Supporting Information

Small Peptides for Inhibiting Serum Amyloid A Aggregation

Asis K. Jana‡, Augustus B. Greenwood‡ and Ulrich H. E. Hansmann*

Department of Chemistry & Biochemistry, University of Oklahoma, Norman, OK 73019, USA

‡ These authors contributed equally to this work.

* To whom corresponding should be addressed: uhansmann@ou.edu

Table of Contents, Supporting Information

1. Supporting Figures

1.1 Figure S1: *Intra/Inter-peptide side-chain distance map measured during control simulations in the presence of 1st residue Arginine at 150 mM ionic solution*. S3

1.2 Figure S2: *Time evolution of the global RMSD measured in simulations of the fibril in presence of the R5S inhibitor at varying ionic solution and in presence/absence of 1st residue Arginine.* S4

1.3 Figure S3: *Intra/Inter-peptide side-chain distance map measured in simulations of SAA fibrils in presence of either L4F or F5M inhibitor.* S5

1.4 Figure S4: *Residue-wise binding probability of the L4F inhibitor.* S6

1.5 Figure S5: *Residue-wise binding probability of the F5M inhibitor*. S7

1.6 Figure S6: *Intra/Inter-peptide side-chain distance map measured in simulations of SAA fibrils in presence of DRI-R5S inhibitor.* S8

1.7 Figure S7: *Residue-wise binding probability of the DRI-R5S inhibitor.* S9

2. Tables

2.1 Table S1: *Details of all simulated systems and simulation length.* S10

2.2 Table S2: *Interpeptide binding free energy between SAA chains and contributions of each component, measured in simulations of SAA fibrils in presence of the R5S inhibitor.* S11

2.3 Table S3: *Interpeptide binding free energy between SAA chains and contributions of each component, measured in control simulations of SAA fibrils.* S12

3. Materials and Methods

3.1 *System Preparation* S13

3.2 *Simulation Protocol* S14

3.3 *Trajectory analysis* S15

4. **References** S16

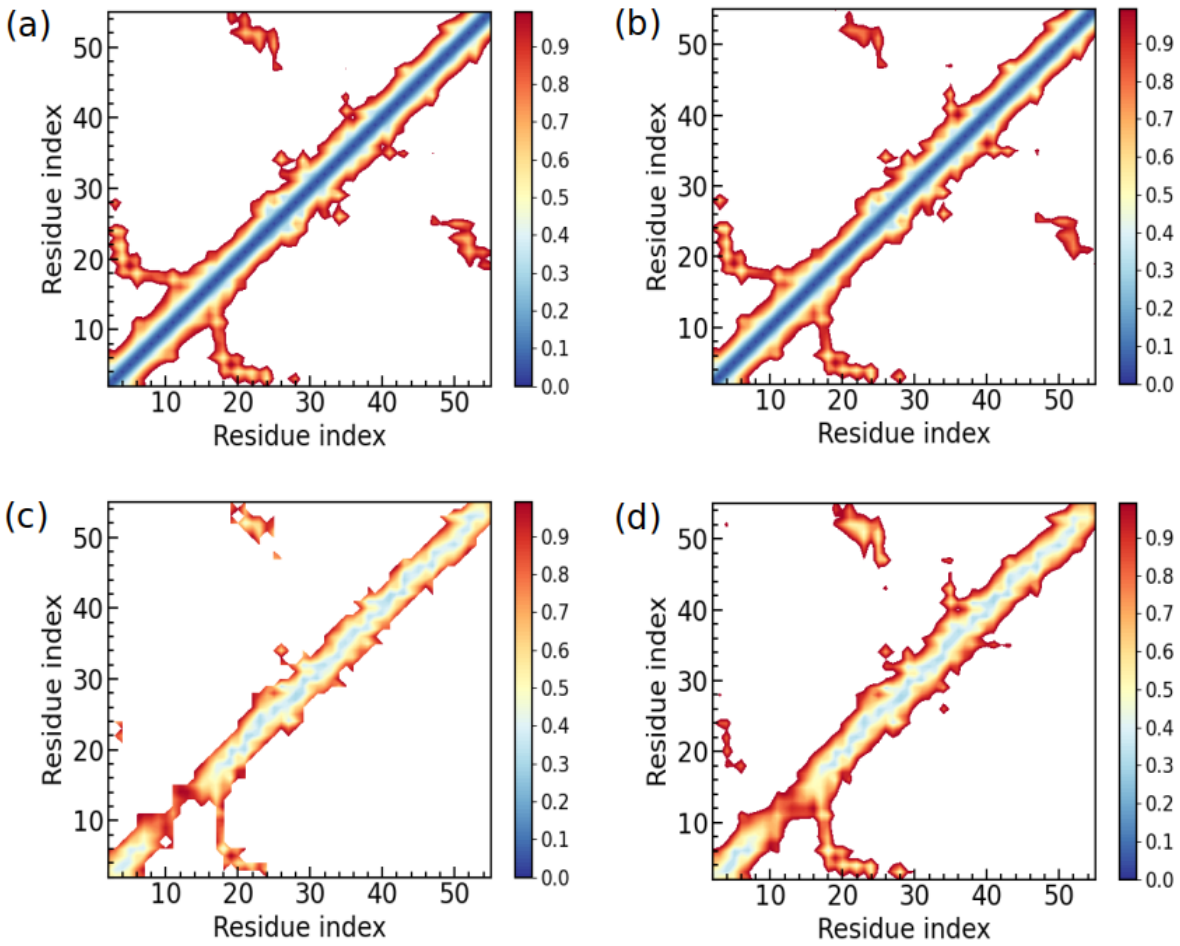

Figure S1: (a) Intra-peptide side-chain distance map measured during simulations of SAA fibrils in the presence of 1st residue Arginine at 150 mM ionic solution. The corresponding inter-peptide side-chain distance map is shown in (c). For a comparison, we have also shown (b) intrapeptide and (d) interpeptide side-chain distance maps measured during simulations of SAA fibrils in the absence of 1st residue Arginine. The data are averaged over three trajectories. Residue pairs whose average contact distance is more than 1.0 nm are excluded.

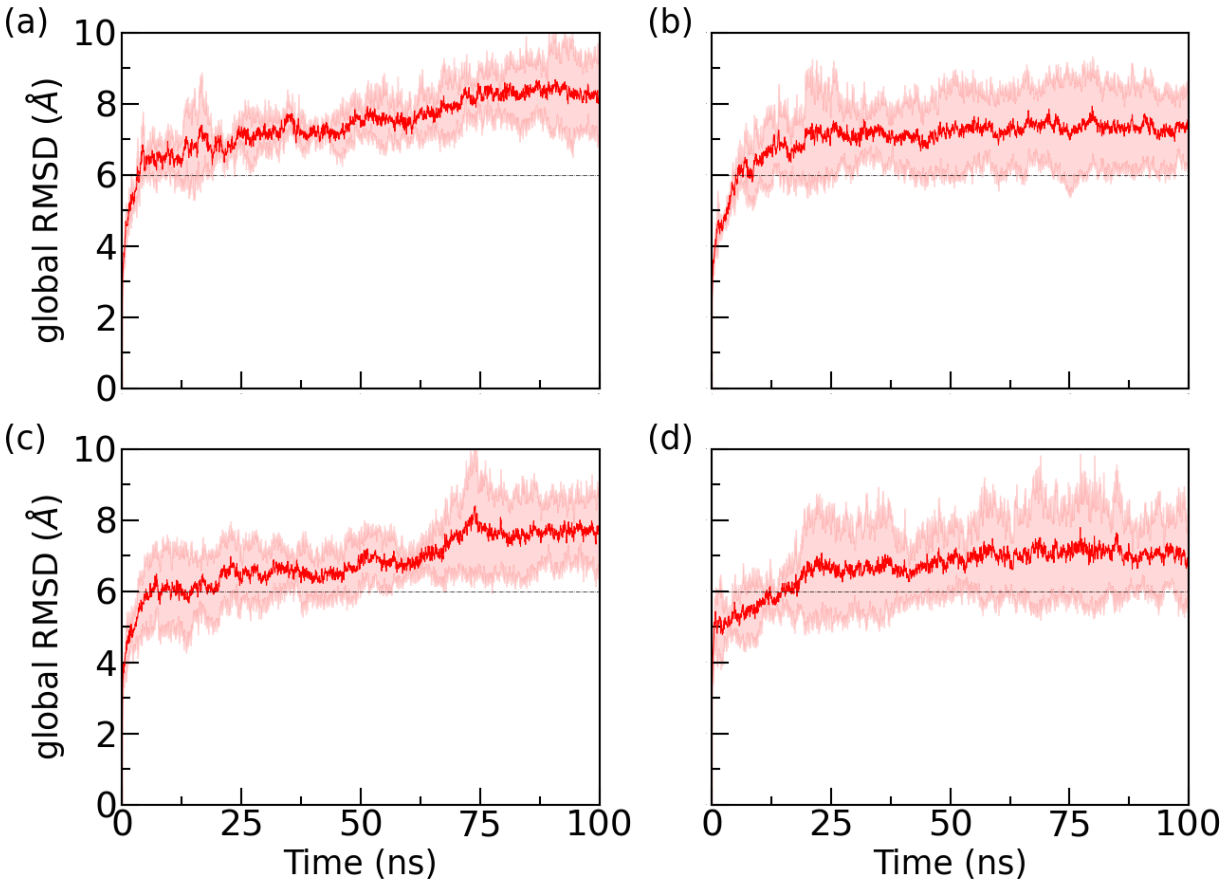

Figure S2: Time evolution of the global RMSD measured in simulations of the fibril in presence of the R5S inhibitor. RMSD values are calculated in reference to the experimentally resolved structure (PDB ID: 6MST) considering only backbone atoms in all chains of the respective fibril model. The data are averaged over three trajectories, with the shaded region indicating the standard deviation of the average. Systems displayed are those with arginine at 0 mM (a), with arginine at 150 mM NaCl (b), without arginine at 0 mM (c) and without arginine at 150 mM (d).

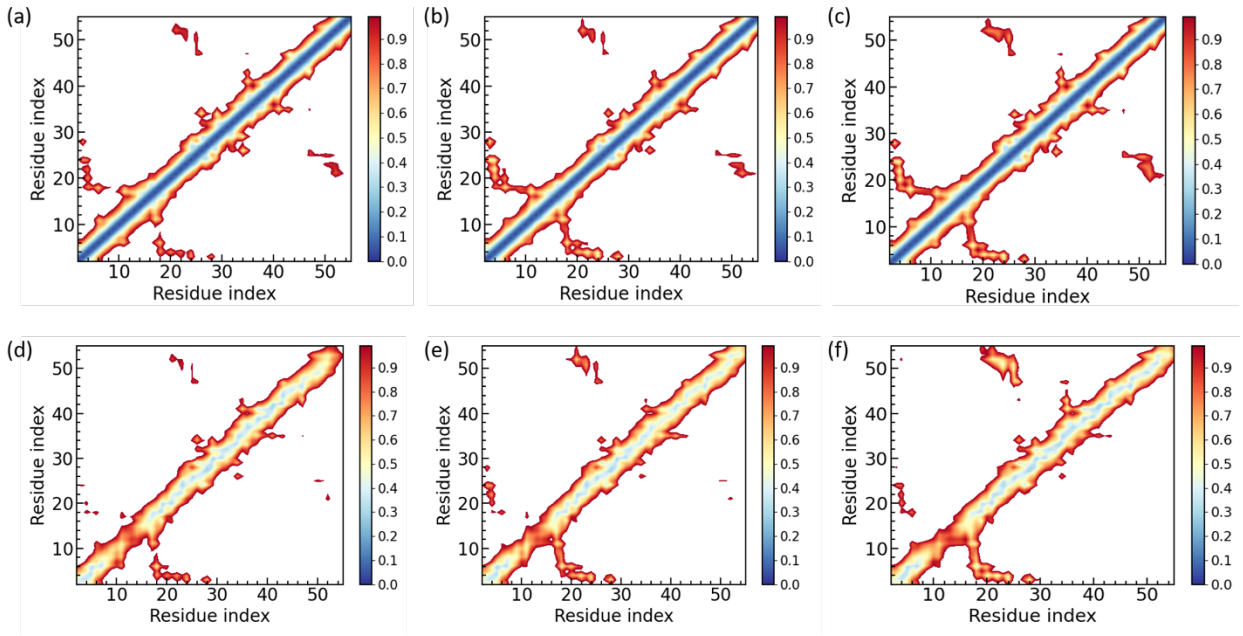

Figure S3: Intrapeptide side-chain distance map measured in simulations of SAA fibrils in presence of either (a) L4F or (b) F5M inhibitor. The corresponding interpeptide side-chain distance maps are shown in (d) and (e), respectively. For a comparison, we have also shown (c) intrapeptide and (f) interpeptide side-chain distance maps obtained from control simulation. Residue pairs whose average contact distance is more than 1.0 nm are excluded. Distance maps were obtained using the last 100 ns of the three independent trajectories for each system.

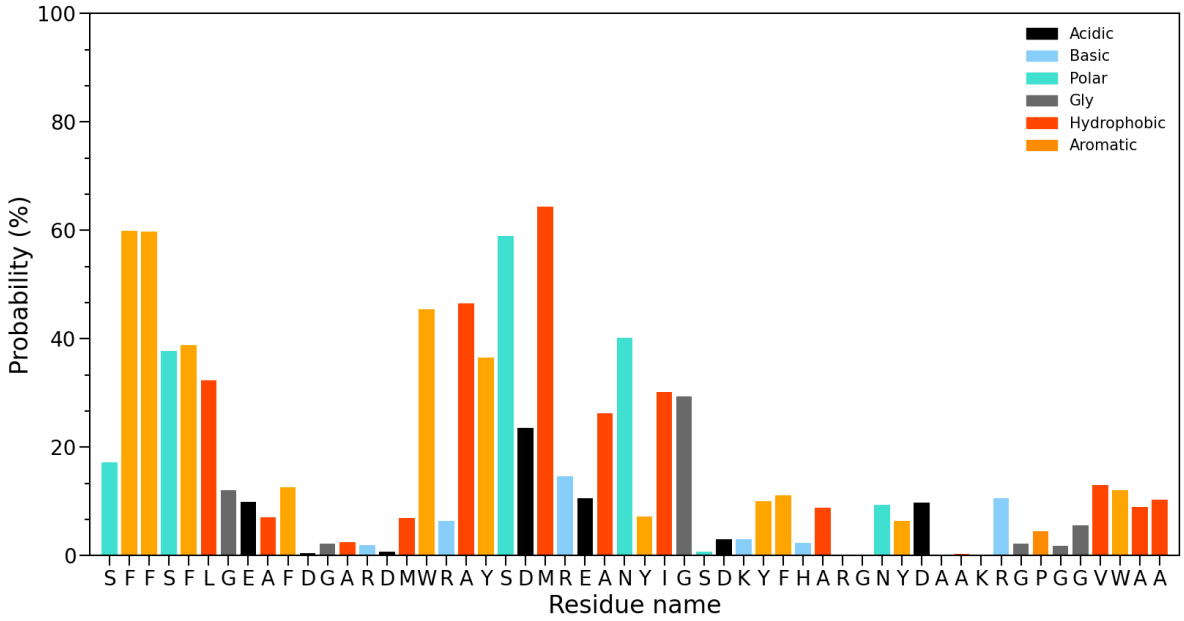

Figure S4: Residue-wise binding probability (normalized) of the L4F inhibitor to the SAA fibril. Data are averaged over the final 100 ns of all three independent runs.

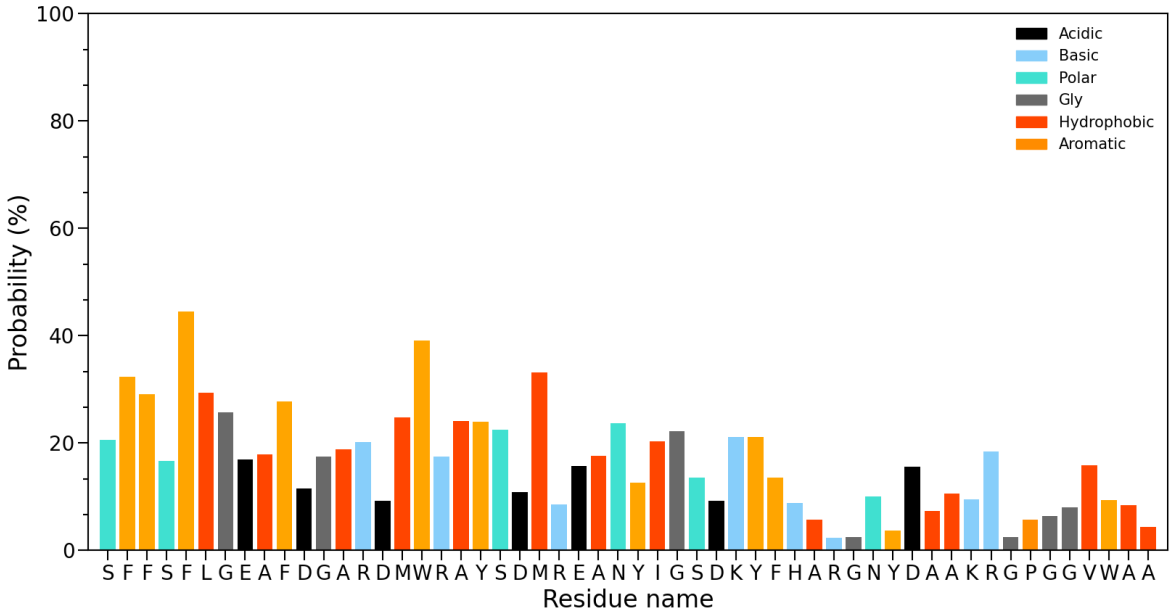

Figure S5: Residue-wise binding probability (normalized) of the F5M inhibitor to the SAA fibril. Data are averaged over the final 100 ns of all three independent runs.

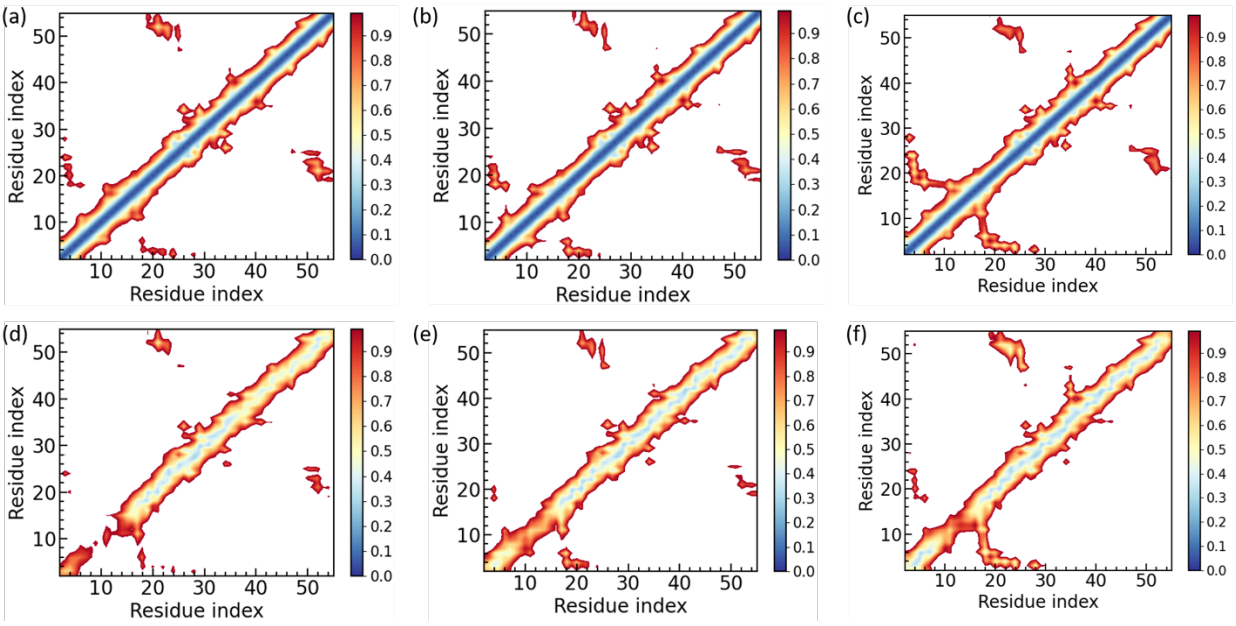

Figure S6: Intrapeptide side-chain distance map measured in simulations of SAA fibrils in presence of either (a) DRI-R5S or (b) D-R5S inhibitor. The corresponding interpeptide side-chain distance maps are shown in (d) and (e), respectively. For a comparison, we have also shown (c) intrapeptide and (f) interpeptide side-chain distance maps obtained from control simulation. Residue pairs whose average contact distance is more than 1.0 nm are excluded. Distance maps were obtained using the last 100 ns of the three independent trajectories for each system.

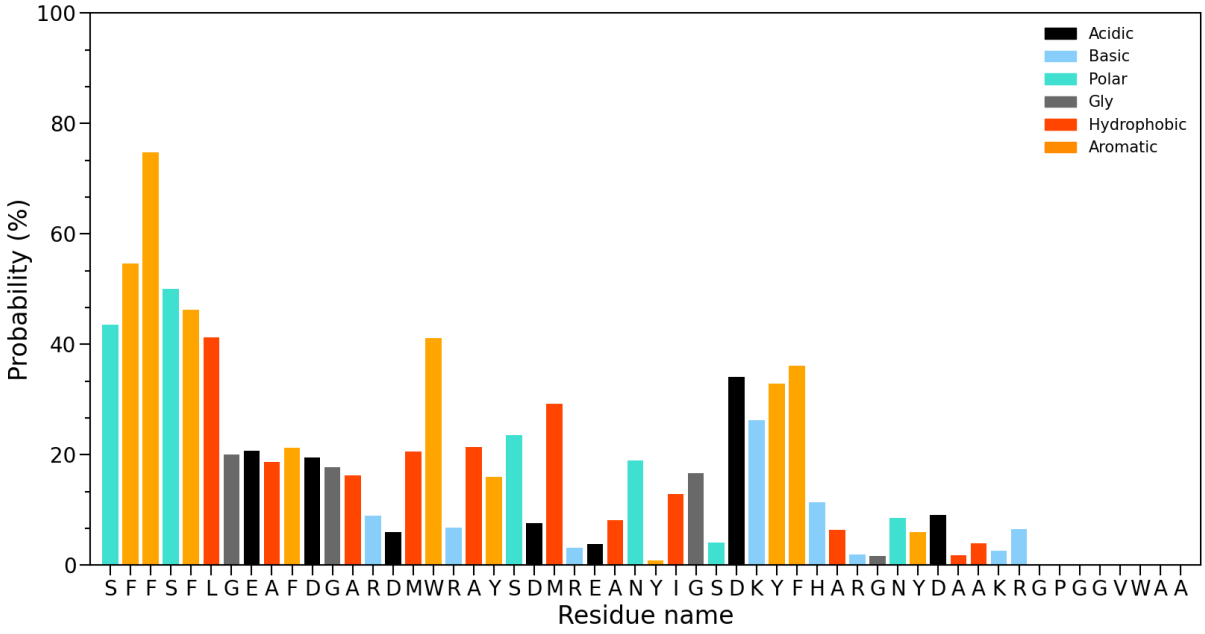

Figure S7: Residue-wise binding probability (normalized) of the DRI-R5S inhibitor to the SAA fibril. Data are averaged over the final 100 ns of all three independent runs.

Table S1: Details of all simulated systems and simulation length.

System			Atoms	Water molecules	Independent runs	Simulation length (ns)	Simulation time (μs)
No Inhibitor	2F3L$_{1-55}$	0mM	133,569	42,890	3	100	0.3
		150mM	133,073	42,642	3	100	0.3
	2F3L$_{2-55}$	0mM	114,481	36,568	3	100	0.3
		150mM	114,111	36,384	3	200	0.6
	2F4L$_{2-55}$	150mM	116,164	36,466	3	200	0.9
R5S	2F3L$_{1-55}$	0mM	150,774	48,443	3	100	0.3
		150mM	150,222	48,167	3	100	0.3
	2F3L$_{2-55}$	0mM	150,684	48,463	3	100	0.3
		150mM	150,132	48,187	3	250	0.75
	2F4L$_{2-55}$	150mM	154,330	48,922	3	200	0.9
L4F	2F3L$_{2-55}$	150mM	150,015	48,172	3	250	0.75
F5M	2F3L$_{2-55}$	150mM	149,937	48,431	3	200	0.6
DRI-R5S	2F3L$_{2-55}$	150mM	150,060	48,163	3	250	0.75
D-R5S	2F3L$_{2-55}$	150mM	150,066	48,165	3	250	0.75
All trajectories							7.8 μs

Table S2: Interpeptide binding free energy between SAA chains and contributions of each component, measured in simulations of SAA fibrils in presence of the R5S inhibitor. Standard deviations are provided within braces. Free energies are listed in Kcal/mol.

$2F3L_{2-55}$ in presence of R5S			trajectory-1		trajectory-2		trajectory-3		average	
ΔG_{bind}	ΔH_{MM}	ΔH_{elec}	-1476(55)	-571(49)	-1583(67)	-716(63)	-1647(54)	-675(56)	-1569(92)	-654(83)
		ΔH_{vdw}		-905(25)		-867(23)		-971(23)		-915(49)
	$\Delta G_{solv-pol}$		1342(83)		1422(72)		1491(82)		1418(100)	
	$\Delta G_{solv-np}$		-104(2)		-100(2)		-108 (2)		-104(4)	
	Total		-239(61)		-262(63)		-263(53)		-254(60)	
ΔG_{pack}	ΔH_{MM}	ΔH_{elec}	-460(33)	-407(31)	-306(40)	-279(36)	-471(39)	-405(40.0)	-412(84)	-363(70)
		ΔH_{vdw}		-53(7)		-28(11)		-66(8)		-49(18)
	$\Delta G_{solv-pol}$		420(41)		292(53)		438(51)		383(81)	
	$\Delta G_{solv-np}$		-10(1)		-6(2)		-10(1)		-9(2)	
	Total		-50 (37)		-21(37)		-43(33)		-38(38)	
ΔG_{elong}	ΔH_{MM}	ΔH_{elec}	-681(47)	-219(42)	-807(54)	-361(50)	-790(43)	-304(45)	-760 (74)	-295(74)
		ΔH_{vdw}		-462(16)		-446(18)		-487(22)		-465(25)
	$\Delta G_{solv-pol}$		623(67)		730(63)		698(72)		684(81)	
	$\Delta G_{solv-np}$		-49(1)		-49(2)		-50(2)		-49(2)	
	Total		-108(54)		-126(50)		-143(53)		-125(54)	

Table S3: Interpeptide binding free energy between SAA chains and contributions of each component, measured in simulations of SAA fibrils in the absence of any inhibitor. See the method sections for details. Standard deviations are provided within braces. Free energies are listed in Kcal/mol.

$2F3L_{2-55}$ with no inhibitor			trajectory-1		trajectory-2		trajectory-3		average	
ΔG_{bind}	ΔH_{MM}	ΔH_{elec}	-1612 (89)	-629(51)	-1920(70)	-813(68)	-1812(56)	-754(50)	-1781(147)	-732(96)
		ΔH_{vdw}		-984(74)		-1107(20)		-1058(24)		-1050(69)
	$\Delta G_{solv-pol}$		1480(70)		1692(72)		1589(78)		1587(113)	
	$\Delta G_{solv-np}$		-110(2)		-123(2)		-115(2)		-116(5)	
	Total		-243(88)		-351(43)		-339(51)		-311(80)	
ΔG_{pack}	ΔH_{MM}	ΔH_{elec}	-580(48)	-528(30)	-704(33)	-605(34)	-708(36)	-612(35)	-664(72)	-582(50)
		ΔH_{vdw}		-52(41)		-99(8)		-96(8)		-83(33)
	$\Delta G_{solv-pol}$		482(40)		524(48)		582(46)		530(61)	
	$\Delta G_{solv-np}$		-12(1)		-15(1)		-15(1)		-14(1)	
	Total		-1010 (54)		-195 (33)		-141 (31)		-148 (54)	
ΔG_{elong}	ΔH_{MM}	ΔH_{elec}	-737(58)	-242(44)	-885(73)	-322(63)	-827 (57)	-303(59)	-816(87)	-289(66)
		ΔH_{vdw}		-495(49)		-563(18)		-524(15)		-527(42)
	$\Delta G_{solv-pol}$		671(61)		790(72)		717 (64)		726 (82)	
	$\Delta G_{solv-np}$		-50(1)		-57(2)		-51 (2)		-53(4)	
	Total		-116(63)		-152(38)		-162(44)		143(53)	

Materials and Methods

System Preparation

We have evaluated the ability of three peptides to inhibit Serum Amyloid A (SAA) amyloid formation. Three candidates were selected after screening recent publications[1] and are listed in **Table 1** of the manuscript. A posteriori, two additional peptides were considered where we replaced the L-amino acids of the best performing candidate by D-amino acids, either in the same or with reverse sequence. Such D or DRI peptides are difficult to proteolyze and therefore have longer life times in the cell.

Assuming that potential inhibitors destabilize the fibril architecture, we have compared all-atom molecular dynamics simulations of SAA fibril fragments in the presence of one of these candidate inhibitors with such where the inhibitor candidates are absent. The fibril fragment is derived from the cryo-EM structure as deposited in the Protein Data Bank (PDB) under identifier 6MST[2], presently the only available experimentally derived human SAA fibril model. While this model describes a fibril made of $SAA_{2\text{-}55}$ fragments instead of the more commonly found $SAA_{1\text{-}76}$ fragments, we did not attempt to add additional residues by homology modeling. This is because we had found in our earlier work[3] that the disordered C-terminal tail of residues 56-76 does not alter by itself the stability of the fibril. Similarly guided by our previous work[3] we choose a two-fold-three-layer (2F3L) hexamer as model for our simulations. This is because this hexamer is above the critical threshold for a stable SAA fibril fragment (which we found in our earlier work to be the two-fold-two-layer (2F2L) tetramer). Note that in some exploratory simulations we also considered the hexamer with the N-terminal Arginine add to all six chains by Chimera[4], and the resulting configuration minimized under the constraint that the experimentally resolved parts of the structure stay unchanged.

We then used Autodock Vina[5] to identify initial binding sites on the fibril for the inhibitors, before preparing start configurations for our simulations where the respective inhibitor candidates are bound to the SAA chains in the fibril at a 1:1 ratio. In the so-obtained configuration are the SAA chains capped by NH_3^+ and COO^-, while the N- and C- termini of inhibitors are capped by NH_2 and $CONH_2$ to counteract the strong electrostatic repulsion between the terminal groups. The resulting systems are then put into a box with periodic boundary conditions and with a box size, large enough that there is at least 15 Å distance between any atom on the system and a box side. The box is filled with TIP3P[6] water molecules and a suitable number of Na_+ and Cl_- ions to obtain the desired ion concentration, in our case either 0 M or 150 mM. We list all systems, together with the number of water molecules, total number of atoms, and ion concentration in **Table S1**.

Simulation Protocol

For our simulations we use the GROMACS 2018 simulation package[7], while configurations were visualized using Visual Molecular Dynamics (VMD)[8]. SAA fibrils and peptide inhibitors are parametrized with the CHARMM 36m all-atom force-field[9], which is consistent with the TIP3P water model[6] used by us. Initially, each system is energy minimized by the steepest-descent algorithm to remove bad contacts between the solute and solvent, before equilibrated over 200 ps in a (NVT) ensemble at a constant temperature of 310 K, and additional 200ps in an isothermal–isobaric (NPT) ensemble at a constant pressure of 1 atm. The nonhydrogen (heavy) atoms of fibril and inhibitors are held fixed during equilibration with a force constant of 1,000 kJ mol^{-1} nm^{-2}. The resulting configurations are the start point for the subsequent production runs performed at 310 K and 1atm, with the temperature and pressure controlled by the v-rescale thermostat[10] (with a coupling constant of 0.1 ps) and the Parrinello-Rahman barostat[11] (with a pressure relaxation time

of 2 ps). Non-water bonds involving hydrogen atoms are constrained using the LINCS algorithm[12], and water molecules kept rigid by the SETTLE algorithm,[13] allowing an integration time step of 2 fs. As we use periodic boundary conditions, we employ the particle-mesh Ewald (PME) method[14] for calculating long-range electrostatic interactions. The cutoff for Van der Waal interactions is set to 12 Å, with the smoothing starting at 10.5 Å.

Trajectory analysis

For most of our analysis we use GROMACS tools[7] such as gmx_rms which calculates the root-mean-square deviation (RMSD) with respect to an initial configuration. For visualization we use the VMD software, which we use also to calculate the solvent-accessible surface area (SASA), selecting a spherical probe of 1.4 Å radius.[8] We have estimated interpeptide binding free energy (ΔG) between SAA chains using the Molecular Mechanics Poisson–Boltzmann surface area (MMPBSA)[15] approach as implemented in the GROMACS package.[16] In this method, ΔG is approximated by:

$$\Delta G = \Delta H_{MM} + \Delta G_{solv-pol} + \Delta G_{solv-np}$$

$$\Delta H_{MM} = \Delta H_{elec} + \Delta H_{vdw},$$

Here, ΔH_{MM} represents the potential energy difference between fibril and the free chains in vacuum, which is the sum of electrostatic (ΔH_{elec}) and van der Waals energy (ΔH_{vdw}) differences. $\Delta G_{solv\text{-}pol}$ represents the change in polar solvation free energy and is estimated by solving the nonlinear Poisson–Boltzmann equation at the solvent dielectric constant of water at 310 K, i.e., $\varepsilon = 78$. $\Delta G_{solv\text{-}np}$ represents the change in non-polar solvation free energy and is assumed to be proportional to the solvent accessible surface area (SASA). As in many recent studies[17,18], we also did

not consider the entropic contribution in the binding free energy calculation. In the MMPBSA approach, entropy is estimated by computationally expensive normal mode analysis. Inclusion of this term based on normal mode analysis does not improve the prediction, as it introduces significant statistical uncertainty in the result[17]. ΔG is computed over the three independent trajectories for each system, considering only the final 100 ns.

References

(1) Sosnowska, M.; Skibiszewska, S.; Kamińska, E.; Wieczerzak, E.; Jankowska, E. Designing Peptidic Inhibitors of Serum Amyloid A Aggregation Process. *Amino Acids* **2016**, *48* (4), 1069–1078.

(2) Liberta, F.; Loerch, S.; Rennegarbe, M.; Schierhorn, A.; Westermark, P.; Westermark, G. T.; Hazenberg, B. P. C.; Grigorieff, N.; Fändrich, M.; Schmidt, M. Cryo-EM Fibril Structures from Systemic AA Amyloidosis Reveal the Species Complementarity of Pathological Amyloids. *Nat. Commun.* **2019**, *10* (1), 1104.

(3) Wang, W.; Hansmann, U. H. E. Stability of Human Serum Amyloid A Fibrils. *J. Phys. Chem. B* **2020**, *124* (47), 10708–10717.

(4) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612.

(5) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461.

(6) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.

(7) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.

(8) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

(9) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* (1), 71–73.

(10) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126* (1), 14101.

(11) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.

(12) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.

(13) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13* (8), 952–962.
(14) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103* (19), 8577–8593.
(15) Luo, R.; David, L.; Gilson, M. K. Accelerated Poisson–Boltzmann Calculations for Static and Dynamic Systems. *J. Comput. Chem.* **2002**, *23* (13), 1244–1253..
(16) Kumari, R.; Kumar, R.; Lynn, A. G_mmpbsa—A GROMACS Tool for High-Throughput MM-PBSA Calculations. *J. Chem. Inf. Model.* **2014**, *54* (7), 1951–1962.
(17) Homeyer, N.; Gohlke, H. Free Energy Calculations by the Molecular Mechanics Poisson−Boltzmann Surface Area Method. *Mol. Inform.* **2012**, *31* (2), 114–122.
(18) Jana, A. K.; Tiwari, M. K.; Vanka, K.; Sengupta, N. Unraveling Origins of the Heterogeneous Curvature Dependence of Polypeptide Interactions with Carbon Nanostructures. *Phys. Chem. Chem. Phys.* **2016**, *18* (8), 5910–5924.