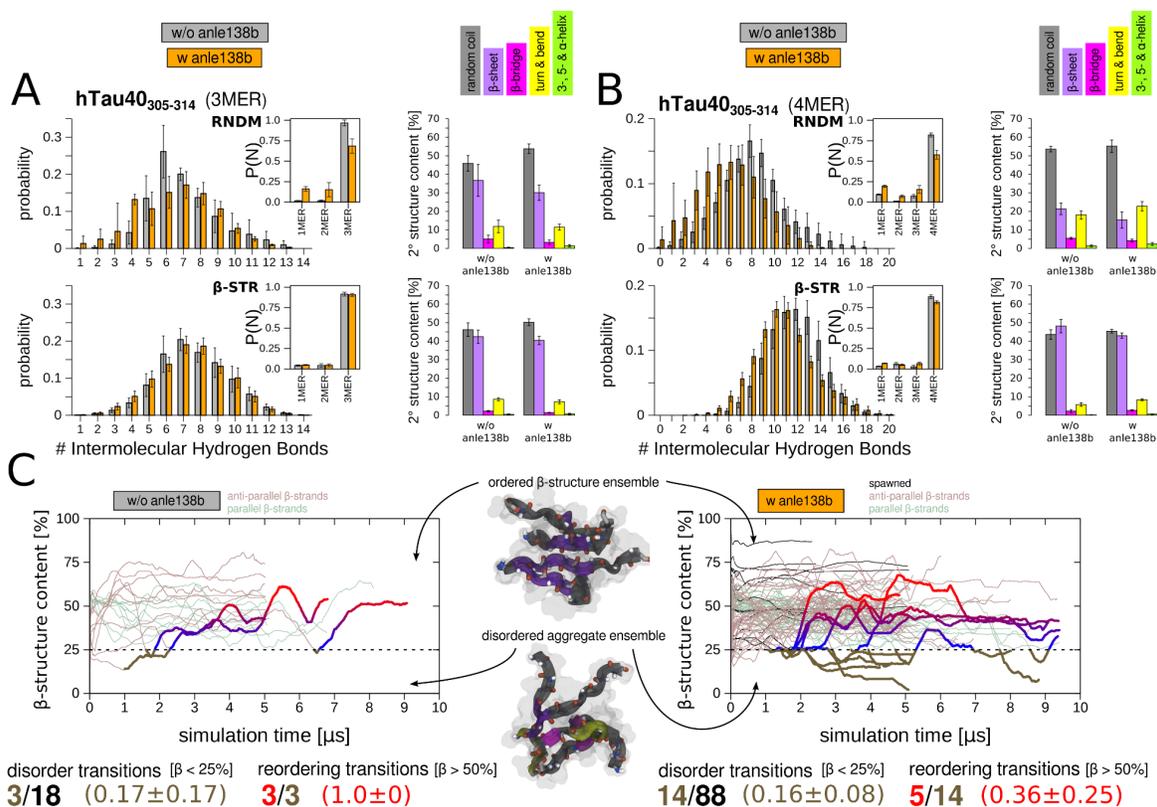


Figure S3: Effect of anle138b on trimeric and tetrameric hTau40 aggregates. Probabilities of interpeptide backbone hydrogen bonds, the sampled aggregate size distributions and the populations of secondary structure elements reported for (A) hTau40 trimer and (B) hTau40 tetramer simulations with and without anle138b: Top panel shows the results for the RNDM setup; bottom panel shows the results for the β -STR setup. Error bars indicate the standard deviation of the individual simulations from the average calculated over all trajectories from simulation subsets with a similar starting configuration. Time dependent β -structure content shown as 500 ns block average for simulations of (E) hTau40 trimers and tetramers started from extended β -sheet structures with and without anle138b. Disorder transitions of aggregates throughout the trajectories are highlighted by thick lines and brown colors, respectively. Red colored lines indicate trajectories with a successful reordering event (β -structure content ≥ 0.5) after an initial disordering.

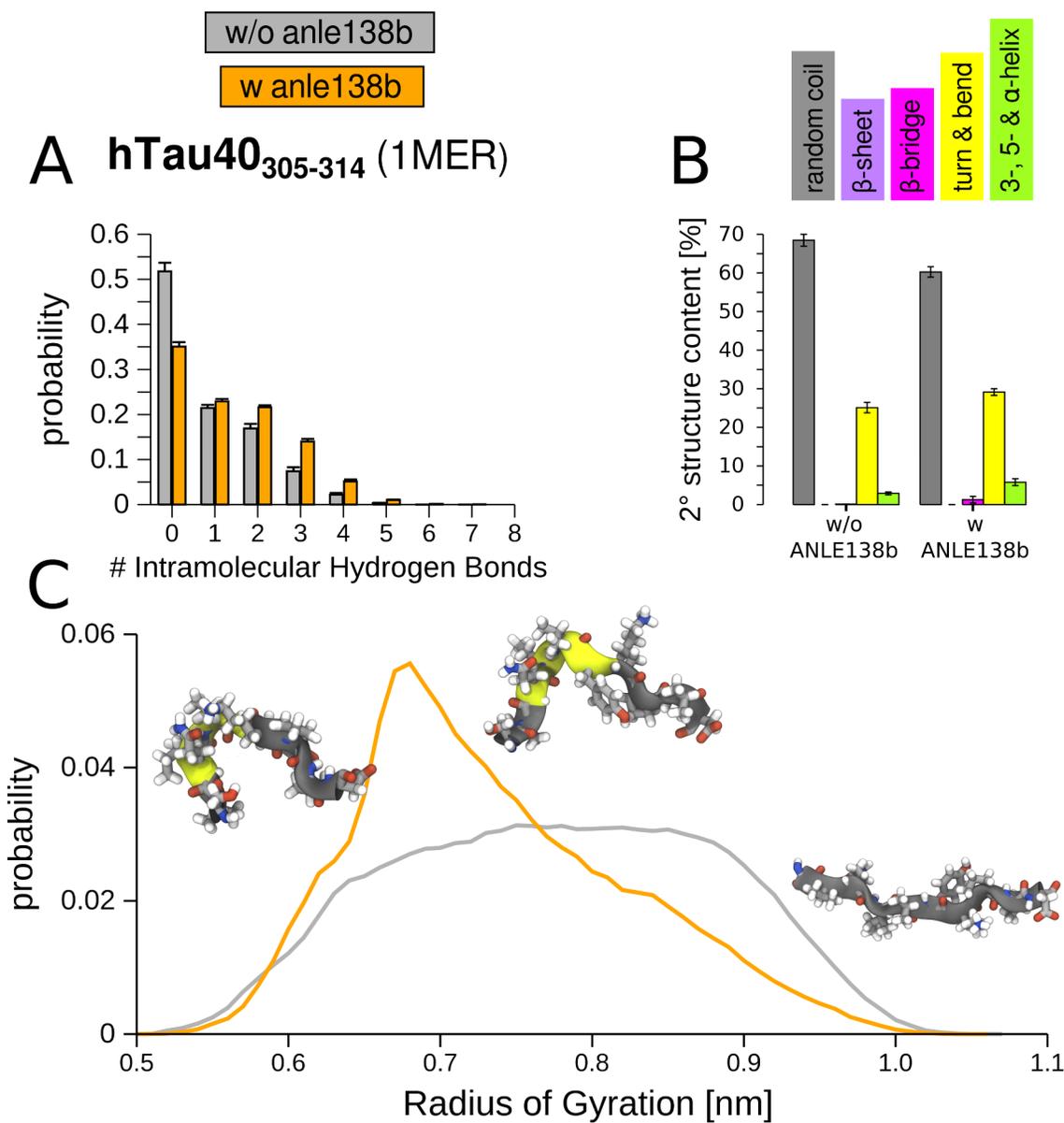


Figure S4: Effect of anle138b on the hTau40 monomer. (A) Probabilities of intrapeptide backbone hydrogen bonds, (B) the populations of secondary structure elements are reported for the hTau40 monomer simulations with and without anle138b. Errorbars indicate the standard deviation of the individual simulations from the average calculated over all trajectories. (C) Population densities of hTau40 monomer conformations with and without anle138b as a function of radius of gyration.

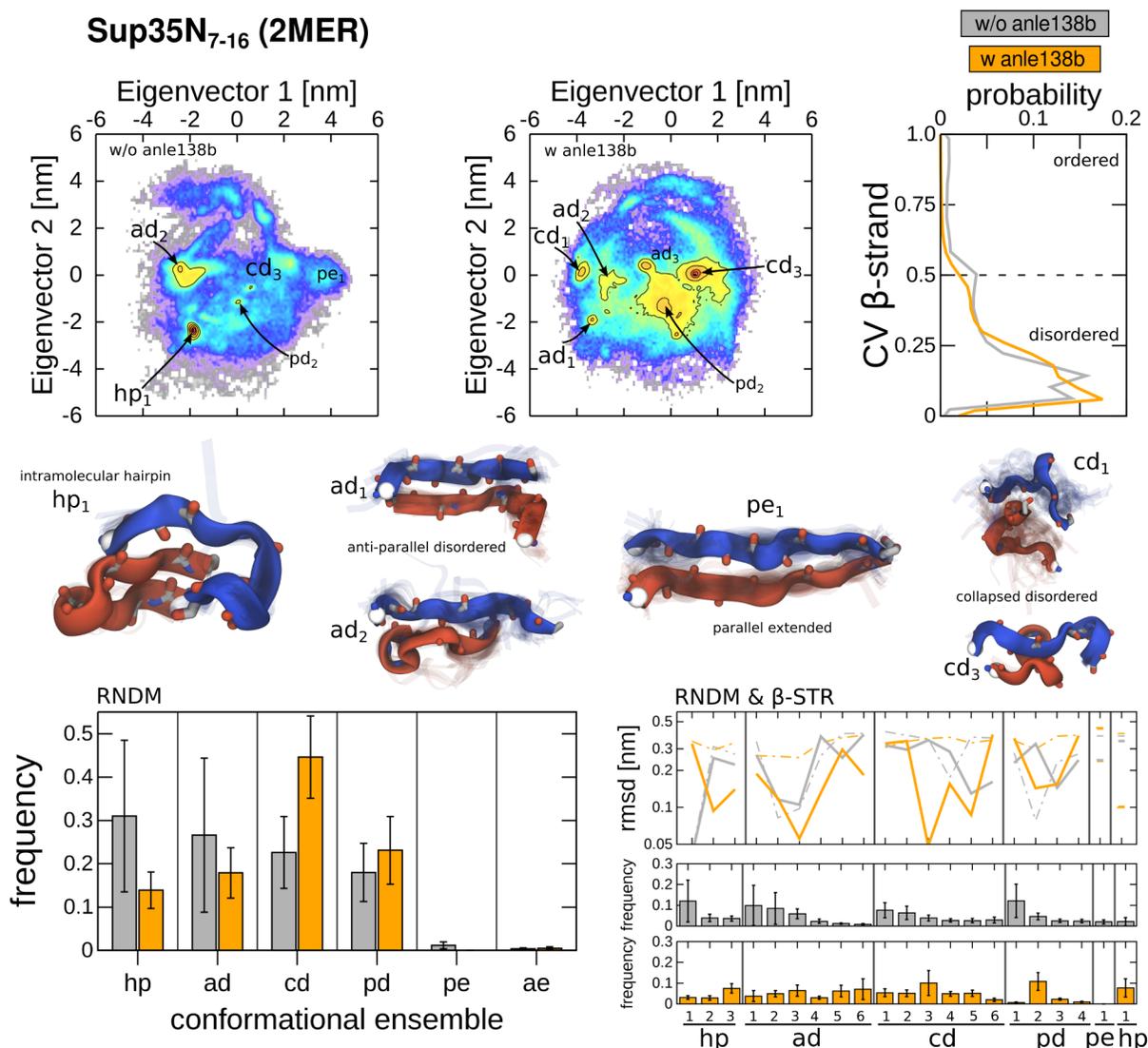


Figure S5: Conformational ensemble of dimeric Sup35N aggregates with and without anle138b. (A) PCA projections of the apo and holo Sup35N dimer simulations starting from random, dispersed structures are shown together with representative cluster center structures as identified by k-means clustering. (B) The fraction of ordered β -strand conformations and (C) population frequency of conformational 'supra-states'. (D) RMSD curves as a function of conformational cluster for two intervals of simulation time (broken lines $t < 2.5 \mu s$, continuous lines $t > 2.5 \mu s$) indicate whether the individual substates were sampled during the simulations.

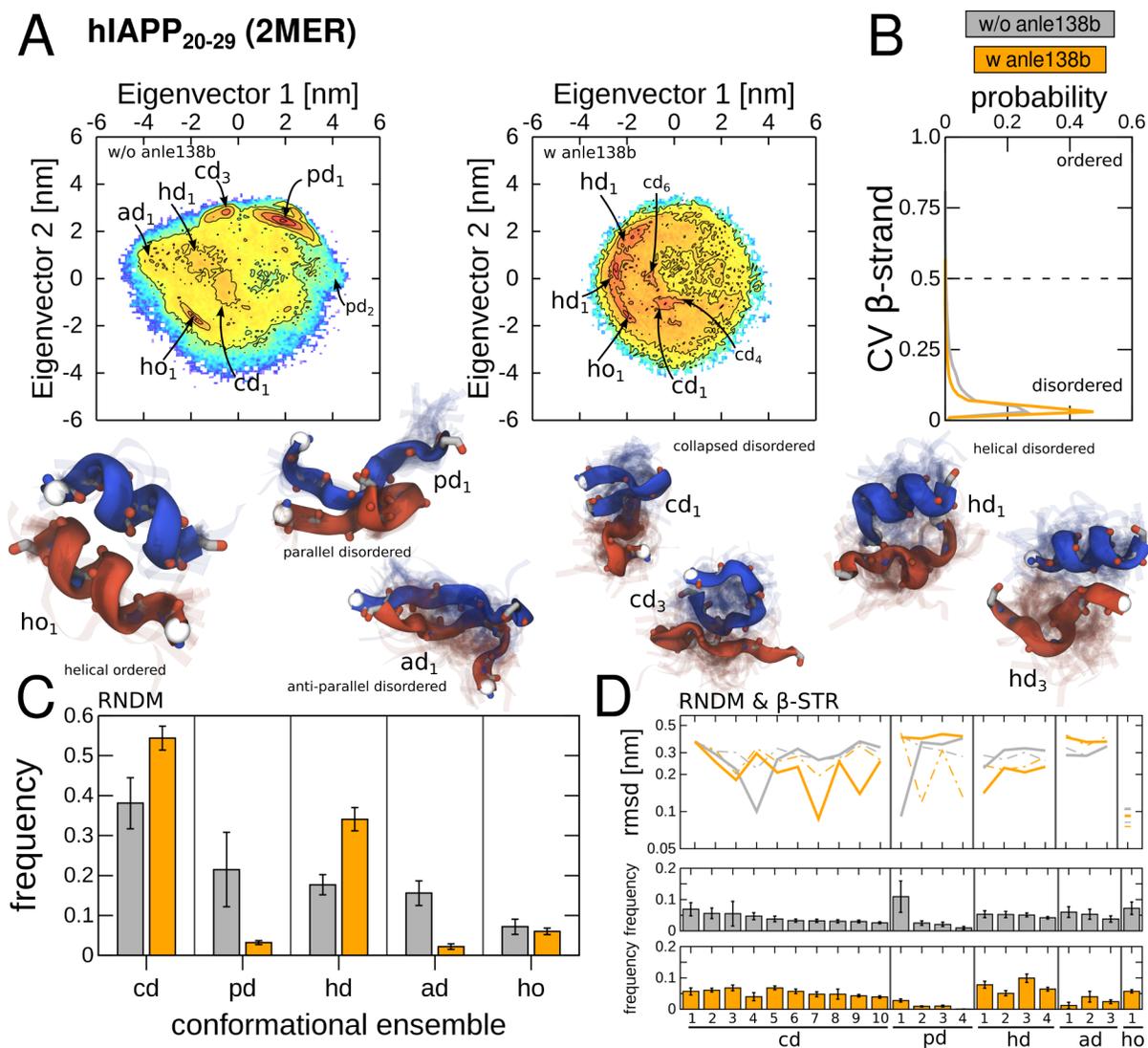
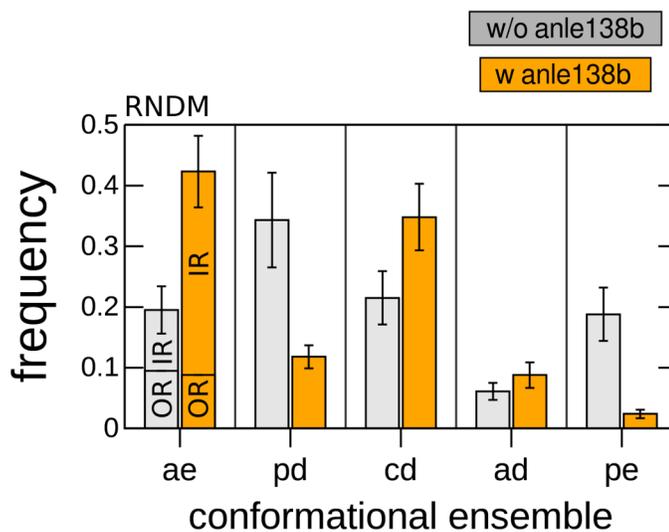
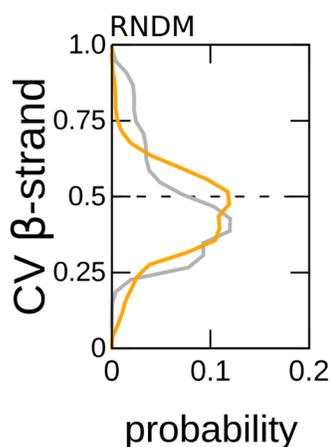


Figure S6: Conformational ensemble of dimeric hIAPP aggregates with and without anle138b. (A) PCA projections of the apo and holo hIAPP dimer simulations starting from random, dispersed structures are shown together with representative cluster center structures as identified by k-means clustering. (B) The fraction of ordered β -strand conformations and (C) population frequency of conformational 'supra-states'. (D) RMSD curves as a function of conformational cluster for two intervals of simulation time (broken lines $t < 2.5 \mu s$, continuous lines $t > 2.5 \mu s$) indicate whether the individual substates were sampled during the simulations.

hTau40₃₀₅₋₃₁₄ (3MER)



hTau40₃₀₅₋₃₁₄ (4MER)

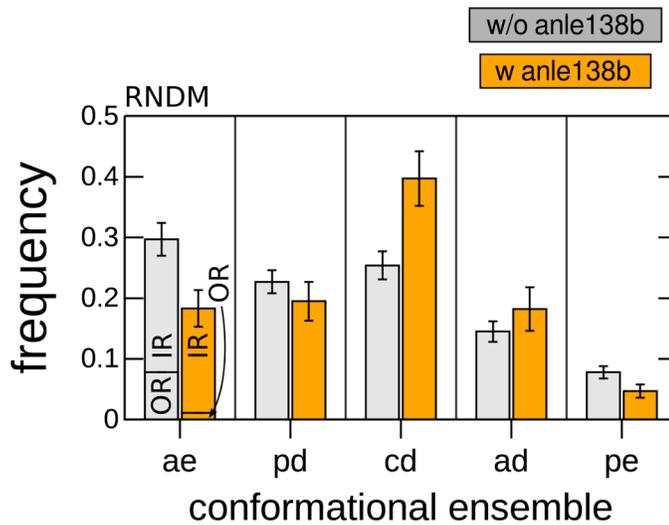
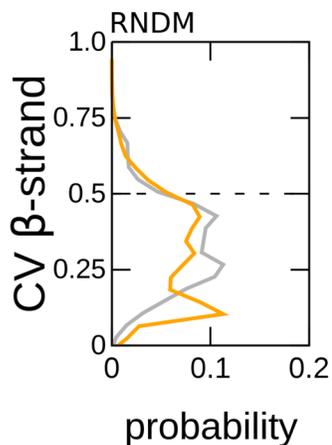


Figure S7: Conformational ensemble trimeric and tetrameric hTau40 aggregates with and without anle138b. Fraction of ordered β -strand conformations are shown for the apo and holo hTau40 trimer and tetramer simulations starting from random, dispersed structures. For the same set of simulations the population frequency of conformational 'supra-states' in hTau40 aggregate trimers and tetramers in terms of dimeric substructure motifs is presented.

As illustrated in Figure S7, an analysis over the individual chain conformations within the larger oligomers, namely the hTau40 trimers and tetramers also shows a higher fraction of disordered versus extended β -strands, complementing the results of the dimeric state. To compare the conformational ensemble of the dimer with the larger oligomers, a previously

established protocol (7) was used, decomposing trimeric and tetrameric aggregates into unique peptide chain pairs (see Supplementary Methods). Similarly to the trends observed for the dimer, an overall higher relative sampling frequency of partially disordered β -strands and completely unstructured peptide conformations are found in the oligomers simulated in the presence of anle138b (Figure S7). Whereas the population of extended anti-parallel β -strands in the holo aggregate ensemble are increased and fractions of pe and ae conformations are more balanced between apo and holo ensembles as compared to the dimeric state

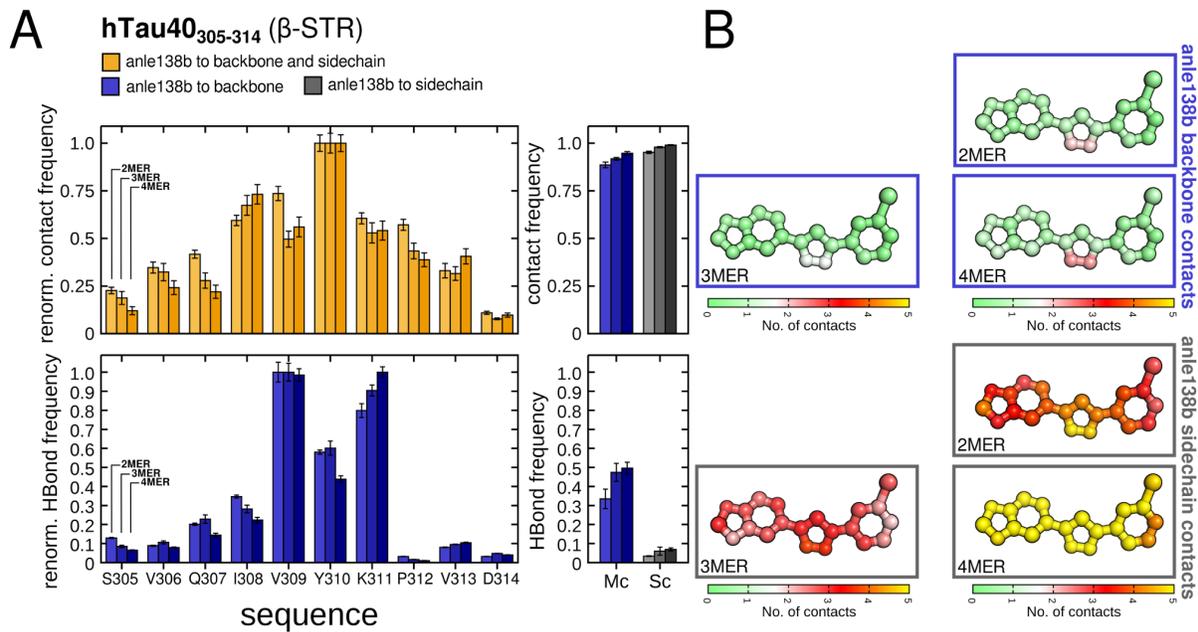


Figure S8: Contact probability and hydrogen bonding interaction of anle138b in hTau40 dimer to tetramer simulations starting from ordered β -sheet conformations. (A) Anle138b contact to peptide heavy atoms and hydrogen bond probability are shown as a function of residue index and aggregate size. (B) Contact frequency of peptide heavy atoms (main and side chain) with anle138b atoms mapped onto chemical structure of the 'trans' conformer.

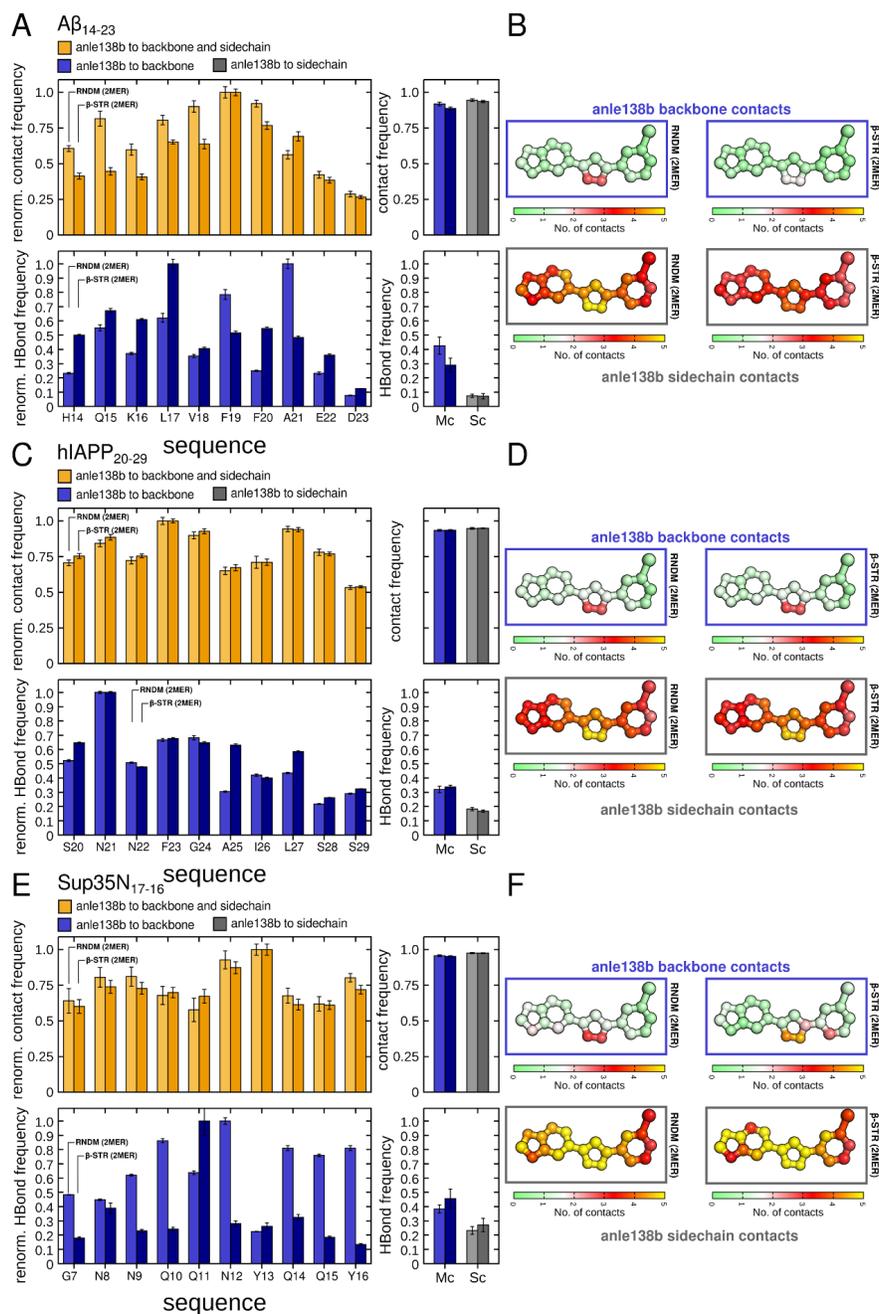


Figure S9: Contact probability and hydrogen bonding interaction of anle138b with $A\beta$, hIAPP and Sup35N dimers. (A) Anle138b contact to peptide heavy atoms and hydrogen bond probability are shown as a function of residue index and simulations setup for $A\beta$, hIAPP and Sup35N dimers. (B) Contact frequency of peptide heavy atoms (main and side chain) with anle138b atoms mapped onto chemical structure of the 'trans' conformer.

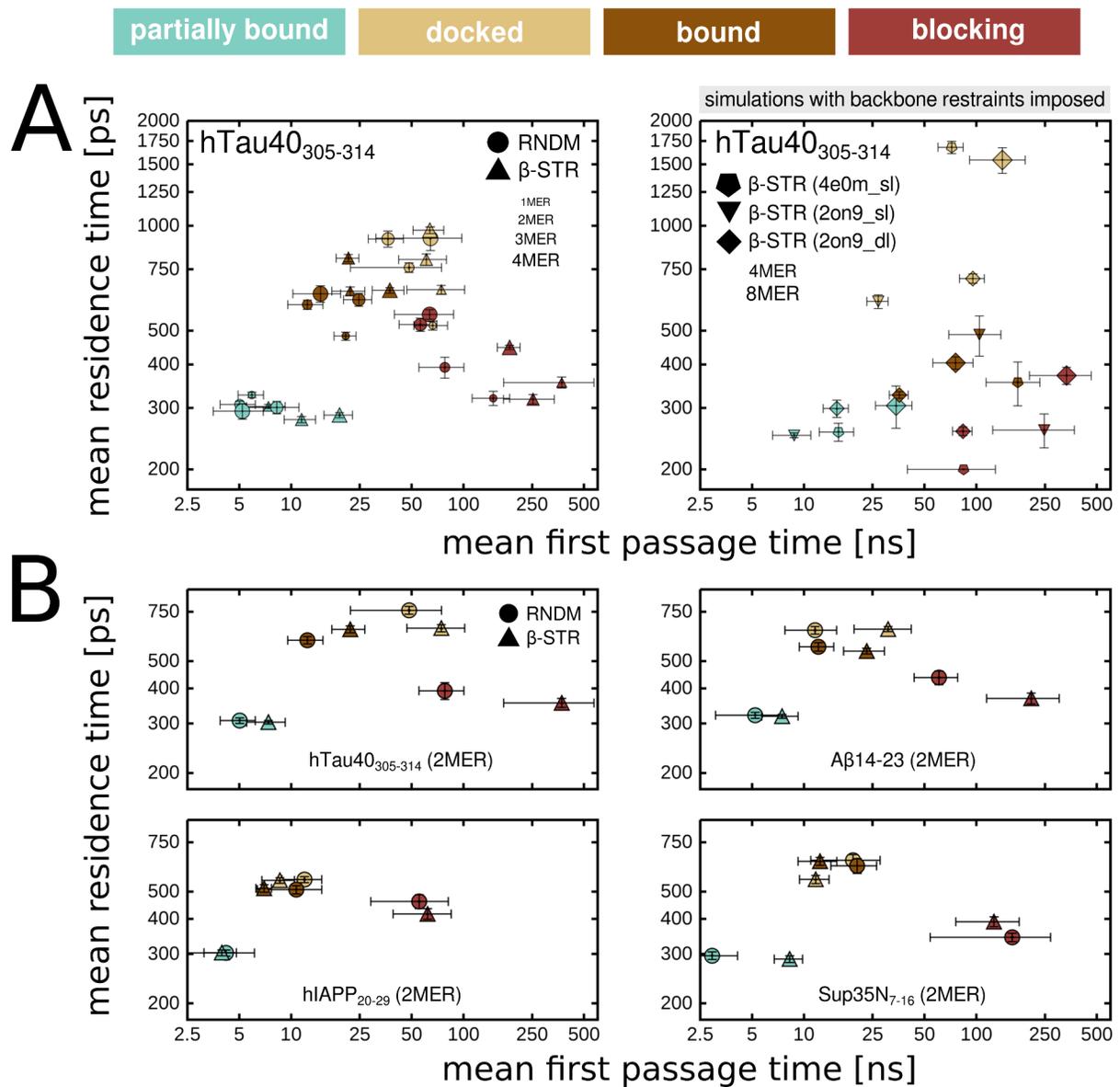


Figure S10: Dynamics of anle138b binding modes. (A) Mean first passage and residence time for the main anle138b binding modes are shown as a function of simulation setup and aggregate size shown over all hTau40 trajectories. (B) Mean first passage and residence time for the main anle138b binding modes are shown as a function of simulation setup over all hTau40, A β , hIAPP and Sup35N dimer trajectories.

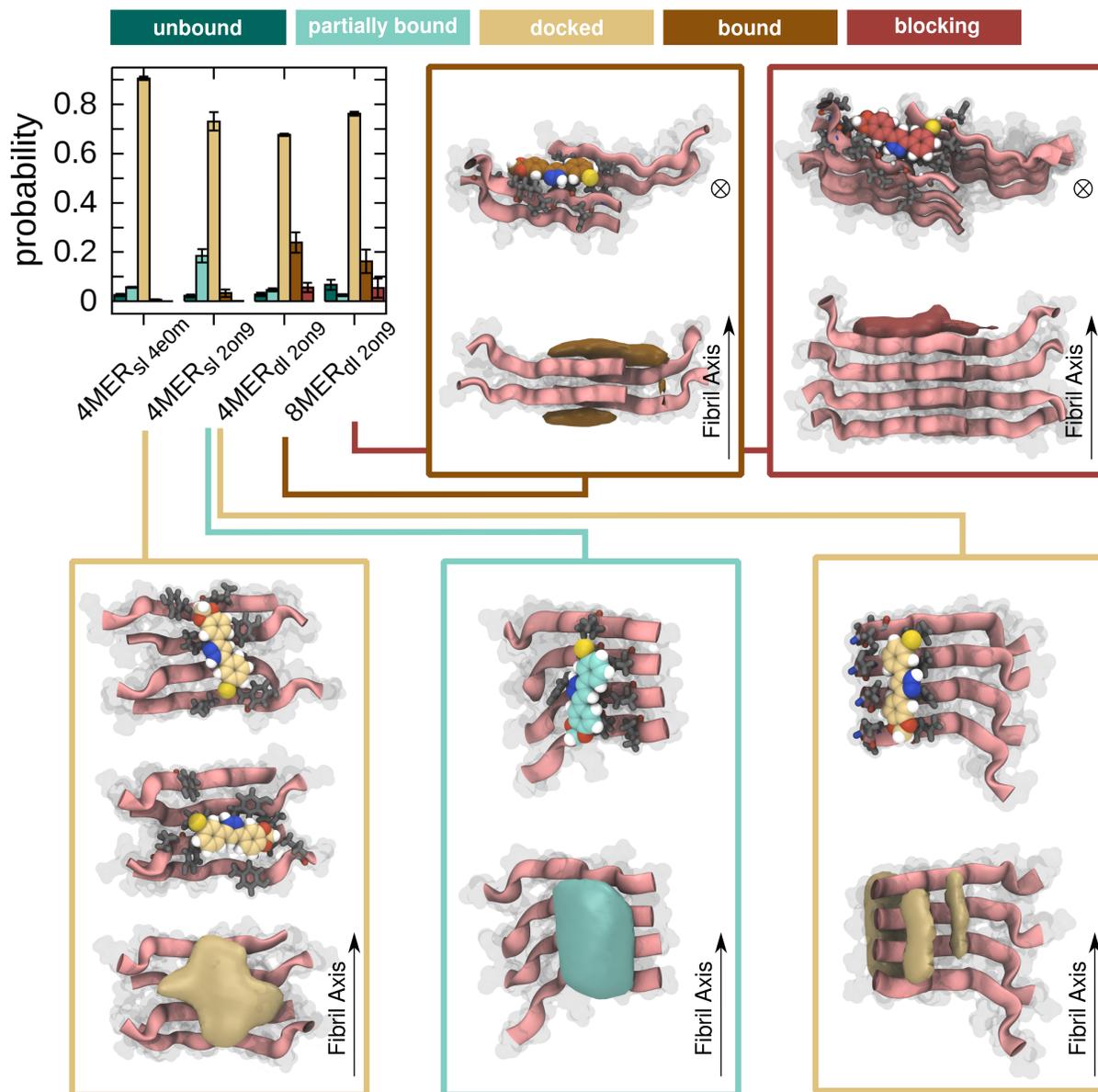


Figure S11: Anle138b binding modes observed in simulations of higher-order oligomers and fibrillar structure templates. Population of binding modes in hTau40 simulations with conformational restraints applied to the protein atoms as a function of simulation system in terms of aggregate size (4MER, 8MER) and β -sheet structure. For each simulation system panels of selected binding modes are shown with representative simulation snapshots (top), together with an iso-contour of the anle138b atomic density at an iso-value of 0.1 (bottom).

Supporting Tables

Table S1: Summary of performed simulations. For each configuration, the total number of independent MD simulation replica with and without anle138b/anle234b is listed including the cumulative length of the trajectories [time in μ s]. The number of peptides and compound molecules is given (N). Simulations with position restraints imposed on the peptide backbone are marked by an asterisk (*).

Configuration		hTau40 ₃₀₅₋₃₁₄		$A\beta_{14-23}$		hIAPP ₂₀₋₂₉		Sub35N ₇₋₁₆	
N_{Peptide}	N_{anle138b}	replica	time	replica	time	replica	time	replica	time
1	0	20	20.0	-	-	-	-	-	-
1	1	30	30.0	-	-	-	-	-	-
2	0	11	63.8	10	73.6	11	68.8	10	56.3
2	1	25	125.3	15	113.4	12	70.2	12	81.5
3	0	11	75.5	-	-	-	-	-	-
3	1	24	193.1	-	-	-	-	-	-
4	0	27	137.7	-	-	-	-	-	-
4	1	70	348.5	-	-	-	-	-	-
4*	1	30	30.0	-	-	-	-	-	-
8*	1	10	10.0	-	-	-	-	-	-
N_{Peptide}	N_{anle234b}	replica	time	replica	time	replica	time	replica	time
2	1	7	38.2	-	-	-	-	-	-

Table S2: Summary of simulated peptide sequences.

Fragment	Polypeptide precursor	Sequence	β -sheet models
hTau40 ₃₀₅₋₃₁₄	human tau protein	SVQIVYKPVD	2ON9 (11) 3Q9G (12) 4E0M (13)
$A\beta_{14-23}$	amyloid- β peptide	HQKLVFFAED	2BEG (14) 2LNQ (15)
hIAPP ₂₀₋₂₉	human islet amyloid polypeptide	SNNFGAILSS	3DGJ (16)
Sup35N ₇₋₁₆	prion-determining region (NM) of Sup35p yeast prion	GNNQQNYQQY	1YJP (11)

References

1. Frisch, M. J. et al. Gaussian09 Revision C.01. 2009; Gaussian Inc. Wallingford CT.
2. Wang, J., Wang, W., A., K. P., and Case, D. A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* *25*, 247–260.
3. Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* *97*, 10269–10280.
4. Sousa da Silva, A. W., and Vranken, W. F. (2012) ACPYPE - AnteChamber PYthon Parser interface. *BMC Res. Notes* *5*.
5. Clark, T., Hennemann, M., Murray, J. S., and Politzer, P. (2007) Halogen bonding: the sigma-hole. *J. Mol. Model* *13*, 291–296.
6. Ibrahim, M. (2011) Molecular mechanical study of halogen bonding in drug discovery. *J. Comput. Chem.* *32*, 2564–2574.
7. Matthes, D., Gapsys, V., Brennecke, J. T., and de Groot, B. L. (2016) An Atomistic View of Amyloidogenic Self-assembly: Structure and Dynamics of Heterogeneous Conformational States in the Pre-nucleation Phase. *Sci. Rep.* *6*.
8. Gapsys, V., and de Groot, B. L. (2013) Optimal Superpositioning of Flexible Molecule Ensembles. *Biophys. J.* *104*, 196–207.
9. Pietrucci, F., and Laio, A. (2009) A Collective Variable for the Efficient Exploration of Protein Beta-Sheet Structures: Application to SH3 and GB1. *J. Chem. Theor. Comput.* *5*, 2197–2201.
10. Blau, C., and Grubmüller, H. (2013) g_contacts: Fast contact search in bio-molecular ensemble data. *Comput. Phys. Commun.* *184*, 2856–2859.

11. Sawaya, M. R., Sambashivan, S., Nelson, R., Ivanova, M. I., Sievers, S. A., Apostol, M. I., Thompson, M. J., Balbirnie, M., Wiltzius, J. J. W., McFarlane, H. T., Madsen, A. O., Riek, C., and Eisenberg, D. (2007) Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* *447*, 453–457.
12. Liu, C., Sawaya, M. R., Cheng, P.-N., Zheng, J., Nowick, J. S., and Eisenberg, D. (2011) Characteristics of Amyloid-Related Oligomers Revealed by Crystal Structures of Macrocyclic beta-Sheet Mimics. *J. Am. Chem. Soc.* *133*, 6736–6744.
13. Liu, C., Zhao, M., Jiang, L., Cheng, P.-N., Park, J., Sawaya, M. R., Pensalfini, A., Gou, D., Berk, A. J., Glabe, C. G., Nowick, J., and Eisenberg, D. (2012) Out-of-register beta-sheets suggest a pathway to toxic amyloid aggregates. *Proc. Natl. Acad. Sci. USA* *109*, 20913–20918.
14. Lührs, T., Ritter, C., Adrian, M., Riek-Loher, D., Bohrmann, B., Dbeli, H., Schubert, D., and Riek, R. (2005) 3D structure of Alzheimer’s amyloid- β (1–42) fibrils. *Proc. Natl. Acad. Sci. USA* *102*, 17342–17347.
15. Qiang, W., Yau, W.-M., Luo, Y., Mattson, M. P., and Tycko, R. (2012) Antiparallel β -sheet architecture in Iowa-mutant β -amyloid fibrils. *Proc. Natl. Acad. Sci. USA* *109*, 4443–4448.
16. Wiltzius, J. J., Sievers, S. A., Sawaya, M. R., Cascio, D., Popov, D., Riek, C., and Eisenberg, D. (2008) Atomic structure of the cross- β spine of islet amyloid polypeptide (amylin). *Prot. Sci.* *17*, 1467–1474.