Acta Mech. Sin. (2012) 28(3):891–903
DOI 10.1007/s10409-012-0112-9

RESEARCH PAPER

Free energy calculation of single molecular interaction using Jarzynski's identity method: the case of HIV-1 protease inhibitor system

De-Chang Li · Bao-Hua Ji

Received: 20 February 2012 / Revised: 5 April 2012 / Accepted: 5 May 2012

Abstract Jarzynski' identity (JI) method was suggested a promising tool for reconstructing free energy landscape of biomolecular interactions in numerical simulations and experiments. However, JI method has not yet been well tested in complex systems such as ligand-receptor molecular pairs. In this paper, we applied a huge number of steered molecular dynamics (SMD) simulations to dissociate the protease of human immunodeficiency type I virus (HIV-1 protease) and its inhibitors. We showed that because of intrinsic complexity of the ligand-receptor system, the energy barrier predicted by JI method at high pulling rates is much higher than experimental results. However, with a slower pulling rate and fewer switch times of simulations, the predictions of JI method can approach to the experiments. These results suggested that the JI method is more appropriate for reconstructing free energy landscape using the data taken from experiments, since the pulling rates used in experiments are often much slower than those in SMD simulations. Furthermore, we showed that a higher loading stiffness can produce higher precision of calculation of energy landscape because it yields a lower mean value and narrower bandwidth of work distribution in SMD simulations.

Keywords Molecular dynamics simulation · Single molecular interaction · Molecular biomechanics · Steered molecular dynamics · Free energy calculation

The project was supported by the National Science Foundation of China (10732050, 10872115 and 11025208), Excellent Young Scholars Research Fund of Beijing Institute of Technology.

De-Chang Li · Bao-Hua Ji (✉)
Biomechanics and Biomaterials Lab,
Department of Applied Mechanics,
Beijing Institute of Technology,
Key Laboratory of Dynamics and Control of Flight Vehicle,
Ministry of Education, 100081 Beijing, China
e-mail: bhji@bit.edu.cn

1 Introduction

Specific molecular interaction is a distinguishing aspect of the life sciences, which is involved in nearly every facet of the cell cycle [1–3]. Examples are molecular recognition between ligand and receptor, antibody and antigen, and complementary strands of DNA/RNA, which are thought to be key modulations in cell adhesion, immunity and genetic processes, respectively. Such interactions are generated by multiple weak bonds between geometrically complementary surfaces and some of recognition sites, which are short range, non-covalent, but can be very strong. The knowledge of these interactions is helpful for the understanding of cell life cycle at molecular scale.

With significant progresses in micro- and nanotechnology, single-molecule manipulation technique is developing rapidly. Atomic force microscope (AFM), biomembrane force probe (BFP), laser optical tweezer (LOT), and magnetic tweezer (MT) are typical methods evolving in the study on interactions in single-molecule systems, e.g. the dissociation processes of molecular complexes (i.e. force-distance plot) and the rupture forces between ligand and its receptor [4–7]. Molecular dynamics (MD) simulations, e.g. steered molecular dynamics (SMD) are often used to mimic these kind of manipulations for studying ligand-receptor interactions [8, 9], protein unfolding [10] and antigen-antibody recognition [11]. In addition, the single-molecule dynamic force spectroscopy (DFS) method was developed for studying the energy landscape of molecular interactions based on the rupture processes in single-molecule experiments [4–7]. Although there are a few existing analytical models for extracting the parameters of energy landscape from the force spectroscopy [5, 12–16], these methods can only obtain limited parameters of the energy landscape, such as the energy barrier ΔG_{off}, the dissociation rate constant k_{off} and the width of the energy well x_β, while more information of energy landscape could not be obtained.

To obtain the entire information of energy landscape, more precise methods were developed. Umbrella sampling

(US), as a typical one [17–19], is to enhance local sampling along a certain pathway by adding a harmonic potential to the intrinsic energy landscape of the system. US method is widely used in the study on molecular mechanics, such as assessing the stability of Alzheimer's amyloid protofibrils [20], calculating the free energy of protein conformational transitions [21, 22] and the energy landscape of ligand-receptor systems [9, 23]. However, the calculation procedure of US method is very complicated and it is also infeasible for experiments. As a substitute, an identity given by Jarzynski [24] can be used for calculating the free energy landscape from both equilibrium and non-equilibrium processes. The Jarzynski's identity (JI) reads

$$\exp\left(-\beta\Delta G\right) = \left\langle \exp\left(-\beta W\right)\right\rangle_{\infty}, \tag{1}$$

where $\beta = 1/k_{\mathrm{B}}T$, k_{B} is the Boltzmann constant and T is the absolute temperature. $\langle ...\rangle_{\infty}$ denotes the mean value of work for infinite times of switch processes between two states. From JI method, we can obtain the free energy difference between two states as

$$\Delta G = \lim_{n\to\infty}\left\{-\frac{1}{\beta}\ln\left\langle \exp\left(-\beta W\right)\right\rangle_{n}\right\}. \tag{2}$$

Compared to US method, the calculation procedure of JI method is much simpler, in which one needs only to simply repeat the switch processes multiple times. Hammer and Sazbo [25] further showed that not only the free energy differences between two states, but also the free energy landscape can be reconstructed from single-molecule pulling experiments using the JI method. However, Hummer [26] and Park et al. [27, 28] showed that only in the study of simple systems the efficiency of JI method is comparable with US method, such as the separation of two methane molecules in water and the helix-coil transition of deca-alanine in vacuum. Cuendet and Michielin [11] showed that the free energy calculation by JI method in a protein-protein complex system is largely underestimated compared to the experimental results. Therefore, JI method has not been well tested for more complex systems in previous simulations.

In this paper, we applied JI method to study a ligand-receptor system of the protease of human immunodeficiency type I virus (HIV-1 protease) and its inhibitors (see Fig. 1). HIV-1 protease is a key protein in the life cycle of HIV-1 virus. It is a symmetric homo-dimeric aspartyl protease with a binding pocket for the substrates (see Figs. 1a–1c), which cleaves the gag and pol viral polyproteins at its active site to process the viral maturation [29]. Due to its indispensability for the infection of the virus, the HIV-1 protease is one of the primary targets of anti-AIDS therapy [30]. When it is bound with inhibitors, the function of the protease will be blocked, and the virus can not mature and consequently becomes noninfectious [30]. The dynamics of HIV-1 protease and its interactions with inhibitors were well studied by atomistic MD simulations [9, 31, 32], as well as coarse-grained methods [33–36]. In addition, there are a lot of experiment data of the binding free energy between the protease and varied inhibitors [37]. Thus, it is a good system

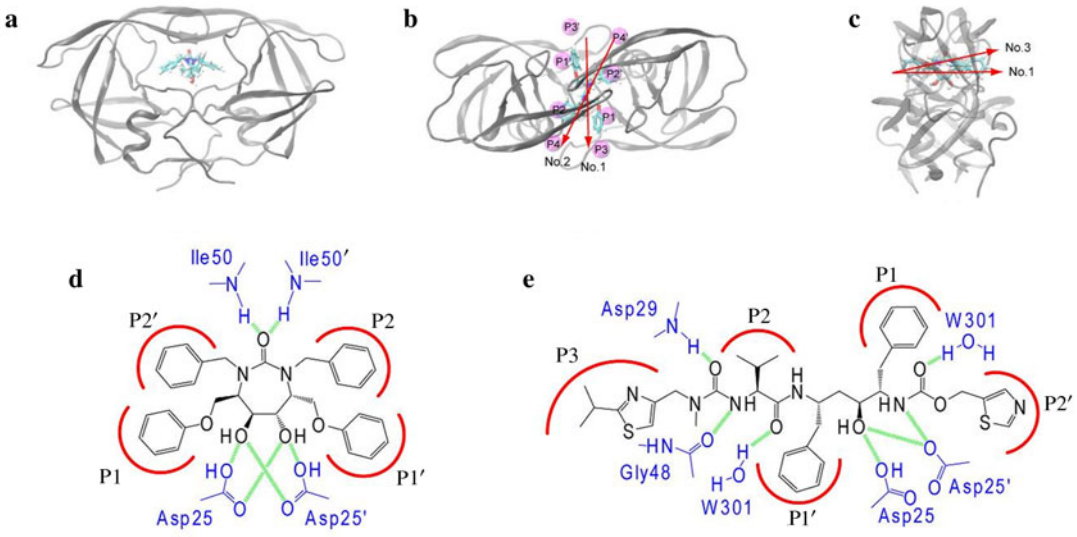

Fig. 1 Molecular structure of HIV-1 protease and its inhibitors. **a, b** and **c** 3D views of HIV-1 protease bound with inhibitor AHA001. The pink spheres in **b** represent the subsites of inhibitors and their corresponding positions. The arrows in **b** and **c** represent the pulling directions in SMD simulations. The structures and interactions of inhibitors with HIV-1 protease are illustrated in **d** AHA001 and **e** ABT538. Sidechains of the inhibitors are labeled as P1, P2, etc. The residues of protease are labeled in blue in **e**. The possible H-bonds between the protease and inhibitors are labeled in green lines

for studying the ligand-receptor like interactions. In this paper, two drug inhibitors AHA001 and ABT538 were chosen which have significantly different interaction strength with HIV-1 protease. AHA001 is a type of cyclic urea inhibitor (see Fig. 1d) that has a weak interaction with HIV-1 protease (i.e., $k_{\text{off}} \sim 100\ \text{s}^{-1}$) [37], while ABT538 is an inhibitor already approved in clinical therapy of AIDS (see Fig. 1e) that has much stronger interaction with HIV-1 protease (i.e. $k_{\text{off}} \sim 10^{-3}\ \text{s}^{-1}$) [37]. The significantly different interaction strength of the two inhibitors with HIV-1 protease allows us to evaluate the validity of different analysis methods. We applied a huge number of SMD simulations to dissociate the HIV-1 protease-inhibitor complexes, the total simulation time is as long as 12.2 μs. Our aims is to calculate the free energy landscapes of the receptor-ligand system using JI method, with focus laid on how to improve the JI method for complex molecular systems.

2 Methods

2.1 JI method

To reconstruct the energy landscape along a reaction coordinate, Hammer and Sazbo [25] rewrote the Jarzynski's identity as

$$\Delta G(z) = \lim_{n\to\infty}\left\{-\frac{1}{\beta}\ln\left\langle\delta(z-z_t)\exp(-\beta W)\right\rangle_n\right\}, \tag{3}$$

where z is the reaction coordinate and z_t is the position of the inhibitors at time t. According to Eq. (3) we can obtain the free energy differences along the reaction coordinate by calculating the work done by external forces.

2.1.1 JI method with block-average analysis

To achieve the asymptotical result $\langle\cdots\rangle_n \to \langle\cdots\rangle_\infty$ when we can only do finite number of times of SMD simulations, the block-average method [38] is often used. Suppose we have obtained N work values $\Omega_N = (W_1, W_2, \cdots, W_N)$ by N times of SMD simulations. The block-average method can be done as follows.

(1) Choose n different values randomly from Ω_N, generating a subset with relabeled indices $(W_1, W_2, \cdots, W_n)$. Let i denote the index for a particular subset, i.e. $i = 1$ for the first subset, $i = 2$ for the second one and so on, each containing n values.

(2) Use JI method to obtain a free energy estimation for one subset,

$$\Delta G(z)_{n,i} = -\frac{1}{\beta}\ln\left[\frac{1}{n}\sum_{j\in\text{subset i}}\delta(z-z_t)\exp\left(-\beta W_j\right)\right]. \tag{4}$$

(3) Repeat Steps 1 and 2 until m subsets (usually $m \sim 100\times N/n$) have been obtained. Then the average ΔG_n and its standard deviation σ_n can be obtained using

$$\Delta G(z)_n = \frac{1}{m}\sum_{i=1}^{m}\Delta G(z)_{n,i} = \frac{1}{m}\sum_{i=1}^{m}\left\{-\frac{1}{\beta}\ln\left[\frac{1}{n}\sum_{j\in\text{subset i}}\delta(z-z_t)\exp\left(-\beta W_j\right)\right]\right\} \tag{5}$$

$$\sigma(z)_n^2 = \frac{1}{m}\sum_{i=1}^{m}\left(\Delta G(z)_{n,i} - \Delta G(z)_n\right)^2. \tag{6}$$

Note that when $n = 1$, we got the free energy difference as $\Delta G_1 = \langle W\rangle$, but when $n = N$ we can directly get ΔG_N by $\exp(-\beta\Delta G_N) = \langle\exp(-\beta W)\rangle_N$. Following the block-average steps 1 to 3, we got N values of free energy difference as $(\Delta G_1, \Delta G_2, \cdots, \Delta G_N)$. Ytreberg and Zuckerman [38] suggested that the energy difference ΔG_∞ can be estimated by the intercept of the linear extrapolation of $\Delta G_n \sim (1/n)^{0.5}$ plot. Moreover, one can also examine the non-linear extrapolation of $\Delta G_n \sim (1/n)^\tau$ (the exponent τ is also a fitting parameter) to estimate the energy difference ΔG_∞.

2.1.2 JI method with cumulant expansion analysis

Alternatively, the right side of JI theory (Eq. (2)) can be expanded as

$$\ln\langle\exp(-\beta W)\rangle = -\beta\langle W\rangle + \frac{\beta^2}{2}\left(\left\langle W^2\right\rangle - \left\langle W\right\rangle^2\right) - \frac{\beta^3}{3!}\left(\left\langle W^3\right\rangle - 3\left\langle W^2\right\rangle\left\langle W\right\rangle + 2\left\langle W\right\rangle^3\right) + \cdots \tag{7}$$

If the work distribution is Gaussian, the third and higher cumulants are identically zero [39]. Thus the free energy difference can be simply rewritten as

$$\Delta G = \langle W\rangle - \frac{1}{2}\beta\sigma_w^2, \tag{8}$$

where σ_w^2 is

$$\sigma_w^2 = \left\langle W^2\right\rangle - \left\langle W\right\rangle^2. \tag{9}$$

The cumulant expansion method has advantage of reducing the statistical error when we can only obtain a limited number of samples.

2.2 SMD simulations

In this paper, we chose HIV-1 protease and its inhibitors as the system under study. The structures for the inhibitor bound complexes were retrieved from Protein Data Bank with PDB codes: 1AJX [40] for AHA001 and 1HXW [41] for ABT538 bound complexes. The catalytic Asp side chains of the bound complex were protonated according to the experiments and theoretical calculations, i.e. both side chains of Asp25/Asp25′ were protonated for AHA001 bound complex [42], and one of the Asp25/Asp25′ was protonated for ABT538 bound complex [43].

The simulations were performed using Gromacs package [44, 45] with the AMBER force field of ffamber99 [46], in which the all-atom force field parameters of inhibitors were obtained by the ANTECHAMBER module

and GAFF [47, 48] with AM1-BCC [49] charges in AMBER package [50].

Each system was solved in a 90 nm × 80 nm × 80 nm TIP3P [51] water box, with about 15 000 water molecules. Appropriate chlorine ions were added to neutralize the system. Particle Mesh Ewald (PME) [52] method was used to calculate the long-range electrostatic interactions. The systems were minimized firstly using steepest descent algorithm by 10 000 steps. Then, it was gradually heated from 200 to 300 K in 200 ps, while positional restraints were used for the heavy atoms of the protease and inhibitor. The restraint force constants were gradually decreased from 239 kcal/(mol · nm^2) to 0 kcal/(mol · nm^2) (1 kcal = 4.182 kJ) in a few stages. The equilibrium MD simulations were conducted for 40 ns, fully unrestrained at 300 K with a pressure of 0.1 MPa with the Berendsen algorithm [53]. The average structures from the trajectory of each 40 ns MD simulations were used to generate the starting structures for SMD simulations. In addition, the LINCS algorithm [54] was applied to constrain the covalent bonds with H-atoms. The time step of the simulations is 2.0 fs. The cut-off of the non-bonded interactions was set to 1 nm. The non-bonded pairs were updated at every 10 steps.

To get the free energy landscape by JI method first we need to obtain N work values along a certain reaction coordinate using SMD simulation. The steered "dummy" atom was attached via a spring to the center of mass (COM) of the inhibitors. The spring is to mimic the elasticity of the loading device, e.g. the probe beam of AFM. The force applied to the inhibitor can be represented as

$$f = K_{\text{spring}}\,(vt - x_{\text{I}})\,, \tag{10}$$

where K_{spring} is the spring constant, v is the pulling velocity and t is the simulation time, x_{I} is the position of the inhibitor. The pulling rate can be obtained as

$$\dot{f} = K_{\text{spring}} v. \tag{11}$$

Because Sadiq et al. [55] showed that the inhibitor tends to laterally escape the binding pocket, and Trylska et al. [56] also showed that the peptide product would laterally slide out from the binding pocket, we chose the pulling direction being consistent with the lateral direction, indicated by the vector from the mass center of residue Arg8 to that of residue Arg8′, which is approximately pointed from the subsite P3′ to P3 (direction No. 1, see Figs. 1b and 1c). To test the sensitivity of the rupture forces and the dissociation processes in different pulling directions, we also tested another two directions. One is indicated by the vectors from the mass center of residue Asp29 to Asp29′, which is approximately pointed from P4′ to P4 (direction No. 2, see Fig. 1b). The other one is labeled by direction No. 3, pointed from the position of Cα atom in Arg8 to Cγ atom in Arg8′ (see Fig. 1c). To prevent translation and rotation of the protease molecule upon application of steering forces, several Cα atoms of the protease were held by positional restraints, including Cα atoms of N- and C-termini residues Pro1, Phe99, Pro1′ and Phe99′, as well as Arg8, Leu23, Pro81 and Asp30′ at the back end of the protease. The restraint force constants were set to 239 kcal/(mol · nm^2). In order to get a correct ensemble of the simulations, all of the SMD simulations used Nosé-Hoover [57, 58] temperature and Parrinello-Rahman [59] pressure coupling.

A huge number of SMD simulations were applied to the Pr-ABT538 and Pr-AHA001 complexes, with different pulling rates (see Table 1). Each pulling rate was designed to perform SMD simulations of similar time scale with different number of switch times. The total time scale of the SMD simulations is 12.2 μs. The pulling direction in JI simulations is direction No. 1.

Table 1 The SMD simulation parameters and the number of simulations performed to construct Jarzynski's identity

Inhibitors	SMD simulation parameters		Simulation time N/ns	Total simulation time of each system/μs	Work distribution $\langle W\rangle \pm \sigma$/(kcal · mol^{-1})	Energy barrier in experiment ΔG_{off}/(kcal · mol^{-1}) [37]
	K/(pN · nm^{-1})	v/(nm · ns^{-1})				
ABT538	6 947.7	0.50	5×300	1.5	82.45±9.60	21.08
		0.02	80×20	1.6	38.31±7.48	
		0.01	150×10	1.5	32.26±6.21	
	694.8	0.50	5×300	1.5	106.90±19.83	
AHA001	6 947.7	0.50	5×300	1.5	54.88±7.32	14.79
		0.02	80×20	1.6	25.02±5.66	
		0.01	150×10	1.5	23.35±6.41	
	694.8	0.5	5×300	1.5	54.97±8.39	

2.3 US method

To compare with the JI results, we also performed US analysis of the same systems. From the SMD simulation trajectories, snapshots were taken to generate the starting configurations for the umbrella sampling. An asymmetrical distribution of sampling windows was used, in which the window spacing used was about 0.05 nm when the COM separation between the inhibitor and the protease active site (i.e. residues Asp25 and Asp25′) was below 1 nm, while the window spacing was about 0.1 nm when the COM separation

was above 1 nm. Such spacing allowed for detailed sampling at smaller COM distance, which resulted in about 40 windows. In each window, 5 ns of MD simulation was performed which results in a total simulation time of ~ 200 ns in each system. The results were analyzed with the weighted histogram analysis method (WHAM) [60].

3 Results and discussion

3.1 Rupture force

The rupture force of the bound complexes was calculated using the SMD simulation under different pulling rates (see Methods section). As shown in Fig. 2, the rupture force of Pr-ABT538 system is larger than that of Pr-AHA001 system. Figure 3 shows force-time plots of the typical dissociation process. In addition, Fig. 2 shows that the rupture forces are

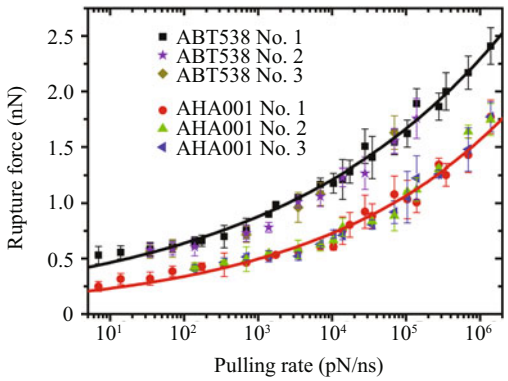

Fig. 2 Rupture forces of inhibitor bound complexes calculated by the SMD simulations in term of pulling rates at different pulling directions. The rupture force was defined as the highest peak loading force during the pulling process in SMD simulations. The solid lines are the exponential fits according to previous studies [61, 62] to guide the view

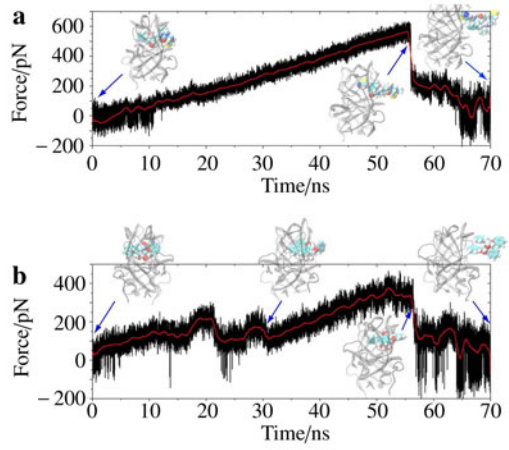

Fig. 3 The force-time plots and snapshots of conformation of complexes at several critical times in SMD simulations, **a** Pr-ABT538 and **b** Pr-AHA001

insensitive to the pulling directions (i.e. approximately same rupture forces at different pulling direction). Therefore, we chose direction No. 1 for constructing the free energy landscape of Pr-inhibitor complexes using JI and US method.

3.2 Work distributions

The work done by the external force during the dissociation/rupture process along the reaction coordinate was calculated using the SMD simulation results. The work distribution was then obtained by performing a number of SMD simulations at a specific pulling rate. Figures 4 and 5 show the comparison of work distribution between Pr-ABT538 and Pr-AHA001 systems at high and low pulling rates, respectively. The average works and standard deviations were given in Table 1.

As expected, the works for dissociating the tighter complex Pr-ABT538 are larger than those for the looser one Pr-AHA001. We noted that a higher loading rate (e.g. a higher loading stiffness at the same pulling rate) led to a larger rupture force for dissociating the Pr-inhibitor system, as shown in Fig. 2. However, a higher loading stiffness (i.e. K = 6947.7 pN/nm) induced a smaller works than the lower one (i.e. K = 694.8 pN/nm), see Fig. 4 and Table 1. This behavior can be understood as follows. The work done by the pulling force is calculated by $\Delta W = F \times \Delta d$, where F is the pulling force and Δd is the displacement of the inhibitor. Although a higher loading stiffness resulted in a larger rupture force, at the same time, it induced a narrower energy well of the complex (i.e., smaller displacement Δd) [5, 10, 12]. As a result, the works done by pulling force using a higher loading stiffness were lower than those using a softer loading stiffness. Furthermore, the bandwidth of the work distributions (denoted by σ in Table 1) using stiffer spring is smaller than those using softer one. We noted that the average works were much larger than the energy barrier of dissociation ΔG_{off} from experiments. These results indicated that the dissociation processes under above pulling rates were far away from the equilibrium.

To bring the pulling process closer to the equilibrium processes, we also did SMD simulations at much slower pulling rates, i.e. v = 0.02 nm/ns and 0.01 nm/ns, which were 25 and 50 times slower than the previous pulling rates. Because of the smaller pulling rates, the switch times of SMD simulations would be much less, see Table 1. As shown in Fig. 5, lower pulling rates yielded much smaller works for dissociating the complexes, also see Table 1. Indeed, some values of the works done in the SMD simulations are closer to the experiment results of the dissociation barriers (~20 kcal/mol for Pr-ABT538 in Fig. 5a and ~ 10 kcal/mol for Pr-AHA001 in Fig. 5c) (see Table 1). However, we noted that the pulling rates were still much faster than that in experiments (e.g. ~ 100 ns in simulation Vs. ~ 10 ms of spontaneous dissociation [37]) so the average works were still larger than the dissociation barriers obtained in experiments, see Table 1.

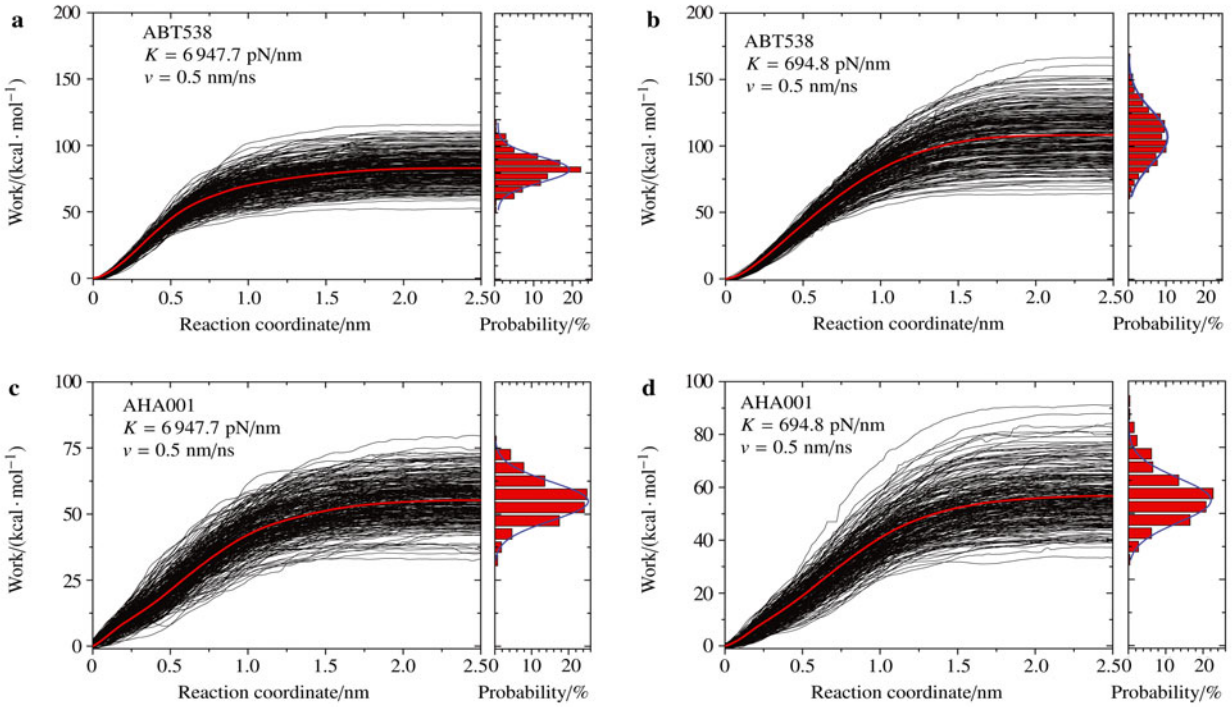

Fig. 4 The works done by the pulling forces and their distributions in SMD simulations at high pulling rates

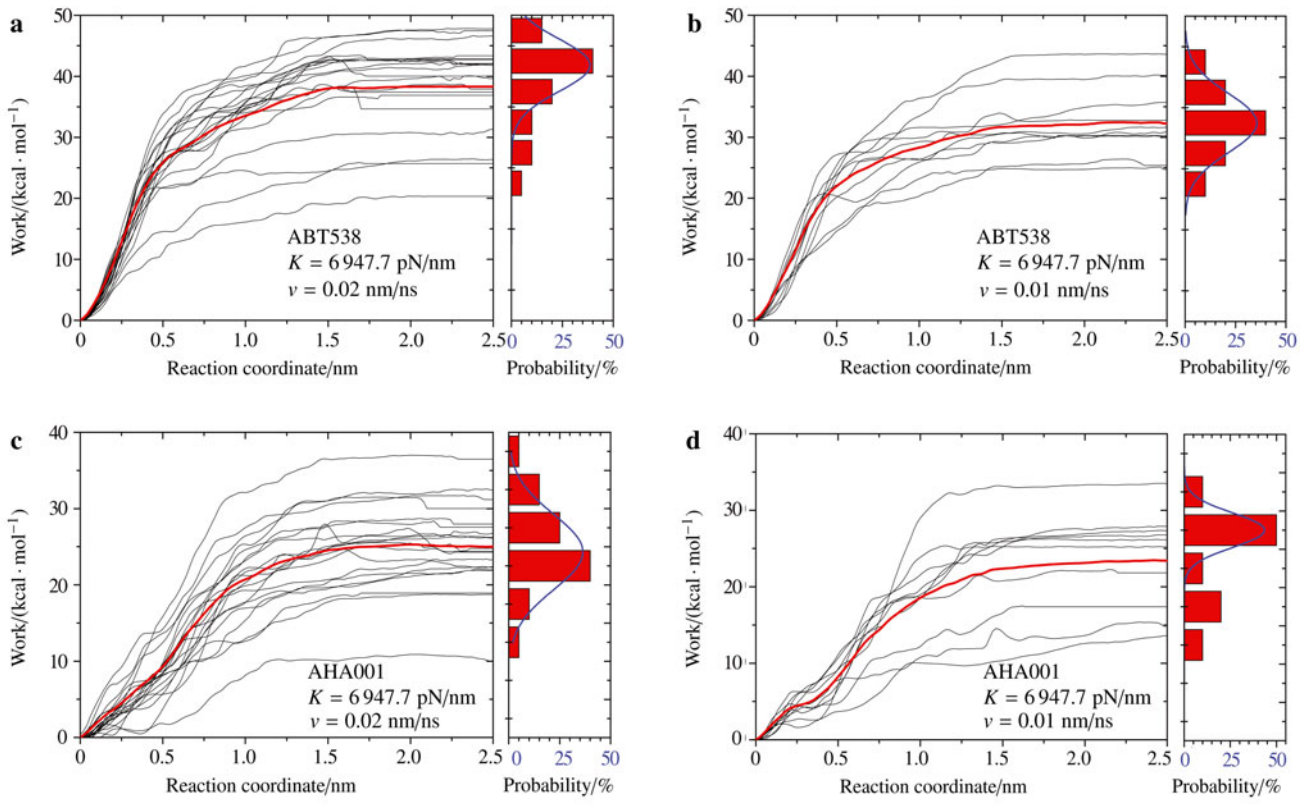

Fig. 5 The works done by the pulling forces and their distributions in SMD simulations at slow pulling rates

We note that the slower the pulling rates, the closer the works value to the experimental results, because the dissociation process under slower pulling rates were closer to the equilibrium processes. In addition, a stiffer spring seemed to be more helpful for improving the predications of simulation, because it induces smaller bandwidth of work distributions (e.g. 6 947.7 pN/nm vs. 694.8 pN/nm).

3.3 Jarzynski's identity results under high loading rates

3.3.1 JI with Block-average analysis

Once we have the work distribution, the free energy landscape was first calculated by using the Block-average analysis. Here the linear and non-linear extrapolation was used to improve the predictions. Figure 6 illustrates the linear and non-linear extrapolation in block-average method for Pr-ABT538 and Pr-AHA001 systems.

Figure 7 showed the block-average results of the energy landscape of Pr-ABT538 and Pr-AHA001 complexes under the pulling velocity v = 0.5 nm/ns with different spring constants K = 6 947.7 pN/ns and K = 694.8 pN/ns. As can

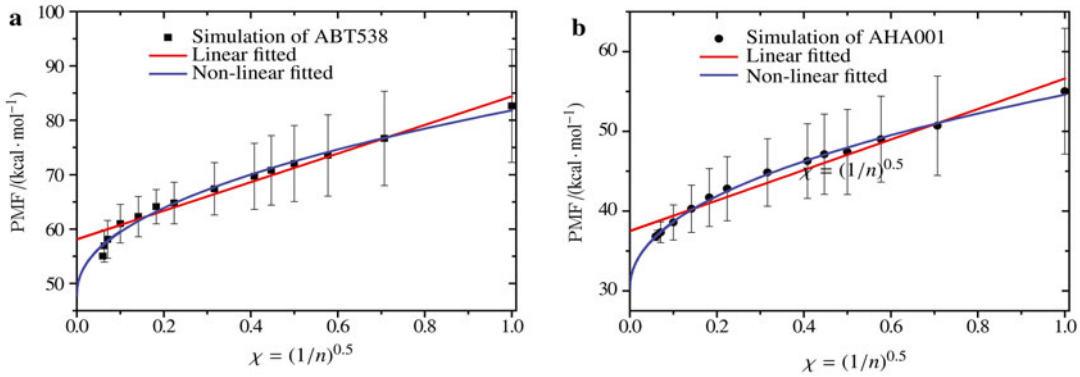

Fig. 6 The illustration of linear and non-linear extrapolation in the block-average analysis. The simulation data were extracted from the SMD simulations at the reaction coordinate 25 nm of **a** ABT538 and **b** AHA001 with K = 6 947.7 pN/nm and v = 0.5 nm/ns. The predictions of free energy decrease with the increase of n (i.e. the decreasing of $\chi = (1/n)^{0.5}$)

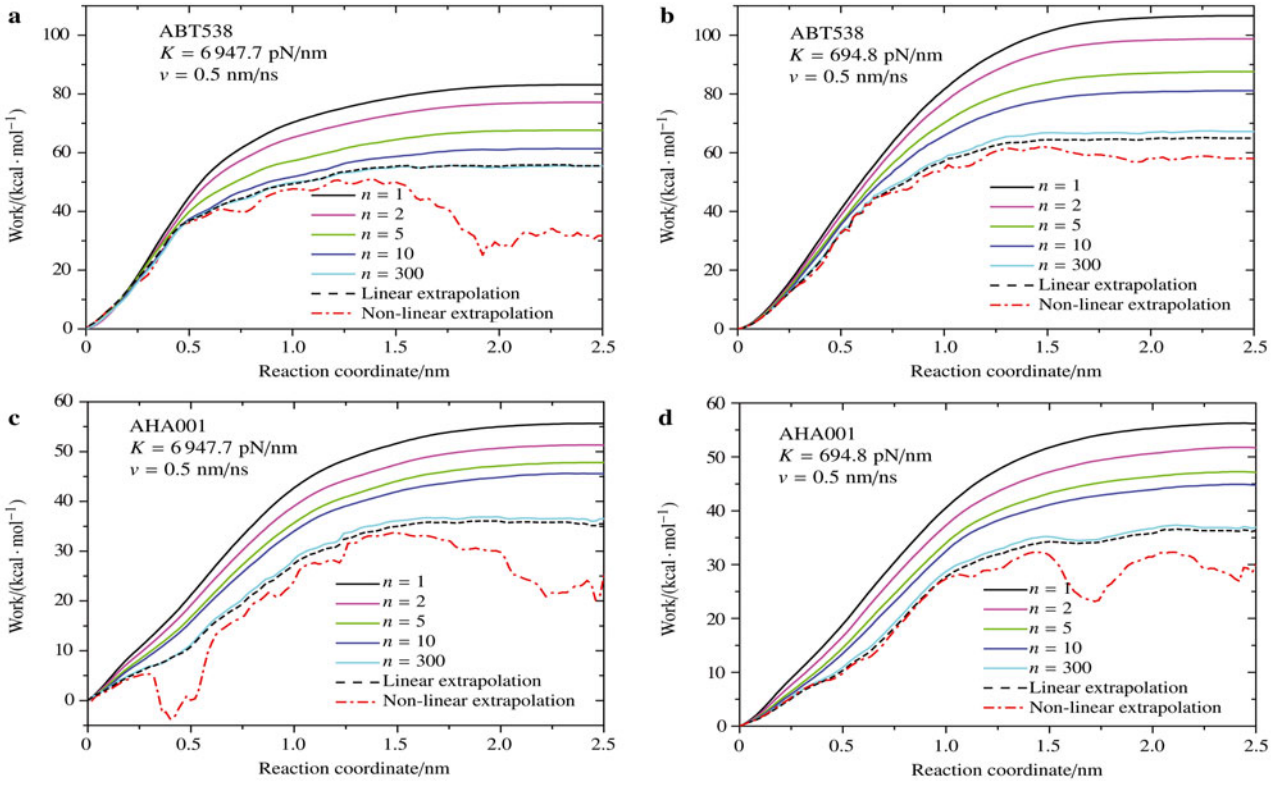

Fig. 7 The PMF calculated by JI method at high pulling rates with linear and non-linear extrapolation using the block-average analysis

be seen, with the increase of n value the prediction of PMF was gradually reduced. But the energy barriers were still overestimated when $n = 300$ (i.e., $n = N$, the total number of SMD simulations). We see that the linear and non-linear extrapolation can help to improve the estimation, and the non-linear extrapolation was better than the linear one. Furthermore, the stiffer spring led to better results, because it led to not only narrower bandwidths of the work distribution, but also smaller work averages, consistent with Park et al.'s result [27, 28]. The width of the energy well x_β can be estimated as 1.1–1.4 nm, which is well consistent with the SMD results. As shown in Fig. 8, when the reaction coordinate increased to 1.1–1.4 nm the inhibitor was just pulled out of the binding pocket. However, the energy barrier predicated by the block-average method, especially at high pulling rates, were still largely overestimated compared to the experimental results, because the tail of the work distribution (i.e. the minimum work) was still much larger than the energy barrier in experiments due to the finite sampling of the analysis.

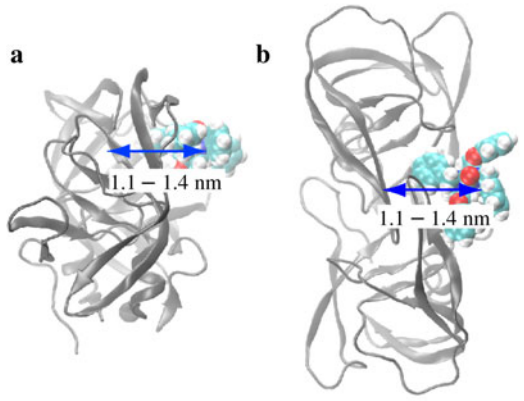

Fig. 8 The illustration of the width of energy well x_β. The protease was represented by gray carton draw, and the inhibitor was represented by spheres

3.3.2 *Cumulant expansion analysis*

To overcome the limitation of block-average method, the cumulant expansion method was adopted. As shown in Figs. 9a and 9b, the energy barrier for Pr-ABT538 estimated by

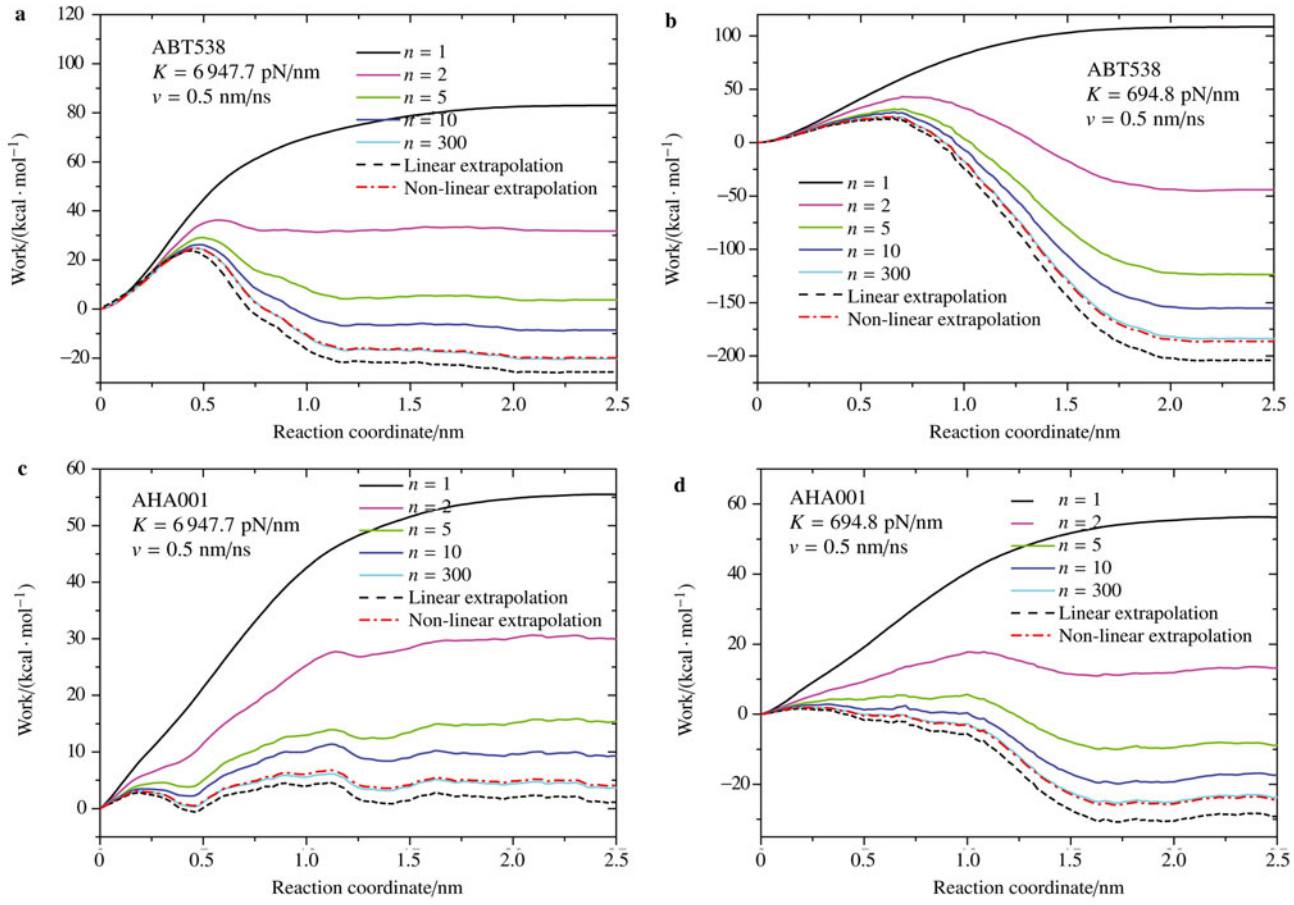

Fig. 9 The PMF calculated by JI method at high pulling rates using the cumulant expansion method

cumulant expansion method was about 24 kcal/mol, which was better than the block-average results. However, the width of energy well x_β was estimated only 0.5–0.7 nm, which was lower than the SMD result. Furthermore, the free energy differences between the bound and unbound state (e.g. the reaction coordinate large to 2 nm) were −19.79 kcal/mol and −186.28 kcal/mol for K = 6947.7 pN/nm and K = 694.8 pN/nm, respectively, which were largely underestimated as compared to the experimental results (12.59 kcal/mol for Pr-ABT538 [37]). Similarly, the free energy differences between the bound and unbound state of Pr-AHA001 complex were also underestimated as compared to the experimental results (10.79 kcal/mol for Pr-AHA001 [63]), see Figs. 9c and 9d. For the results from K = 694.8 pN/nm and v = 0.5 nm/ns in Pr-AHA001 system, the energy barrier estimated nearly vanished for Pr-AHA001, see Fig. 9d.

These phenomena are similar to Cuendet and Michielin's [11] calculation results for the protein-protein complex dissociation. Cuendet and Michielin [11] also observed the large underestimation of free energy, as well as the vanishing of energy barriers for the mutant complex (worse interaction strength). Because the cumulant expansion method assumes the works distributions are Gaussian, the higher cumulants are supposed to be identically zero [39]. However, the Gaussian distribution of works can be satisfied only for near-equilibrium processes [64, 65]. For non-Guassian distribution, dropping the higher cumulants of Eq. (7) will bring in systematic error in the free energy estimation. Generally, the statistical error can be indicated by the standard deviation σ of the distributions due to insufficient sampling. For example, as shown in Fig. 10, some of the work distributions have the standard deviation σ as large as ~ 10 kcal/mol. As a result, the second term $-\beta\sigma^2/2$ in Eq. (7) is as large as ~ 80 kcal/mol so that the free energy difference is largely underestimated. The values of skewness in Fig. 10 indicate that the work distributions were asymmetric [66].

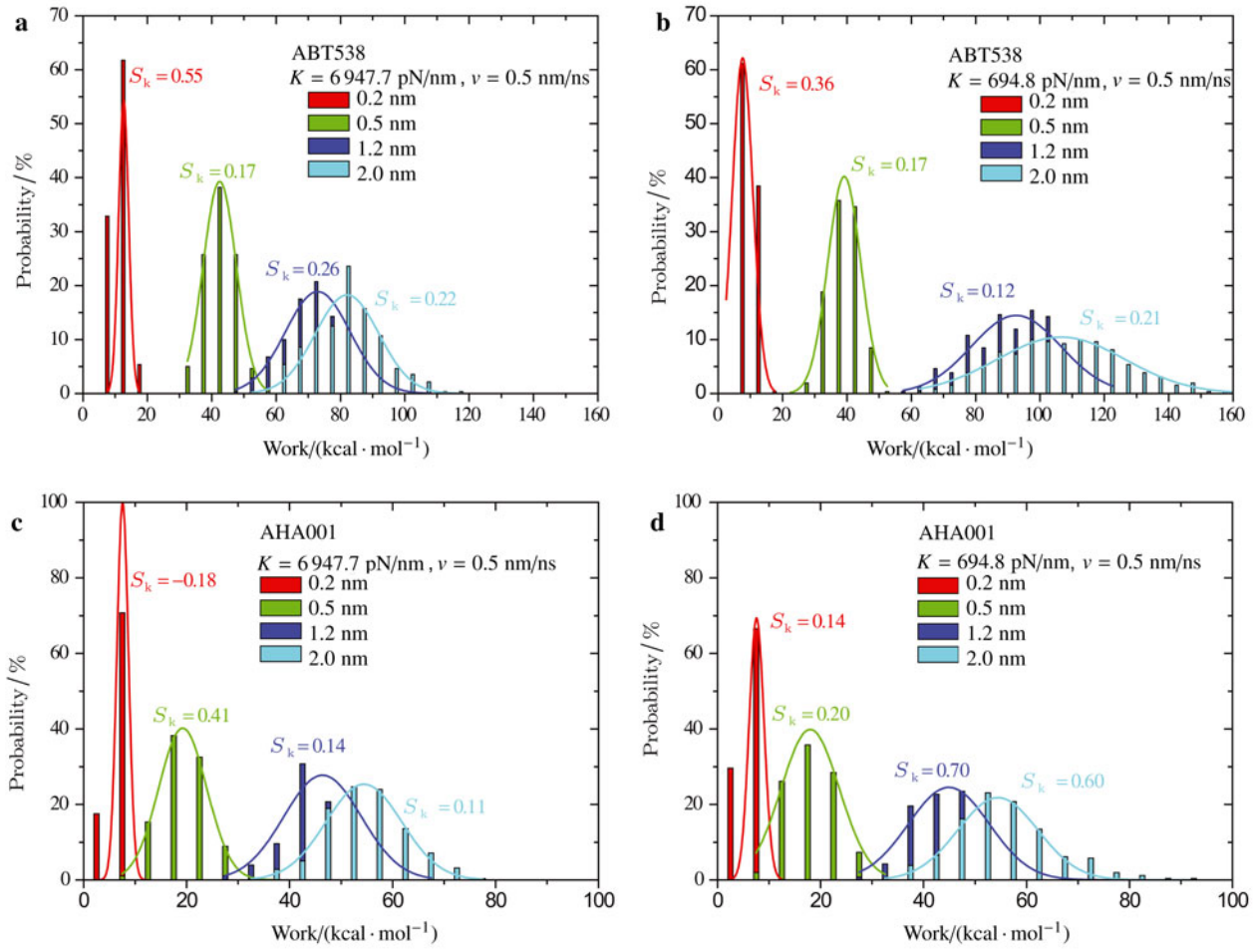

Fig. 10 The work distributions at different reaction coordinates. $S_k = \sum_{j=1}^{N} \left[W_j - \langle W \rangle / \sigma\right]^3 / N$ characterizes the skewness of the distribution. Skewness is a measure of the asymmetry of the probability distribution of a random variable [66]

Figure 11 shows several selected force-time plots of dissociation processes of Pr-ABT538 and Pr-AHA001 complexes, respectively, in which the external forces have done the maximum and minimum works among all the SMD simulations at specific pulling rates (see Figs. 4 and 5). We can see that these dissociation processes were significantly different (e.g. regarding peak force and rupture displacement). The maximum work corresponds to the largest rupture force and largest rupture displacement, indicating that more functional groups or interactions cooperatively resist the dissociation. In contrast, the minimum work corresponds to the smallest rupture force and smallest rupture displacement. Therefore, a wide range of work distribution is intrinsic for a complex system, since there are many different dissociation modes (e.g. different dissociation sequence of groups or interactions). The free energy landscape is determined by the ensemble of all possible dissociation modes and pathways. For dynamic process of a simple system, e.g. separation of two methane molecules in water [26] and the helix-coil transition of deca-alanine in vacuum [27, 28], as well as K^+ ion getting across a carbon nanotube [67], the degrees of freedom of the process are limited so that the standard deviation of the work distributions satisfied $\sigma \sim k_B T$ quite well, therefore free energy converged. In contrast, for the process of complex systems, e.g. the dissociation of protein-protein complex [11] and ligand-receptor systems as in this study, the standard deviation $\sigma \gg k_B T$ at some reaction coordinates so that the free energy would be largely underestimated by the cumulant expansion method.

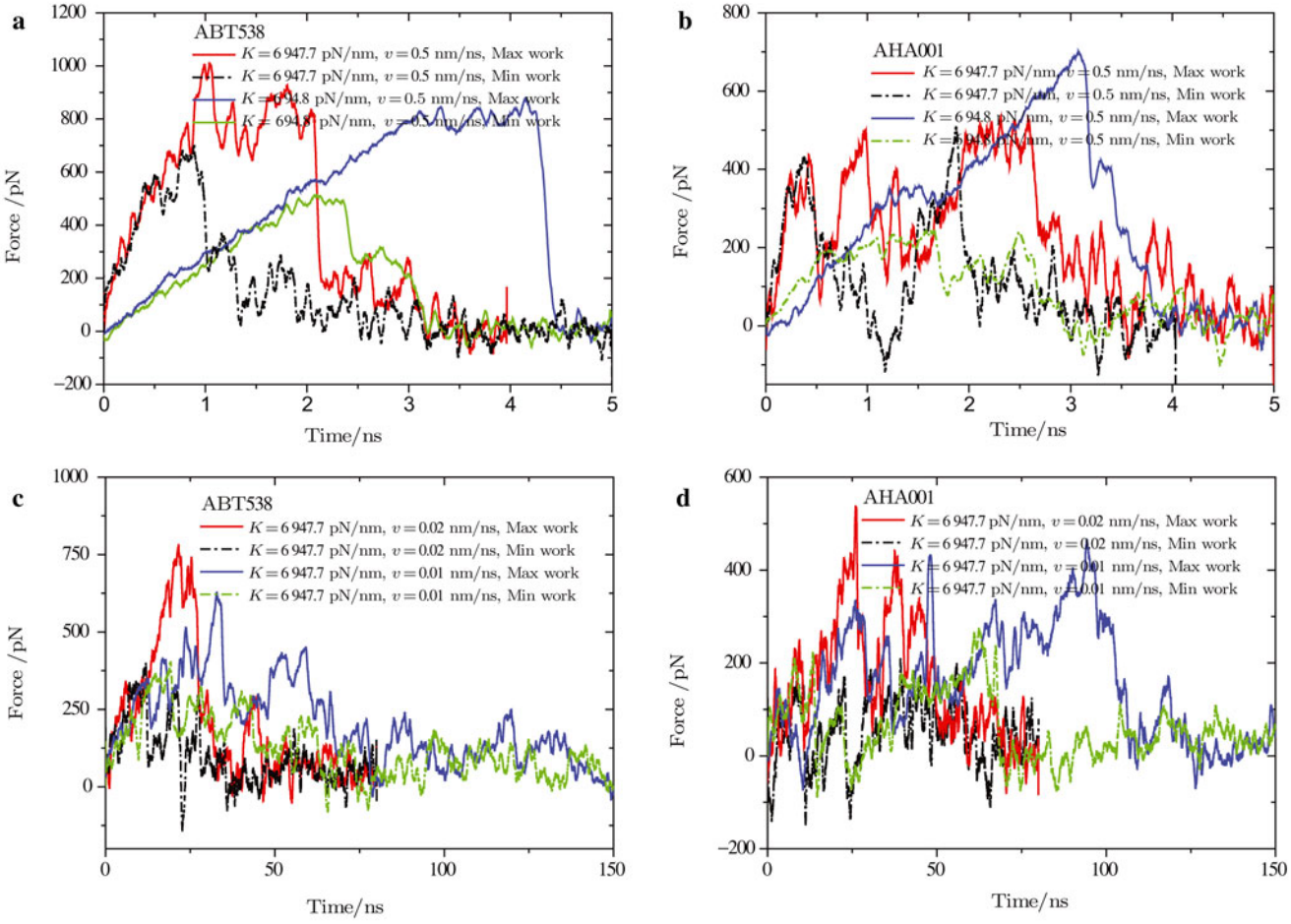

Fig. 11 The selected typical force-time plots of Pr-ABT538 and Pr-AHA001 dissociation processes, in which the external forces have done the maximal and minimal works among all the SMD simulations

3.4 JI's results under very slow pulling rates

Figure 12 shows predictions of the free energy of Pr-ABT538 and Pr-AHA001 complexes under very slow pulling rates, $v = 0.02$ and 0.01 nm/ns, which are 25 and 50 times slower than that of the above analyses and previous studies [11, 27, 28, 68]. As can be seen, the results from slower pulling rates with fewer switch times give much better results, in which the energy barriers, as well as the local minimum are well reconstructed (see Fig. 12b). The predictions of potential of mean forces (PMF) of Pr-ABT538 and Pr-

AHA001 complexes are very close to the results obtained by US method. We note that the energy barriers calculated by US method are in good agreement with the experiments [9]. For example, the energy barrier for dissociation of Pr-ABT538 is $\Delta G_{\text{off}} \sim 20$ kcal/mol, while that of Pr-AHA001 is $\Delta G_{\text{off}} \sim 15$ kcal/mol (the corresponding values are 21.08 kcal/mol and 14.79 kcal/mol in experiments [37], see Table 1).

The reason for JI method to offer better predictions is that the slower pulling rates increase the possibility of sampling rare event with small work values (see Fig. 5) which dominates the results of JI method. This phenomenon would be more significant in complex systems, because each group of the complex at slow pulling rates has longer time for relaxation and searching the natural pathway to dissociate. This study might provide important guidelines for optimizing simulation procedures, in which a slower pulling rate with fewer switch times in an affordable computational time scale should be used for an accurate result. Furthermore, because the pulling rates in experiments are much slower than those in the simulations (e.g. 10^5 fold slower), JI method will be more promising in the analysis of the single-molecule experimental results [69].

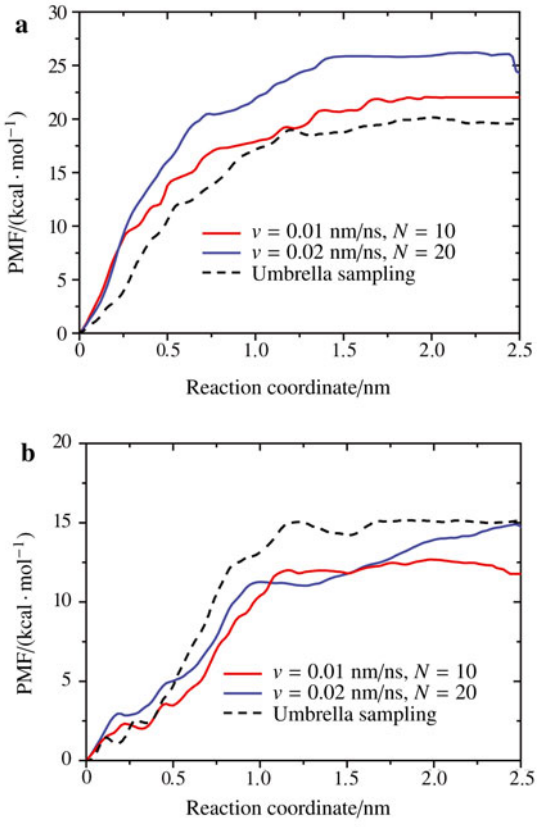

Fig. 12 The PMF calculated from simulations with very slow pulling rates. **a** Pr-ABT538 and **b** Pr-AHA001. *N* denotes the total number of simulations at specific pulling rates

4 Conclusions

The JI method has not yet been well tested for complex system such as ligand-receptor pairs in previous studies. In this paper, we calculated the energy barrier for dissociating a ligand-receptor system (HIV-1 protease and its inhibitors) using the JI method. The huge number of SMD simulations enables us to study the efficiency of JI method in free energy calculations. We showed that at high pulling rates JI method would overestimate the energy barrier of dissociation using the block-average analysis, while it will underestimate the energy differences between the bound and unbound states using the cumulant expansion analysis. However, we found that JI method can achieve more accurate results by SMD simulations with slower pulling rate with fewer switch times, compare to higher pulling rate with more switch times. Furthermore, we found that JI method can give a better result using a stiffer spring in the SMD simulations than a softer one. These results might be helpful for optimizing simulation procedures as to how to choose the pulling rate and the pulling stiffness.

Although the JI method is not as accurate as the US method for the ligand-receptor system (which is more complex compared to those in previous studies [26–28]), we showed that JI method is much simpler in its simulation procedure, in which one needs only to repeat the SMD simulations straightforwardly multiple times. As a matter of fact, the JI method is ideally suited to massively parallel computation because each repeated SMD simulation is independent. More importantly, JI method is easier to perform in the experiments, while the US method is infeasible. With the much slower pulling rates in experiments, JI method will yield results with good accuracy.

References

1 Ji, B., Bao, G.: Cell and molecular biomechanics: perspectives and challenges. Acta Mechanica Solida Sinica **24**, 27–51 (2011)
2 Bustamante, C., Chemla, Y. R., Forde, N. R., et al.: Mechanical processes in biochemistry. Annual Review of Biochemistry **73**, 705–748 (2004)
3 Bao, G.: Mechanics of biomolecules. Journal of the Mechanics and Physics of Solids, **50**, 2237–2274 (2002)
4 Kumar, S., Li, M. S.: Biomolecules under mechanical force. Physics Reports-Review Section of Physics Letters **486**, 1–74 (2010)
5 Evans, E.: Probing the relation between force - lifetime - and chemistry in single molecular bonds. Annual Review of Biophysics and Biomolecular Structure **30**, 105–128 (2001)
6 Hugel, T., Seitz, M.: The study of molecular interactions by AFM force spectroscopy. Macromolecular Rapid Communications **22**, 989–1016 (2001)
7 Moffitt, J. R., Chemla, Y. R., Smith, S. B., et al.: Recent advances in optical tweezers, in Annual Review of Biochemistry **77**, 205–228 (2008)

8 Colizzi, F., Perozzo, R., Scapozza, L., et al.: Single-molecule pulling simulations can discern active from inactive enzyme inhibitors. Journal of the American Chemical Society **132**, 7361–7371 (2010)

9 Li, D., Ji, B., Hwang, K. C., et al.: Strength of hydrogen bond network takes crucial roles in the dissociation process of inhibitors from the HIV-1 protease binding pocket. PLoS ONE, **6**, e19268 (2011)

10 Ackbarow, T., Chen, X., Keten, S., et al.: Hierarchies, multiple energy barriers, and robustness govern the fracture mechanics of alpha-helical and beta-sheet protein domains. Proceedings of the National Academy of Sciences of the United States of America **104**, 16410–16415 (2007)

11 Cuendet, M. A., Michielin, O.: Protein-protein interaction investigated by steered molecular dynamics: The TCR-pMHC complex. Biophysical Journal **95**, 3575–3590 (2008)

12 Bell, G. I.: Models for specific adhesion of cells to cells. Science **200**, 618–627 (1978)

13 Hummer, G., Szabo, A.: Kinetics from nonequilibrium single-molecule pulling experiments. Biophysical Journal **85**, 5–15 (2003)

14 Dudko, O. K., Filippov, A. E., Klafter, J., et al.: Beyond the conventional description of dynamic force spectroscopy of adhesion bonds. Proceedings of the National Academy of Sciences of the United States of America **100**, 11378–11381 (2003)

15 Dudko, O. K., Hummer, G., Szabo, A.: Intrinsic rates and activation free energies from single-molecule pulling experiments. Physical Review Letters **96**, 108101 (2006)

16 Maitra, A., Arya, G.: Model Accounting for the Effects of Pulling-Device Stiffness in the Analyses of Single-Molecule Force Measurements. Physical Review Letters **104**, (2010)

17 Patey, G. N., Valleau, J. P.: Free-Energy of spheres with dipoles - monte-carlo with multistage sampling. Chemical Physics Letters **21**, 297–300 (1973)

18 Torrie, G. M., Valleau, J. P.: Monte-carlo free-energy estimates using non-boltzmann sampling - application to subcritical lennard-jones fluid. Chemical Physics Letters **28**, 578–581 (1974)

19 Torrie, G. M., Valleau, J. P.: Non-physical sampling distributions in monte-carlo free-energy estimation - umbrella sampling. Journal of Computational Physics **23**, 187–199 (1977)

20 Lemkul, J. A., Bevan, D. R.: Assessing the stability of alzheimer's amyloid protofibrils using molecular dynamics. Journal of Physical Chemistry B **114**, 1652–1660 (2010)

21 Arora, K., Brooks, III, C. L.: Large-scale allosteric conformational transitions of adenylate kinase appear to involve a population-shift mechanism. Proceedings of the National Academy of Sciences of the United States of America **104**, 18496–18501 (2007)

22 Rick, S. W., Erickson, J. W., Burt, S. K.: Reaction path and free energy calculations of the transition between alternate conformations of HIV-1 protease. Proteins-Structure Function and Bioinformatics **32**, 7–16 (1998)

23 Buch, I., Kashif Sadiq, S., Fabritiis, G. De: Optimized potential of mean force calculations for standard binding free energies. Journal of Chemical Theory and Computation **7**, 1765–1772 (2011)

24 Jarzynski, C.: Nonequilibrium equality for free energy differences. Physical Review Letters **78**, 2690–2693 (1997)

25 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 **98**, 3658–3661 (2001)

26 Hummer, G.: Fast-growth thermodynamic integration: Error and efficiency analysis. Journal of Chemical Physics **114**, 7330–7337 (2001)

27 Park, S., Khalili-Araghi, F., Tajkhorshid, E., et al.: Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. Journal of Chemical Physics **119**, 3559–3566 (2003)

28 Park, S., Schulten, K.: Calculating potentials of mean force from steered molecular dynamics simulations. Journal of Chemical Physics **120**, 5946–5961 (2004)

29 Katz, R. A., Skalka, A. M.: The Retroviral Enzymes. Annual Review of Biochemistry **63**, 133–173 (1994)

30 Kohl, N. E., Emini, E. A., Schleif, W. A., et al.: Active human immunodeficiency virus protease is required for viral infectivity. Proceedings of the National Academy of Sciences of the United States of America **85**, 4686–4690 (1988)

31 Hornak, V., Okur, A., Rizzo, R. C., et al.: HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. Proceedings of the National Academy of Sciences of the United States of America **103**, 915–920 (2006)

32 Li, D. C., Ji, B. H., Hwang, K., et al.: Crucial roles of the subnanosecond local dynamics of the flap tips in the global conformational changes of HIV-1 protease. Journal of Physical Chemistry B **114**, 3060–3069 (2010)

33 Tozzini, V., McCammon, J. A.: A coarse grained model for the dynamics of flap opening in HIV-1 protease. Chemical Physics Letters **413**, 123–128 (2005)

34 Li, D. C., Ji, B. H., Hwang, K. C., et al.: Coarse grained modeling of biopolymers and proteins: methods and applications. International Journal of Applied Mechanics **1**, 113–136 (2009)

35 Li, D. C., Liu, M. S., Ji, B. H., et al.: Coarse-grained molecular dynamics of ligands binding into protein: The case of HIV-1 protease inhibitors. Journal of Chemical Physics **130**, 215102 (2009)

36 Cheng, Y., Li, D. C., Ji, B. H., et al.: Structure-based design of carbon nanotubes as HIV-1 protease inhibitors: Atomistic and coarse-grained simulations. Journal of Molecular Graphics & Modelling **29**, 171–177 (2010)

37 Markgren, P. O., Schaal, W., Hamalainen, M., et al.: Relationships between structure and interaction kinetics for HIV-1 protease inhibitors. Journal of Medicinal Chemistry **45**, 5430–5439 (2002)

38 Ytreberg, F. M., Zuckerman, D. M.: Efficient use of nonequilibrium measurement to estimate free energy differences for molecular systems. Journal of Computational Chemistry **25**, 1749–1759 (2004)

39 Marcinkiewicz, J.: Sur une propriété de la loi de Gauβ Mathematische Zeitschrift **44**, 612–618 (1939)

40 Backbro, K., Lowgren, S., Osterlund, K., et al.: Unexpected binding mode of a cyclic sulfamide HIV-1 protease inhibitor. Journal of Medicinal Chemistry **40**, 898–902 (1997)

41 Kempf, D. J., Marsh, K. C., Denissen, J. F., et al.: Abt-538 is a potent inhibitor of human-immunodeficiency-virus protease and has high oral bioavailability in humans. Proceedings of the National Academy of Sciences of the United States of America **92**, 2484–2488 (1995)

42 Yamazaki, T., Nicholson, L. K., Torchia, D. A., et al.: Nmr and x-ray evidence that the hiv protease catalytic aspartyl groups are protonated in the complex formed by the protease and a nonpeptide cyclic urea-based inhibitor. Journal of the American Chemical Society **116**, 10791–10792 (1994)
43 Wittayanarakul, K., Hannongbua, S., Feig, M.: Accurate prediction of protonation state as a prerequisite for reliable MM-PB(GB)SA binding free energy calculations of HIV-1 protease inhibitors. Journal of Computational Chemistry **29**, 673–685 (2008)
44 Berendsen, H. J. C., Vanderspoel, D., Vandrunen, R.: Gromacs - a Message-Passing Parallel Molecular-Dynamics Implementation. Computer Physics Communications **91**, 43–56 (1995)
45 Lindahl, E., Hess, B., van der Spoel: GROMACS 3.0: a package for molecular simulation and trajectory analysis. Journal of Molecular Modeling **7**, 306–317 (2001)
46 Sorin, E.J., Pande, V. S.: Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. Biophysical Journal **88**, 2472–2493 (2005)
47 Junmei, W., Wei, W., Kollman, P. A., et al.: Automatic atom type and bond type perception in molecular mechanical calculations. Journal of Molecular Graphics & Modelling **25**, 247–260 (2006)
48 Wang, J. M., Wolf, R. M., Caldwell, J. W., et al.: Development and testing of a general amber force field. Journal of Computational Chemistry **25**, 1157–1174 (2004)
49 Jakalian, A., Bush, B. L., Jack, D. B., et al.: Fast, efficient generation of high-quality atomic Charges. AM1-BCC model: I. Method. Journal of Computational Chemistry **21**, 132–146 (2000)
50 Case, D. A., Cheatham, T. E., Darden, T., et al.: The Amber biomolecular simulation programs. Journal of Computational Chemistry **26**, 1668–1688 (2005)
51 Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., et al.: Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics **79**, 926–935 (1983)
52 Essmann, U., Perera, L., Berkowitz, M. L., et al.: A smooth particle mesh ewald method. Journal of Chemical Physics **103**, 8577–8593 (1995)
53 Berendsen, H. J. C., Postma, J. P. M., Vangunsteren, W. F., et al.: Molecular-dynamics with coupling to an external bath. Journal of Chemical Physics **81**, 3684–3690 (1984)
54 Hess, B.: P-LINCS: A parallel linear constraint solver for molecular simulation. Journal of Chemical Theory and Computation **4**, 116–122 (2008)
55 Sadiq, A. K., Wan, S., Coveney, P. V.: Insights into a mutation-assisted lateral drug escape mechanism from the HIV-1 protease active site. Biochemistry **46**, 14865–14877 (2007)
56 Trylska, J., Tozzini, V., Chang, C. A., et al.: HIV-1 protease substrate binding and product release pathways explored with coarse-grained molecular dynamics. Biophysical Journal **92**, 4179–4187 (2007)
57 Nose, S.: A molecular-dynamics method for simulations in the canonical ensemble. Molecular Physics **52**, 255–268 (1984)
58 Hoover, W. G.: Canonical dynamics - equilibrium phase-space distributions. Physical Review A **31**, 1695–1697 (1985)
59 Parrinello, M., Rahman, A.: Polymorphic transitions in single-crystals - a new molecular-dynamics method. Journal of Applied Physics **52**, 7182–7190 (1981)
60 Kumar, S., Bouzida, D., Swendsen, R. H., et al.: The weighted histogram analysis method for free-energy calculations on biomolecules .1. the method. Journal of Computational Chemistry **13**, 1011–1021 (1992)
61 Qin, Z., Kreplak, L., Buehler, M. J.: Nanomechanical properties of vimentin intermediate filament dimers. Nanotechnology **20**, 425101 (9 page) (2009)
62 Evans, E., Ritchie, K.: Dynamic strength of molecular adhesion bonds. Biophysical Journal **72**, 1541–1555 (1997)
63 Hulten, J., Bonham, N. M., Nillroth, U., et al.: Cyclic HIV-1 protease inhibitors derived from mannitol: Synthesis, inhibitory potencies, and computational predictions of binding affinities. Journal of Medicinal Chemistry **40**, 885–897 (1997)
64 Hermans, J.: Simple analysis of noise and hysteresis in (slow-growth) free-energy simulations. Journal of Physical Chemistry **95**, 9029–9032 (1991)
65 Wood, R. H., Muhlbauer, W. C. F., Thompson, P. T.: Systematic-errors in free-energy perturbation calculations due to a finite-sample of configuration space - sample-size hysteresis. Journal of Physical Chemistry **95**, 6670–6675 (1991)
66 Press, W. H., Flannery, B. P., Teukolsky, S. A., et al.: Numerical Recipes in C: The Art of Scientific Computing 1992: Cambridge University Press.
67 Bastug, T., Kuyucak, S.: Application of Jarzynski's equality in simple versus complex systems. Chemical Physics Letters **436**, 383–387 (2007)
68 Zhang, D. Q., Gullingsrud, J., McCammon, J. A.: Potentials of mean force for acetylcholine unbinding from the alpha7 nicotinic acetylcholine receptor ligand-binding domain. Journal of the American Chemical Society **128**, 3019–3026 (2006)
69 Liphardt, J., Dumont, S., Smith, S. B., et al.: Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. Science **296**, 1832–1835 (2002)