PNAS
www.pnas.org

Supplementary Information for

Multi-*e*GO: an *in-silico* lens to look into protein aggregation kinetics at atomic resolution

Emanuele Scalone[1], Luca Broggini[1,2], Cristina Visentin[1,2], Davide Erba[1], Fran Bačić Toplek[1], Kaliroi Peqini[3], Sara Pellegrino[3], Stefano Ricagno[1,2,*], Cristina Paissoni[1,*], and Carlo Camilloni[1,*]

[1]Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy

[2]Institute of Molecular and Translational Cardiology, IRCCS Policlinico San Donato, San Donato Milanese, Italy

[3]DISFARM, Dipartimento di Scienze Farmaceutiche, Sezione Chimica Generale e Organica, Università degli Studi di Milano, Milano, Italy

*Stefano Ricagno, Cristina Paissoni, Carlo Camilloni

Email: stefano.ricagno@unimi.it, cristina.paissoni@unimi.it, carlo.camilloni@unimi.it

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Supplementary methods

Simulations details: $TTR_{105-115}$ monomer in explicit solvent

$TTR_{105-115}$ peptide in solution does not possess a unique well-defined structure being too short, consequently we performed a reference simulation using the a99SB-*disp* (1) force field. The starting structure was obtained from pdb 4TLT and the simulation was set to resemble the experimental conditions of $TTR_{105-115}$ aggregation, therefore, to be at pH 2.0 we protonated the N- and C- termini. The molecule was solvated with 3408 water molecules in a dodecahedron box initially 0.9 nm larger than the chain in each direction. One chlorine ion was added to neutralize the charges. Short range interactions were cut-off at 1 nm, with long range electrostatic handled using the particle mesh Ewald scheme (2). Lincs constraints were applied only to bonded hydrogens (3).

Following energy minimization, a positional restraint step was performed under NVT conditions at 300K temperature for 500 ps using the velocity-rescale thermostat (4); then the cell-rescale barostat (5) was used to equilibrate the system in the NPT ensemble to the target pressure of 1 atm for 1000 ps. A production MD simulation under NPT ensemble was run for 1.5 μs. This reference simulation was then analyzed to obtain the monomer native pairs and their corresponding interaction strength. These were defined as all the pairs of atoms, more than one residue apart, being closer than 5.5 Å with a population *P* larger than a threshold $P_{threshold}$, 0.09 in this case.

Simulations details: Multi-eGO $TTR_{105-115}$ fibril

The strength of the native pair interactions ε is the only parameter to be set empirically, this value is then used as it is for all native pairs obtained from the fibril structure, while it is scaled as in equation (1) for native pairs obtained from the simulation of the monomer. To find an optimal value for $TTR_{105-115}$, we performed multiple simulations at fixed temperature using a pre-formed fibril. The fibril model was obtained extracting 64 chains from pdb 2M5K (6). Multiple simulations were performed testing a range of ε values between 0.265 and 0.295 kJ/mol at 315 K using a Langevin dynamics with an inverse friction constant of 50 ps and a timestep of 5 fs. The structure was placed in a cubic box 10 nm larger than the fibril in each direction. After the energy minimization and a positional restraint simulation of 1 ns, production simulations were performed for 200 ns. The optimal ε value was chosen to be the highest at which the structure of the fibril was stable with fibril ends monomers showing some flexibility and ability to partially detach from the fibril. An ε value of 0.275 kJ/mol has then been used for all subsequent simulations.

Simulations details: Multi-eGO $TTR_{105-115}$ aggregation kinetics

To setup the simulations for aggregation kinetics a monomer model was extracted from Y105 to S115 using pdb 4TLT (7) as reference and then, the C-terminal oxygen was added. Four thousand molecules were randomly placed in a cubic box whose side length depends on the desired concentration as $len\,[nm] = 10^{8}\sqrt[3]{\frac{nmol}{N_A[C]}}$, where *nmol* is the number of molecules to add in the system, N_A is the Avogadro Number, and [C] is the concentration. The resulting volumes are in the range of 1 attoliter. The system energy was minimized and then thermalized at a constant temperature of 310 K using positions restraint for 1 ns. Simulations were then performed at 5 different concentration, 13 mM, 11.5 mM, 10 mM, 8.5 mM, and 7 mM in triplicate and evolved long enough to form stable fibrils including most of the monomers. Furthermore, three seeded simulations were performed at 7 mM using a structure of a 10 molecules protofilament obtained at 13 mM as a seed. To analyze the simulations, we used a modified version of the GROMACS *clustsize* tool and homemade python scripts based on MDTraj (8) and MDAnalysis (9, 10). *clustsize* provides a matrix of clusters-sizes as a function of time. The kinetic curves were built by multiplying the number of clusters at every frame by the cluster dimension. Four multi-GO aggregation kinetics were initialized similarly, but using 2000 monomers and concentrations of 2, 1, 0.7 and 0.5 mM, respectively.

$TTR_{105-115}$ peptide synthesis

$TTR_{105-115}$ peptide (YTIAALLSPYS) was prepared by microwave assisted Fmoc solid-phase peptide synthesis on Wang resin (0.4 mmol/g loading) using the CEM Liberty Blue synthesizer. The

coupling reaction was performed with 5 eq excess of the amino acid (0.2 M in DMF), DIC (0.5 M in DMF) and Oxyma Pure (1M in DMF) as coupling reagents. The MW cycle was as following: 15s at 75°C, 170W followed by 110s at 90°C, 40W. N-Fmoc deprotection was performed using 20% piperidine in DMF with a MW cycle of 15s at 75°C, 155 W, followed by 60 s at 90°C, 50 W. Full-cleavage from the resin was performed by shaking the resin for 3 hours in a mixture of TFA/TIPS/H2O/phenol 90:2.5:2.5:5. After cleavage, the peptide was precipitated and washed using ice-cold diethyl ether. $TTR_{105-115}$ was purified in reverse-phase by using RP-HPLC with a ADAMAS C-18 column from Sepachrom (10 µm, 250 × 21.2 mm, phase A 0.1% TFA in water, phase B: 1% TFA in ACN with a gradient of 20-100% phase B over 40 min at a flow rate of 20 mL/min.).

$TTR_{105-115}$ aggregation assays

Lyophilized $TTR_{105-115}$ peptide was dissolved at 13 mM in 10% acetonitrile/water solution and pH adjusted to 2.5 with HCl. The solution was sonicated on ice for 15 min and then centrifugated at 4°C for 15 min at 20800 *x g*. The stock solution was eventually diluted to the final concentrations for the experiments (13 mM, 10 mM, and 7 mM) in 10% acetonitrile/water solution. Freshly prepared ThT was added to a final concentration of 20 µM. 50 µL of each condition was then pipetted into black polystyrene 96-well half-area plates with clear bottoms and polyethylene glycol coating (Corning). Each condition was performed in duplicate in each experiment. Plates were sealed to prevent evaporation and incubated at 37 °C under quiescent conditions in a Varioskan Lux plate reader (Thermo Fisher Scientific). Upon excitation at 450 nm, fluorescence at 480 nm was recorded through the bottom of the plate every 120 min. All the experiments were performed in triplicates, except for the 10 mM condition that was done in duplicate. The mean ThT fluorescence values from the independent experiments were normalized and subjected to nonlinear regression analysis, using Boltzmann sigmoidal equation.

TEM analysis on $TTR_{105-115}$ fibrils

Freshly prepared $TTR_{105-115}$ fibrils were analyzed by TEM. 4-µl droplet of sample was applied onto a 400-mesh copper carbon-coated grids (Agar Scientific) glow discharged for 30 s at 30 mA using a GloQube system (Quorum Technologies). After 1-min incubation, excess of sample was removed and the grid was stained with 2% (wt/v) uranyl acetate solution, blotted dry, and imaged on a Talos L120C transmission electron microscope (Thermo Fisher Scientific) operating at 120 kV. Morphological characterization of the fibrils was performed using the software ImageJ (11).

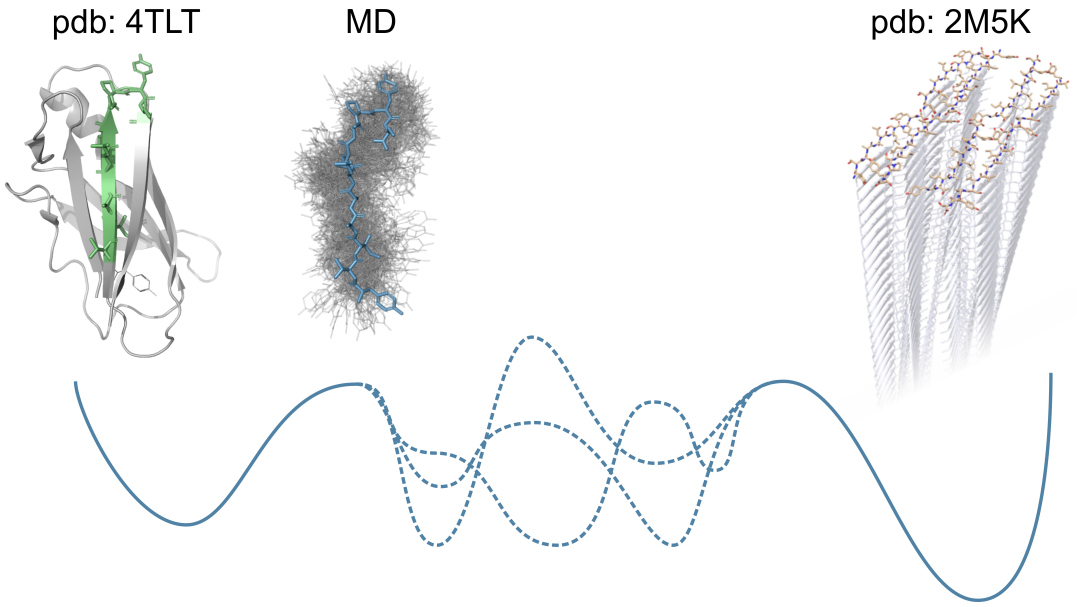


Figure S1. Schematic representation of multi-GO and multi-*e*GO models. Multi-GO model is based completely on two structures, one representing the native structure and one the fibril structure. Multi-*e*GO model instead includes information about the conformational dynamics of the native state and the fibril structure, furthermore multi-*e*GO local geometry is described using a transferable potential.

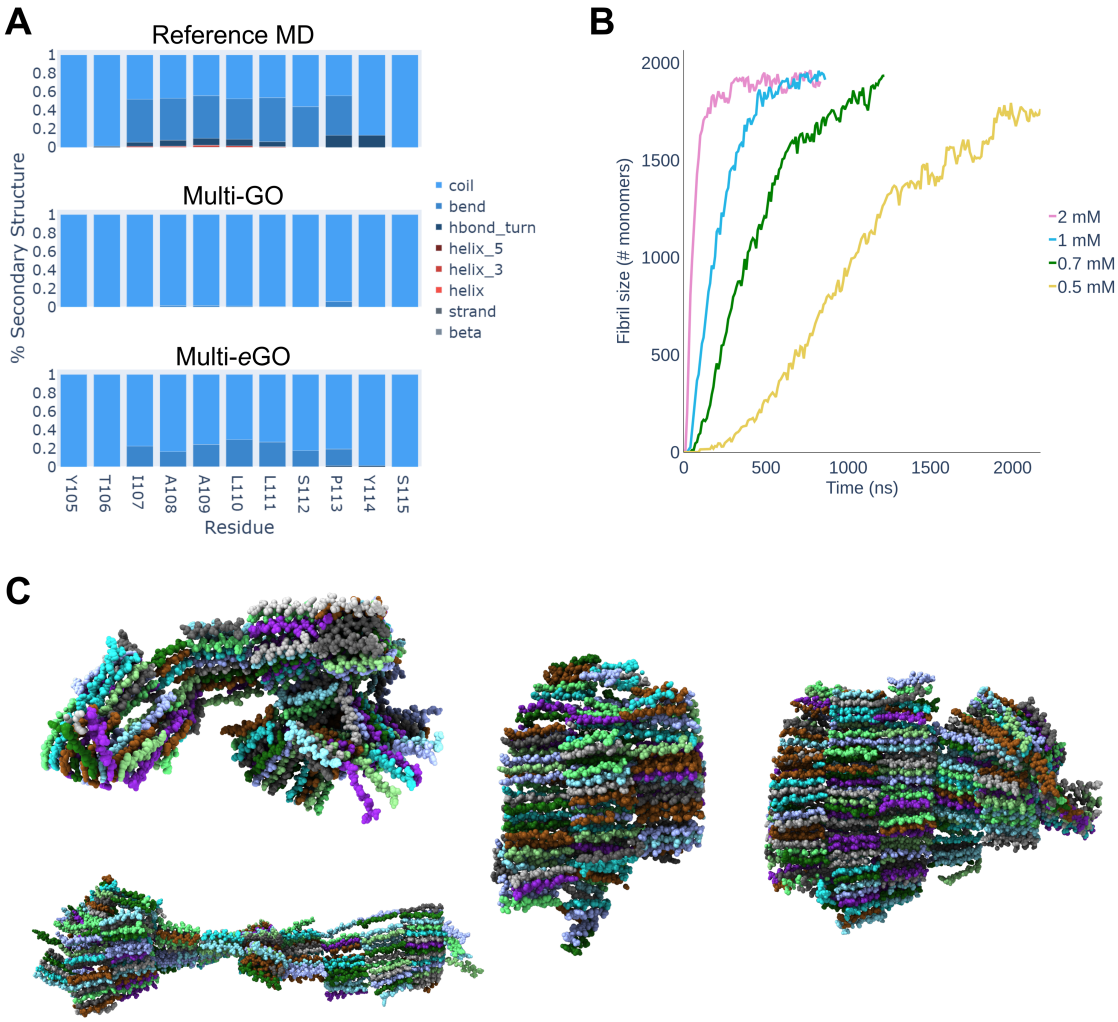


Figure S2. (A) Secondary structures populations for the $TTR_{105\text{-}115}$ monomer simulated using the a99SB-*disp* (top), multi-GO (middle) and multi-*e*GO (bottom) force fields. (B) Simulated aggregation kinetics using multi-GO. Curves represent the number of monomers involved in an aggregate of at least 10 monomers as a function of nominal simulation time. (C) Representative structure of amyloid fibrils formed at the end of the multi-GO aggregation kinetics.

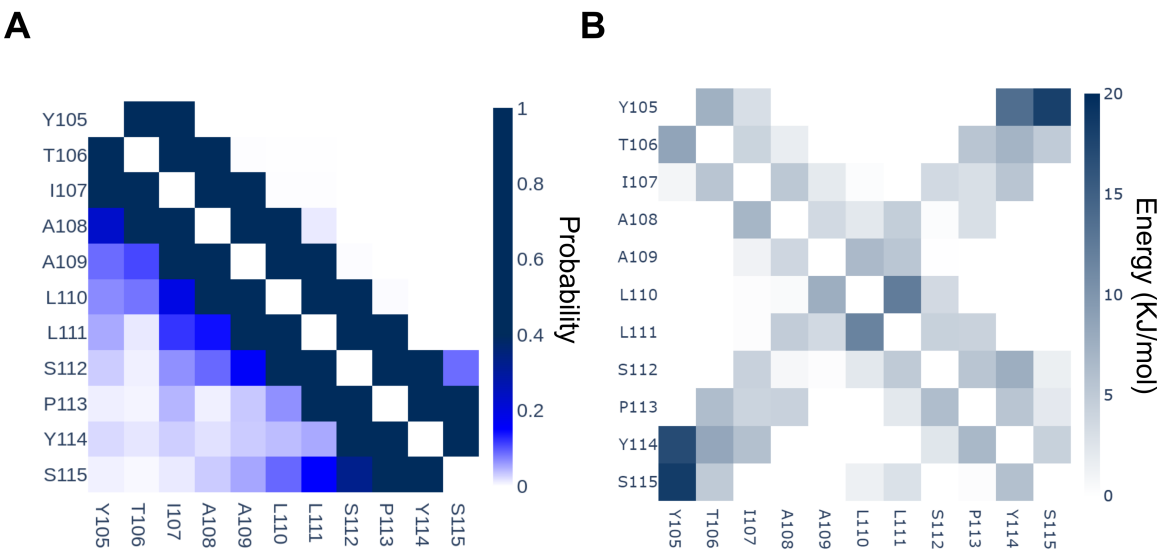


Figure S3. (A) Per residue probability contact map for the $TTR_{105\text{-}115}$ peptide conformational ensemble according to a MD simulation with a99SB-*disp* (lower left) and multi-GO (upper right). (B) Lower left multi-GO, Upper right multi-*e*GO.

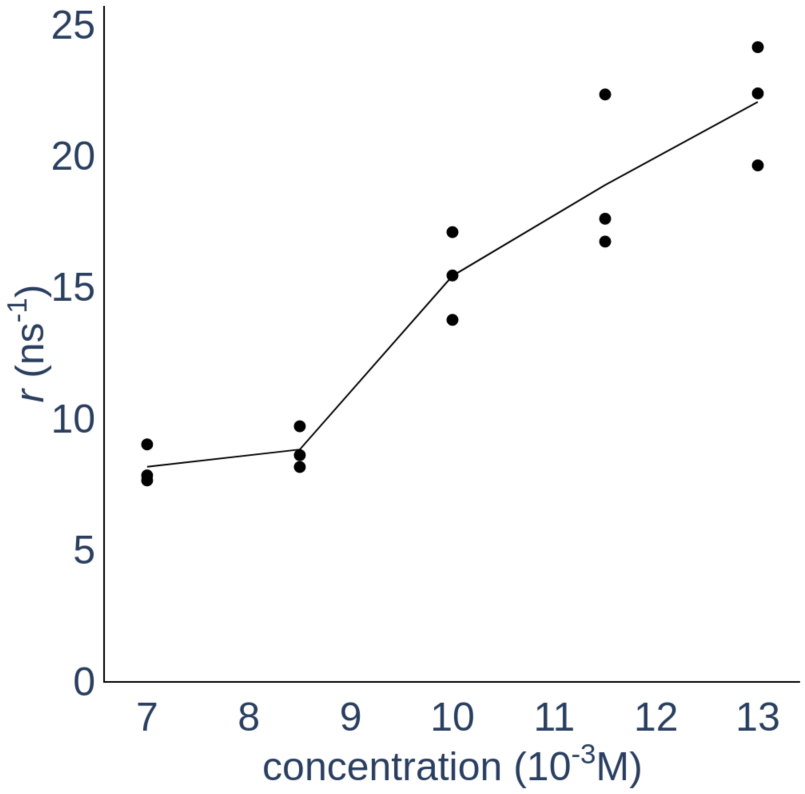


Figure S4. Growth rates, *r*, as a function of the initial monomer concentration, for the 15 simulated aggregation kinetics.

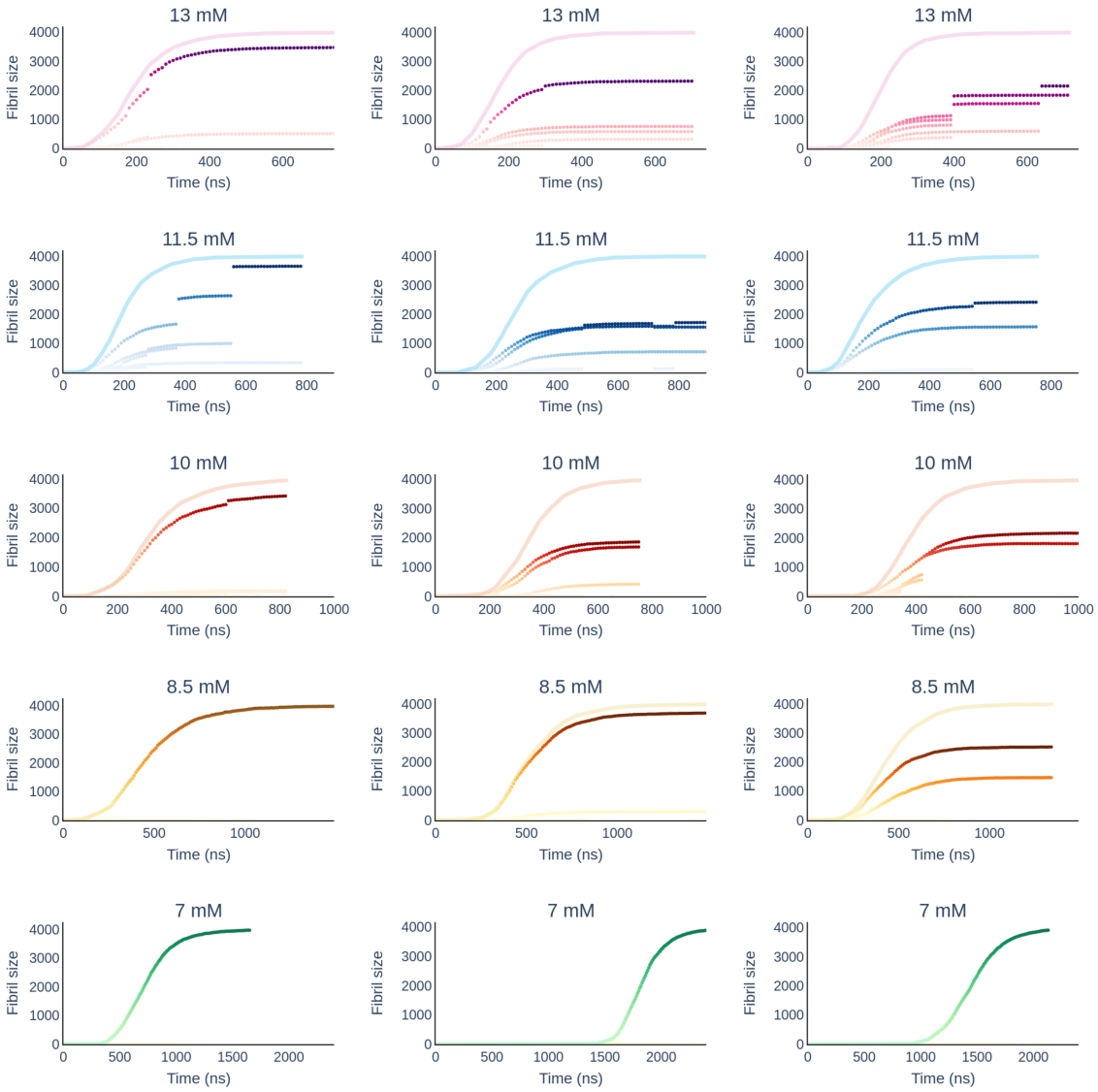
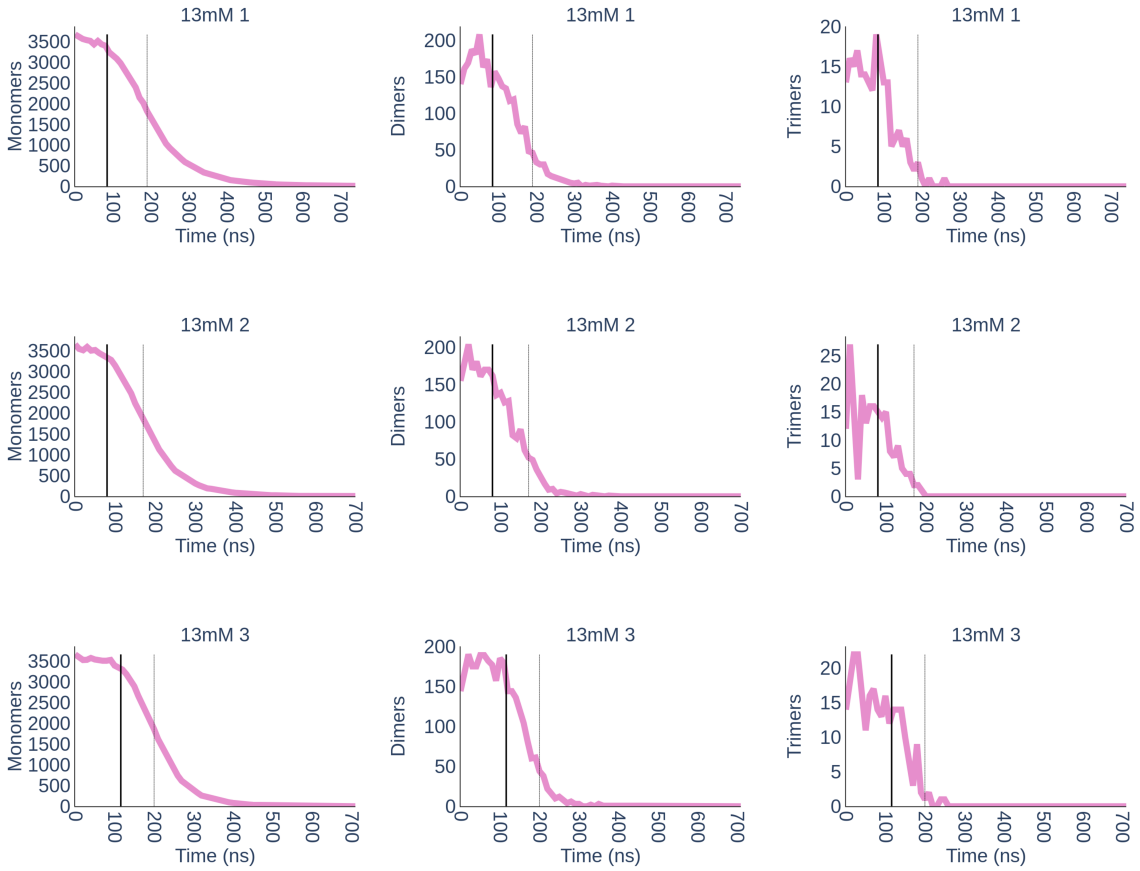
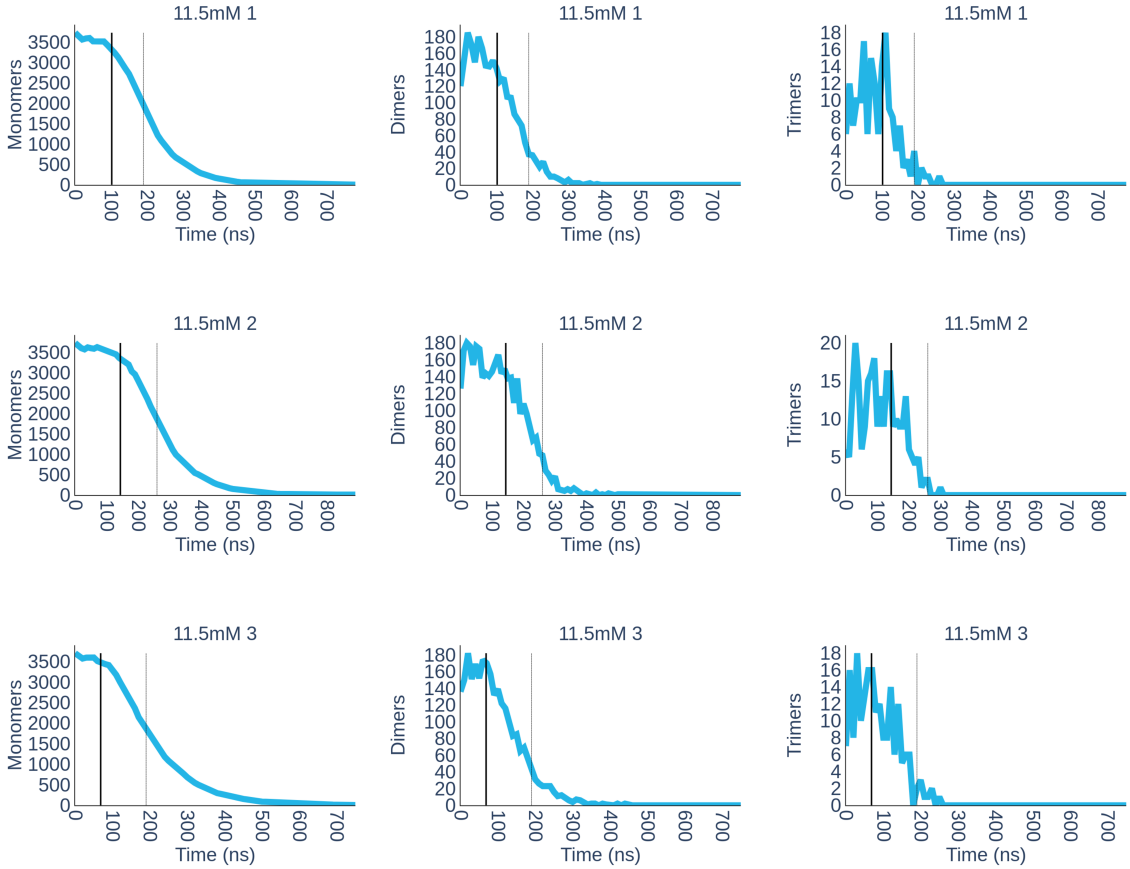
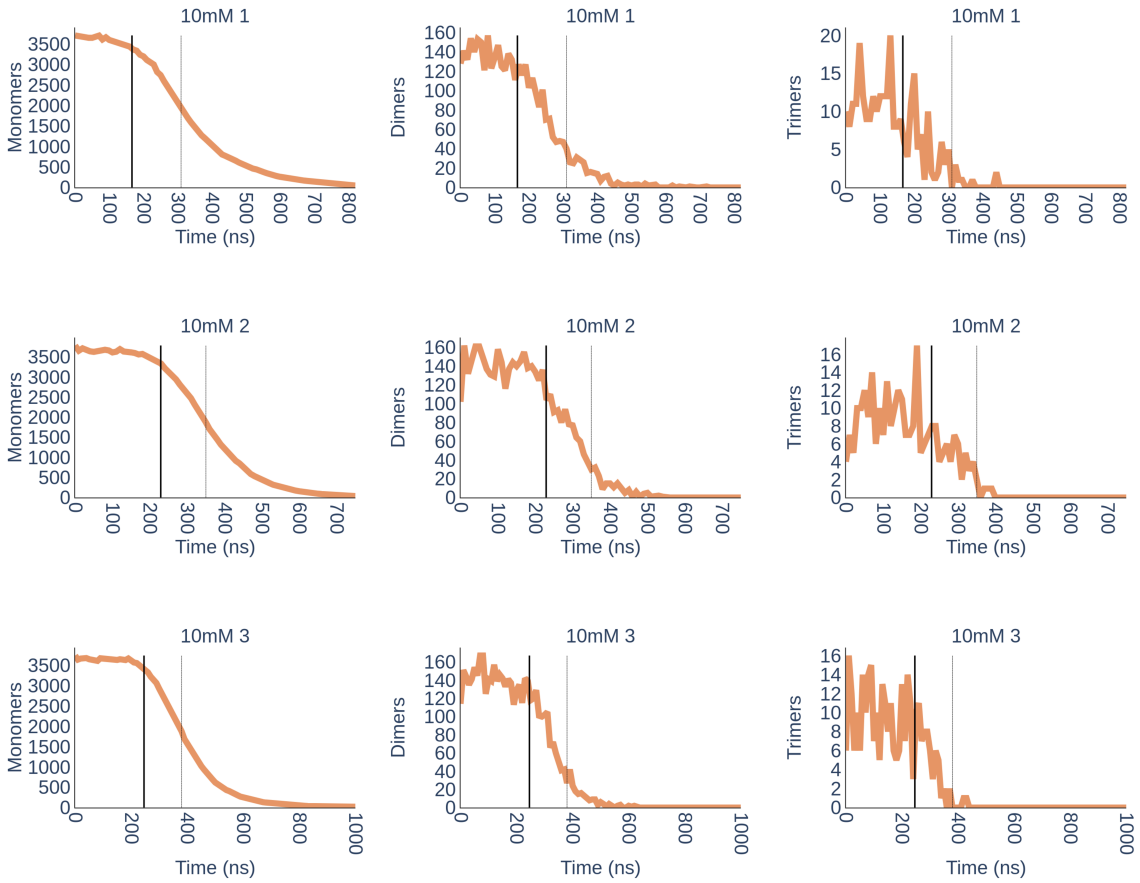
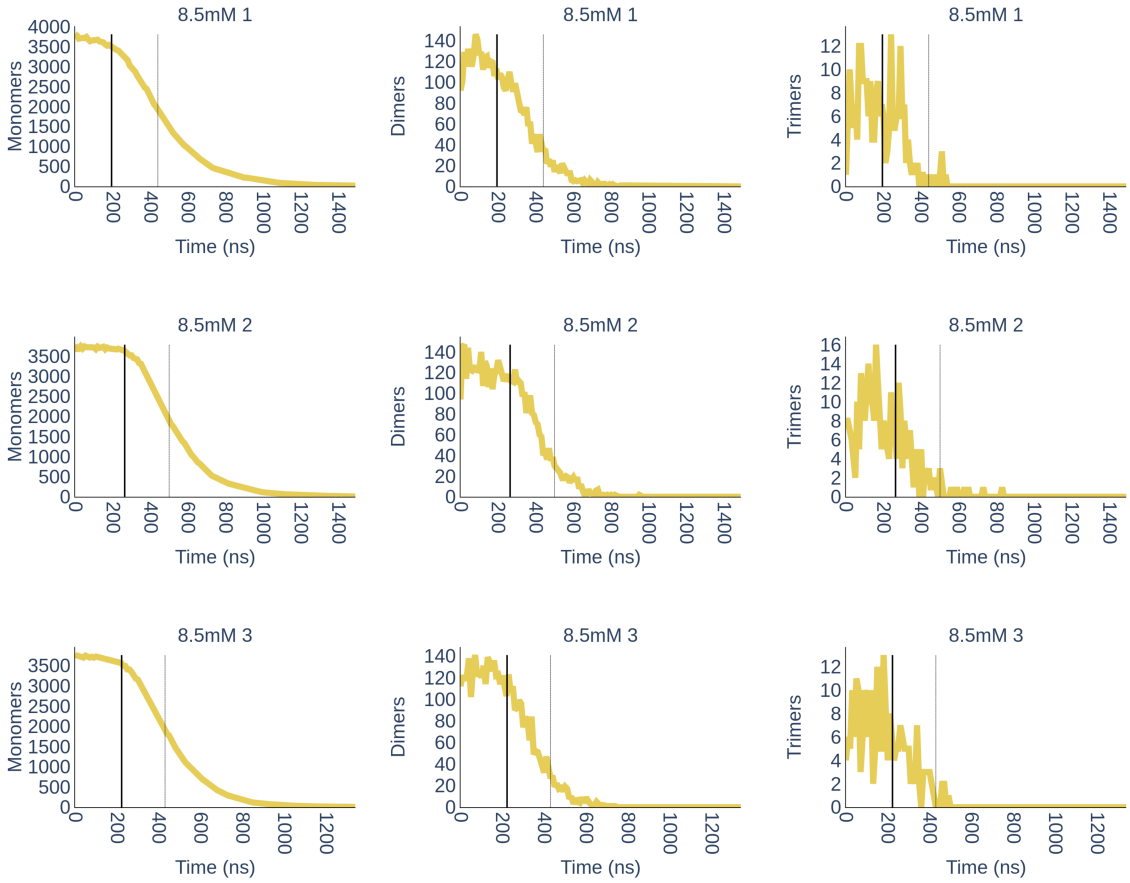


Figure S5. Fibrils grow for all simulations. The size of every fibril formed in the simulation is defined by the number of monomers. At high concentration different fibrils are formed and during the simulation they merge into a single fibril as pointed by the gaps.


13mM 1
Monomers
Time (ns)
13mM 1
Dimers
Time (ns)
13mM 1
Trimers
Time (ns)
13mM 2
Monomers
Time (ns)
13mM 2
Dimers
Time (ns)
13mM 2
Trimers
Time (ns)
13mM 3
Monomers
Time (ns)
13mM 3
Dimers
Time (ns)
13mM 3
Trimers
Time (ns)


11.5mM 1
Monomers
Time (ns)
11.5mM 1
Dimers
Time (ns)
11.5mM 1
Trimers
Time (ns)
11.5mM 2
Monomers
Time (ns)
11.5mM 2
Dimers
Time (ns)
11.5mM 2
Trimers
Time (ns)
11.5mM 3
Monomers
Time (ns)
11.5mM 3
Dimers
Time (ns)
11.5mM 3
Trimers
Time (ns)


10mM 1
Monomers
Time (ns)
10mM 1
Dimers
Time (ns)
10mM 1
Trimers
Time (ns)
10mM 2
Monomers
Time (ns)
10mM 2
Dimers
Time (ns)
10mM 2
Trimers
Time (ns)
10mM 3
Monomers
Time (ns)
10mM 3
Dimers
Time (ns)
10mM 3
Trimers
Time (ns)


8.5mM 1
Monomers
Time (ns)
8.5mM 1
Dimers
Time (ns)
8.5mM 1
Trimers
Time (ns)
8.5mM 2
Monomers
Time (ns)
8.5mM 2
Dimers
Time (ns)
8.5mM 2
Trimers
Time (ns)
8.5mM 3
Monomers
Time (ns)
8.5mM 3
Dimers
Time (ns)
8.5mM 3
Trimers
Time (ns)

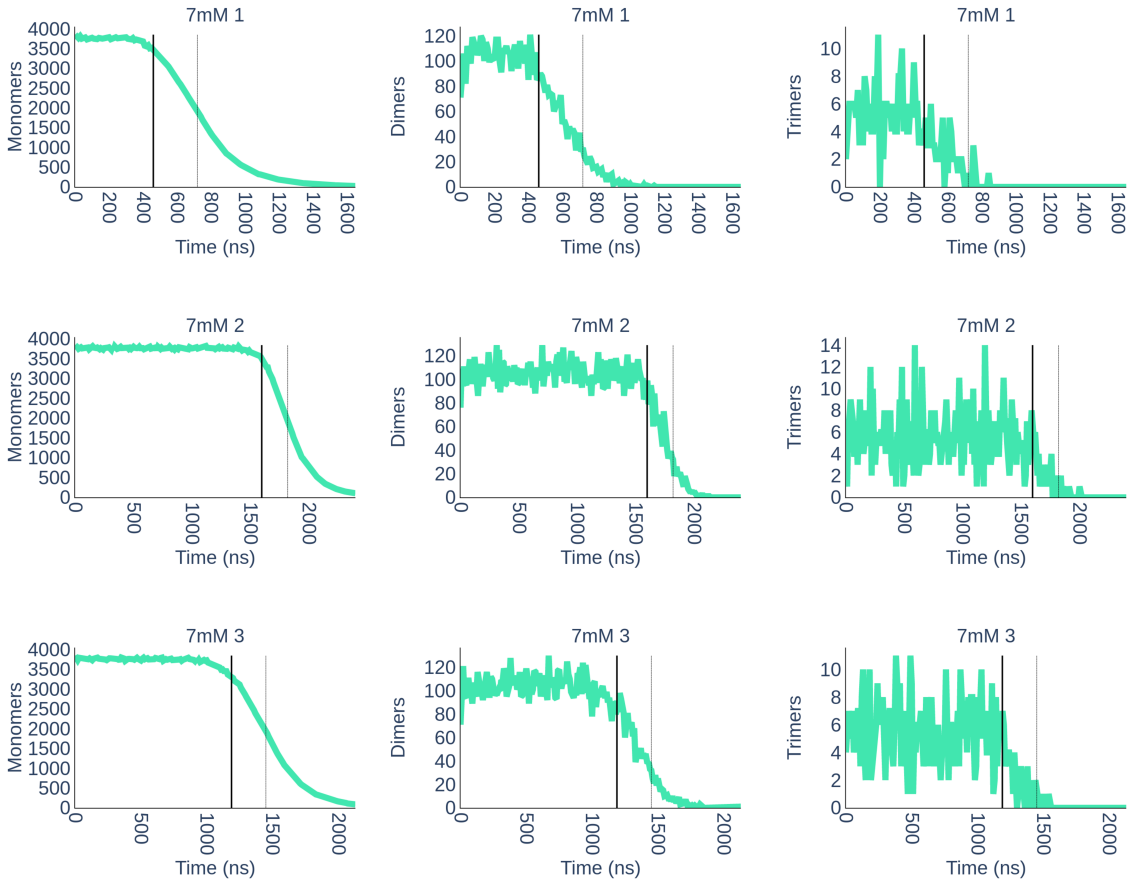


Figure S6. Number of monomers, dimers and trimers formed over time for each simulation.

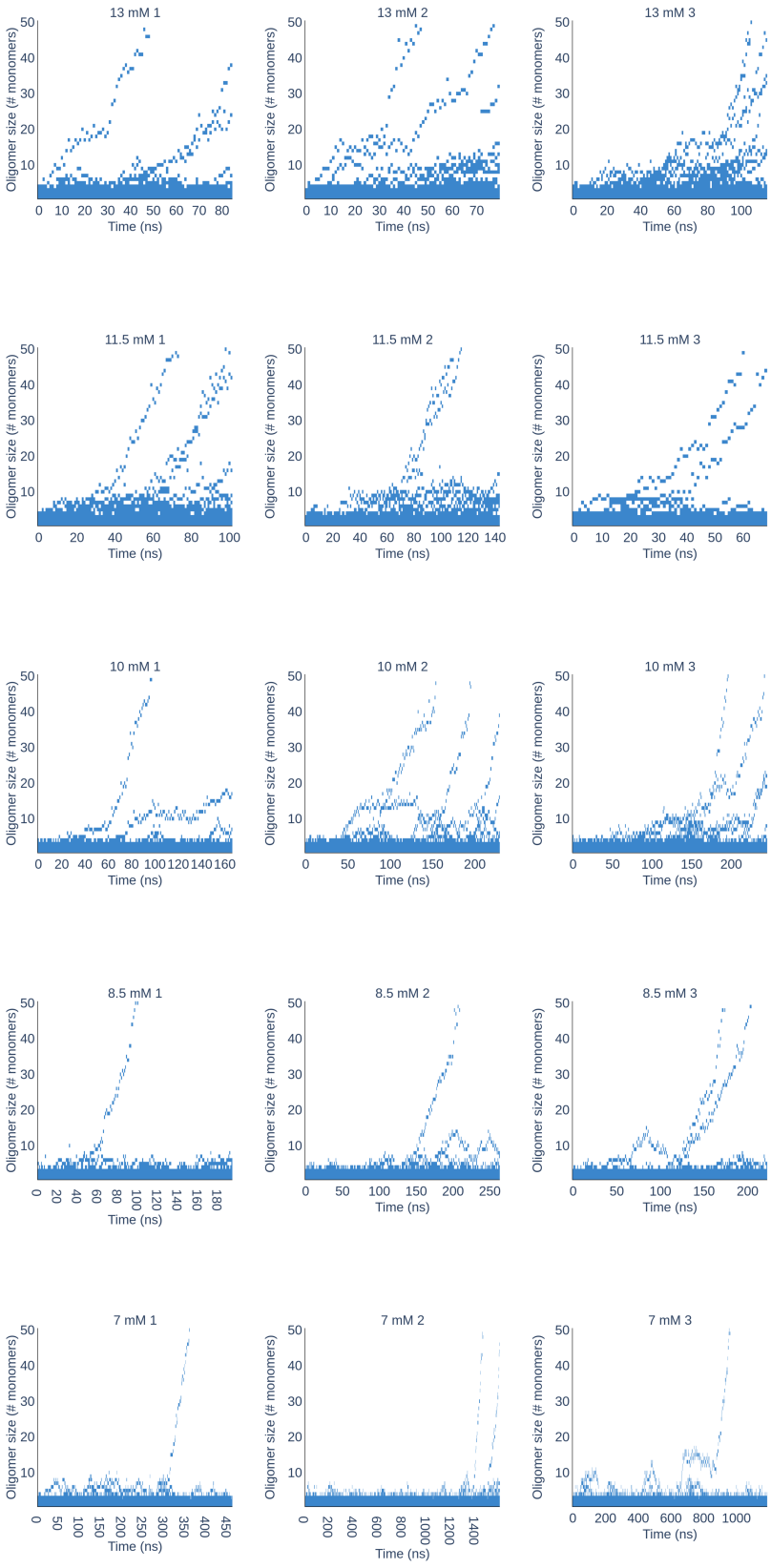
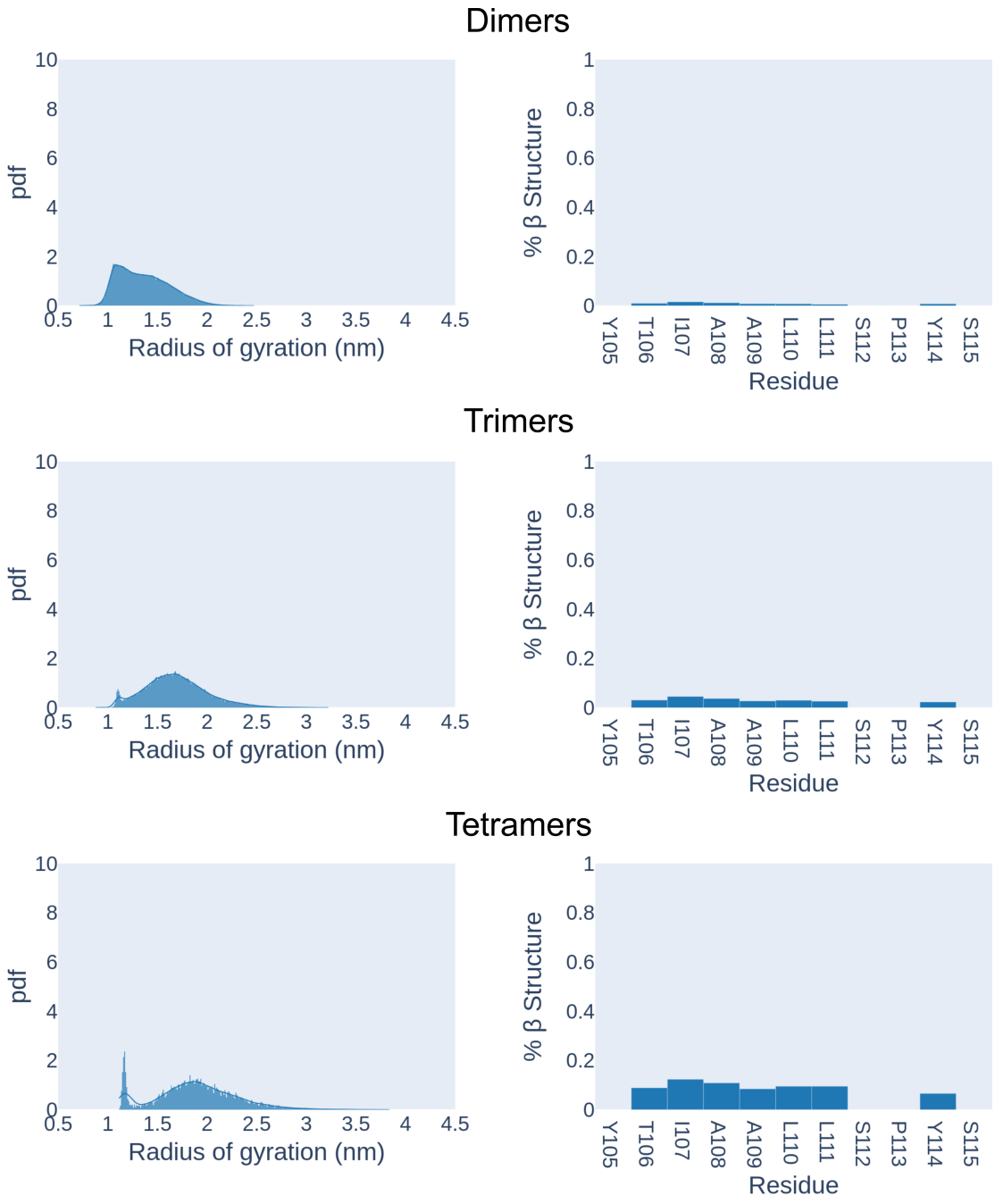
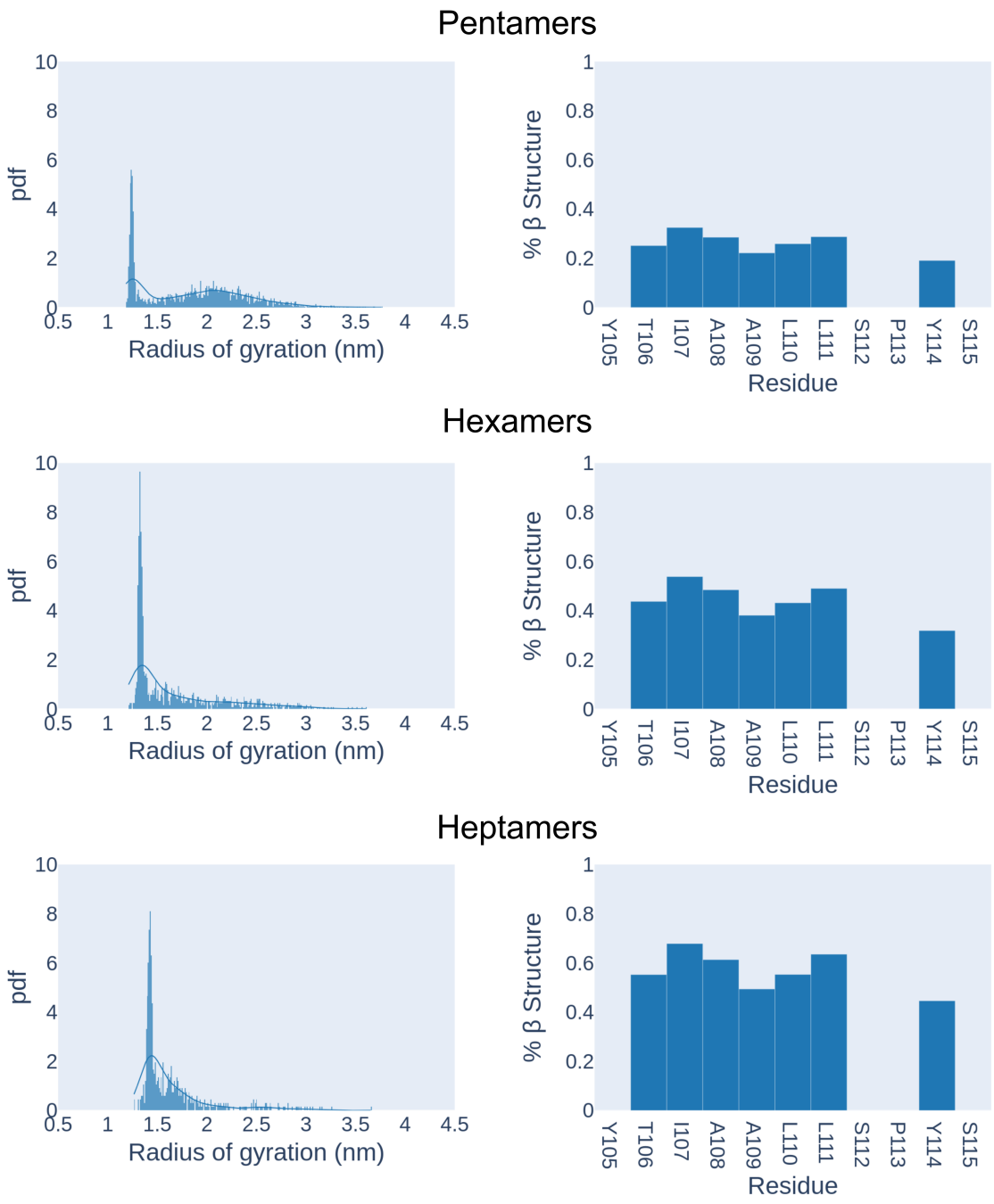


Figure S7. Time resolved evolution of oligomers order for all simulations before t_{lag}.


Dimers
pdf
Radius of gyration (nm)
% β Structure
Residue
Y105
T106
I107
A108
A109
L110
L111
S112
P113
Y114
S115
Trimers
pdf
Radius of gyration (nm)
% β Structure
Residue
Tetramers
pdf
Radius of gyration (nm)
% β Structure
Residue


Pentamers
pdf
Radius of gyration (nm)
% β Structure
Residue
Y105
T106
I107
A108
A109
L110
L111
S112
P113
Y114
S115
Hexamers
pdf
Radius of gyration (nm)
% β Structure
Residue
Heptamers
pdf
Radius of gyration (nm)
% β Structure
Residue

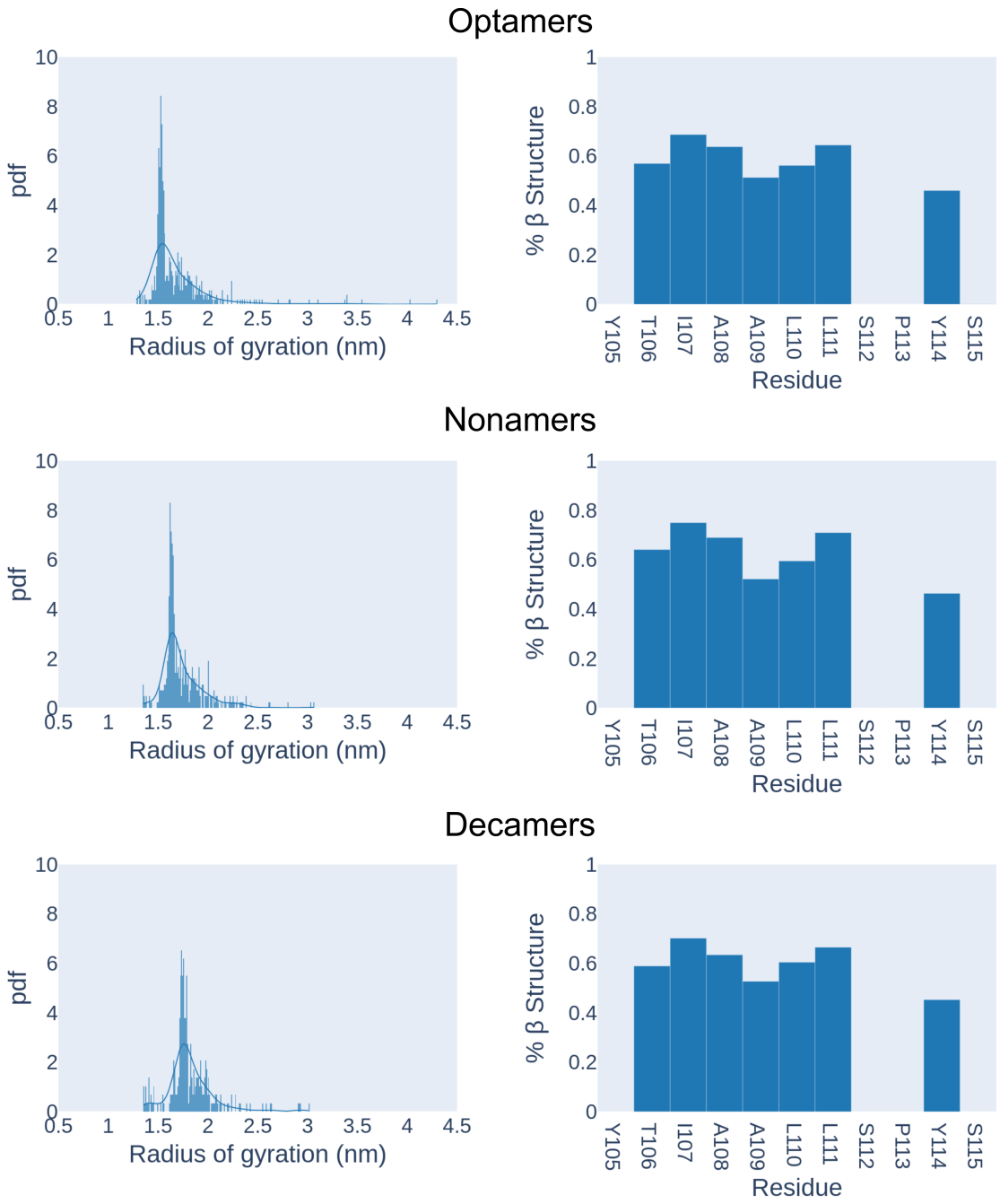


Figure S8. Distribution of the radius of gyration (left) and per residue average beta structure (right) for the oligomers, from dimers to decamers. populated during the multi-*e*GO aggregation kinetics until t_{lag}.

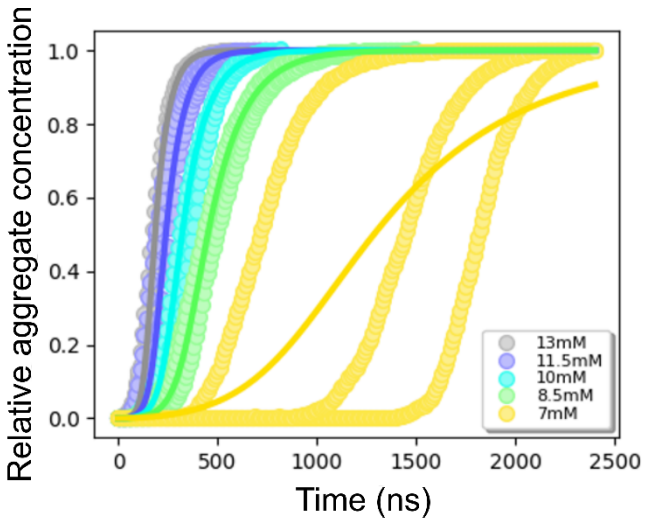


Figure S9. Chemical kinetics analysis by Amylofit of the 15 simulated aggregation kinetics traces. The data are fitted using the multi-step secondary nucleation, unseeded model.

Table S1. Detailed description of fibrils morphologies obtained *in-silico*.

	Replica	# Fibrils	Length (Å)	# β-sheets in filaments	# Filaments	Twist (°)	Number of monomers
13 mM	1	4	368	9	5	-0.22	2100
			232	9	2	-0.29	505
			211	5	3	-0.43	465
			300	7	2	-0.54	512
	2	7	320	10	5	-0.59	1513
			198	4	2	-0.58	206
			175	6	3	-0.73	336
			100	5	1	-4.97	86
			246	10	2	-0.50	324
			221	9	2	-0.32	584
			163	4	3	-0.56	324
	3	5	380	7	4	-0.47	1143
			190	8	2	-0.38	377
			190	8	3	-0.18	602
			255	9	4	-0.33	959
			215	7	4	-0.5	862
11.5 mM	1	4	390	6	4	-0.46	961
			280	12	4	-0.50	1717
			220	8	3	-0.23	909
			210	5	2	-0.85	336
	2	3	330	11	4	-0.24	1717
			300	11	3	-0.24	1560
			250	5	4	-0.30	713
	3	2	260	13	3	-0.33	1574
			310	12	6	-0.33	2420

10 mM	1	1	415	9	5	-0.70	3425
	2	3	366	5	6	-0.79	1690
			270	8	4	-0.23	1860
			214	5	3	-0.31	417
	3	2	372	9	4	-0.37	1811
			315	7	6	-0.24	2168
8.5 mM	1	1	515	9	3	-0.54	3980
	2	1	500	9	4	-0.15	3688
	3	2	355	14	3	-0.19	2524
			370	10	4	-0.34	1469
7 mM	1	1	435	17	5	-0.11	3980
	2	1	480	7	6	-0.48	3359
	3	1	490	13	4	-0.19	3313
7 mM seeded	1	1	443	13	5	-0.35	3959
	2	1	460	14	5	-0.10	3907
	3	1	485	13	5	-0.80	3929

Table S2. Detailed description of *in vitro* fibrils morphologies.

	Crossover (Å)*	Estimated twist (°)	Maximum width (Å)[Δ]	Estimated number of filaments in a fibril[•]	Width at crossover (Å)
I	1041 ± 24	-0.81	115 ± 11	3	39 ± 6
II	1112 ± 57	-0.76	237 ± 28	6	36 ± 3
III	1121 ± 33	-0.75	171 ± 9	5	36 ± 2
IV	1185 ± 43	-0.71	326 ± 21	8	36 ± 4
V	1060 ± 20	-0.80	247 ± 1	6	35 ± 1
VI	1115 ± 39	-0.76	191 ± 1	5	ND

* Crossover is defined as the length between two consecutive twists.
˜ Estimated twist is calculated by dividing 180° by the estimated number of molecules in a crossover (calculated considering an average distance between molecules of 4.7 Å).
Δ Width is defined as the maximum diameter at the largest point within a crossover.
• Calculated considering a peptide length of 38 Å

Movie S1 (separate file).
Multi-*e*GO MD trajectory for the aggregation kinetics of the first replica at 13 mM.
Available in Dataset S1, https://dx.doi.org/10.5281/zenodo.6125995

Movie S2 (separate file).
Multi-*e*GO MD trajectory for the aggregation kinetics of the second replica at 7 mM.
Available in Dataset S1, https://dx.doi.org/10.5281/zenodo.6125995

Movie S3 (separate file).
Multi-*e*GO MD trajectory for the aggregation kinetics of the first seeded replica at 7 mM.
Available in Dataset S1, https://dx.doi.org/10.5281/zenodo.6125995

Dataset S1 (separate file).

DOI: 10.5281/zenodo.6125995

https://dx.doi.org/10.5281/zenodo.6125995

Molecular dynamics simulation trajectories of TTR peptide monomers and aggregation kinetics:

- Amber99sb_disp: TTR monomer in explicit solvent using amber99disp force field.
- multi-GO-monomer: TTR monomer simulation using the multi-GO force field
- multi-eGO-monomer: TTR monomer simulation using the multi-*e*GO ensemble force field.
- multi-eGO-XXmM-Y: aggregation kinetics simulations of TTR using the multi-*e*GO force field at 13 mM, 11.5 mM and 10 mM concentration; replicate Y
- .ipynb files: analysis scripts employed for the aggregation kinetics simulations.

Dataset S2 (separate file).

DOI: 10.5281/zenodo.6125424

https://dx.doi.org/10.5281/zenodo.6125424

Molecular dynamics simulation trajectories of TTR peptide aggregation kinetics:

- multi-eGO-XXmM-Y: aggregation kinetics simulations of TTR using the multi-*e*GO force field at 8.5 mM, 7 mM and 7 mM seeded; replicate Y.

Dataset S3 (separate file).

DOI: 10.5281/zenodo.6414572

https://dx.doi.org/10.5281/zenodo.6414572

Multi-GO Molecular dynamics simulation trajectories of TTR peptide and aggregation kinetics; multi-*e*GO oligomer structures and trajectories:

- multi-GO-XXmM: aggregation kinetics simulations of TTR using the multi-*e*GO force field at XXmM concentration.
- Ensembles of all the oligomers, from dimers to decamers, sampled in each multi-*e*GO aggregation kinetics simulation.

SI References

1. P. Robustelli, S. Piana, D. E. Shaw, Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A* **115**, E4758–E4766 (2018).
2. U. Essmann, *et al.*, A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **103**, 8577 (1998).
3. B. Hess, P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation* **4**, 116–122 (2007).
4. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *The Journal of Chemical Physics* **126**, 014101 (2007).
5. M. Bernetti, G. Bussi, Pressure control using stochastic cell rescaling. *The Journal of Chemical Physics* **153**, 114107 (2020).
6. A. W. P. Fitzpatrick, *et al.*, Atomic structure and hierarchical assembly of a cross-β amyloid fibril. *Proc Natl Acad Sci U S A* **110**, 5468–5473 (2013).
7. L. Saelices, *et al.*, Uncovering the mechanism of aggregation of human transthyretin. *Journal of Biological Chemistry* **290**, 28932–28943 (2015).
8. R. T. McGibbon, *et al.*, MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophysical Journal* **109**, 1528–1532 (2015).
9. N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, O. Beckstein, MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *Journal of Computational Chemistry* **32**, 2319–2327 (2011).
10. R. Gowers, *et al.*, MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations in *Proceedings of the 15th Python in Science Conference*, (SciPy, 2016), pp. 98–105.
11. C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nature Methods 2012 9:7* **9**, 671–675 (2012).