Supporting Information

The Effect of Environmental Factors on the Catalytic Activity of Intramembrane Serine Protease

Mojgan Asadi, Gabriel Oanca, and Arieh Warshel*

Department of Chemistry, University of Southern California, Los Angeles, California 90089-1062, United States

Computational Methods and Computational Studies

1.1 The Empirical Valence Bond (EVB) Method

The enzyme's catalytic activity was explored by the EVB approach used in many of our previous works[1–4]. The EVB method is a QM/MM method where the chemical reactions are described by mixing the relevant diabatic states. The diabatic states that constitute the EVB Hamiltonian are expressed as

$$H_{ii} = \alpha_{gas}^{i} + U_{intra}(r,q) + U_{Ss}(r,\mathrm{q},r',s) + U_{ss}(r',q'), \quad \text{(S1)}$$

where r and q are, respectively, the atomic coordinates and partial charges of a region of the reacting fragments (solute) in the i^{th} diabatic state, r' and q' are the positions and charges of the rest of the atoms in the system, α_{gas}^{i} represent the gas phase energy, $U_{intra}(r,q)$ is the intramolecular potential of the solute system, $U_{Ss}(r,q,r',q')$ represents the interaction between the solute (S) atoms and the surrounding (s) solvent, and $U_{ss}(r',q')$ represents the potential energy of the protein/solvent system. The off-diagonal elements $\left(H_{ij}\right)_{i\neq j}$ are assumed to be constant or represented by a simple function in the gas phase in the solution and the protein.

The adiabatic ground-state energy E_g and the corresponding eigenvector C_g are obtained by solving the secular equation

$$H_{EVB}C_g = E_gC_g. \quad \text{(S2)}$$

Morse potentials give the intramolecular contributions from the bonding interaction

$$V_b = D_M\left[1 - e^{a(b-b_0)}\right]^2, \quad \text{(S3)}$$

where b is the bond length. A harmonic potential expresses all the other bonds in the solute. Figures S1 and S2 illustrate the charge states of the region I atoms during the EVB calculations, and Tables S1-S4 list the parameters used for this calculation. The other intramolecular parameters such as bond angles, proper dihedral, and improper dihedral angles were represented by a MM force field

where the values of the parameters were taken from the ENZYMIX force field[5]. The non-bonding interactions among the solute atoms are taken in two separate ways depending on whether a pair of atoms: (a) never form bonds in any diabatic states, or (b) form bonds only in one of the diabatic states. The potential used for van der Waals interactions are described by

$$V_{nb} = Ce^{-kr_{ij}}. \tag{S4}$$

Since diabatic states represent the end-states of the chemical reaction, the system is driven from the reactant state to the product state in an attempt to calculate a complete reaction profile. The EVB free energy surfaces are determined by running MD simulation on a mapping potential ϵ_m, which is a linear combination of the diabatic potentials of the starting state of the reaction (state 1), and the final state (state 2). Thus, for a two-state representation of the reaction, the mapping potential takes the form:

$$\epsilon_m = \lambda_m \epsilon_1 + (1 - \lambda_m)\epsilon_2, \tag{S5}$$

where the mapping parameter λ_m varies between 0 and 1 in *N* windows.

The associated change in the free energy can be calculated using the free energy perturbation formalism. In this case, the free energy functional $\Delta G(\lambda_n)$ at a given λ_n can be defined as:

$$\Delta G(\lambda_n) = \Delta G(\lambda_0 \rightarrow \lambda_n) = \sum_{i=0}^{n-1} \delta G\ (\lambda_i \rightarrow \lambda_{i+1}), \tag{S6}$$

where each element is calculated as

$$\delta G(\lambda_i \rightarrow \lambda_{i+1}) = -\left(\frac{1}{\beta}\right) \ln\left[\langle e^{-\beta(\epsilon_{i+1} - \epsilon_i)} \rangle_i\right], \tag{S7}$$

where $\beta = \frac{1}{K_B T}$, and K_B is the Boltzmann constant, and T is the temperature, kept constant throughout the simulations. The angular bracket ($\langle \cdot \rangle_i$) operator averages concerning the mapping potential ϵ_m, but our aim is to evaluate the ground state free energy $E_g(X)$. Thus, we obtain the corresponding free energy by an umbrella sampling approach, using:

$$\exp[-\beta \Delta G(X^n)] = \exp[-\beta \Delta G(\lambda_m)] \left\langle \exp\left[-\beta(\epsilon_m(X^n) - E_g(X^n))\right] \right\rangle_m, \tag{S8}$$

where the reaction coordinate X^n is taken to be the energy gap $\epsilon_2 - \epsilon_1$.

The EVB parameters are given in tables S1-S4.

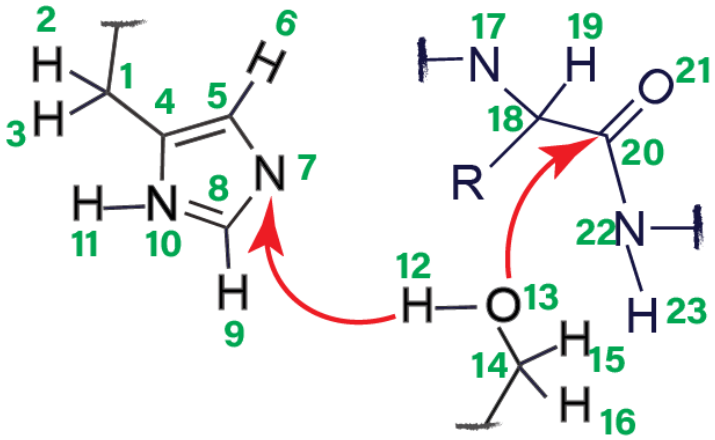

Figure S1. A scheme for the first proton transfer and nucleophilic attack. The EVB energy profile for this step is considered in the main text.

Table S1. Atomic charges were used for the first proton transfer from serine 201 to histidine 254 and the nucleophilic attack between the oxygen of serine 201 and the carbonyl carbon of the ligand.

Systems	Atom	RESP charges in RS state	Atom type (RS)*	RESP charges in PS state	Atom type (PS)*
Histidine	1	-0.327	C0	-0.277	C0
	2	0.132	H0	0.076	H0
	3	0.131	H0	0.119	H0
	4	0.292	C0	0.207	C0
	5	0.142	C+	0.289	C0
	6	0.031	H0	-0.084	H0
	7	-0.569	N0	-0.116	N+
	8	0.190	C0	0.395	C0
	9	0.111	H0	-0.045	H0
	10	-0.454	N0	-0.194	N0
	11	0.322	H0	0.105	H0
Serine	12	-0.431	O0	-0.419	O0
	13	0.194	H0	0.096	H0
	14	0.150	C0	0.213	C0
	15	0.042	H0	-0.070	H0
	16	0.092	H0	0.212	H0
Substrate	17	-0.119	N0	-0.313	N0
	18	0.294	C0	0.134	C0
	19	0.150	H0	0.094	H0
	20	-0.240	C0	-0.140	C+
	21	-0.440	O0	-0.861	O-
	22	0.288	N0	0.311	N0
	23	-0.192	H0	-0.390	H0

*RS and PS designate, respectively, the reactant and product state.

Table S2. All EVB parameters used for the first proton transfer and the nucleophilic attack steps (path 1).

Morse potential			
Atom pair	D(kcal/mol)	$a(\text{Å}^{-1})$	$r_0(\text{Å})$
O0-H0	94.00	1.40	1.345
N0-H0	100.00	2.00	0.88
C0-C0	96.00	1.54	0.80
C0-N0	95.00	1.40	2.00
C20-O21(RS)	93.00	1.25	2.00
C20-O21(PS)	94.00	1.50	0.80
Non-bonding potential			
Atom Type	C(kcal/mol)	$k(\text{Å}^{-1})$	
N+	60.0	2.52	
N0	60.0	2.5	
H0	5.0	2.5	
O0	53.0	2.5	
O-	90.0	2.5	
C+	91.0	2.5	
C0	91.0	2.5	
Off-diagonal $\boldsymbol{H_{ij} = Ae^{-\mu(r-r_0)}}$		Gas phase shift (kcal/mol)	
A	-54.56	-158.98	
$\boldsymbol{\mu}$	2.5		

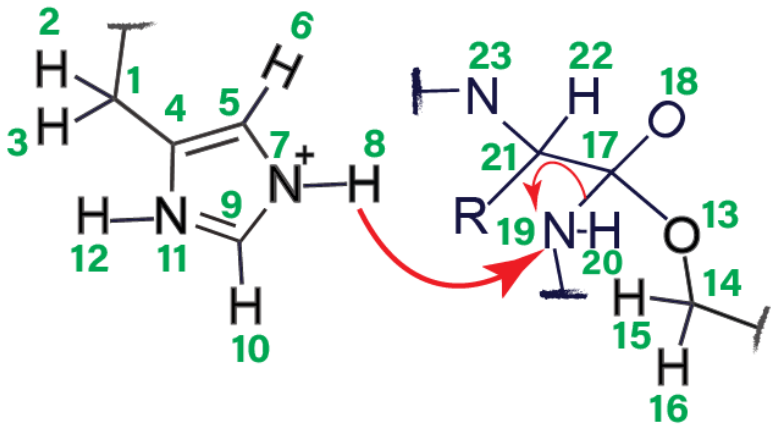

Figure S2. A scheme for path 2, second proton transfer, and the bond breaking. The EVB energy profile for this step is presented in the main text.

Table S3. Atomic charges for the second proton transfer from histidine 254 to the nitrogen of the scissile bond.

Systems	Atom	RESP charges in RS state	Atom type (RS)*	RESP charges in PS state	Atom type (PS)*
Histidine	1	-0.277	C0	-0.131	C0
	2	0.076	H0	0.098	H0
	3	0.119	H0	0.076	H0
	4	0.207	C0	0.243	C0
	5	0.289	C0	0.318	C0
	6	-0.084	H0	0.117	H0
	7	-0.116	N+	-0.465	N0
	8	0.123	H0	0.204	H0
	9	0.395	C0	0.288	C0
	10	-0.045	H0	-0.133	H0
	11	-0.194	N0	-0.367	N0
	12	0.105	H0	0.171	H0
Substrate	13	-0.419	O0	-0.111	O0
	14	0.213	C0	0.317	C0
	15	-0.070	H0	-0.087	H0
	16	0.212	H0	0.011	H0
	17	-0.140	C+	-0.196	C0
	18	-0.861	O-	-0.311	O0
	19	0.311	N0	-0.421	N0
	20	-0.390	H0	-0.085	H0
	21	0.134	C0	0.311	C0
	22	0.094	H0	0.104	H0
	23	-0.313	N0	-0.423	N0

*RS and PS designate, respectively, the reactant and product states.

Table S4. The EVB parameters used for the second proton transfer and the breaking steps (path 2).

Morse potential			
Atom pair	D(kcal/mol)	$a(\text{Å}^{-1})$	$r_0(\text{Å})$
O0-H0	94.00	1.40	1.345
N0-H0	100.00	2.00	0.988
C0-C0	96.00	1.54	0.80
C0-N0	95.00	1.40	2.00
C0-O0	93.00	1.50	0.80
C17-O18(RS)	93.00	1.25	2.00
C17-O18(PS)	94.00	1.50	0.80
Non-bonding potential			
Atom Type	C(kcal/mol)	$k(\text{Å}^{-1})$	
N+	60.0	2.52	
N0	60.0	2.5	
H0	5.00	2.5	
O0	53.0	2.5	
O-	90.0	2.5	
C+	91.0	2.5	
C0	91.0	2.5	

Off-diagonal $\boldsymbol{H_{ij} = Ae^{-\mu(r-r_0)}}$		Gas phase shift (kcal/mol)
A	-81.11	-131.01
$\boldsymbol{\mu}$	2.5	

1.2 Renormalizing the Energetics of the landscape of the Enzymatic Reaction.

The role of the dynamics of conformational changes has been a problem of significant interest. In particular, there has been significant research on whether this dynamical impact is crucial for enzyme function[6–11]. Since the rhomboid protease is a unique integral serine protease, the idea of rhomboid conformational and substrate dynamics is of particular interest. To explore this question, we used the renormalization method. We try to obtain the best correspondence between the full explicit model and a reduced 2-D model in this approach. Here we focus on the 2-D map of the rhomboid conformational change and the substrate movement. The rhomboid starts with the binding of the substrate to the open state configuration, undergoes a conformational change to reach a closed state and begin the catalytic process, and then finally opens to release the catalytic product. The substrate movement coordinate is taken as the distance between the oxygen atom of the alanine in the P1 position of the substrate and the active site residues S201 and H254. The induced closed to open configuration and substrate binding profiles are shown in Figure 7. To model the reaction in the protein, we use EVB, as explained previously.

The 2-D surface is taken as the ground state of an EVB type Hamiltonian of the form:

$$H_{lm,lm} = \epsilon_{lm} = \frac{\hbar\omega_Q}{2}\left(Q - \delta_Q^{lm}\right)^2 + \frac{\hbar\omega_R}{2}\left(R - \delta_R^{lm}\right)^2 + \alpha_{lm}, \tag{S9}$$

where the Q and R are the conformational and the substrate dimensionless effective coordinates. The dimensionless and dimensional coordinates of conformational and substrate are related by: $Q' = l_Q Q$, $\delta_Q' = l_Q Q$ and $R' = l_R R$, $\delta_R' = l_R R$, where the dimensional length scales are

$$l_Q^{-1} = \sqrt{\frac{\omega_Q m_Q^*}{\hbar}} \quad \text{and} \quad l_R^{-1} = \sqrt{\frac{\omega_R m_R^*}{\hbar}}, \tag{S10}$$

and ω_Q and ω_R are effective frequencies, δ_Q^{lm} and δ_R^{lm} are the locations of the minima, and proper α_{lm} are the minima shifts. The δ terms are related to the corresponding effective reorganization energies are defined as

$$\lambda_Q = \frac{\hbar}{2}\omega_Q\delta_Q^2 \quad \text{and} \quad \lambda_R = \frac{\hbar}{2}\omega_R\delta_R^2, \tag{S11}$$

so that the diabatic surfaces behave as a harmonic oscillator with frequency ω and mass m^*,

$$\epsilon_{lm} = \frac{m_Q^*\hbar^2\omega_Q^2}{2}\left(Q' - \delta'^{lm}_Q\right)^2 + \frac{m_R^*\hbar^2\omega_R^2}{2}\left(R' - \delta'^{lm}_R\right)^2 + \alpha_{lm}, \tag{S12}$$

where the effective mass m_Q^* in equation S10 is estimated using the relations $\frac{1}{2}m\dot{Q}^2 = \frac{1}{2}k_BT$ for the conformational degree of freedom, $m_Q^* = \frac{K_BT}{\langle \dot{Q} \rangle}$.

The 2-D potential surface was used in a Langevin Dynamics (LD) simulation[11] using

$$m_k^* \ddot{Q}_k' = -m_k^* \gamma_q \dot{Q'}_k - \frac{\partial \Delta G_s}{\partial Q'_k} + A'_k(t), \tag{S13}$$

where k runs over the component of the vector q, γ_Q is the effective friction in ps^{-1} and m^* is the effective mass of the protein coordinate, ΔG_s is the energy surface, and $A'_k(t)$ is a random force that satisfies the fluctuation dissipation theorem at temperature T.

Now in the renormalization model, we fit the 2-D model to the full model by forcing both models to have similar dynamics under applied forces. Specifically, we run MD simulation of the explicit model applying different pulling forces for the conformational coordinates. A restraint potential of the form

$$U_{con}(r) = k_{con}(r - r_0)^2, \tag{S14}$$

where the r is the distance between the sulfur heavy atom of residue methionine 247 in both closed and open structures, PDB ID 4NJN and 2IRV, is employed to pull the system closed to open conformation. Here r_0 is the distance between the two residues in the open form. Similarly, we applied a force of the form $U_{con}(z) = k_{con}(z - z_0)^2$ on the substrate motion, with $z_0 = 2000$Å and measured the first passage time using different force constants ($k = 1.0, 1.3, 1.5, 1.7, 2.0 \times 10^{-4}$ kcal/(mol Å^2)). Imposing a large value of z_0 with a small value of k_{con} ensures that the pulling force is approximately constant and can drive the conformational change.

We then run an effective 1-D model with LD using the same pulling forces for each effective conformational and substrate coordinate. Based on the range of the targeted MD coordinates, we estimated $\delta_R = 8.8$ Å for the range of the renormalized dynamics, and δ_Q is evaluated in terms of reorganization energy.

We tuned γ_Q and γ_R so that the velocity autocorrelation qualitatively matched the MD velocity autocorrelation directly and found the best matches with both $\gamma_Q \approx 3\ \mathrm{ps}^{-1}$ and $\gamma_R \approx 7\ \mathrm{ps}^{-1}$. We tuned the height, width, and position of a Gaussian barrier in the 1D so that the initial behavior and first passage times were consistent with MD simulations under the same force constants. From the pulled Langevin simulations, we obtained a barrier $\Delta G^{\ddagger} \approx 18 - 20$ kcal/mol for the

conformational change and $\Delta G^{\ddagger} \approx 13\ \mathrm{kcal/mol}$ substrate movement. The process is illustrated for the conformational coordinate in Figure S3.

The mean square velocity $\langle \dot{Q}'^2 \rangle$ was estimated using the sum of the velocities of the Cα atoms involved in the conformation change, this gave an effective mass of $m_Q^* = 1260$. The substrate effective mass was taken as the total mass of the substrate $m_R^* = 546$, while the reorganization energies λ_Q and λ_R were tuned to match the barriers, which yielded $\lambda_Q = 30.5\ \mathrm{kcal/mol}$, and $\lambda_R = 19.0\ \mathrm{kcal/mol}$. Following our previous study[12]. we selected the frequencies $\omega_Q = 40cm^{-1}$, and $\omega_R = 20cm^{-1}$. The correspondence between the explicit and reduced responses to the applied force is described in Figure S3.

After calibrating the 2-D parameters we ran free LD simulations on the 2D energy surface to explore possible dynamical effects. This study involved first exploring the coupling between the conformational changes and the substrate motion. We started the trajectories in an open conformation and the substrate away from the active site. The simulation results and a typical trajectory are shown in Figure S4. Since the barrier between state I and state II are very high and reaching it would require far too long simulations, we scaled down the surface to lower the barrier from state I to state II to about 10 kcal/mol. This scaling procedure was used and validated in our previous study[6]. As seen from Figure S4 and as found in our early study, the trajectories randomize when they reach the chemical barrier and lose their inertial memory. Thus, we cannot have any dynamical contribution to the barrier crossing event.

The chemistry of the rhomboid-substrate complex was captured using the EVB method. However, although the chemical step has an indirect coupling with the conformational change: the catalysis occurs when the substrate is bound and stabilized to its binding site while the conformation is closed state. For the substrate to depart the binding site, it undergoes a conformational shift.

Table S5. Different observed kinetic parameters in a various *Ec*GlpG and the substrate in a neutral membrane environment.

	Membrane(neutral)	
System	$k_{cat}(s^{-1})$	$\Delta G^{\ddagger}_{cal}$(kcal/mol)
*Hi*GlpG[c]	$1.56 \cdot 10^{-3}$	21.39
*Vc*Rho[c]	$2.88 \cdot 10^{-4}$	22.40
*Pa*Rom[c]	$7.22 \cdot 10^{-5}$	23.22
*Sp*AarA[c]	$2.81 \cdot 10^{-3}$	21.04
*Ps*AarA[c]	$1.37 \cdot 10^{-2}$	20.10
*Aq*ROM[c]	$9.27 \cdot 10^{-2}$	18.96
*Bf*ROM1[c]	$1.15 \cdot 10^{-1}$	18.83
*Bf*ROM2[c]	$3.78 \cdot 10^{-2}$	19.49
AarA[d]	$1.77 \cdot 10^{-2}$	19.94
*Ec*GlpG[d]	$1.77 \cdot 10^{-3}$	21.31
*Hi*GlpG[d]	$1.00 \cdot 10^{-3}$	21.66
Aard[e]	$6.00 \cdot 10^{-3}$	20.59
*Ec*GlpG[e]	$6.10 \cdot 10^{-3}$	20.58
*Hi*GlpG[e]	$6.03 \cdot 10^{-3}$	20.56

[c] Observed kinetic parameters of diverse GlpG in neutral membrane environment[4].

[d] Observed parameters for AarA, *Ec*GlpG, and *Hi*GlpG with FL-casein substrate[34].

[e] Observed kinetic parameters for AarA, *Ec*GlpG, and *Hi*GlpG with psTatA as substrate[34].

Table S6: Activation barriers of the rate determining step of different serine proteases.

Serine protease	$\Delta G^{\ddagger}$(kcal/mol)
Rhomboid	20.40[4]
Chymotrypsin	14.80[11]
Prolyl Oligopeptidase	15.00[12]
Subtilisin	15.30[13]

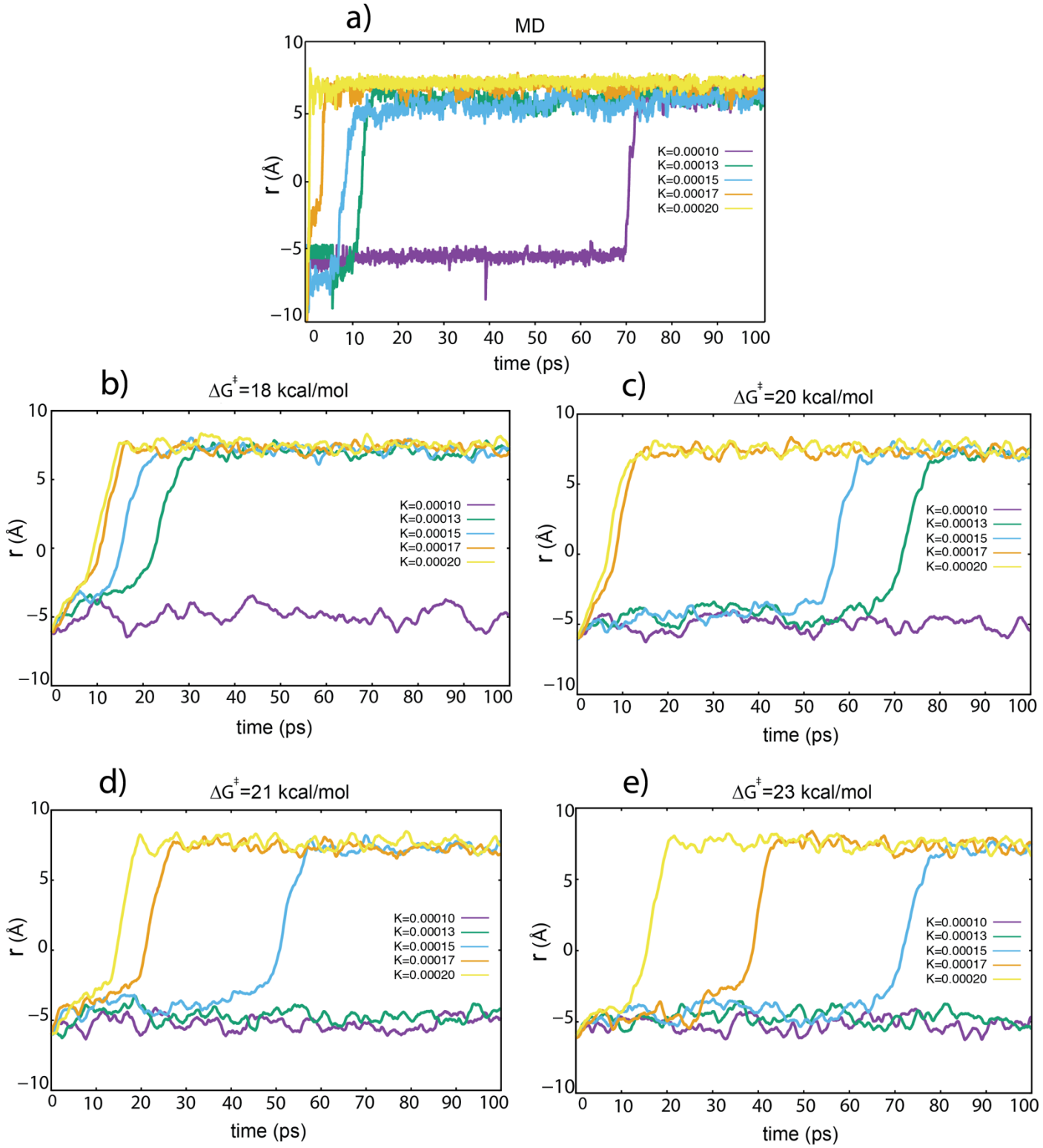

Figure S3. Renormalization method for tuning the barrier of the conformational change. The explicit targeted MD (a) is compared to a simplified 1D LD model with varying barriers (b-e) until the characteristic passage times for each force constant are consistent with those of the MD passage times. We observe that the closest behavior to the explicit model is recreated with a barrier between 18-20 kcal/mol.

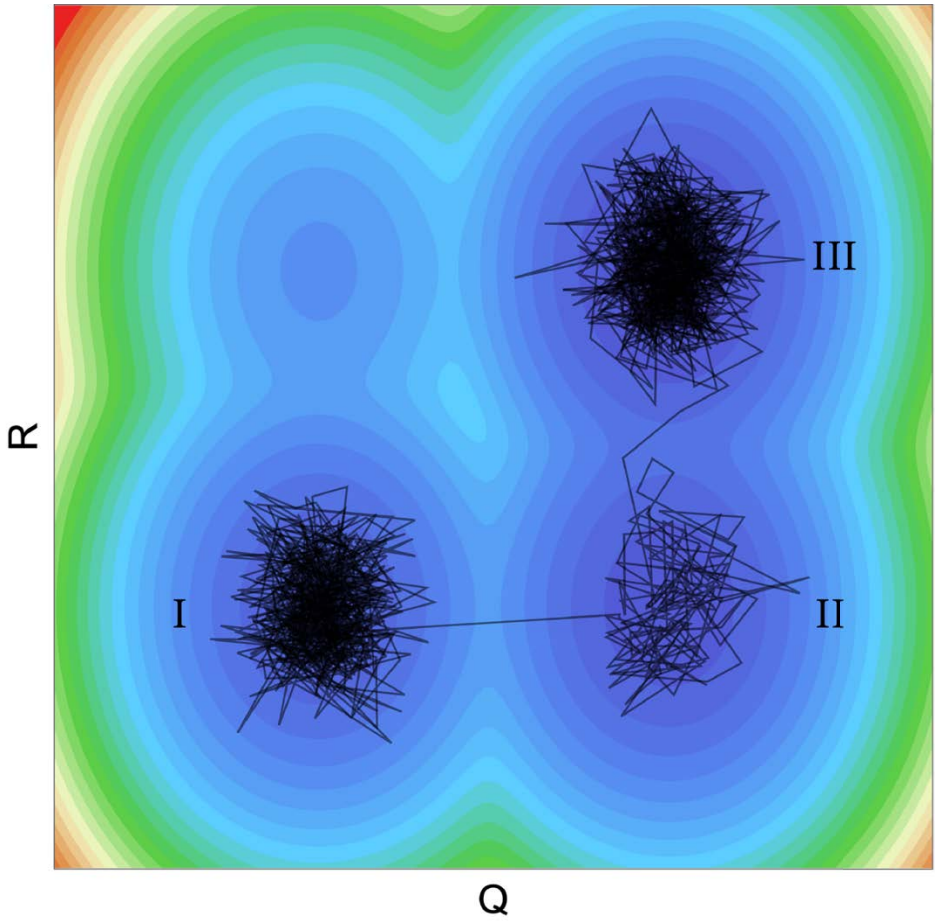

Figure S4. Exploring the possible dynamical coupling between the substrate movement and the conformational change. A typical LD trajectory initialized on the reactant I and the energy barrier from I to II was scaled down to reach its transition.

References:

(1) Warshel, A.; Weiss, R. M. An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. *J. Am. Chem. Soc.* **1980**, *102* (20), 6218–6226. https://doi.org/10.1021/ja00540a008.

(2) Kamerlin, S. C. L.; Warshel, A. The Empirical Valence Bond Model: Theory and Applications. *WIREs Comput. Mol. Sci.* **2011**, *1* (1), 30–45. https://doi.org/10.1002/wcms.10.

(3) Hwang, J. K.; Warshel, A. Microscopic Examination of Free-Energy Relationships for Electron Transfer in Polar Solvents. *J. Am. Chem. Soc.* **1987**, *109* (3), 715–720. https://doi.org/10.1021/ja00237a013.

(4) Braun-Sand, S.; Strajbl, M.; Warshel, A. Studies of Proton Translocations in Biological Systems: Simulating Proton Transport in Carbonic Anhydrase by EVB-Based Models. *Biophys. J.* **2004**, *87* (4), 2221–2239. https://doi.org/10.1529/biophysj.104.043257.

(5) Lee, F. S.; Chu, Z. T.; Warshel, A. Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the POLARIS and ENZYMIX Programs. *J. Comput. Chem.* **1993**, *14* (2), 161–185. https://doi.org/10.1002/jcc.540140205.

(6) Pisliakov, A. V.; Cao, J.; Kamerlin, S. C. L.; Warshel, A. Enzyme Millisecond Conformational Dynamics Do Not Catalyze the Chemical Step. *Proc. Natl. Acad. Sci.* **2009**, *106* (41), 17359–17364. https://doi.org/10.1073/pnas.0909150106.

(7) Levitt, M.; Warshel, A. Computer Simulation of Protein Folding. *Nature* **1975**, *253* (5494), 694–698. https://doi.org/10.1038/253694a0.

(8) Min, W.; Xie, X. S.; Bagchi, B. Two-Dimensional Reaction Free Energy Surfaces of Catalytic Reaction: Effects of Protein Conformational Dynamics on Enzyme Catalysis. *J. Phys. Chem. B* **2008**, *112* (2), 454–466. https://doi.org/10.1021/jp076533c.

(9) Vamvaca, K.; Vögeli, B.; Kast, P.; Pervushin, K.; Hilvert, D. An Enzymatic Molten Globule: Efficient Coupling of Folding and Catalysis. *Proc. Natl. Acad. Sci.* **2004**, *101* (35), 12860–12864. https://doi.org/10.1073/pnas.0404109101.

(10) Kamerlin, S. C. L.; Warshel, A. At the Dawn of the 21st Century: Is Dynamics the Missing Link for Understanding Enzyme Catalysis? *Proteins* **2010**, *78* (6), 1339–1375. https://doi.org/10.1002/prot.22654.

(11) Zhou, H.; Yu, H.; Zhao, X.; Yang, L.; Huang, X. Molecular Dynamics Simulations Investigate the Pathway of Substrate Entry Active Site of Rhomboid Protease. *J. Biomol. Struct. Dyn.* **2019**, *37* (13), 3445–3455. https://doi.org/10.1080/07391102.2018.1517609.