Supporting Information

Constant pH molecular dynamics reveals pH-modulated binding of two small-molecule BACE1 inhibitors

Christopher R. Ellis[1,§], Cheng-Chieh Tsai[1,§], Xinjun Hou[2], and Jana Shen[1,*]

[1]Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, MD.
[2]Neuroscience Worldwide Medicinal Chemistry, Pfizer Worldwide Research and Development, Cambridge, MA.
[§]Joint first authors.
[*]Corresponding author. Phone: (410) 706-4187; Fax: (410) 706-5017; E-mail: jshen@rx.umaryland.edu.

1 Methods and Protocols

Structure preparation and force field parameters

The X-ray crystal structures with PDB ID 4FRS[1] (obtained at pH 7) and 4YBI[2] (obtained at pH 7.4) were used as the starting configurations for simulations of BACE1 in complex with inhibitor **1** and **2**, respectively. All crystal structure waters were retained, and hydrogen atoms were added using the HBUILD facility in CHARMM (version C37b[3]). Each system was solvated in an octahedral box with at least 10 Å between the box edge and the protein. The complex with **1** has a total of 53,498 atoms including 15,779 water atoms, while the complex with **2** has a total of 53,409 atoms including 15,756 water atoms.

BACE1 was represented by the CHARMM22/CMAP all-atom force field,[4,5] and water was represented by the modified TIP3P water model.[3] The force field parameters for inhibitors were optimized following the CHARMM General Force Field (CGenFF) protocol.[6] The initial parameters were estimated using ParamChem[7,8]. Inhibitor geometries were optimized using Gaussian03[9] with the MP2/6-31G* basis set. Partial charges were optimized by matching quantum mechanical (QM) and molecular mechanical (MM) water interaction energies. QM interaction energies were calculated at the HF/6-31G level. The bond lengths and angles were adjusted according to the QM geometries, and the bonded force constants were optimized by matching the molecular vibrational frequency between QM and MM calculations. Finally, the dihedral parameters were optimized by matching QM and MM dihedral potential energy scans. The partial charges are presented in Tables S1 and S2.

Continous constant pH molecular dynamics

All simulations were performed using the PHMD module in CHARMM (version C37b[3]), which implements the generalized-Born (GB)[10,11] and hybrid-solvent versions[12] of the continuous constant pH molecular dynamics (CpHMD) method. Based on the λ-dynamics approach for free energy simulations[13], CpHMD utilizes a set of auxiliary coordinates to represent the changes in protonation states, which are simultaneously sampled with the conformational degrees of freedom.[10,11] An enhanced sampling protocol, e.g., temperature-[14] or pH-based[12] replica-exchange scheme is necessary to enable rapid convergence in protonation-state sampling and p$K_{\rm a}$ convergence. We refer the interested readers to the recent reviews of the development and applications of CpHMD.[15,16]

Dummy hydrogen atoms were added to the titratable sites in the protein (Asp, Glu and His sidechains) and inhibitors (see Figure 1B and 1C of the main text, nitrogens circled in blue) in preparation for CpHMD simulations. Detailed protocol can be found in CHARMM documentation

(http://www.charmm.org/). The solvated system was first energy minimized using the steepest descent method followed by the Newton-Raphson algorithm. The system was then gradually heated over the course of 120 ps from 100 K to 300 K with the protein heavy atoms restrained using a harmonic potential with a force constant of 5 kcal/mol/Å^2 . Following heating, the system was equilibrated for 180 ps under harmonic restraints, where the force constant was gradually reduced from 5 (40 ps), to 1 (40 ps), and 0.1 kcal/mol/Å^2 (100 ps). Lastly, the system was equilibrated for 100 ps without restraints. In the energy minimization and equilibration steps, the pH was set to 7.

Production simulations were performed with the hybrid-solvent CpHMD method with pH replica exchange protocol[12]. Each pH replica was simulated in the NPT ensemble at a temperature of 300 K and pressure of 1 atm. Temperature was maintained using a modified Hoover thermostat method[17], while pressure was maintained using the Langevin piston coupling method[18]. The van der Waals interactions were smoothly switched to zero between 10 and 12 Å. The particle mesh Ewald method[19] was used to calculate long-range electrostatics, with a real-space cutoff of 12 Å and a sixth-order interpolation with 2.7 Å^{-1} grid spacing. The SHAKE algorithm was used to constrain bonds involving hydrogen to enable a 2-fs timestep. An electrostatic solvation free energy calculation using the GBSW generalized-Born implicit-solvent model[20] was invoked every 5 MD steps to update the titration coordinates. In the GBSW calculation, the default settings were used, consistent with our previous work[12]. The GBSW input radii for the protein were taken from Chen et al.[21]. The input radii for carbon and sulfur atoms were set to 2.0 and 2.3 Å.[22] Each system was simulated 26 ns using 20 pH replicas in the pH range 1.3-8.0. The final 18 ns of simulation was used for analysis. The pH conditions of the 20 replicas of each simulation are presented in Figure S1. The average exchange ratio was 36%.

2 Supplementary Tables and Figures

Table S1: Calculated pK_{a} values from the simulations of BACE1 complexed with inhibitor **1** and **2**.

Asp Res	**1** pK_{a}	**2** pK_{a}	Glu Res	**1** pK_{a}	**2** pK_{a}	His Res	**1** pK_{a}	**2** pK_{a}
4	5.5	5.3	1	3.7	3.5	45	7.0	7.1
32	2.5	3.7	17	3.5	3.6	49	7.1	7.0
62	2.6	2.5	77	4.4	4.3	89	8.3	8.0
83	3.9	3.5	79	3.8	3.7	145	6.6	6.7
106	2.8	2.8	104	3.7	3.7	181	7.2	7.1
130	4.1	3.8	116	5.3	5.2	360	4.9	2.8
131	3.3	3.3	125	4.0	3.9	362	>8	>8
138	<1.3	<1.3	134	3.3	3.1			
180	3.8	3.8	165	4.4	4.2			
212	1.4	1.8	196	2.9	3.0			
216	2.8	2.8	200	5.2	5.3			
223	3.6	3.3	207	2.1	2.1			
228	2.3	2.0	219	3.2	3.3			
259	3.0	3.0	242	3.1	3.2			
311	3.6	4.0	255	4.5	4.3			
317	5.3	4.7	265	4.2	4.2			
318	5.6	5.7	290	2.9	2.9			
346	4.0	4.0	310	2.9	2.9			
363	2.2	3.4	339	5.1	5.2			
378	4.9	5.0	364	4.1	4.2			
381	3.4	3.7	371	3.1	3.2			
			380	3.5	3.5			
Inhibitor	**1**	**2**						
	2.9	4.2						

Column names **1** and **2** indicate BACE1 complexed with **1** and **2**, respectively.

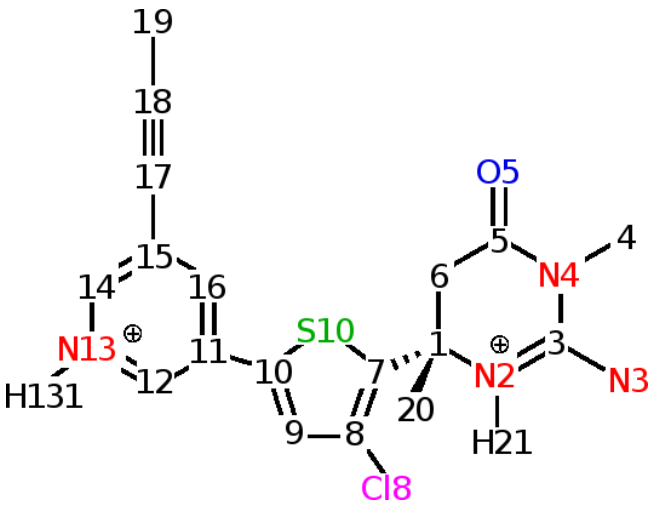

Table S2: Partial atomic charges of inhibitor **1**.

Ring 3			Ring 2			Ring1		
Name	Type	Charge	Name	Type	Charge	Name	Type	Charge
C1	CG301	0.38	C7	CG2R51	0.09	C11	CG2R62	0.097/-0.004
N2	NG2P1	-0.63	C8	CG2R51	0.03	C12	CG2R62	0.160/ 0.177
H21	HGP2	0.36				H121	HGR63	0.190/ 0.122
C3	CG2N1	0.70	Cl8	CLGR1	-0.13	N13	NG2R61	-0.512/-0.602
						H131	HGP2	0.451/ 0.000
N3	NG2P1	-0.64	C9	CG2R51	-0.25	C14	CG2R62	0.207/ 0.178
H31	HGP2	0.36	H91	HGR51	0.23	H141	HGR63	0.193/ 0.122
H32	HGP2	0.36						
N4	NG2P1	-0.09	C10	CG2R51	0.12	C15	CG2R62	0.226/ 0.100
C4	CG334	-0.16	S10	SG2R50	-0.09	C16	CG2R62	-0.093/-0.104
H41	HGA3	0.09				H161	HGR63	0.192/ 0.115
H42	HGA3	0.09						
H43	HGA3	0.09						
C5	CG2O1	0.17				C17	CG1T1	-0.097/-0.103
O5	OG2D1	-0.28						
C6	CG321	0.02				C18	CG1T1	-0.094/-0.080
H61	HGA2	0.09						
H62	HGA2	0.09						
C20	CG331	-0.27				C19		-0.190/-0.191
H201	HGA3	0.09				H191		0.090/ 0.090
H202	HGA3	0.09				H192		0.090/ 0.090
H203	HGA3	0.09				H193		0.090/ 0.090
Total		+1.00			0.00			+1.00/ 0.00

N13 from ring 1 is the titratable nitrogen circled in blue in Figure 1 of the main text.

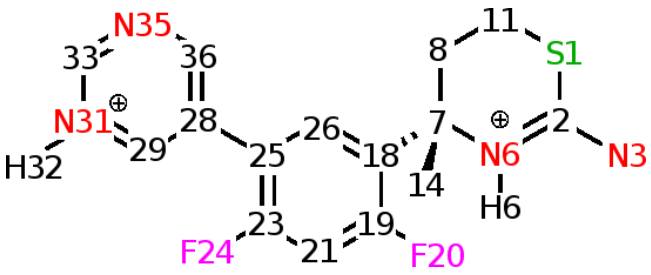

Table S3: Partial atomic charges of inhibitor **2**.

	Ring 3			**Ring 2**			**Ring 1**	
Name	Type	Charge	Name	Type	Charge	Name	Type	Charge
S1	SG311	-0.33	C18	CG2R61	0.07	C28	CG2R76	0.19/ 0.00
C2	CG2N2	0.66						
N3	NG2P1	-0.88	C19	CG2R66	0.22	C29	CG2R62	0.24/ 0.29
H4	HGP2	0.44	F20	FGR1	-0.21	H30	HGR63	0.19/ 0.12
H5	HGP2	0.44						
N6	NG2P1	-0.63	C21	CG2R61	-0.24	N31	NG2R61	-0.53/-0.73
H6	HGP2	0.40	H22	HGR62	0.19	H32	HGP2	0.46/ 0.00
C7	CG301	0.42						
C8	CG321	-0.13	C23	CG2R66	0.28	C33	CG2R62	0.38/ 0.50
H9	HGA2	0.09	F24	FGR1	-0.21	H34	HGR63	0.23/ 0.14
H10	HGA2	0.09						
C11	CG321	0.25	C25	CG2R67	-0.03	N35	NG2R62	-0.49/ -0.73
H12	HGA2	0.09						
H13	HGA2	0.09						
C14	CG331	-0.27	C26	CG2R61	-0.28	C36	CG2R62	0.11/ 0.29
H15	HGA3	0.09	H26	HGR61	0.21	H37	HGR63	0.22/ 0.12
H16	HGA3	0.09						
H17	HGA3	0.09						
Total		+1.00			0.00			+1.00/ 0.00

N31 from ring 1 is the titratable nitrogen circled in blue in Figure 1 of the main text.

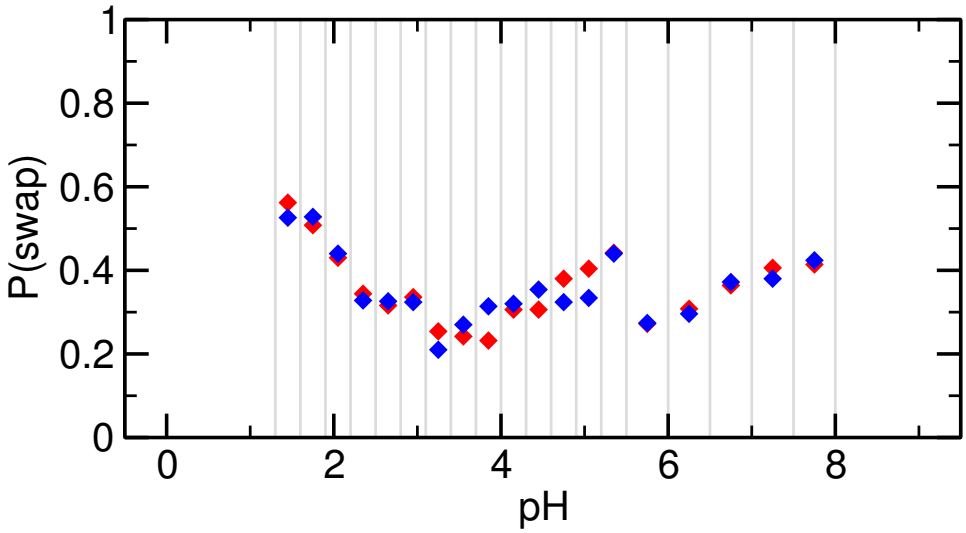

Figure S1: **Acceptance rate of replica-exchange attempts.** The blue and red diamonds represent the probabilities of exchange between adjacent replicas from the simulations of BACE1 complexed with inhibitor **1** and **2**, respectively. The average exchange rates across replicas in each simulation is 36$\pm$9 %. The vertical grey lines represent the 20 pH replicas from each simulation. The replicas are evenly spaced with an increment of 0.3 pH units from pH 1.3 to 5.5 and from pH 5.5-8.0 the increment is 0.5 pH units.

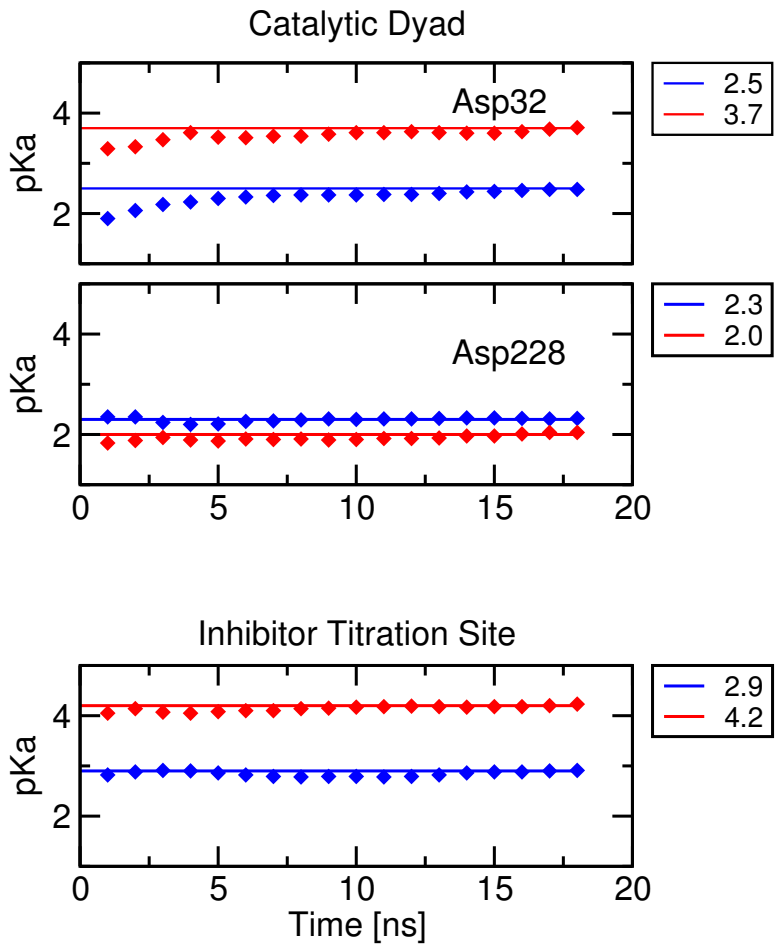

Figure S2: **Convergence of the calculated pK_a's of the catalytic dyad of BACE1 and inhibitors 1 and 2**. The cumulatively calculated pK_a of Asp32 (top), Asp228 (middle), and each inhibitor (bottom) as a function of the simulation time. In each panel, results from simulation of **1** and **2** in complex with BACE1 are presented in blue and red, respectively. The legend reports the final calculated pK_a. The calculated pK_a for the two residues remain nearly constant for the final 10 ns.

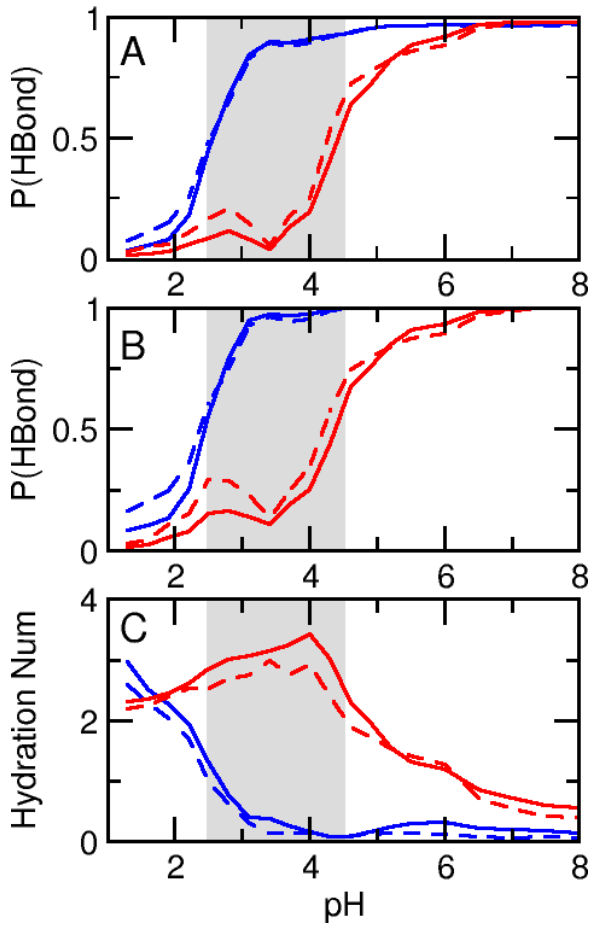

Figure S3: **Convergence of the inhibitor-BACE1 interactions and dyad hydration**. Panels A, B, and C present the hydrogen bond occupancies of exocyclic amine$\cdots$Asp228 and endocyclic amine$\cdots$Asp32, and hydration number of Asp32 as a function of pH from the first half (dashed) and the total simulation (solid). The simulations in the presence of **1** and **2** are shown in blue and red, respectively.

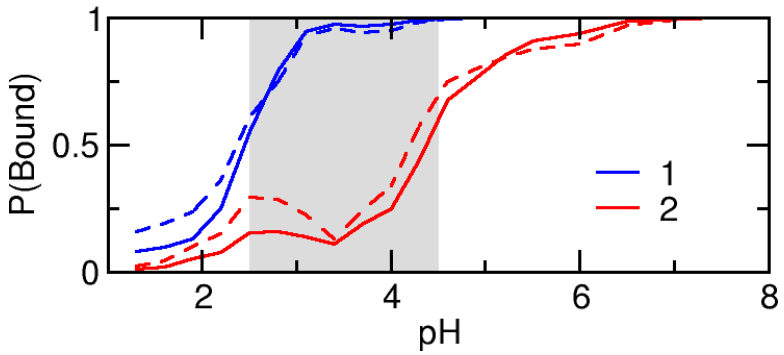

Figure S4: **Convergence of the fraction of bound configurations.** Population of the bound configurations as a function of pH from the first half (dashed) and the total simulation (solid). The simulations in the presence of **1** and **2** are shown in blue and red, respectively.

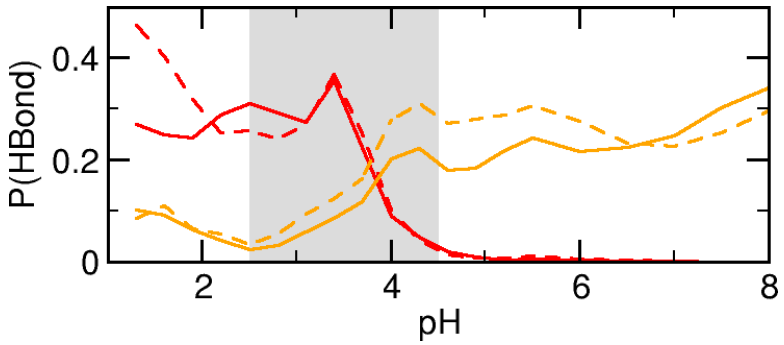

Figure S5: **Convergence of the Thr232 interactions with the 10s loop and inhibitor 2**. Population of the hydrogen bond occupancies of Thr232 with Ser10 of the 10s loop (orange) and inhibitor **2** as a function of pH from the first half (dashed) and the total simulation (solid).

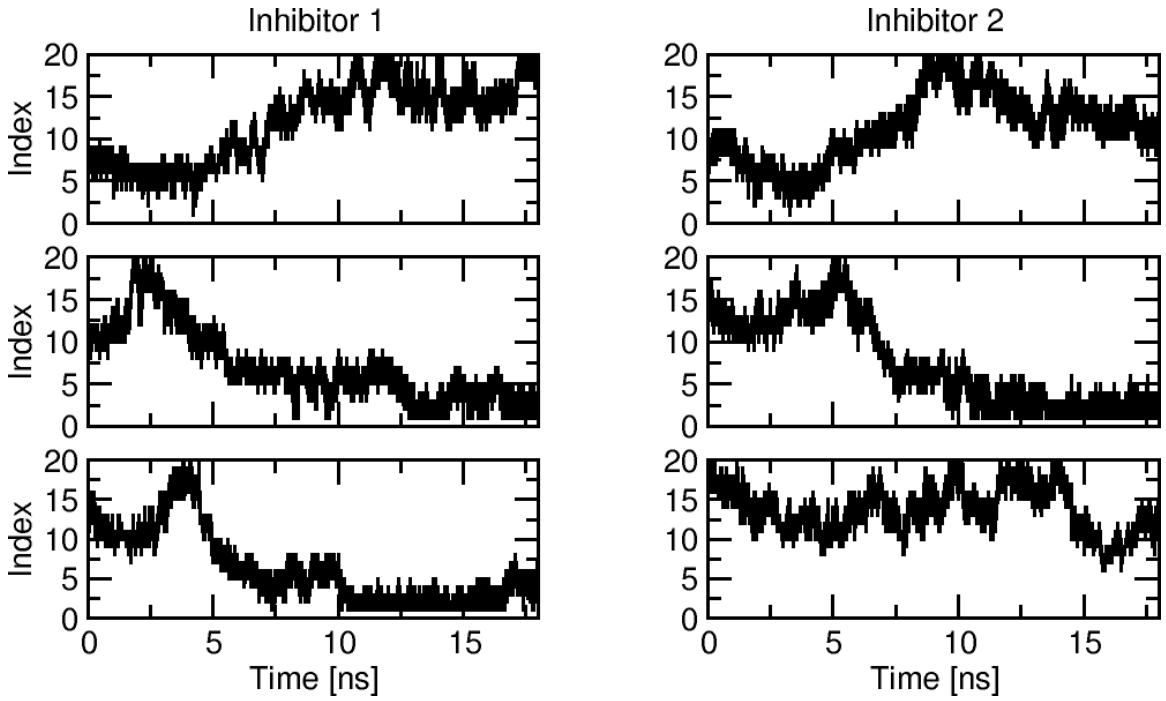

Figure S6: **Replica walk through pH space.** Time series of the replica walk through pH space. Three randomly chosen replicas are shown for simulations in the presence of inhibitor **1** (left) and **2** (right). The index corresponds to 20 pH conditions.

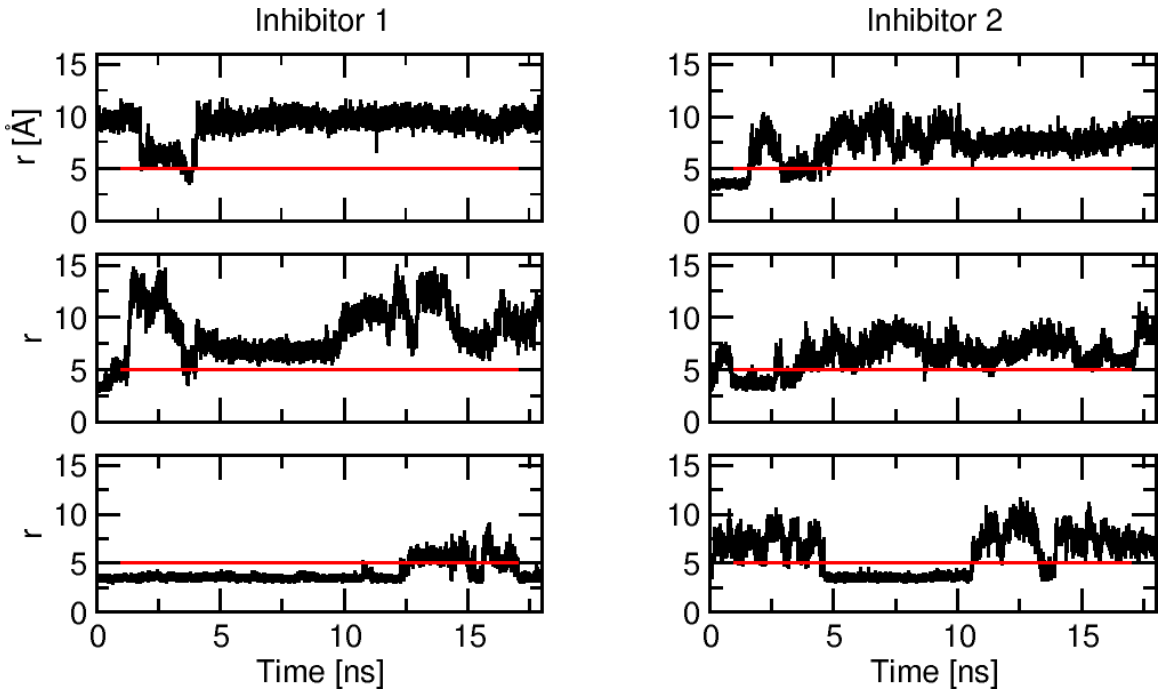

Figure S7: **Binding and unbinding events of individual replicas.** These panels monitor the distance between the endocyclic nitrogen and the CG of Asp32 of individual replicas as they walk through pH space in simulations of BACE1 in complex with inhibitors **1** (left) and **2** (right). Note, in both simulations the inhibitor becomes unbound at low pH. A red line is draw at 5 Å to guide the eye.

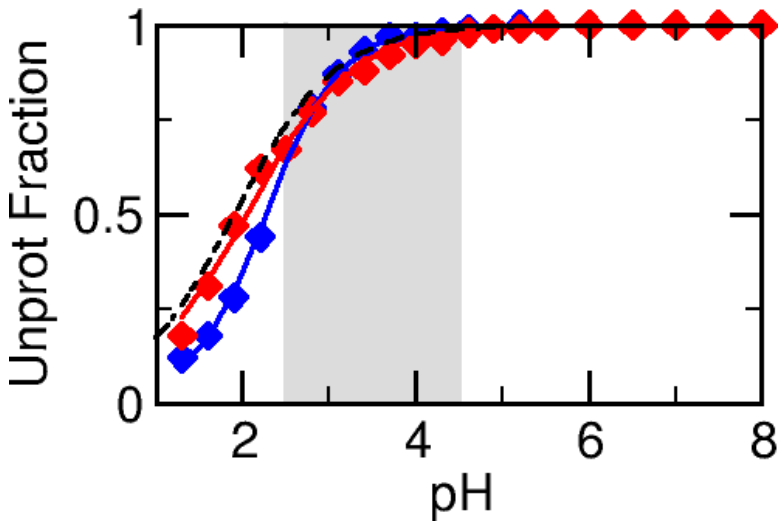

Figure S8: **Titration curve of Asp228 from the catalytic dyad**. Unprotonated fractions of Asp228 at different pH. Simulations of BACE1 complexed with **1** and **2** are shown in blue and red, respectively. The curves are the best fits to the Hill equation. The black dashed curve is from the *apo* simulation.

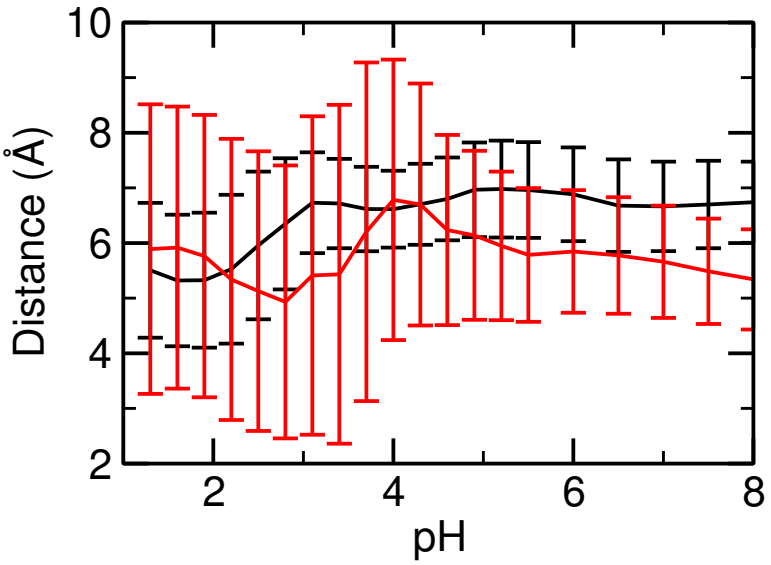

Figure S9: **Interaction between Thr232 and inhibitors**. The average and root-mean-squared fluctuation of the distance from Thr232 (hydroxyl oxygen) to the pyridine/pyrimidine nitrogen of the inhibitor as a function of pH. Black and red represent the data from simulations with inhibitor **1** and **2**, respectively. In contrast to inhibitor **2**, inhibitor **1** is never within hydrogen bond distance to Thr232.

References

1. Stamford, A. W. et al. *ACS Med. Chem. Lett.* **2012**, *3*, 897–902.
2. May, P. C. et al. *J. Neurosci.* **2011**, *31*, 16507–16516.
3. Brooks, B. R. et al. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
4. MacKerell Jr., A. D. et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
5. MacKerell, Jr., A. D.; Feig, M.; Brooks, III, C. L. *J. Am. Chem. Soc.* **2004**, *126*, 698–699.
6. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell Jr., A. D. *J. Comput. Chem.* **2010**, *31*, 671–690.
7. Vanommeslaeghe, K.; Jr., A. D. M. *J. Chem. Inf. Model.* **2012**, *52*, 3144–3154.
8. Vanommeslaeghe, K.; Raman, E. P.; Jr., A. D. M. *J. Chem. Inf. Model.* **2012**, *52*, 3155–3168.
9. Frisch, M. J. et al. Gaussian 03, Revision D.02. Gaussian, Inc., Wallingford, CT, 2004.
10. Lee, M. S.; Salsbury, Jr., F. R.; Brooks III, C. L. *Proteins* **2004**, *56*, 738–752.
11. Khandogin, J.; Brooks III, C. L. *Biophys. J.* **2005**, *89*, 141–157.
12. Wallace, J. A.; Shen, J. K. *J. Chem. Theory Comput.* **2011**, *7*, 2617–2629.
13. Kong, X.; Brooks, III, C. L. *J. Chem. Phys.* **1996**, *105*, 2414–2423.
14. Khandogin, J.; Brooks III, C. L. *Biochemistry* **2006**, *45*, 9363–9373.
15. Wallace, J. A.; Shen, J. K. *Methods Enzymol.* **2009**, *466*, 455–475.
16. Chen, W.; Morrow, B. H.; Shi, C.; Shen, J. K. *Mol. Simulat.* **2014**, *40*, 830–838.
17. Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.
18. Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
19. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
20. Im, W.; Lee, M. S.; Brooks III, C. L. *J. Comput. Chem.* **2003**, *24*, 1691–1702.
21. Chen, J.; Im, W.; Brooks III, C. L. *J. Am. Chem. Soc.* **2006**, *128*, 3728–3736.
22. Nina, M.; Simonson, T. *J. Phys. Chem. B* **2002**, *106*, 3696–3705.