Supporting Information for:

The Rational Discovery of a Tau Aggregation Inhibitor

David W. Baggett, Abhinav Nath

Department of Medicinal Chemistry, University of Washington

Figure S1: ReSA explores relevant conformational space more efficiently than canonical molecular dynamics.

a) In ReSA, rates of change are increased during periods of high temperature, as indicated by higher RMSD values when comparing to structures 50ps prior in the simulation.

b) Root-mean-squared fluctuation (RMSF) values of atoms in ReSA simulation are greater than or equal to the RMSF of those atoms in canonical MD of the same length generated from the same starting structure, indicating that in ReSA the protein is more dynamic.
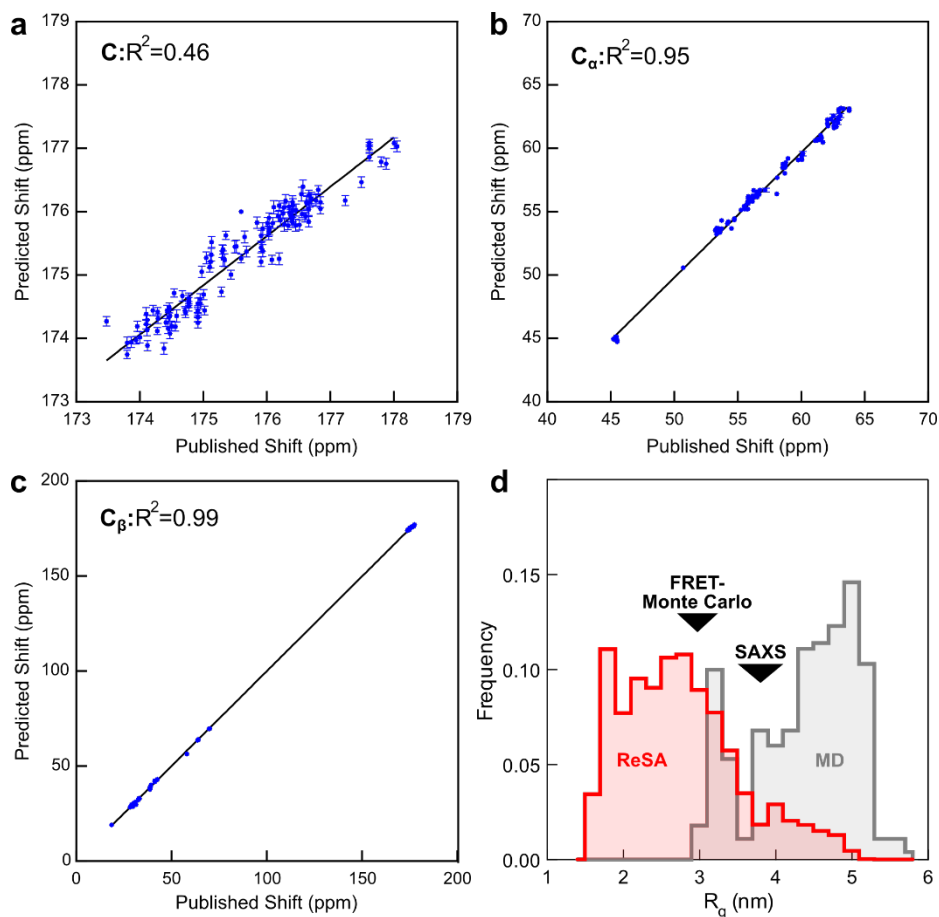
Figure S2: Experimental validation of the ReSA-derived conformational ensemble.

[13]C chemical shifts of carbonyl C (**a**), Cα (**b**), and Cβ (**c**) predicted using SPARTA+[81] from the ReSA ensemble align well with published values.[82]

d) The conformational ensemble generated by ReSA (500 ns) more closely recapitulates radius of gyration (R_g) values derived from FRET-constrained Monte Carlo simulations and small-angle X-ray scattering[74] than that generated by canonical molecular dynamics (50 ns).
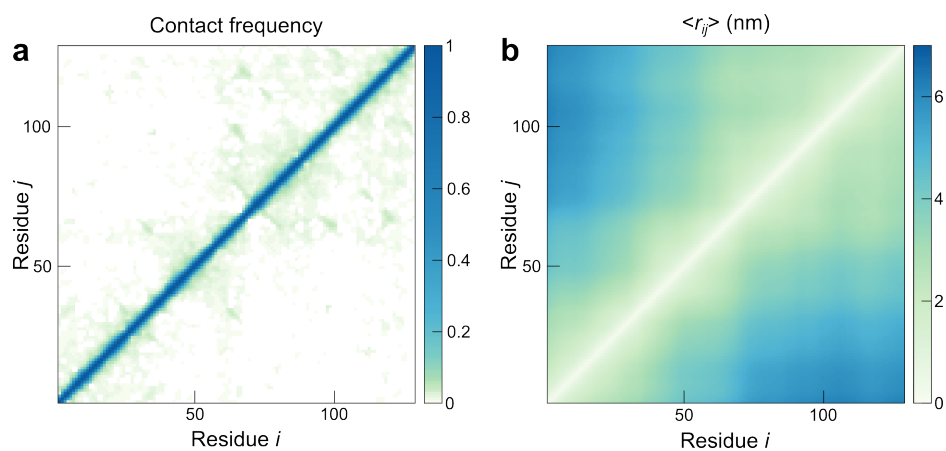
Figure S3: Inter-residue contacts in the ReSA-derived conformational ensemble

a) Frequency of ≤ 0.6 nm contacts between residues and (**b**) average inter-residue distances (nm) for the entire ReSA conformational ensemble show that while long-range contacts do occur, they are rare relative to contacts within 10 residues.
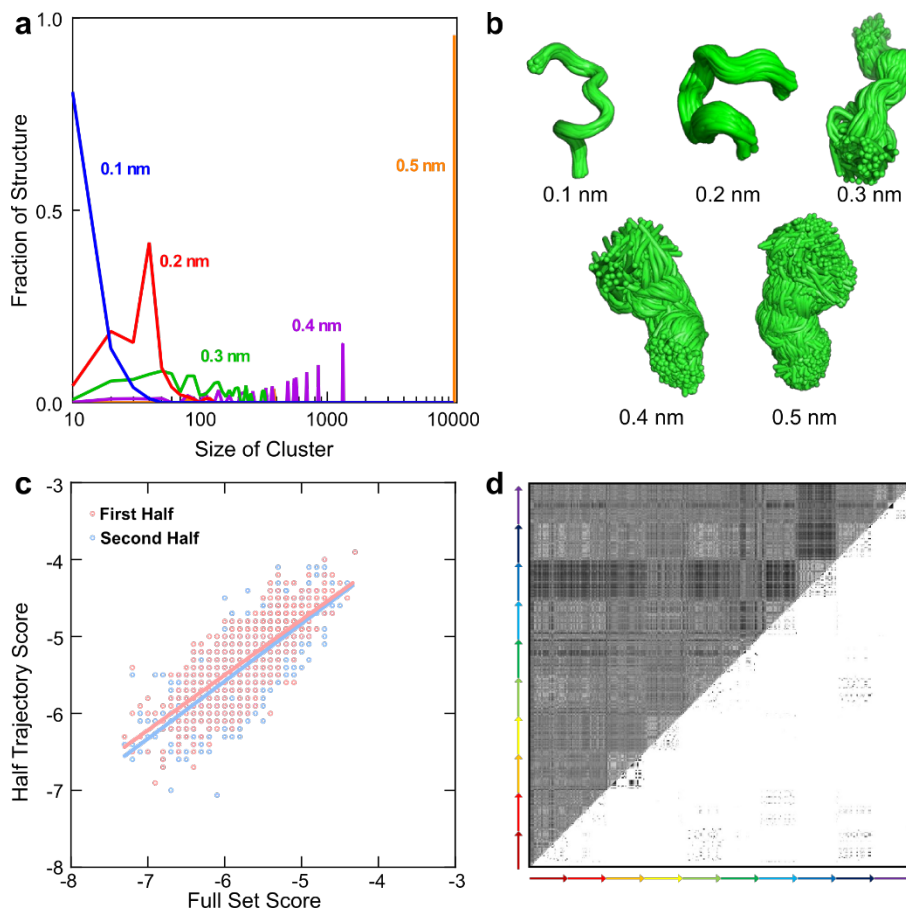
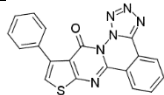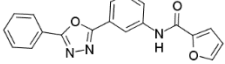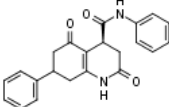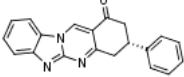**Figure S4: Cluster analysis and demonstrations of convergence**
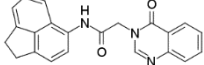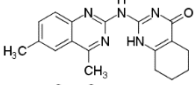
a) Cluster population distributions vary greatly with the RMSD cutoff. Clusters generated with overly large RMSD cutoffs are too large and poorly defined. Cutoffs that are too small do not generate well-populated clusters.

b) Representative clusters with RMSD cutoffs between 0.1 and 0.5 nm.

c) Docking targets generated from either the first or the second half of the conformational ensemble of tau4RD yielded similar docking as targets generated from the entire set of conformations, further suggesting that the ReSA simulations have converged for our purposes. Pearson correlation coefficients are 0.78 and 0.76 for the first and second half targets, respectively.
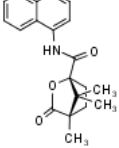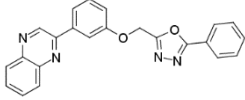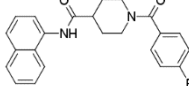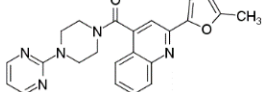
d) An RMSD matrix of Segment 1 compares each structure of all 10 independent ReSA simulations (denoted with different colored arrows) against every other structure. The top half shows structural similarity with darker pixels signifying higher RMSD values. On the bottom half, black pixels represent when two structures fall into the same cluster based on an all-atom RMSD criteria of 0.2 nm. The off-diagonal points on this half show that many clusters are sampled in non-contiguous parts of a given simulation as well as in multiple independent simulations.

Table S1. Compounds selected for in vitro screening, based on a combination of docking score, chemical diversity and sequence coverage.

Compound #	Chembridge ID	PubChem CID	Best Score	Best Target (Segment-Cluster)	Structure
1	7579671	1184213	-8	3-1	
2	6399586	659810	-7.5	21-6	
3	7999316	2988410	-7.4	21-6	
4	7987291	2983350	-7.3	3-1	
5	7358814	1077875	-7.3	21-7	
6	7282131	707162	-6.9	16-9	
7	6952140	3151892	-6.9	19-9	
8	9011432	2994718	-6.9	14-2	
9	6983997	1072807	-6.9	14-2	
10	7233320	1001781	-6.7	10-7	
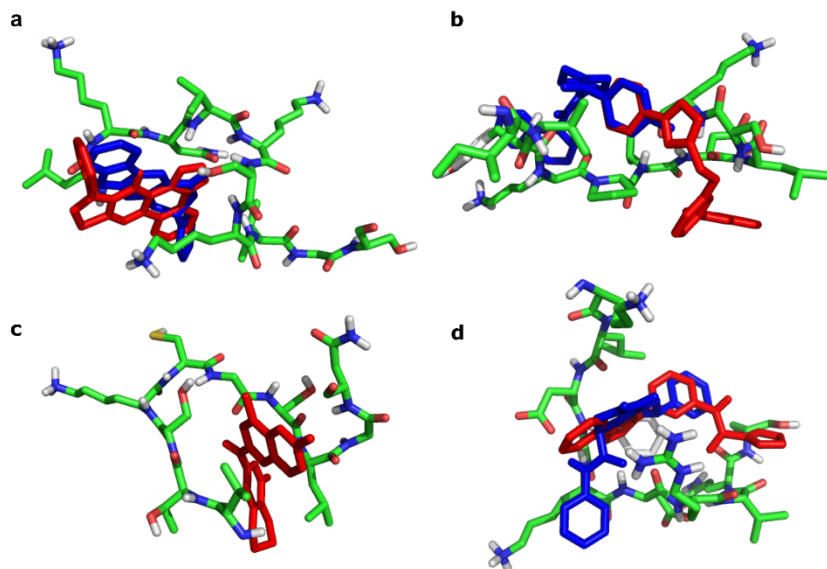
**Figure S5: Examples of compound docking positions on favored targets.**

a) Compound 1 (Red) and Compound 4 (Blue) docked against Segment 3, Cluster 1

b) Compound 8 (Red) and Compound 9 (Blue) docked against Segment 14, Cluster 2

c) Compound 6 (Red) docked against Segment 16, Cluster 9

d) Compound 2 (Red) and Compound 3 (Blue) bound to Segment 21, Cluster 6
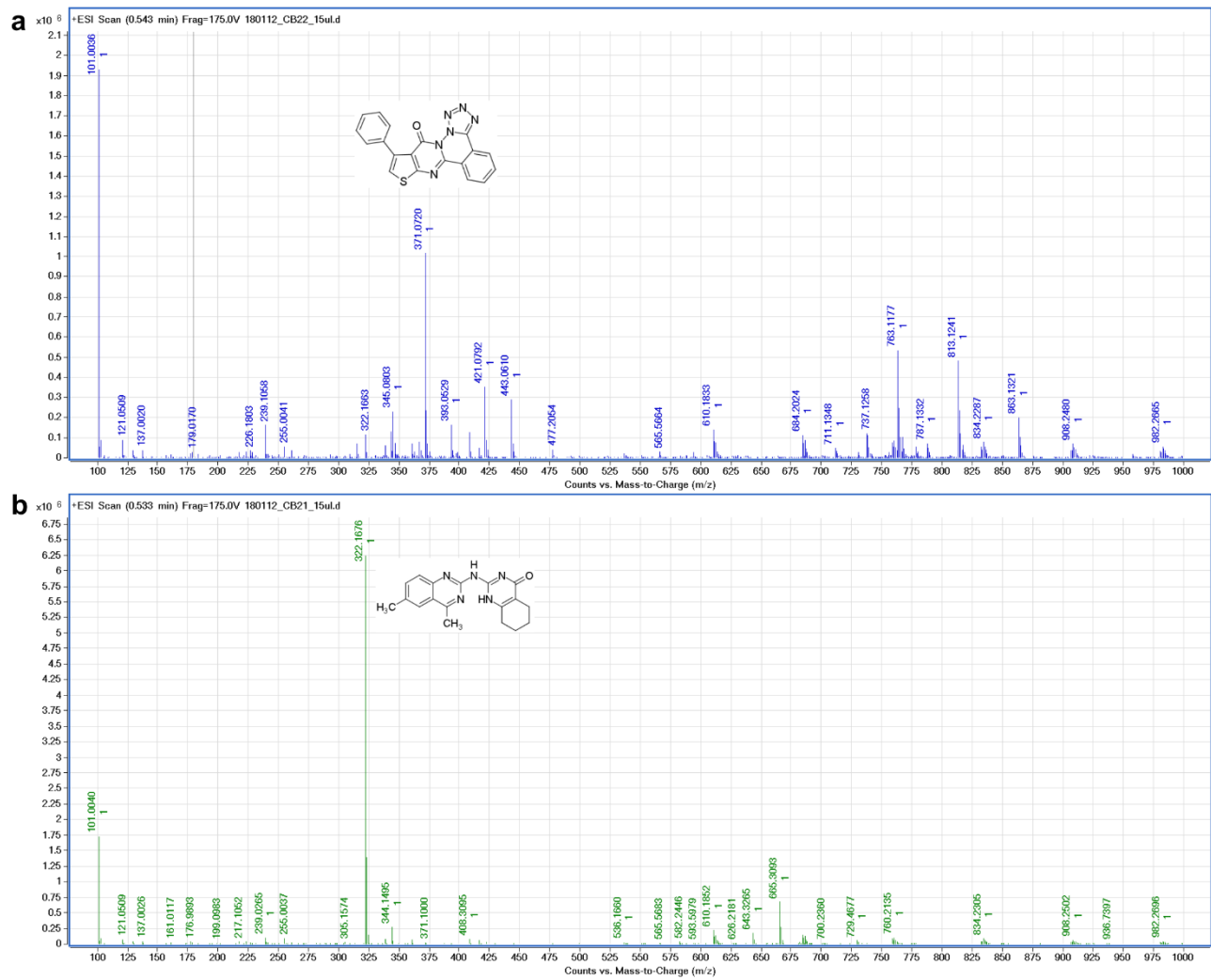
**Figure S6: Validation of small molecules**

Mass spectrometry confirms the identity of Compound 1 (**a**) and Compound 6 (**b**)
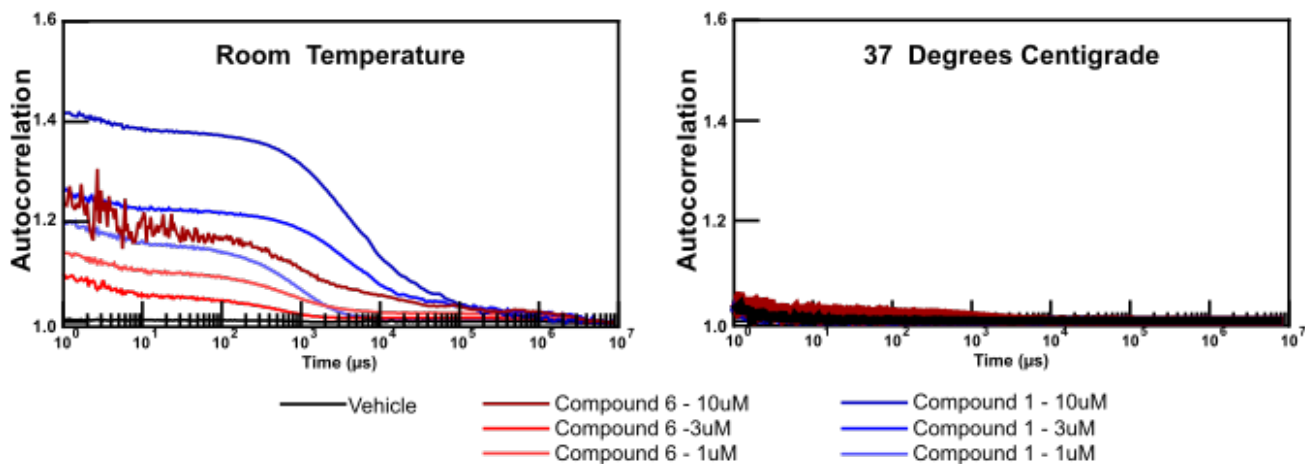
Figure S7: Assessment of compound solubility

DLS shows that at room temperature, particles are present at micromolar concentrations of both compound 1 and compound 6. However, these particles disperse at 37°C, the temperature used for all aggregation assays.
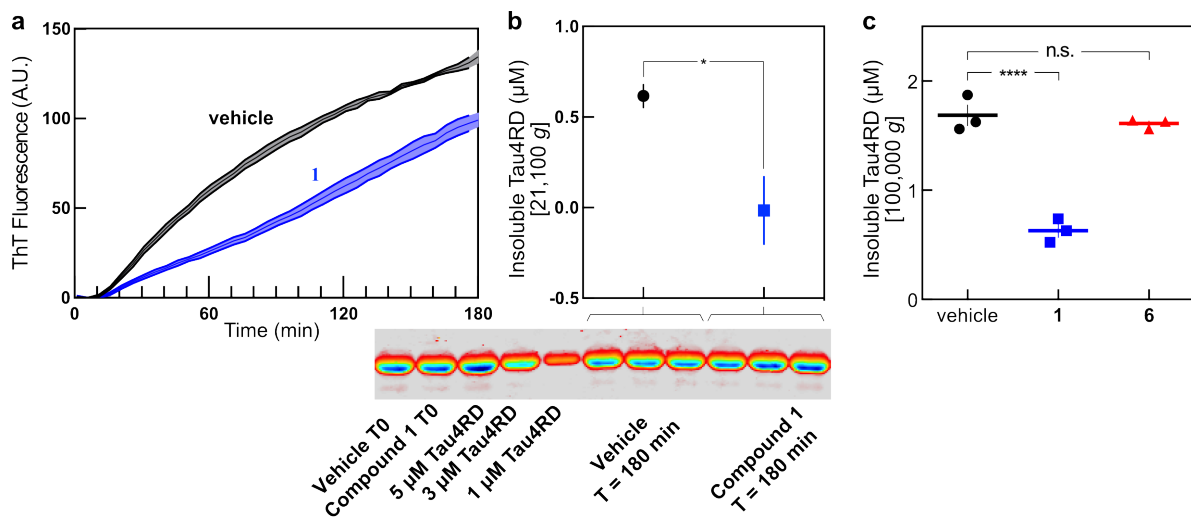
Figure S8: Label-free confirmation of Compound 1 activity

After 3 hours of aggregation monitored by ThT fluorescence (a), centrifugation at 21,100 g followed by SDS-PAGE and gel densitometry (b) confirms that Compound **1** delays aggregation. This agrees well with intrinsic Tyr fluorescence measurements of tau4RD concentration in the supernatant following centrifugation at 21,100 g (see Figure 4d) or at 100,000 g (panel c).
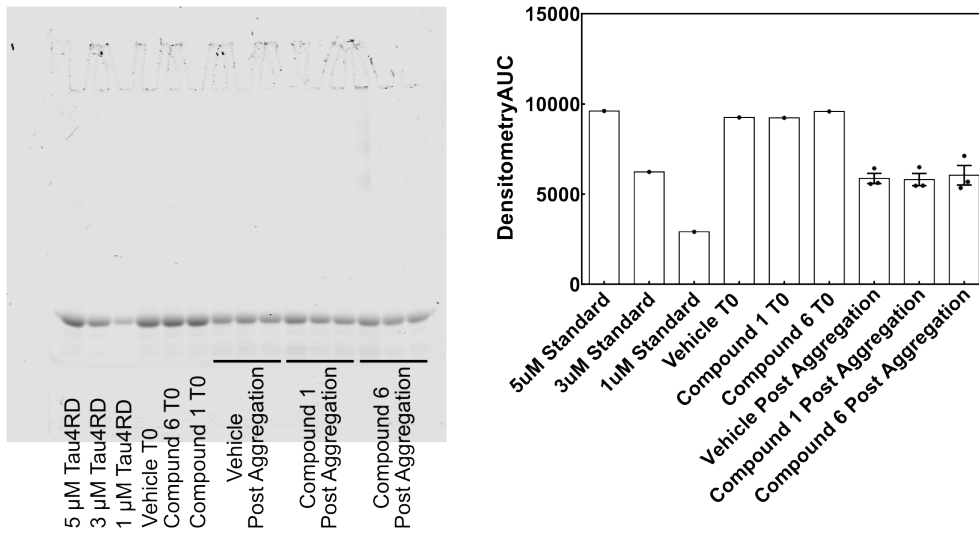
Figure S9: Compound effects on extent of aggregation

After 96 hours of aggregation, SDS-PAGE followed by gel densitometry indicates that the amount of aggregated tau is similar in the presence or absence of **1** or **6**. This suggests that even though **1** affects aggregation kinetics, it does not alter the amount of tau4RD fibrils formed at equilibrium.