Supporting Information

Proton-Coupled Conformational Allostery Modulates the Inhibitor Selectivity for β-secretase

Robert C. Harris,[†] Cheng-Chieh Tsai,[†,‡] Christopher R. Ellis,[†] and Jana Shen[*,†]

† *Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, MD 21201.*

‡ *Joint first author.*

E-mail: jana.shen@rx.umaryland.edu

Methods and Protocols

The absolute binding free energies $\Delta G(\mathrm{pH}^{\mathrm{ref}})$ of the inhibitor LY112811376 to BACE1 and CatD were computed with the double decoupling scheme as described in Boresch et al.[1] In all double decoupling calculations, Asp and Glu sidechains as well as the inhibitor were fixed in the charged state (the titratable site of the inhibitor is neutral).[2,3] For BACE1, His sidechains were fixed in the neutral state, which, when considering the calculated pK_{a}'s (Table S1), corresponds to pH $\sim$9.5 and higher. However, since there are no pK_{a} shifts above pH 7, ΔG is the same as at pH 8. CatD is known to undergo a large conformational transition that relocates the N-terminal residues to the active site at high pH.[4] Although this may not occur in the limited simulation time, to avoid potential drift in the structure, His sidechains were fixed in the charged state, which, when considering the calculated pK_{a}'s (Table S2), corresponds to pH $\sim$6. However, since there are no pK_{a} shifts above pH 6, ΔG is the same as at pH 8.

All FEP calculations were set up with the VMD visualization program[5] and performed with the NAMD molecular dynamics engine.[6] The force field parameters of the inhibitor were obtained previously.[2] The proteins were modeled with the CHARMM22/CMAP force field.[7,8] Unless otherwise stated, all FEP simulations used Langevin thermostats and barostats to maintain constant temperatures and pressures of 300 K and 1 atm. Particle mesh Ewald electrostatics was used with a real-space cutoff of 12 Å. SHAKE was used to restrain all bonds containing a hydrogen atom. All FEP simulations used a 2 fs time step and the default soft core in NAMD. Restraints were maintained with the collective variable methods in NAMD. The final error was estimated by combining with standard error propagation the errors in the following components of the binding free energy.

Inhibitor desolvation free energy. To compute the desolvation free energy of the inhibitor ($\Delta G^{\mathrm{L}}_{\mathrm{dsolv}}$), it was placed in a water box with a 15 Å cushion in each direction using the solvate plugin in the VMD program.[5] A single chloride atom was then added to the

box with the autoionize plugin in VMD. The system was then minimized for 500 steps with NAMD. Copies of this system were then used as the starting points for FEP calculations at many fixed λ values. At each λ value the system was equilibrated by heating it from 25 K to 300 K in increments of 25 K with 1000 steps between each increment. Production runs of 1 ns were then run at each λ value with frames recorded every 100 steps for analysis. Simulations were run at the following λ values: 0, 0.032, 0.064, 0.096, 0.128, 0.16, 0.192, 0.224, 0.256, 0.288, 0.32, 0.352, 0.384, 0.416, 0.448, 0.48, 0.512, 0.544, 0.576, 0.608, 0.64, 0.672, 0.704, 0.736, 0.768, 0.8, 0.832, 0.864, 0.896, 0.928, and 0.96.

The free energy difference across a λ window could then be estimated using FEP:

$$\Delta G_{i\to j} = -kT \ln \left\langle \exp\left[-\beta U_{i\to j}\right]\right\rangle_i , \tag{1}$$

where $\Delta G_{i\to j}$ is the free energy of going from λ_i to λ_j k is Boltzmann's constant, T is the temperature, $\beta = 1/kT$, and $\langle\cdots\rangle_i$ means an expectation (average) value taken over the frames from the simulation at λ_i.

Note that $\Delta G_{i\to j}$ can be computed from either λ_i or λ_j. When $j = i + 1$, forward and backward perturbation estimates of the free energy across a λ window can be obtained. To estimate the errors in the desolvation free energy the difference, δ_k, between the forward and backward FEP estimates in each window was computed and the square root of the sum of the squares of these quantities, $\delta = \sqrt{\sum \delta_k^2}$, was used to estimate the error in the free energy.

Restraint free energy. Following the methods in Boresch et al.,[1] three atoms in the inhibitor had to be selected to define the positional restraints. For each complex three atoms in the receptor were chosen as well. As in the calculation of $\Delta G^{\mathsf{L}}_{\mathsf{dsolv}}$, the complex was then placed in a water box and the system was neutralized by the addition of counterions, energy minimized and equilibrated. The system was then simulated for 5 ns. The equilibrium values of the restraints were taken from the final configuration of this simula-

tion. All angle and dihedral restraints used a force constant of 0.1 kcal/mol/deg^2, and all distance restraints used a force constant of 10 kcal/mol/Å^2. The free energy ($\Delta G^{\text{vac}}_{\text{rest}}$) of imposing the restraints on the inhibitor in vacuum were computed with Equation 14 from Boresch et al.[1]

To compute the free energy ($\Delta G^{\text{comp}}_{\text{rest}}$) of removing the restraints from the inhibitor in the solvated complex, the restraint constants were scaled by λ and the methods used to compute $\Delta G^{\text{inhib}}_{\text{desolv}}$ and its error were used. The initial structures of the complex came from the simulation used to determine the equilibrium value of the restraints. Simulations of 1 ns were run at the following λ values: For BACE1 $\lambda = 0$, 0.00390625, 0.0078125, 0.015625, 0.03125, 0.0625, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 0.875, and 1 and for CatD $\lambda = 0$, 0.00390625, 0.0078125, 0.01171875, 0.013671875, 0.015625, 0.01953125, 0.0234375, 0.02734375, 0.03125, 0.0390625, 0.046875, 0.0625, 0.09375, 0.125, 0.25, 0.3125, 0.375, 0.5, 0.625, 0.75, 0.875, and 1.

Complex coupling free energy. To compute the free energy ($\Delta G^{\text{C}}_{\text{coup}}$) of coupling the ligand to the solvated receptor the same general procedure as for computing the $\Delta G^{\text{L}}_{\text{dsolv}}$ was followed. The initial structures were taken from the simulation used to determine the equilibrium value of the restraints. Simulations of 1 ns were run at the following λ values: For BACE1 $\lambda = 0$, 0.008, 0.016, 0.032, 0.048, 0.064, 0.096, 0.112, 0.128, 0.16, 0.192, 0.224, 0.256, 0.288, 0.32, 0.352, 0.384, 0.416, 0.448, 0.48, 0.512, 0.544, 0.576, 0.608, 0.64, 0.656, 0.672, 0.704, 0.736, 0.768, 0.8, 0.832, 0.864, 0.896, 0.928, 0.96, and 1 and for CatD $\lambda = 0$, 0.016, 0.032, 0.064, 0.096, 0.128, 0.16, 0.192, 0.224, 0.256, 0.288, 0.32, 0.352, 0.384, 0.416, 0.448, 0.48, 0.512, 0.544, 0.576, 0.608, 0.64, 0.672, 0.704, 0.736, 0.768, 0.8, 0.832, 0.864, 0.896, 0.928, 0.96, and 1. For each λ window forward and backward FEP estimates of the free energy were obtained, as described above, as well as an estimate of the free energy from thermodynamic integration,

$$\Delta G^{\mathrm{TI}}_{i\to j} = 1/2\left(\langle U_{i\to j}\rangle_i + \langle U_{j\to i}\rangle_j\right), \quad (2)$$

where $U_{i\to j}$ is the difference between the energy the system would have in state j and the energy it does have in state i, where it is being simulated. The reported error in the coupling free energies was computed by taking the difference between $\Delta G_{i\to j}$ and $\Delta G^{\mathrm{TI}}_{i\to j}$ in each window and taking the square root of the sum of the squares of these differences.

Continuous constant pH molecular dynamics simulations Continuous constant pH molecular dynamics simulations with a hybrid-solvent scheme and the pH replica-exchange protocol[9] were performed on apo and holo forms of BACE1 and CatD.[2,3,10] The simulation of apo BACE1 utilized 24 pH replicas, which occupied the pH conditions 1 to 8 and lasted 21 ns each, while the simulation of holo BACE1 utilized 20 pH replicas, which occupied the pH conditions 1.3 to 8 and lasted 26 ns each. The apo and holo CatD simulations used 24 pH replicas which occupied the pH conditions 1 to 8 and lasted 36 ns each. To enable a more rigorous error analysis, the pK_{a}'s for apo and holo systems were recalculated based on the last 15 ns of all replicas, which correspond to an aggregate sampling time of 300/360 ns per system. Detailed simulation protocols are given in the previous studies.[2,3,10]

Error analysis for the calculated pK_{a}'s and related free energy changes A block error analysis was performed for the calculated pK_{a}'s by dividing the 300/360 ns data into 3 blocks. pK_{a}'s were calculated from each block, and the block standard errors BSE (listed in parentheses of Table S1 and Table S2) were calculated as $\sigma_n/\sqrt{M}$, where σ_n is the standard deviation among the pK_{a}'s of each block and M is the number of blocks.[11] The BSE of the pK_{a} shifts were calculated from the BSE of apo and holo pK_{a}'s using error propagation as BSE(ΔpK_{a})$= \sqrt{\mathrm{BSE}^2_{\mathrm{apo}} + \mathrm{BSE}^2_{\mathrm{holo}}}$. It can be shown that the maximum change in the binding free energy due to a pK_{a} shift is 2.303 RT ΔpK_{a}.[12] Thus, the stan-

dard error in the binding free energy due to the BSE of the pK_{a} shift is 1.36·BSE($\triangle$pK_{a}) kcal/mol at 298 K.

Supplementary Tables

Table S1: Calculated pK_{a} values for apo BACE1 and holo BACE1 complexed with the inhibitor LY2811376[a]

Asp	apo	holo	ΔpK_{a}	Glu	apo	holo	ΔpK_{a}
4	5.3 (0.06)	5.3 (0.03)	0.0 (0.07)	1	3.5 (0.06)	3.5 (0.09)	0.0 (0.10)
32	4.0 (0.14)	3.8 (0.11)	-0.2 (0.18)	17	3.3 (0.01)	3.6 (0.05)	0.3 (0.05)
62	2.6 (0.16)	2.4 (0.18)	-0.2 (0.24)	77	4.4 (0.03)	4.3 (0.04)	-0.1 (0.05)
83	3.7 (0.09)	3.5 (0.12)	-0.2 (0.15)	79	3.6 (0.07)	3.7 (0.01)	0.1 (0.07)
106	2.8 (0.04)	2.8 (0.09)	0.0 (0.10)	104	3.7 (0.05)	3.7 (0.00)	0.0 (0.05)
130	4.3 (0.10)	3.7 (0.09)	-0.6 (0.14)	116	5.1 (0.05)	5.2 (0.07)	0.1 (0.09)
131	3.4 (0.11)	3.3 (0.13)	-0.1 (0.17)	125	4.0 (0.07)	3.9 (0.14)	-0.1 (0.16)
138*	$<$1 (n/d)	1.3 (0.01)	n/d	134	3.3 (0.09)	3.1 (0.02)	-0.2 (0.09)
180	4.3 (0.08)	3.8 (0.23)	-0.5 (0.25)	165	4.3 (0.04)	4.2 (0.06)	-0.1 (0.07)
212*	$<$1 (n/d)	1.9 (n/d)	n/d	196	3.1 (0.07)	3.0 (0.04)	-0.1 (0.08)
216	2.7 (0.10)	2.7 (0.10)	0.0 (0.14)	200	5.2 (0.11)	5.4 (0.07)	0.2 (0.14)
223	4.0 (0.11)	3.3 (0.13)	-0.7 (0.17)	207	2.1 (0.08)	2.2 (0.05)	0.1 (0.10)
228	1.8 (0.01)	2.0 (0.14)	0.2 (0.14)	219	3.3 (0.06)	3.3 (0.04)	0.0 (0.07)
259	2.8 (0.13)	3.0 (0.11)	0.2 (0.17)	242	3.1 (0.02)	3.1 (0.08)	0.0 (0.09)
311	4.1 (0.12)	4.0 (0.10)	-0.1 (0.16)	255	4.4 (0.13)	4.3 (0.02)	-0.1 (0.13)
317	4.6 (0.01)	4.7 (0.11)	0.1 (0.11)	265	4.1 (0.02)	4.2 (0.04)	0.1 (0.05)
318	5.8 (0.05)	5.6 (0.23)	-0.2 (0.23)	290	3.0 (0.03)	2.8 (0.02)	-0.2 (0.03)
346	4.1 (0.06)	4.1 (0.07)	0.0 (0.09)	310	2.7 (0.08)	2.8 (0.05)	0.1 (0.10)
363	3.1 (0.07)	3.5 (0.11)	0.4 (0.14)	339	5.0 (0.07)	5.2 (0.16)	0.2 (0.17)
378	5.0 (0.04)	5.0 (0.06)	0.0 (0.07)	364	4.1 (0.02)	4.1 (0.04)	0.0 (0.04)
381	3.7 (0.03)	3.7 (0.04)	0.0 (0.05)	371	2.9 (0.13)	3.2 (0.07)	0.3 (0.15)
				380	3.3 (0.17)	3.6 (0.05)	0.3 (0.17)

His	apo	holo	ΔpK_{a}
45	6.3 (0.10)	7.1 (0.14)	0.8 (0.18)
49	7.1 (0.03)	6.9 (0.04)	-0.2 (0.05)
89	8.1 (0.07)	8.0 (0.21)	-0.1 (0.22)
145	6.2 (0.06)	6.7 (0.02)	0.5 (0.06)
181	7.2 (0.07)	7.2 (0.05)	0.0 (0.08)
360*	4.2 (n/d)	3.1 (n/d)	n/d
362*	8.6 (n/d)	8.6 (0.53)	0.0 (0.53)
Lilly	3.7 (0.05)	4.3 (0.06)	0.6 (0.06)

[a]The block standard errors are given in parentheses. Residues making statistically significant contributions to the pH-dependent binding free energy are highlighted in bold font. Residues denoted with an asterisk had incomplete titration or not converged pK_{a}'s. Since these pK_{a}'s are outside of the interested pH range (4–8), they are excluded from the binding free energy calculations. The apo pK_{a}'s of Asp138 and Asp212 could not be determined as titration was outside of the simulated pH range and did not converge. The pK_{a} shift of His360 was not determined as the apo and holo pK_{a}'s of His360 were not converged. His362 was not fully deprotonated at the highest simulation pH.

Table S2: Calculated pK_{a} values for apo CatD and holo CatD complexed with the inhibitor LY2811376[a]

Asp	apo	holo	ΔpK_{a}	Glu	apo	holo	ΔpK_{a}
12	4.6 (0.08)	4.7 (0.04)	0.1 (0.09)	5	5.2 (0.10)	5.1 (0.08)	-0.1 (0.13)
33	4.2 (0.10)	2.4 (0.14)	-1.8 (0.17)	18	2.9 (0.02)	2.8 (0.03)	-0.1 (0.04)
50	3.3 (0.07)	3.3 (0.04)	0.0 (0.08)	111	3.9 (0.05)	3.7 (0.01)	-0.2 (0.05)
62	4.2 (0.09)	4.0 (0.04)	-0.2 (0.10)	117	3.3 (0.09)	3.4 (0.08)	0.1 (0.12)
75	4.0 (0.11)	3.4 (0.03)	-0.6 (0.12)	180	3.9 (0.07)	3.8 (0.06)	-0.1 (0.09)
90*	1.8 (n/d)	2.8 (n/d)	n/d (n/d)	214	3.6 (0.06)	3.6 (0.03)	0.0 (0.06)
132*	1.2 (n/d)	$<$1(n/d)	n/d	224	4.0 (0.02)	4.0 (0.02)	0.0 (0.03)
152	1.6 (0.12)	2.2 (0.11)	0.6 (0.16)	227	3.1 (0.14)	3.2 (0.04)	0.1 (0.14)
161	4.0 (0.10)	4.2 (0.02)	0.2 (0.10)	243	5.8 (0.08)	5.8 (0.06)	0.0 (0.09)
172	3.5 (0.05)	3.5 (0.04)	0.0 (0.06)	246	3.5 (0.07)	3.4 (0.04)	-0.1 (0.08)
174	4.9 (0.05)	4.6 (0.03)	-0.3 (0.06)	260	5.5 (0.10)	5.5 (0.06)	0.0 (0.12)
187	3.5 (0.13)	3.5 (0.10)	0.0 (0.17)	266	3.6 (0.00)	3.6 (0.04)	0.0 (0.04)
211	4.7 (0.02)	4.6 (0.09)	-0.1 (0.10)	288	3.5 (0.02)	3.3 (0.09)	-0.2 (0.09)
231	3.5 (0.03)	3.1 (0.02)	-0.4 (0.03)	344	2.7 (0.03)	2.6 (0.14)	-0.1 (0.14)
242	3.0 (0.03)	3.3 (0.16)	0.3 (0.16)				
289	2.5 (0.18)	2.8 (0.08)	0.3 (0.20)				
310	3.5 (0.07)	3.4 (0.03)	-0.1 (0.08)				
323	3.5 (0.07)	3.4 (0.18)	-0.1 (0.19)				
334	2.6 (0.11)	2.9 (0.14)	0.3 (0.18)				
336	3.6 (0.02)	3.5 (0.10)	-0.1 (0.10)				
His	apo	holo	ΔpK_{a}				
45	7.8 (0.16)	7.6 (0.04)	-0.2 (0.17)				
56*	$>$9	$>$9	n/d				
57	7.2 (0.07)	7.1 (0.04)	-0.1 (0.08)				
77	6.5 (0.08)	6.7 (0.14)	0.2 (0.16)				
209*	7.4 (n/d)	8.0 (n/d)	n/d				
Lilly	3.7 (0.05)	4.9 (0.11)	1.2 (0.12)				

[a]The block standard errors are given in parentheses. Residues making statistically significant contributions to the pH-dependent binding free energy are highlighted in bold font. Residues denoted with an asterisk had incomplete titration or not converged pK_{a}'s. Since these pK_{a}'s are outside of the interested pH range (4–8), they are excluded from the binding free energy calculations. The apo and holo pK_{a}'s of Asp90 were not fully converged. The holo pK_{a} of Asp132 was outside of the pH range and not fully converged. The pK_{a} of His56 could not be determined, as the titration was outside of the pH range. The apo and holo pK_{a}'s of His209 were not fully converged.

Supplementary Figures

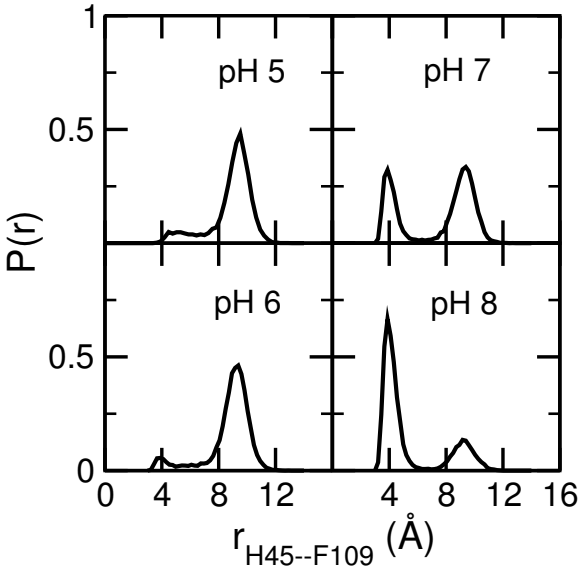

Figure S1: Interaction of His45–Phe109 in the apo BACE1 at different pH.

References

(1) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. *J. Phys. Chem. B* **2003**, *107*, 9535–9551.

(2) Ellis, C. R.; Tsai, C.-C.; Hou, X.; Shen, J. Constant pH Molecular Dynamics Reveals pH-Modulated Binding of Two Small-Molecule BACE1 Inhibitors. *J. Phys. Chem. Lett.* **2016**, *7*, 944–949.

(3) Ellis, C. R.; Tsai, C.-C.; Lin, F.-Y.; Shen, J. Conformational Dynamics of Cathepsin D and Binding to a Small-Molecule BACE1 Inhibitor. *J. Comput. Chem.* **2017**, *38*, 1260–1269.

(4) Lee, A. Y.; Gulnik, S. V.; Erickson, J. W. Conformational Switching in an Aspartic Proteinase. *Nat. Struct. Biol.* **1998**, *5*, 866–871.

(5) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.

(6) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kaleé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(7) MacKerell Jr., A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr., R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S. et al. All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(8) MacKerell, Jr., A. D.; Feig, M.; Brooks, III, C. L. Improved Treatment of the Protein Backbone in Empirical Force Fields. *J. Am. Chem. Soc.* **2004**, *126*, 698–699.

(9) Wallace, J. A.; Shen, J. K. Continuous constant pH Molecular Dynamics in Explicit Solvent with pH-Based Replica Exchange. *J. Chem. Theory Comput.* **2011**, *7*, 2617–2629.

(10) Ellis, C. R.; Shen, J. pH-Dependent Population Shift Regulates BACE1 Activity and Inhibition. *J. Am. Chem. Soc.* **2015**, *137*, 9543–9546.

(11) Grossfield, A.; Zuckerman, D. M. Quantifying Uncertainty and Sampling Quality in Biomolecular Simulations. *Annu. Rep. Comput. Chem.* **2009**, *5*, 23–48.

(12) Shen, J. K. Uncovering Specific Electrostatic Interactions in the Denatured States of Proteins. *Biophys. J.* **2010**, *99*, 924–932.