# Probing Binding Affinity by Jarzynski's Non-equilibrium Binding Free Energy and Rupture Time

Duc Toan Truong<sup>1,2</sup> and Mai Suan Li<sup>1,3,\*</sup>

<sup>1</sup>Institute for Computational Sciences and Technology, SBI building, Quang Trung Software City, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Vietnam

<sup>2</sup>Department of Theoretical Physics, Faculty of Physics and Engineering Physics, Ho Chi Minh University of Science, Vietnam

<sup>3</sup>Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland

\*masli@ifpan.edu.pl

#### **Abstract**

Binding affinity of small ligand to receptor is the important quantity in drug design and it might be characterized by different quantities. The most popular one is the binding free energy which can be estimated by several methods in conventional molecular dynamics simulation. So far in steered molecular dynamics (SMD) one can use either the rupture force or non-equilibrium pulling work as a measure for binding affinity. In this paper we have shown that the non-equilibrium binding free energy  $\Delta G_{\text{neq}}^{\text{Jar}}$ , obtained by Jarzynski's equality at finite pulling speed, has good correlation with experimental data on inhibition constants implying that this quantity can be used as a good scoring function for binding affinity. A similar correlation has been also disclosed for binding and unbinding free energy barriers. Applying the SMD method to unbinding of 23 small compounds from the binding site of beta-lactamase protein, a bacterial produced enzyme, we have demonstrated that the rupture or unbinding time strongly correlates with experimental data with correlation level  $R \approx 0.84$ . As follows from the Jarzynski's equality, the rupture time depends on the unbinding barrier exponentially. We show that  $\Delta G_{\rm neq}^{\rm Jar}$ , the rupture time as well as binding and unbinding free energy barriers are good descriptors for binding affinity. Our observation may be useful for fast screening of potential leads as the SMD simulation is not time consuming. Based on non-equilibrium simulation we disclosed that, in agreement with experiment, the binding time is much longer than the unbinding one.

#### Introduction

Estimation of binding affinity of ligand to receptor is a central problem in drug design. The binding affinity might be described by various scoring functions but the most popular one is the binding free energy  $\Delta G$ . Typical inhibition constant of  $1 \mu M - nM$  corresponds to the binding free energy between -9 and -12 kcal/mol. One of the fastest methods for estimating  $\Delta G$  is the molecular docking<sup>1</sup> in which the entropy contribution is neglected. However, this method is not accurate as several crude approximations such as omission of receptor dynamics and a finite number of trials for ligand position have been adopted. Despite low accuracy the docking method is widely used in virtual screening due to its high speed.

With the help of molecular dynamics (MD) simulation one can use more accurate methods for computing  $\Delta G$  such as free energy perturbation (FEP)<sup>2,3</sup>, thermodynamics integration (TI)<sup>4,5</sup>, linear interaction energy (LIE)<sup>6</sup>, linear response approximation (LRA)<sup>7</sup>, molecular mechanics Poisson-Boltzmann surface area (MM-PBSA)<sup>8</sup>. The exact FEP method is the best one as it can predict not only relative binding affinity but also absolute binding free energy<sup>9</sup>. MM-PBSA method is reliable in estimation of relative binding free energy but, in general, it fails to reproduce experimental data on absolute value of inhibition constant<sup>10</sup>.

Recently it has been shown that steered molecular dynamics (SMD)<sup>11-13</sup>, in which the external force is applied to pull a ligand out from the binding site (Fig. 1), becomes a powerful tool for drug design<sup>14,15</sup> as its accuracy is compatible with MM-PBSA but it is about two orders of magnitude faster<sup>16</sup>. This rationalizes its frequent use to refine docking results in virtual screening of potential drugs from huge databases<sup>17-19</sup>. In SMD the binding affinity is characterized by the rupture force  $F_{\text{max}}$  in the force-extension/time profile (Fig. 1) in such a way that the higher  $F_{\text{max}}$  the stronger binding<sup>14,15</sup>. Previously we have shown that non-equilibrium pulling work  $W_{\text{pull}}$  even better correlates with experimental data than  $F_{\text{max}}$  because the latter is defined at a single time, while the pulling work is computed over the whole process<sup>20</sup>.

Despite success of using different scoring functions like  $F_{\rm max}$  and  $W_{\rm pull}$ , a search for new quantities that might improve correlation with experiment is important not only from application prospect but also from the point of view of theory development. Motivated by this, the first question we ask is can one use the popular Jarzynski's equality<sup>21</sup> to estimate relative binding affinity for a large number of protein-ligand complexes. Because the estimation of the equilibrium binding free energy by this method is very time consuming<sup>22</sup> we propose to estimate non-equilibrium binding free energy  $\Delta G_{\rm neq}^{\rm Jar}$  by SMD at finite pulling speed<sup>23</sup> and use it for ranking binding affinity.

From the dependence of  $\Delta G$  on displacement/time (Fig. 1)<sup>23</sup> one can extract the unbinding  $(\Delta G^{\ddagger}_{unbind})$  and binding  $(\Delta G^{\ddagger}_{bind})$  free energy which characterizes the barrier crossing process. Here  $\Delta G^{\ddagger}_{unbind}$  is the difference between  $\Delta G_{TS}$  of the transition state (TS) corresponding to maximal  $\Delta G$  and  $\Delta G$  of the bound state, while  $\Delta G^{\ddagger}_{bind}$  is the barrier between the TS and unbound state. We have shown that both  $\Delta G^{\ddagger}_{unbind}$  and  $\Delta G^{\ddagger}_{bind}$  can be used to measure ligand binding affinity.

In SMD experiment one can consider a ligand as unbound after reaching the maximum force in the force-time curve (Fig. 1). Therefore, it is reasonable to assume that the larger rupture time  $t_{\text{max}}$  the stronger binding of ligand to a given receptor and our second goal is to study the relationship between  $t_{\text{max}}$  and the experimentally measured binding free energy.

In order to illustrate the main idea, we have considered binding of 23 small compounds to beta-lactamase (BL) protein, a bacterial produced enzyme. Experimental data on inhibition constants of these compounds are available for comparison with our simulation results. Using SMD simulation and Jarzynski's equality we obtained  $\Delta G_{\text{neq}}^{\text{Jar}}$  which is highly correlated with experimental data (correlation level R=0.89). We have also demonstrated that having the high fit level with experiment,  $t_{\text{max}}$ ,  $\Delta G_{\text{bind}}^{\ddagger}$  and  $\Delta G_{\text{unbind}}^{\ddagger}$  can be also used as reliable scores for relative binding affinity.

#### **Material and Methods**

#### Receptors and ligands

One of the major mechanisms to kill bacteria is to use beta-lactam since it attacks and breaks the bacterial cell wall synthesis. Based on BL structures, four molecular classes (A, B, C and D) were reported 30-33. Currently, only a few potential drugs such as meropenem, doripenem, ertapenem are on clinical trial 34,35, while many synthesized compounds can bind into the active site of BL at a widely range of inhibition constant  $K_i^{36-39}$ . Here we focus on 23 small compounds that have the highest and lowest  $K_i$  value of  $1\mu M$  and  $19000\mu M$ , respectively (Table S1 in the Supporting Information (SI)). The 2D structures and chemical formula of these compounds are listed in Table S1. In SMD simulation for each BL-ligand complex we used its PDB structure (Table S1) as the starting configuration.

#### Docking simulation

After removing the ligand and water molecules from the binding site of PDB structure 1XGJ by PyMol package, the protein coordinate file was opened with AutoDocktools (ADT) program to merge all non-apolar hydrogens and it was saved into PDBQT. The grid 60\*50\*50 (ų), placed in the center of the BL binding pocket, was large enough to cover the entire binding site. Twenty three small compounds with experimentally measured Ki, after ADT preparing into PDBQT file, were docked to the receptor and their binding energy was estimated by AutodockVina Program¹ with the exhaustiveness set at 800.

#### Steered molecular dynamics simulation

Steered molecular dynamic (SMD) is a non-equilibration MD simulation in which an external force is applied to the ligand in the binding site through a dummy atom connected with a ligand atom by a spring with spring constant k. This setup allows one to probe binding affinity of a given ligand to receptor. Results of SMD simulation, in general, depend on pulling direction which may be obtained by different approaches. We have used the Caver method<sup>40</sup> to determine the pulling pathway through the widest tunnel and this guarantees that there is no collision between the ligand and enzyme during simulation. The PyMol and Caver

 $3.0^{40}$  package were employed to rotate the protein in such a way that the ligand unbinding pathway is along the z-axis (Fig. 1).

The topology files and full atomistic structure of BLs were prepared by using pdb2gmx package with CHARMM 27 force field<sup>41</sup> and the TIP3P water model<sup>42</sup>. After adding hydrogen atoms to ligand by Chimera 1.11.2 and saving as Mol2 type, the ligand topology and coordinate files were separated by Swiss Param. All simulations were performed using GROMACS package version 5.1.3. The center of protein-ligand complex was placed at 4.7, 4.7 nm in a cubic box with the edges of 9.5, 9.5, 11 nm. The periodic boundary condition (PBC) was applied to minimize size effects. The box volume is around 990 nm<sup>3</sup> to ensure that the smallest distance from the protein to the boundaries is 1.4 nm and the protein would not interact with itself. The system was solvated with more than 25000 TIP3P water molecules, three Cl<sup>-</sup> ions were added to keep the system neutral and the total number of atoms was about 80000.

Three equilibration steps were carried out at 300K including energy minimization, 500ps of positional restraint in canonical ensemble (NVT) and 500ps of isothermal-isobaric ensemble (NPT). The cut-off radius for van-der-Waals potentials and Coulomb interaction was set at 1.4 nm and 1.0 nm, respectively.

The last snapshot of the NPT step was used as the initial structure for SMD simulation. To prevent protein to drift with ligand  $C_{\alpha}$  atoms of BLs were restrained by a weak harmonic force. The force with constant pulling speed v was applied to the dummy atom along the z-direction. The ligand experiences force F = k(vt - z), where z is displacement from the initial position of ligand atom connected with the dummy atom via a spring. We chose the spring constant  $k = 600 \text{ kJ/mol/nm}^2$  which is typical in AFM experiment, and v=5 m/s. The force and the ligand center of mass position were recorded every 2 fs. For each protein-ligand complex, 20 independent runs were performed starting from the same initial configuration but with different random seed numbers generating different initial velocity fields.

The pulling work is calculated as follows

$$W = \int F(x)dx = \frac{1}{2} \sum_{i=1}^{n-1} (f_{i+1} + f_i)(x_{i+1} - x_i)$$
(1)

Here for two successive simulation steps i and i+1 the force which drives ligand motion is assigned as an averaged value, and  $\Delta x_i$  is a ligand displacement in step i. The work calculated in each step is positive if the ligand moves in the pulling direction and negative otherwise. The shorter is the time interval of data collection, the more accurate the external work. We have carried out 450 ps SMD simulation which corresponds to n=22500 steps of 2 fs each.

#### **Results and Discussion**

Docking results

For preparing the receptor in docking simulation, we used the structure of BL complex from PDB bank with ID 1XGJ (Fig. 1) because for this structure is similar to the remaining 22 PDB structures shown in Table S1. The docking simulation has been performed using AutodockVina Program<sup>1</sup>. The binding energies obtained in the best docking mode have the correlation level R=0.73 with experimental data (Fig. S1 in SI). Given a low accuracy of the docking method due to crude approximations such as omission of receptor dynamics and finite number of ligand configuration trials<sup>1</sup>, this correlation should be considered as high.

Correlation of unbinding time  $t_{max}$  with binding affinity

In order to calculate mean  $t_{\text{max}}$  we extracted it from individual SMD runs (Fig. 1) and then averaged over all runs. Note that the rupture time defined by this way is a bit different from what followed from the time/displacement dependence of mean F(t) (Figure 2).

The 1XGJ complex is most stable having the smallest inhibition constant Ki= 1  $\mu$ M<sup>24</sup> (Table S1 in SI) and consequently it has unbinding time  $t_{\rm max}$  larger than that of 1XGI, 1L2S, 4KZ6, 3GR2 and 2HDR (Fig. 2 and Table 1). Complexes 2R9W, 2R9X, 4JXS, 4JXW, and 4JXV<sup>25</sup> have lower binding affinity compared to the most potent compound HTC in 1XGJ complex (Table 1) but within error these complexes have the same  $t_{\rm max}$ . The most unstable 2HDR system apparently has the shortest unbinding time of about 84 ps.

For the 23 studied complexes the correlation level between  $t_{\rm max}$  and  $\Delta G_{\rm exp}$  is R=0.84 (Fig. 3). For comparison we have computed the average rupture force  $F_{\rm max}$  and non-equilibrium pulling work  $W_{\rm pull}$  (Fig. 2).  $W_{\rm pull}(t)$  becomes non-zero only after some time as the ligand atom connected to the dummy atom moves back and forth at the beginning, while at large enough time  $W_{\rm pull}(t)$  saturates. Here the binding affinity is defined by the non-equilibrium pulling work at the end of simulation, i.e.  $W_{\rm pull} = W_{\rm pull}(t_{\rm end})$ .

As evident from Fig. S2 in SI, the correlation level with experiment is R = 0.82 and 0.84, for  $F_{\rm max}$  and  $W_{\rm pull}$ , respectively. In accord with our previous study<sup>20</sup>, the non-equilibrium work correlates with experiment better than the rupture force. Because  $t_{\rm max}$  has the same correlation level as  $W_{\rm pull}$  it can be used as a good scoring function for predicting relative binding affinity of small compounds to protein. This result is reasonable as the tighter binding the larger time needed to pull ligand out from the binding pocket.

Estimation of Jarzynski's non-equilibrium binding free energy by SMD

Extending the Jarzynski's equality<sup>21</sup> to the case when the external force with constant pulling velocity v is applied Hummer and Szabo have shown that Jarzynski's binding free energy  $\Delta G$  may be calculated using the following equation<sup>23</sup>

$$\exp\left(\frac{-\Delta G}{k_B T}\right) = \left\langle \exp\left(\frac{-\left[W - \frac{1}{2}k(z_t - vt)^2\right]}{k_B T}\right)\right\rangle_{N}$$
(2)

where  $<...>_N$  stands for averaging over N trajectories where a ligand is pulled from the binding site to unbound state,  $z_t$  is the displacement of ligand atom, which is connected with the dummy atom via a spring, from its initial position. Non-equilibrium work W given by Eq. (1) can be estimated from force-displacement/time profile. The equilibrium binding free energy is obtained in the low pulling speed and  $N \to \infty$  limit. In our all-atom SMD simulation with finite N and high pulling speed v, the obtained binding free energy is, therefore, non-equilibrium.

Using Eq. (2) and collected snapshots one can obtain the dependence of  $\Delta G$  on displacement and time. Typical dependence of Jarzynski's binding free energy on displacement in a single SMD trajectory is shown in Fig. 1 for 1XGJ complex. One can show that the maximum which corresponds to TS occurs at rupture time  $t_{\rm max}$ . The existence of TS is observed in all studied complexes as shown in Fig. 4 where the binding free energy has been averaged over 20 independent SMD runs.

The bound state appeared at the beginning of simulation has  $\Delta G_{\text{bound}}=0$  (Fig. 4). The binding free energy of the unbound state, which occurs at long time (large displacement), is defined at the end of simulation, i.e.  $\Delta G_{\text{unbound}}=\Delta G(t_{\text{end}})$ . As follows from Eq. (2),  $\Delta G(t_{\text{end}})$  can be used for describing binding/unbinding affinity. To stress that this quantity is obtained from the Jarzynski's equality and is not at equilibrium, in what follows it will be referred to as  $\Delta G_{\text{neq}}^{\text{Jar}}$  ( $\Delta G_{\text{neq}}^{\text{Jar}}=\Delta G(t_{\text{end}})$ ). For 23 studied ligands the absolute value of this quantity is large varying from 62.9 (2PU2) to 7.5 kcal/mol (2HDU) as it have been computed at high pulling speed (Table 1). In general,  $\Delta G_{\text{neq}}^{\text{Jar}}$  decreases as v is lowered but this problem is beyond the scope of the present paper. In all cases  $|\Delta G_{\text{neq}}^{\text{Jar}}| > W_{\text{pull}}$  and the difference between these quantities is partially due to  $k(z_t - vt)^2/2$  term in Eq. (2).

#### Ligand unbinding/binding is a barrier crossing process

The bound and unbound states are separated by the transition state (TS) implying that the ligand unbinding is a barrier crossing event (Fig. 4). The unbinding barrier  $\Delta G^{\ddagger}_{unbind} = \Delta G_{TS}$  abound, while the binding barrier  $\Delta G^{\ddagger}_{bind} = \Delta G_{TS} - \Delta G_{unbound}$ . Here  $\Delta G_{TS}$  refers to the binding free energy of TS. One has to stress that the barriers (Fig. 4) were obtained out of equilibrium and they are higher than the equilibrium ones. For all studied ligands  $\Delta G^{\ddagger}_{bind} > \Delta G^{\ddagger}_{unbind}$  (Table 1) and, as evident from Table S2, the ratio  $\Delta G^{\ddagger}_{bind}/\Delta G^{\ddagger}_{unbind}$  varies between 2.5 (ligand 3GV9) and 4.0 (ligand 3GRJ) implying that the unbinding time is shorter than the binding one. This observation is in line with experiment showing that the ligand binding process is considerably slower than the unbinding one  $^{43}$ .

Because the unbinding is a barrier crossing event the rupture time  $t_{\text{max}} \sim \exp(c\Delta G^{\ddagger}_{\text{unbind}})$ , where c is constant. The exponential dependence is indeed observed in our simulation (Fig. 5) with high accuracy (R=0.94).

Correlation of  $\Delta G_{\text{neq}}^{\text{Jar}}$ ,  $\Delta G_{\text{bind}}^{\ddagger}$ , and  $\Delta G_{\text{unbind}}^{\ddagger}$  with experimental binding affinity

Despite the simulation has been conducted at high pulling speed, the non-equilibrium binding free energy  $\Delta G_{\text{neq}}^{\text{Jar}}$  well fits with experimental data with R=0.89 (Fig. 3). This correlation

level is, as expected, much higher than the docking one (R= 0.73, Fig. S1).  $\Delta G_{\rm neq}^{\rm Jar}$  describes experiment even better than the non-equilibrium work  $W_{\rm pull}$  and  $t_{\rm max}$  (R=0.84) having higher correlation level, but it is not clear if this holds for other complexes. On the other hand, because the Jarzynski's non-equilibrium free energy is defined through the pulling work (Eq. 2) one can expect that these quantities have about the same fit level. Thus, together with the well-studied quantity  $W_{\rm pull}$ ,  $\Delta G_{\rm neq}^{\rm Jar}$  can serve as a good descriptor for discerning good binders from the bad ones.

The correlation level with experiment is R=0.82 and 0.80, for barriers  $\Delta G^{\ddagger}_{bind}$  and  $\Delta G^{\ddagger}_{unbind}$ , respectively (Fig. 3). This fit is worse than  $\Delta G_{neq}^{Jar}$  (R=0.89) but compatible with  $F_{max}$ ,  $W_{pull}$  and  $t_{max}$ . Because  $\Delta G^{\ddagger}_{bind}$  and  $\Delta G^{\ddagger}_{unbind}$  are closely related to  $\Delta G_{neq}^{Jar}$  (Figures 1 and 4) these three quantities are expected to have the same correlation level with experiment. Thus, we have shown that together with the rupture fore and pulling work,  $\Delta G_{neq}^{Jar}$  as well as binding and unbinding barriers are useful in refining docking results in virtual screening.

#### Conclusion

Nowadays SMD is widely used in computer-aided drug design<sup>44-47</sup> necessitating its deeper understanding. In the present paper we have made a step towards this direction. Although an accurate estimation of equilibrium binding free energy by Jarzynski's equality computationally remains problematic, for the first time we have shown that the non-equilibrium binding free energy  $\Delta G_{\rm neq}^{\rm Jar}$  is a good scoring function for relative binding affinity. Because  $\Delta G_{\rm neq}^{\rm Jar}$  is easily computed by fast pulling simulation one can use it for efficient virtual screening in drug design. For describing unbinding process the rupture time  $t_{\rm max}$  has been introduced and investigated through comparison with experimental data on inhibition constants of 23 small compounds bound to BL protein. It turns out that together with  $F_{\rm max}$ ,  $W_{\rm pull}$ ,  $\Delta G_{\rm bind}^{\dagger}$ ,  $\Delta G_{\rm unbind}^{\dagger}$  and  $\Delta G_{\rm neq}^{\rm Jar}$  it may be used as a reliable indicator for binding strength that the larger  $t_{\rm max}$  the higher propensity to binding. It would be interesting to check our main conclusion for other systems. Even in high pulling speed SMD we were able to disclose that, in agreement with experiment, the binding process is much slower than the unbinding one.

It has been shown that at high pulling speeds  $F_{\rm max}$  decreases with v linearly (viscosity regime), while the logarithmic dependence occurs at low pulling speeds (barrier crossing regime)<sup>48</sup>. The similar behavior is expected to hold for the binding free energy and barriers. Our work in this direction is in progress.

Finally, in this paper the ligand was pulled along a single direction and this choice probably is not as accurate as pulling along multiple directions <sup>45,49</sup> to minimize the resistance from the receptor. It would be interesting to check if our main conclusions remain valid if the force direction is not fixed.

ASSOCIATED CONTENT

Supporting Information. Details on receptors and ligand are presented. Table S1 shows

inhibition constant, molecular weight, chemical formula and 2D structure of the 23 studied

ligands. The ratio of binding and unbinding barriers is given in Table S2. The correlation

between the docking binding energy and experimental data is shown in Fig. S1. The

correlation between pulling work  $W_{\text{pull}}$  and rupture force  $F_{\text{max}}$  with experimental data is

presented in Fig. S2.

**AUTHOR INFORMATION** 

Corresponding author

Mai Suan Li

\*Email: masli@ifpan.edu.pl

Phone: +48 22 843 66 01 (ext. 3326)

Address: Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668

Warsaw, Poland

**Author Contributions** 

DTT and MSL conceived the experiments. DTT conducted the experiment. DTT and MSL

analyzed the results. DTT and MSL wrote the paper. All authors reviewed the manuscript.

**Notes:** The authors declare no competing financial interest.

**ACKNOWLEDGMENT** 

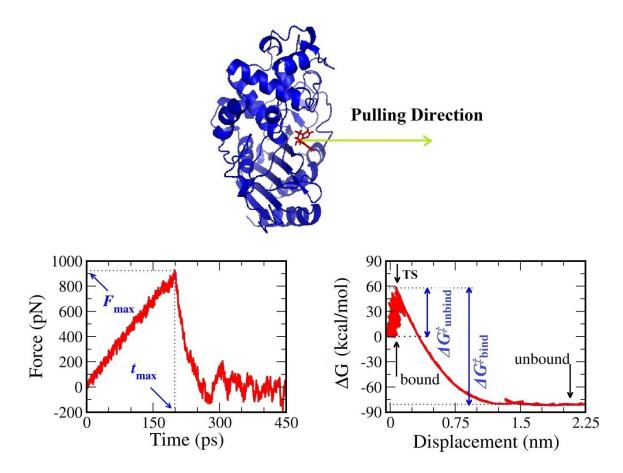
This work was supported by Department of Science and Technology at Ho Chi Minh city,

Vietnam, and the Polish NCN grant 2015/19/B/ST4/02721, Poland. We thank Le Thi Kim Thu

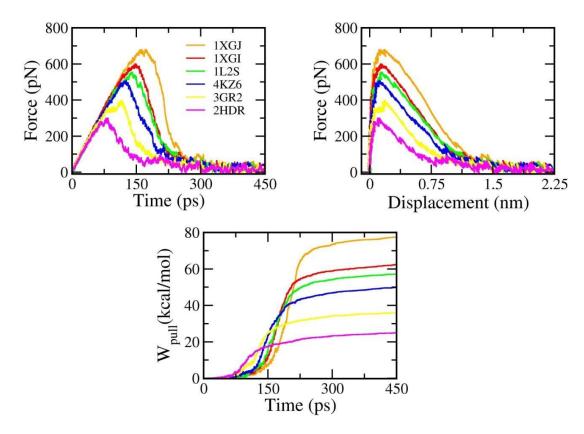
for critical reading of the manuscript.

8

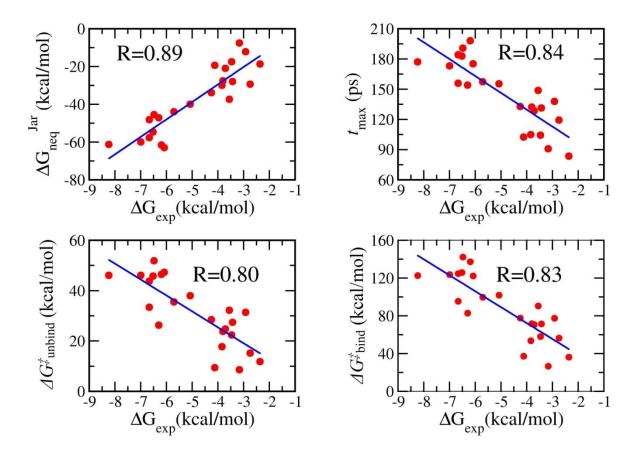
# **FIGURES**



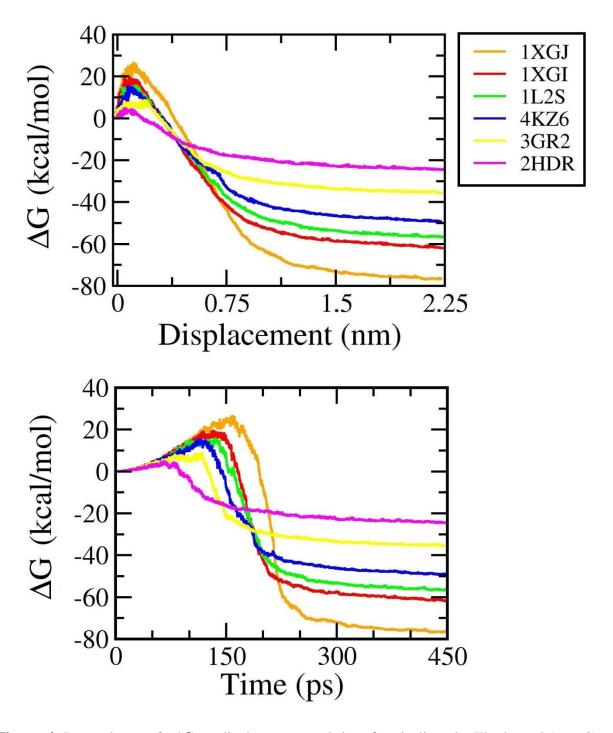
**Figure 1**: Upper panel : beta-lactamase (AmpC) protein in complex with ligand and the pulling force directed along the z-direction. Lower panel: Typical dependence of pulling force on time (left) and of Jarzynski's binding free energy on displacement (right), where data were obtained in one SMD trajectory for 1XGJ complex. Here  $t_{\rm max}$  refers to time when the force experienced by ligand reaches maximum. TS, bound and unbound refer to the transition, bound and unbound states, respectively.



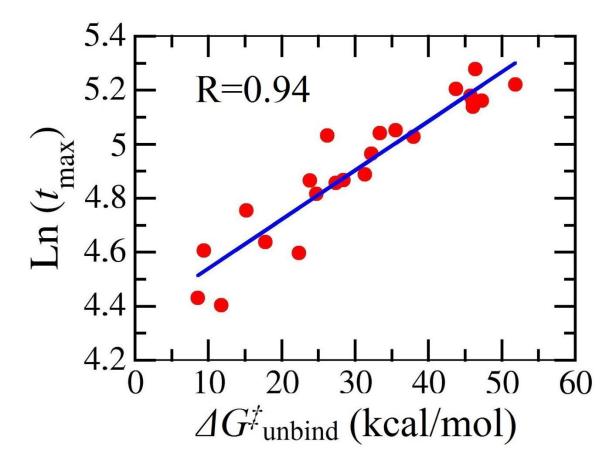
**Figure 2:** Shown are force-time (upper-left), force-displacement (upper-right) and pulling work-time profiles (lower) of six chosen complexes. The results were obtained from 20 independent MD runs.



**Figure 3.** Correlation of Jarzynski's non-equilibrium binding free energy, rupture time,  $\Delta G^{\ddagger}_{\text{unbind}}$  and  $\Delta G^{\ddagger}_{\text{bind}}$  with experimental data obtained for 23 compounds. Correlation level R is also shown.



**Figure 4.** Dependence of  $\Delta G$  on displacement and time for six ligands. The bound (at t=0) and unbound (at  $t_{end}$ ) are seaparated by the TS corresponding to maximal  $\Delta G$ . Results were averaged over 20 independent SMD runs.



**Figure 5.** Dependence of unbinding time  $t_{\text{max}}$  (in ps ) on unbinding barrier  $\Delta G^{\ddagger}_{\text{unbind}}$ . Results were obtained for 23 ligands (see also Table 1). Correlation level R=0.94.

Table 1: The non-equilibrium binding free energy  $\Delta G_{\rm neq}^{\ \ Jar}$ , work of external force  $W_{pull}$ , rupture force  $F_{\rm max}$ , unbinding time  $t_{\rm max}$ ,  $\Delta G^{\ddagger}_{\rm unbind}$  and  $\Delta G^{\ddagger}_{\rm bind}$  were obtained in 20 independent SMD trajectories for 23 complexes. Experimental binding free energy  $\Delta G_{\rm exp}$  was taken from different sources mentioned in the text.

N°	PDB ID	$\Delta G_{\rm exp}$ (kcal/mol)	$\Delta G_{ m neq}^{ m Jar}$ (kcal/mol)	W <sub>pull</sub> (kcal/mol)	F <sub>max</sub> (pN)	t <sub>max</sub> (ps)	$\Delta G^{\ddagger}_{\text{unbind}}$ (kcal/mol)	$\Delta G^{^{\ddagger} \text{bind}}$ (kcal/mol)
1	1XGJ	-8.2	$-61.3 \pm 1.0$	$76.5 \pm 8.0$	$819.7 \pm 103.6$	$176.9 \pm 23.4$	$46.1 \pm 11.2$	$122.6 \pm 18.1$
2	2R9W	-7.0	$-60.0 \pm 1.0$	$77.6 \pm 12.7$	$811.6 \pm 114.6$	$173.1 \pm 21.8$	$46.0 \pm 14.9$	$123.7 \pm 26.9$
3	1XGI	-6.7	$-48.1 \pm 1.1$	$61.9 \pm 6.4$	$707.7 \pm 73.4$	$155.9 \pm 13.7$	$33.4 \pm 10.0$	$95.4 \pm 15.4$
4	2R9X	-6.7	$-57.6 \pm 1.1$	$80.9 \pm 12.6$	$820.8 \pm 150.9$	$184.2 \pm 20.1$	$43.7 \pm 19.4$	$124.6 \pm 31.2$
5	4JXS	-6.5	$-54.7 \pm 1.1$	$80.1 \pm 15.5$	$818.7 \pm 114.3$	$182.9 \pm 24.6$	$45.7 \pm 16.5$	$125.8 \pm 27.1$
6	4JXW	-6.5	$-45.4 \pm 1.1$	$90.3 \pm 22.8$	$864.1 \pm 206.1$	$190.8 \pm 40.4$	$51.9 \pm 28.5$	$142.2 \pm 41.1$
7	1L2S	-6.3	$-47.2 \pm 1.0$	$56.6 \pm 6.7$	$653.2 \pm 69.3$	$153.9 \pm 11.8$	$26.2 \pm 8.9$	$82.8 \pm 13.8$
8	4JXV	-6.2	$-61.6 \pm 1.1$	$90.7 \pm 14.4$	$859.7 \pm 140.5$	$198.1 \pm 22.3$	$46.4 \pm 19.7$	$137.1 \pm 34.2$
9	2PU2	-6.1	$-62.9 \pm 1.0$	$74.9 \pm 7.7$	$826.1 \pm 60.2$	$175.2 \pm 14.1$	$47.3 \pm 7.0$	$122.2 \pm 13.4$
10	4KZ4	-5.7	$-43.9 \pm 1.1$	$64.1 \pm 12.3$	$725.4 \pm 97.3$	$157.3 \pm 14.1$	$35.5 \pm 13.9$	$99.6 \pm 24.3$
11	4KZA	-5.1	$-39.9 \pm 1.1$	$63.9 \pm 12.2$	$728.3 \pm 120.1$	$155.2 \pm 23.4$	$37.9 \pm 12.9$	$101.8 \pm 23.8$
12	4KZ6	-4.3	$-33.8 \pm 1.1$	$49.1 \pm 8.1$	$623.3 \pm 97.9$	$132.8 \pm 18.1$	$28.4 \pm 9.5$	$77.6 \pm 17.0$
13	3GRJ	-4.1	$-19.3 \pm 1.1$	$27.8 \pm 4.8$	$392.7 \pm 39.8$	$102.4 \pm 16.1$	$9.3 \pm 3.6$	$37.2 \pm 6.0$
14	4KZ8	-3.8	$-29.9 \pm 1.0$	$35.9 \pm 3.4$	$492.1 \pm 48.7$	$104.8 \pm 12.5$	$17.7 \pm 4.5$	53.6 ± 6.9
15	4KZ3	-3.8	$-27.5 \pm 1.1$	$47.8 \pm 8.8$	$595.7 \pm 86.4$	$132.1 \pm 20.7$	$23.8 \pm 8.9$	$71.6 \pm 17.5$
16	3GSG	-3.7	$-21.0 \pm 1.1$	$45.7 \pm 12.5$	$578.7 \pm 132.7$	$128.4 \pm 26.8$	$24.7 \pm 12.7$	$70.5 \pm 24.7$
17	3GVB	-3.5	$-37.4 \pm 1.0$	$58.2 \pm 19.9$	$661.3 \pm 175.7$	$148.7 \pm 34.2$	$32.3 \pm 19.7$	$90.4 \pm 37.7$
18	3GR2	-3.5	$-17.5 \pm 1.1$	$35.6 \pm 8.6$	$524.7 \pm 105.3$	$104.2 \pm 23.7$	$22.3 \pm 7.9$	57.9 ± 16.2
19	4KZ7	-3.4	$-27.9 \pm 1.0$	$43.5 \pm 9.3$	$607.1 \pm 111.2$	$131.3 \pm 17.4$	$27.4 \pm 10.8$	$71.3 \pm 19.3$
20	2HDU	-3.2	$-7.5 \pm 1.1$	$17.9 \pm 4.7$	$337.8 \pm 51.4$	$90.7 \pm 26.9$	$8.7 \pm 3.4$	$26.5 \pm 7.1$
21	3GV9	-2.9	$-12.2 \pm 1.1$	$46.1 \pm 14.5$	$638.9 \pm 166.8$	$137.8 \pm 29.2$	$31.3 \pm 15.6$	$77.4 \pm 30.1$
22	4KZ5	-2.7	-29.3 ± 1.0	$41.2 \pm 9.7$	$484.8 \pm 100.1$	$119.3 \pm 21.4$	$15.2 \pm 9.6$	$56.4 \pm 19.0$
23	2HDR	-2.4	$-18.6 \pm 1.0$	$24.4 \pm 3.3$	$388.7 \pm 104.6$	$83.5 \pm 10.8$	$11.8 \pm 3.1$	$36.1 \pm 5.8$

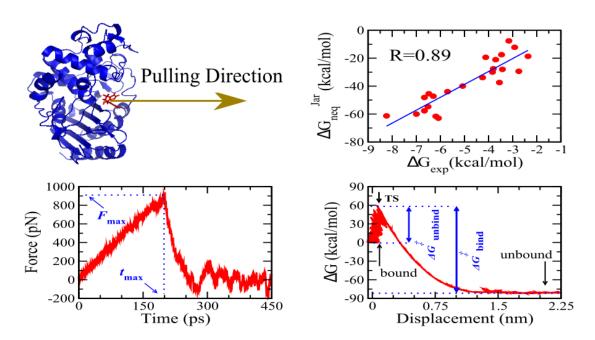
### **References**

- (1) Trott, O.; Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **2010**, *31*, 455-461.
- (2) Zwanzig, R. W. High-temperature equation of state by a perturbation method. I. nonpolar gases. *The Journal of Chemical Physics* **1954**, *22*, 1420-1426.
- (3) Ngo, S. T.; Mai, B. K.; Hiep, D. M.; Li, M. S. Estimation of the Binding Free Energy of AC1NX476 to HIV-1 Protease Wild Type and Mutations Using Free Energy Perturbation Method. *Chemical biology & drug design* **2015**, *86*, 546-558.
- (4) Kirkwood, J. G. Statistical mechanics of fluid mixtures. *The Journal of Chemical Physics* **1935**, *3*, 300-313.
- (5) Leonis, G.; Steinbrecher, T.; Papadopoulos, M. G. A contribution to the drug resistance mechanism of Darunavir, Amprenavir, Indinavir, and Saquinavir complexes with HIV-1 protease due to flap mutation I50V: A systematic MM–PBSA and thermodynamic integration study. *Journal of chemical information and modeling* **2013**, *53*, 2141-2153.
- (6) Åqvist, J.; Medina, C.; Samuelsson, J.-E. A new method for predicting binding affinity in computer-aided drug design. *Protein Engineering, Design and Selection* **1994**, *7*, 385-391.
- (7) Lee, F. S.; Chu, Z.-T.; Bolger, M. B.; Warshel, A. Calculations of antibody-antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603. *Protein Engineering, Design and Selection* **1992**, *5*, 215-228.
- (8) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research* **2000**, *33*, 889-897.
- (9) Wang, L.; Wu, Y. J.; Deng, Y. Q.; Kim, B.; Pierce, L.; Krilov, G.; Lupyan, D.; Robinson, S.; Dahlgren, M. K.; Greenwood, J.; Romero, D. L.; Masse, C.; Knight, J. L.; Steinbrecher, T.; Beuming, T.; Damm, W.; Harder, E.; Sherman, W.; Brewer, M.; Wester, R.; Murcko, M.; Frye, L.; Farid, R.; Lin, T.; Mobley, D. L.; Jorgensen, W. L.; Berne, B. J.; Friesner, R. A.; Abel, R. Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *Journal of the American Chemical Society* **2015**, *137*, 2695-2703.
- (10) Hou, T. J.; Wang, J. M.; Li, Y. Y.; Wang, W. Assessing the Performance of the Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated From Docking. *Journal of Computational Chemistry* **2011**, *32*, 866-877.
- (11) Grubmuller, H.; Heymann, B.; Tavan, P. Ligand binding: Molecular mechanics calculation of the streptavidin biotin rupture force. *Science* **1996**, *271*, 997-999.
- (12) Isralewitz, B.; Gao, M.; Schulten, K. Steered molecular dynamics and mechanical functions of proteins. *Current Opinion in Structural Biology* **2001**, *11*, 224-230.
- (13) Kumar, S.; Li, M. S. Biomolecules under mechanical force. *Physics Reports-Review Section of Physics Letters* **2010**, *486*, 1-74.
- (14) Mai, B. K.; Viet, M. H.; Li, M. S. Top leads for swine influenza A/H1N1 virus revealed by steered molecular dynamics approach. *Journal of chemical information and modeling* **2010**, *50*, 2236-2247.
- (15) Colizzi, F.; Perozzo, R.; Scapozza, L.; Recanatini, M.; Cavalli, A. Single-Molecule Pulling Simulations Can Discern Active from Inactive Enzyme Inhibitors. *Journal of the American Chemical Society* **2010**, *132*, 7361-7371.
- (16) Suan Li, M.; Khanh Mai, B. Steered molecular dynamics-a promising tool for drug design. *Current Bioinformatics* **2012**, *7*, 342-351.

- (17) Nguyen, T. T.; Tran, D. P.; Huy, P. D. Q.; Hoang, Z.; Carloni, P.; Pham, P. V.; Nguyen, C.; Li, M. S. Ligand binding to anti-cancer target CD44 investigated by molecular simulations. *Journal of Molecular Modeling* **2016**, *22*.
- (18) Thai, N. Q.; Tseng, N. H.; Vu, M. T.; Nguyen, T. T.; Linh, H. Q.; Hu, C. K.; Chen, Y. R.; Li, M. S. Discovery of DNA dyes Hoechst 34580 and 33342 as good candidates for inhibiting amyloid beta formation: in silico and in vitro study. *Journal of Computer-Aided Molecular Design* **2016**, *30*, 639-650.
- (19) Thai, N. Q.; Nguyen, H. L.; Linh, H. Q.; Li, M. S. Protocol for fast screening of multi-target drug candidates: Application to Alzheimer's disease. *Journal of Molecular Graphics and Modelling* **2017**, *77*, 121-129.
- (20) Van Vuong, Q.; Nguyen, T. T.; Li, M. S. A New Method for Navigating Optimal Direction for Pulling Ligand from Binding Pocket: Application to Ranking Binding Affinity by Steered Molecular Dynamics. *Journal of Chemical Information and Modeling* **2015**, *55*, 2731-2738.
- (21) Jarzynski, C. Nonequilibrium equality for free energy differences. *Physical Review Letters* **1997**, *78*, 2690-2693.
- (22) Park, S.; Schulten, K. Calculating potentials of mean force from steered molecular dynamics simulations. *Journal of Chemical Physics* **2004**, *120*, 5946-5961.
- (23) Hummer, G.; Szabo, A. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 3658-3661.
- (24) Tondi, D.; Morandi, F.; Bonnet, R.; Costi, M. P.; Shoichet, B. K. Structure-based optimization of a non- $\beta$ -lactam lead results in inhibitors that do not up-regulate  $\beta$ -lactamase expression in cell culture. *Journal of the American Chemical Society* **2005**, *127*, 4632-4639.
- (25) Babaoglu, K.; Simeonov, A.; Irwin, J. J.; Nelson, M. E.; Feng, B.; Thomas, C. J.; Cancian, L.; Costi, M. P.; Maltby, D. A.; Jadhav, A. Comprehensive mechanistic analysis of hits from high-throughput and docking screens against  $\beta$ -lactamase. *Journal of medicinal chemistry* **2008**, *51*, 2502-2511.
- (26) Powers, R. A.; Morandi, F.; Shoichet, B. K. Structure-based discovery of a novel, noncovalent inhibitor of AmpC β-lactamase. *Structure* **2002**, *10*, 1013-1023.
- (27) Barelier, S.; Eidam, O.; Fish, I.; Hollander, J.; Figaroa, F.; Nachane, R.; Irwin, J. J.; Shoichet, B. K.; Siegal, G. Increasing chemical space coverage by combining empirical and computational fragment screens. *ACS chemical biology* **2014**, *9*, 1528-1535.
- (28) Teotico, D. G.; Babaoglu, K.; Rocklin, G. J.; Ferreira, R. S.; Giannetti, A. M.; Shoichet, B. K. Docking for fragment inhibitors of AmpC  $\beta$ -lactamase. *Proceedings of the National Academy of Sciences* **2009**, *106*, 7455-7460.
- (29) Babaoglu, K.; Shoichet, B. K. Deconstructing fragment-based inhibitor discovery. *Nature chemical biology* **2006**, *2*, 720-723.
- (30) Drawz, S. M.; Bonomo, R. A. Three decades of  $\beta$ -lactamase inhibitors. *Clinical microbiology reviews* **2010**, 23, 160-201.
- (31) Drawz, S. M.; Papp-Wallace, K. M.; Bonomo, R. A. New  $\beta$ -lactamase inhibitors: a therapeutic renaissance in an MDR world. *Antimicrobial agents and chemotherapy* **2014**, *58*, 1835-1846.
- (32) Farina, D.; Spyrakis, F.; Venturelli, A.; Cross, S.; Tondi, D.; Paola Costi, M. The inhibition of extended spectrum  $\beta$ -lactamases: hits and leads. *Current medicinal chemistry* **2014**, *21*, 1405-1434.
- (33) Ambler, R. The Structure of \$\ beta \$-Lactamases. Philosophical Transactions of the Royal Society B: Biological Sciences **1980**, 289, 321-331.
- (34) Iaconis, J. P.; Pitkin, D. H.; Sheikh, W.; Nadler, H. L. Comparison of antibacterial activities of meropenem and six other antimicrobials against Pseudomonas aeruginosa isolates from North American studies and clinical trials. *Clinical infectious diseases* **1997**, *24*, S191-S196.
- (35) Lucasti, C.; Popescu, I.; Ramesh, M. K.; Lipka, J.; Sable, C. Comparative study of the efficacy and safety of ceftazidime/avibactam plus metronidazole versus

- meropenem in the treatment of complicated intra-abdominal infections in hospitalized adults: results of a randomized, double-blind, Phase II trial. *Journal of Antimicrobial Chemotherapy* **2013**, *68*, 1183-1192.
- (36) Sgrignani, J.; De Luca, F.; Torosyan, H.; Docquier, J.-D.; Duan, D.; Novati, B.; Prati, F.; Colombo, G.; Grazioso, G. Structure-based approach for identification of novel phenylboronic acids as serine-β-lactamase inhibitors. *Journal of computer-aided molecular design* **2016**, *30*, 851-861.
- (37) McKinney, D. C.; Zhou, F.; Eyermann, C. J.; Ferguson, A. D.; Prince, D. B.; Breen, J.; Giacobbe, R. A.; Lahiri, S.; Verheijen, J. C. 4, 5-disubstituted 6-aryloxy-1, 3-dihydrobenzo [c][1, 2] oxaboroles are broad-spectrum serine  $\beta$ -lactamase inhibitors. *ACS Infectious Diseases* **2015**, *1*, 310-316.
- (38) Er, M.; Isildak, G.; Tahtaci, H.; Karakurt, T. Novel 2-amino-1, 3, 4-thiadiazoles and their acyl derivatives: Synthesis, structural characterization, molecular docking studies and comparison of experimental and computational results. *Journal of Molecular Structure* **2016**, *1110*, 102-113.
- (39) Parthasarathy, K.; Praveen, C.; Saranraj, K.; Balachandran, C.; Kumar, P. S. Synthesis, antimicrobial and cytotoxic evaluation of spirooxindole [pyrano-bis-2H-l-benzopyrans]. *Medicinal Chemistry Research* **2016**, *25*, 2155-2170.
- (40) Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; Biedermannova, L.; Sochor, J.; Damborsky, J. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *Plos Computational Biology* **2012**, *8*.
- (41) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry* **2004**, *25*, 1400-1415.
- (42) 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, 926-935.
- (43) Bruce, N. J.; Ganotra, G. K.; Kokh, D. B.; Sadiq, S. K.; Wade, R. C. New approaches for computing ligand–receptor binding kinetics. *Current Opinion in Structural Biology* **2018**, *49*, 1-10.
- (44) Li, M. S. Ligand migration and steered molecular dynamics in drug discovery: Comment on "Ligand diffusion in proteins via enhanced sampling in molecular dynamics" by Jakub Rydzewski and Wieslaw Nowak. *Physics of Life Reviews* **2017**, *22-23*, 79-81.
- (45) Gu, J.; Li, H.; Wang, X. A Self-Adaptive Steered Molecular Dynamics Method Based on Minimization of Stretching Force Reveals the Binding Affinity of Protein–Ligand Complexes. *Molecules* **2015**, *20*, 19236.
- (46) Nguyen, H.; Nguyen, H. L.; Linh, H. Q.; Nguyen, M. T. Binding affinity of the L-742,001 inhibitor to the endonuclease domain of A/H1N1/PA influenza virus variants: Molecular simulation approaches. *Chemical Physics* **2018**, *500*, 26-36.
- (47) Truong, D. T.; Nguyen, M. T.; Vu, V. V.; Ngo, S. T. Fast pulling of ligand approach for the design of  $\beta$ -secretase 1 inhibitors. *Chemical Physics Letters* **2017**, *671*, 142-146.
- (48) Heymann, B.; Grubmuller, H. AN02/DNP-hapten unbinding forces studied by molecular dynamics atomic force microscopy simulations. *Chemical Physics Letters* **1999**, 303, 1-9.
- (49) Bu, L. T.; Beckham, G. T.; Shirts, M. R.; Nimlos, M. R.; Adney, W. S.; Himmel, M. E.; Crowley, M. F. Probing Carbohydrate Product Expulsion from a Processive Cellulase with Multiple Absolute Binding Free Energy Methods. *Journal of Biological Chemistry* **2011**, *286*, 18161-18169.

# **TOC Graphic**



# Probing Binding Affinity by Jarzynski's Non-equilibrium Binding Free Energy and Rupture Time

Duc Toan Truong<sup>1,2</sup> and Mai Suan Li<sup>1,3,\*</sup>

# SUPPORTING INFORMATION

<sup>&</sup>lt;sup>1</sup>Institute for Computational Sciences and Technology, SBI building, Quang Trung Software City, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Vietnam

<sup>&</sup>lt;sup>2</sup>Department of Theoretical Physics, Faculty of Physicsand Engineering Physics, Ho Chi Minh University of Science, Vietnam

<sup>&</sup>lt;sup>3</sup>Institute of Physics, Polish Academy of Sciences, Al. Lotnikow 32/46, 02-668 Warsaw, Poland

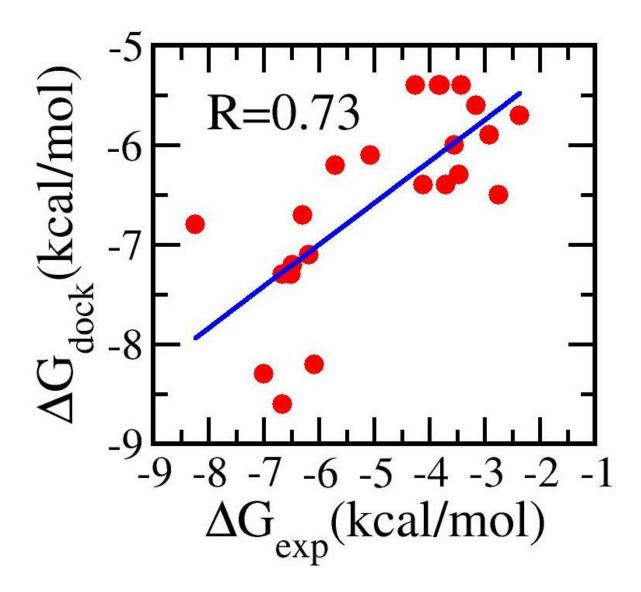
<sup>\*</sup>masli@ifpan.edu.pl

#### 1. Receptor and ligands

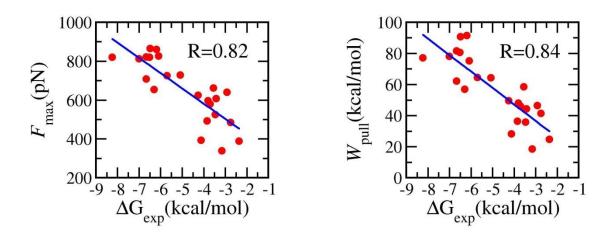
From the protein data bank (PDB) , we retrieved three dimensional structures of BL protein in complex with 23 small compounds. Their PDB identify codes as:  $1XGJ^1$ ,  $2R9W^2$ ,  $1XGI^1$ ,  $2R9X^2$ ,  $4JXS^2$ ,  $4JXW^2$ ,  $1L2S^3$ ,  $4JXV^2$ ,  $2PU2^2$ ,  $4KZ4^4$ ,  $4KZA^4$ ,  $4KZ6^4$ ,  $3GRJ^5$ ,  $4KZ8^4$ ,  $4KZ3^4$ ,  $3GSG^5$ ,  $3GVB^5$ ,  $3GR2^5$ ,  $4KZ7^4$ , 2HDU <sup>6</sup>,  $3GV9^5$ ,  $4KZ5^4$ , 2HDR <sup>6</sup>. The chosen complexes have the inhibition constant  $K_i$  ranging from  $1\mu M$  to  $19000\mu M$ . ID,  $K_i$ , chemical formula and 2D structures of all 23 ligands are listed in Table S1.

For SMD simulation we have used the PDB structures of all 23 complexes as starting configurations are almost identical. However, in docking simulation, where the receptor structure is rigid, the 1XGJ structure <sup>1</sup> (Fig. 1 in the main text) was chosen as a common receptor for all ligands.

- (1) Tondi, D.; Morandi, F.; Bonnet, R.; Costi, M. P.; Shoichet, B. K. Structure-based optimization of a non- $\beta$ -lactam lead results in inhibitors that do not up-regulate  $\beta$ -lactamase expression in cell culture. *Journal of the American Chemical Society* **2005**, *127*, 4632-4639.
- (2) Babaoglu, K.; Simeonov, A.; Irwin, J. J.; Nelson, M. E.; Feng, B.; Thomas, C. J.; Cancian, L.; Costi, M. P.; Maltby, D. A.; Jadhav, A. Comprehensive mechanistic analysis of hits from high-throughput and docking screens against  $\beta$ -lactamase. *Journal of medicinal chemistry* **2008**, *51*, 2502-2511.
- (3) Powers, R. A.; Morandi, F.; Shoichet, B. K. Structure-based discovery of a novel, noncovalent inhibitor of AmpC β-lactamase. *Structure* **2002**, *10*, 1013-1023.
- (4) Barelier, S.; Eidam, O.; Fish, I.; Hollander, J.; Figaroa, F.; Nachane, R.; Irwin, J. J.; Shoichet, B. K.; Siegal, G. Increasing chemical space coverage by combining empirical and computational fragment screens. *ACS chemical biology* **2014**, *9*, 1528-1535.
- (5) Teotico, D. G.; Babaoglu, K.; Rocklin, G. J.; Ferreira, R. S.; Giannetti, A. M.; Shoichet, B. K. Docking for fragment inhibitors of AmpC β-lactamase. *Proceedings of the National Academy of Sciences* **2009**, *106*, 7455-7460.
- (6) Babaoglu, K.; Shoichet, B. K. Deconstructing fragment-based inhibitor discovery. *Nature chemical biology* **2006**, *2*, 720-723.



**Figure S1.** Correlation between the binding energy, obtained by the docking method for 23 ligands, and the experimental data. The correlation level R=0.73.



**Figure S2**: Correlation between  $F_{\text{max}}(R=0.82)$  and  $W_{\text{pull}}(R=0.84)$  with experimental binding free energy  $\Delta G_{\text{exp}}$ . Results were obtained by averaging over 20 SMD runs.

**Table S1:** Inhibition constant, molecular weight, chemical formula and 2D structure of the 23 studied ligands. Their ID is given in the third column, while the second column refers to PDB ID of their complex with BL protein. Experimental data on inhibition constant were taken from different groups <sup>1-6</sup>.

N <sup>o</sup>	PDB ID	Ligand ID	Inhibition constant $\operatorname{Ki}(\mu M)$	Molecular weight $(g/mol)$	Chemical formula	2D structure
1	1XGJ	НТС	1	343.3	C <sub>12</sub> H <sub>29</sub> NO <sub>7</sub> S <sub>2</sub>	HO OH OH
2	2R9W	23C	8	389.3	C <sub>22</sub> H <sub>15</sub> NO <sub>6</sub>	HO OH
3	1XGI	NST	14	328.3	$C_{11}H_8N_2O_6S_2$	TE SE
5	4JXS	18U	18	341.3	$C_{13}H_{11}NO_6S_2$	HO OH
6	4JXW	1MW	19	369.4	C <sub>15</sub> H <sub>15</sub> NO <sub>6</sub> S <sub>2</sub>	N N N N N N N N N N N N N N N N N N N
7	1L2S	STC	26	317.7	C <sub>11</sub> H <sub>8</sub> ClNO <sub>4</sub> S <sub>2</sub>	HOO
8	4JXV	1MU	31	335.4	C <sub>14</sub> H <sub>13</sub> NO <sub>6</sub> S <sub>2</sub>	HO S N N N N N N N N N N N N N N N N N N

9	2PU2	DK2	37	355.3	C <sub>18</sub> H <sub>13</sub> NO <sub>7</sub>	но он
10	4KZ4	4A1	70	243.3	C <sub>10</sub> H <sub>13</sub> NO <sub>4</sub> S	OH HIN
11	4KZA	NZ9	200	247.3	C <sub>8</sub> H <sub>9</sub> NO <sub>4</sub> S <sub>2</sub>	HIN S
12	4KZ6	ZB6	800	231.3	C <sub>10</sub> H <sub>17</sub> NO <sub>3</sub> S	HS OHOME
13	3GRJ	G14	1000	188.2	C <sub>10</sub> H <sub>8</sub> N <sub>2</sub> O <sub>2</sub>	T Z
14	4KZ8	1U6	1600	200.3	C <sub>8</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub> S	H <sub>3</sub> C S N N O CH <sub>3</sub> O
15	4KZ3	1U1	1700	241.7	C₅H₄CINO₄S₂	H <sub>2</sub> N — S — OH
16	3GSG	GF1	2000	237.3	C <sub>12</sub> H <sub>15</sub> NO <sub>4</sub>	······································
17	3GVB	3GV	2600	221.2	C <sub>11</sub> H <sub>11</sub> NO <sub>4</sub>	ОН
18	3GR2	GF4	3000	194.2	C <sub>7</sub> H <sub>10</sub> N <sub>6</sub> O	H <sub>3</sub> C CH <sub>3</sub>

19	4KZ7	1U5	3200	198.2	$C_{10}H_{14}O_4$	H <sub>3</sub> C CH <sub>3</sub> OH
20	2HDU	F12	5000	185.2	C <sub>7</sub> H <sub>7</sub> NO₃S	OH N CH <sub>3</sub>
21	3GV9	GV9	7500	185.2	C <sub>7</sub> H <sub>7</sub> NO <sub>3</sub> S	OH CH <sub>3</sub>
22	4KZ5	1U3	10000	294.7	C <sub>13</sub> H <sub>11</sub> ClN <sub>2</sub> O <sub>4</sub>	HO NH H <sub>3</sub> C
23	2HDR	4A3	19000	153.1	C <sub>7</sub> H <sub>7</sub> NO <sub>3</sub>	HO OH NH <sub>2</sub>

Table S2: Ratio  $\Delta G^{\sharp}_{\text{bind}}/\Delta G^{\sharp}_{\text{unbind}}$  for 23 ligands. Results were obtained in SMD simulation.

N°	PDB ID	$\Delta G^{\sharp}_{\mathrm{bind}}/\!\Delta G^{\sharp}_{\mathrm{unbind}}$
1	1XGJ	$2.7 \pm 0.3$
2	2R9W	$2.7 \pm 0.3$
3	1XGI	$2.9 \pm 0.4$
4	2R9X	$2.9 \pm 0.5$
5	4JXS	$2.8 \pm 0.3$
6	4JXW	$2.7 \pm 0.1$
		$2.7 \pm 0.1$
7	1L2S	$3.2 \pm 0.5$
8	4JXV	$3.0 \pm 0.2$
9	2PU2	$2.6 \pm 0.1$
10	4KZ4	$2.8 \pm 0.2$
11	4KZA	$2.7 \pm 0.3$
12	4KZ6	$2.7 \pm 0.2$
13	3GRJ	$4.0\pm0.9$
14	4KZ8	$3.0 \pm 2.0$
15	4KZ3	$3.0 \pm 0.2$
16	3GSG	$2.9 \pm 0.5$
17	3GVB	$2.8 \pm 0.5$
18	3GR2	$2.6 \pm 0.2$
19	4KZ7	$2.6 \pm 0.4$
20	2HDU	3.1 ± 0.4
21	3GV9	$2.5 \pm 0.1$
22	4KZ5	$3.7 \pm 0.8$
23	2HDR	$3.1 \pm 0.3$
<u> </u>	1	