ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2021, *66*(10-11), 835–844
DOI: 10.33224/rrch.2021.66.10-11.07

Dedicated to the memory of
Prof. Petre T. FRANGOPOL (1933-2020)

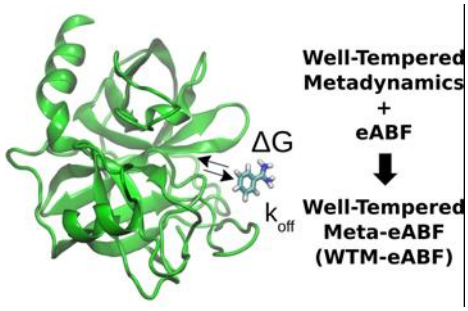
STUDYING THE ASSOCIATION AND DISSOCIATION OF A TRYPSIN-BENZAMIDINE COMPLEX WITH ENHANCED SAMPLING MOLECULAR DYNAMICS

Sharon E. STONE,[a] Dhiman RAY[a] and Ioan ANDRICIOAEI[a,b,*]

[a]Department of Chemistry, University of California Irvine, Irvine CA 92697, SUA
[b]Department of Physics and Astronomy, University of California Irvine, Irvine CA 92697, SUA

Received September 18, 2021

We showcase the application of an enhanced sampling scheme, the well-tempered meta-eABF (WTM-eABF) method, to compute the free energy landscape and explore the pathways of dissociation and association of a trypsin-benzamidine complex. We compared our results with previous computational studies and experimental data. While free energy estimates are within 2-3 kcal/mol from the experimental value, the accuracy of our prediction of the unbinding rate constant (k_{off}) is comparable to that obtained from two orders of magnitude longer and more extensive molecular dynamics and Markov state modeling data. Moreover, we identified two intermediate states, and two different binding pathways, depending on the orientation of the approaching ligand. Electrostatic interactions have been found to have the dominant contribution in stabilizing the bound conformation. We conclude that the WTM-eABF approach, in conjunction with simple distance-based reaction coordinates, can be used to study protein-ligand interactions with reasonable accuracy and potentially find application in computer-aided drug design.



INTRODUCTION

Protein-ligand interaction plays an important role in the action of therapeutic drugs, and consequently is one of the key areas of interest in the field of computer-aided drug design (CADD). With the advent of fast computing facilities involving supercomputer clusters and graphical processing unit (GPU) based hardware and the rapid development of new efficient algorithms, it is now possible to computationally screen a large library of molecules for their pharmacological properties. One of the major computational techniques used in this process is the molecular dynamics (MD) simulations, which use Newton's equations of motion to propagate trajectories of molecules and provide a unique opportunity to study the process of protein-ligand recognition and to understand the mechanism in atomistic detail.[1] Most of the physiologically relevant biomolecular recognition processes occur at a timescale beyond milliseconds, and involve 10^4-10^6 atoms including the solvent molecules to account for the aqueous environment.[2] But the time steps for the numerical integration of the equations of motion should be of the order of fs to correctly simulate the fastest vibrational motions in the system. This would necessitate multiple billions of integration time steps for a millisecond scale process, which is beyond the scope of currently available

* Corresponding author: andricio@uci.edu

computational resources, which can accommodate up to a few microseconds.

This limitation led to the development of the enhanced sampling methods which use external biasing forces accelerating transitions across large energy barriers and increasing the possibility of observing slow processes which evade the brute-force molecular dynamics approach. The oldest of such methods is the umbrella sampling approach,[3,4] which is now widely used to compute the free energy landscape of a biomolecular process.[5] A more recent alternative is the use of metadynamics, in which Gaussian hills are deposited along a chosen degree of freedom, called reaction coordinate (RC), thereby filling up low energy minima and accelerating barrier-crossing transitions.[6] An improved version of metadynamics is the well-tempered metadynamics, where the hill size is adapted during the simulation to ensure a better convergence at the long time limit.[7] Both of these methods have found widespread application in studying protein-ligand binding problems.[8–14]

The adaptive biasing force (ABF) is another class of methods to compute the free energy landscape of complex biomolecular processes.[15,16] In ABF, a force equal and opposite to the potential energy gradient is applied along the reaction coordinate, pulling the system across high free energy barriers. Although this method is extremely useful there are strict limitations on the type of reaction coordinates that can be used in ABF; not only the RC has to be mutually orthogonal to the other degrees of freedom in the system, the gradient of the potential energy along the RC needs to be perpendicular to the same along other coordinates.[17] This makes it very difficult to apply ABF directly to the protein-ligand recognition problem, where multiple protein degrees of freedom are coupled to the ligand unbinding coordinate. The extended system adaptive biasing force (eABF) avoids these stringent criteria by applying the pulling force to a fictitious particle coupled to the reaction coordinate with a strong harmonic spring.[17] The eABF approach, albeit being more useful than the regular ABF, is slow to converge and often requires stratification of the reaction coordinate into multiple windows, which can lead to inadequate exploration of the free energy landscape.[18]

Very recently, Fu *et al.* proposed to combine metadynamics with eABF to accelerate the exploration of the free energy landscape and increase the rate at which the converged PMF is obtained during the simulation.[18] In this approach, the Gaussian hills are deposited along the reaction coordinate, and a pulling force is applied against the potential energy gradient simultaneously, to sample configurations separated by a high energy barrier and quickly converge the free energy landscape for complex systems. We used this approach previously to study the formation of Hoogsteen base pairing in DNA and RNA. We could accurately reproduce the free energy landscape and the free energy difference between the WC and HG base-pairing forms in DNA and RNA.[19] Our simulations converged within 200 ns which is orders of magnitude less than the computational time required for studying these systems using metadynamics (40 μs[20]) and umbrella sampling approach (6 μs[21]). Recently Fu *et al*. have introduced the well-tempered meta-eABF method where the metadynamics part of the meta-eABF is replaced by well-tempered metadynamics, to ensure a smooth convergence of the free energy landscape.[22,23] They successfully applied this method to compute the binding free energy of the p41:Abl-SH3 protein-ligand system.[22]

In the current work, we aim to study the binding and unbinding mechanism of the benzamidine ligand to the trypsin protein. The trypsin-benzamidine complex is a ligand receptor system, which is very well studied both experimentally and computationally.[8,9,11,24–30] The dissociation constant K_d of this complex is $1.2 \pm 0.1 \times 10^{-5}$ M, which makes benzamidine a very potent ligand.[24] The experimental residence time of the ligand is 1.7 ms, making it impossible to study the dissociation events from continuous unbiased molecular dynamics simulation. Also, the presence of a well-defined binding pocket makes it a suitable system for testing new sampling methods for ligand binding (Figure 1A). Large scale computational studies using conventional MD and Markov State Modeling were performed to study the unbinding process of trypsin-benzamidine complex by Buch *et al.*[25] and by Plattner *et al*.[26] But such studies are extremely computationally expensive, requiring ~50 and ~150 microseconds of simulation, respectively. The binding free energy predicted by Buch *et al.* was -5.2 ± 0.4 kcal/mol which is 1.5 kcal/mol away from the experimental number, -6.7 ± 0.05 kcal/mol.[24] The prediction by Plattner *et al.* (-6.05 ± 1 kcal/mol) was more accurate and within one kcal/mol of the experimental value. However, the unbinding rate constants were overestimated by almost two orders of magnitude in both of these

methods. Funnel metadynamics study by Limongeli *et al.*[11] could predict a free energy of -8.5 ± 0.7 kcal/mol, which is 1.8 kcal/mol different from the experimental data. But, using a novel machine learning based adaptive reaction coordinate, Brotzakis *et al*. could make more accurate predictions of the free energy (-6.7 and -7.5 kcal/mol in two independent trials)[9] with the funnel metadynamics approach. Votapka *et al.* (2017) and Jagger *et al*.[30] used the milestoning[31–33] and the Markovian milestoning with Voronoi tessellation (MMVT)[34] approaches, respectively, to compute the kinetics and free energy of trypsin-benzamidine binding. Their kinetics results were within one order of magnitude of the experimental value and the binding free energy predictions were -7.4 ± 0.1 kcal/mol and -8.8 ± 0.07 kcal/mol respectively. Still, the computational cost was beyond microseconds (19 μs and 2.3-4.4 μs, respectively).

We developed the novel weighted ensemble milestoning methodology (WEM)[35] which combines two path sampling methods, weighted ensemble[36–38] and milestoning,[31–33] to efficiently compute the kinetics and free energy change in biomolecular rare events, particularly in ligand-receptor binding.[39] We recently improved our WEM method by replacing the milestoning with Markovian milestoning with voronoi tessellation (MMVT)[34,40] to develop the Markovian weighted ensemble milestoning (M-WEM) and were able to reproduce the binding free energy and dissociation and association kinetics of the trypsin-benzamidine system from less than 1 μs cumulative of trajectory data.[41] Although the milestoning based approaches are quite successful in calculating experimental observables, it is difficult to elucidate the ligand binding or release pathways and identify key intermediates due to the lack of continuous trajectories starting from bound state to unbound state and vice-versa. For that, we need to use traditional enhanced sampling methods or weighted ensemble based methods with non-equilibrium steady state, both of which can be computationally expensive and predicted results can have limited accuracy. In this work, we employ the newly developed WTM-eABF approach to study the thermodynamics, kinetics and the molecular mechanism of the binding and unbinding process of the trypsin-benzamidine complex. We particularly evaluate the quality of our predictions from limited trajectory data and compare our results with that from extensive MD simulation studies and experiments.

COMPUTATIONAL METHODS

For the trypsin-benzamidine complex system, the structure, parameter, and topology files were obtained from Votapka *et al.*, and was set up identically to their previous work.[29] The atomic coordinates of the protein ligand system were obtained from the Protein Data Bank (PBD), with a PDB ID: 3PTB.[42] The protonated residues, Asp, Glu, and His, were determined to have a pH of 7.7, which was used in this study in order to replicate the experimental work.[24] For modeling the protein structure and ligand, the AMBER ff14SB force field[43] and Generalized Amber Force Field (GAFF),[44] respectively, were used. The following structure was solvated in an octahedron box of TIP4Pew,[45] along with eight Cl^- ions to neutralize the system. Overall, the system contains about 23,000 atoms. All molecular dynamics (MD) simulations were performed using the NAMD 2.14b2 package[46] and a time step of 2 fs. The Langevin integrator with a 5 ps^{-1} damping coefficient was used to ensure a constant temperature at 298 K. A Langevin piston was used to keep the pressure constant at 1 atm. The bound state structure was equilibrated for 10 ns in the NPT ensemble before proceeding with the enhanced sampling simulation.

The one dimensional free energy surface of ligand unbinding for the trypsin-benzamidine complex was calculated using well tempered meta-eABF.[18,22,23] In the meta-aABF technique, metadynamics and extended system adaptive biasing force (eABF) are used simultaneously to explore the conformational space in a short amount of computational time. For the well-tempered meta-eABF method metadynamics, metadynamics is replaced with well tempered metadynamics, which results in a smoother convergence of the free energy surface. In the current work, we analyzed the WTM-eABF trajectories generated in our previous work,[41] where further details on the simulation protocol can be found. Here, we provide a brief outline. The reaction coordinate is defined by the Euclidean distance between the center of mass of the heavy atoms of the benzamidine ligand and the center of mass of the C-α atoms of the following residues in the binding pocket: 190, 191,192, 195, 213, 215, 216, 219, 220, 224, and 228 (Figure 1B). (The residues are numbered according to PDB: 3PTB.) We refer to this RC as the “protein-ligand distance” in the rest of the paper. The equilibrated bound state was used

at the start of the simulation. Gaussian hills with width 0.6 Å were deposited for every 2 ps along the distance between the trypsin binding pocket and the benzamidine ligand. The height of Gaussian hills, initially set at 0.2 kcal/mol, was slowly reduced during the simulation with a bias factor of 6.0. The adaptive biasing force was applied on a fictional particle coupled to the original reaction coordinate with a stiff spring of force constant of 1500 kcal mol^{-1} $Å^{-2}$ along with an oscillation period of 200 fs. The WTM-eABF simulation was performed using the colvars module[47] patched with NAMD 2.14b2,[46] and was propagated for about 550 ns. The convergence of the free energy landscape was monitored by calculating the root-mean-square deviation (RMSD) of the curve with respect to the last curve. The RMSD was under 1 kcal/mol for the last half of the WTM-eABF simulation, and below 0.25 kcal/mol during the last 50 ns.[41]

The free energy landscape was recorded at 8 ns intervals for the last 32 ns, leading to 5 samples. The average free energy profile is reported herein and the error bars are computed as the 95% confidence interval of the 5 samples. The electrostatic and van der Waals (vdW) interaction energies between trypsin and benzamidine ligand were calculated using the NAMD.[48] The number and occupancy of the hydrogen bonds between trypsin and benzamidine were analyzed using the VMD program.[49]

RESULTS AND DISCUSSIONS

In the free energy landscape of the protein-ligand interaction, calculated from the WTM-eABF approach, there is a deep free energy minimum at 0.5Å value of the protein-ligand distance or the collective variable (CV) (Figure 2). This minimum corresponds to the bound state of the benzamidine ligand in the binding pocket of the trypsin host. At a CV value of 5.6 Å, we observe a free energy barrier of 11.8 ± 0.3 kcal/mol, followed by a shallow minimum corresponding to the unbound state at ~ 10 Å. The free energy difference between the bound and the unbound state is 9.5 ± 0.2 kcal/mol. This value is ~ 2.8 kcal/mol different from the experimentally obtained binding free energy of 6.7± 0.05 kcal/mol.[24] This discrepancy can arise due to multiple factors, ranging from deficiencies in the force field to inadequate sampling. But the most prominent effect is that our approach does not take into account the spatial entropic contribution to the free energy, which can be as large as 3.8 kcal/mol.[9] Nevertheless, the topology of the free energy profile is in agreement with previous studies by Votapka *et al.*[29] and Brotzakis *et al*.[9] Our free energy results is also within 1 kcal/mol agreement with the MMVT SEEKR minimum simulation estimate binding free energy of -8.8 ± 0.07 kcal/mol[30] where the same topology and parameter files were used as of the current study. The MMVT SEEKR simulation took more than 5 times longer simulation time (~2.9 μs) compared to our WTM-eABF study (~0.55 μs). Nevertheless, using other milestoning based approaches, Amaro and co-workers have computed more accurate binding free energies of this complex but at the expense of larger computational cost, such as -7.4 ± 0.1 kcal/mol using SEEKR approach with ~ 19 μs of simulation.

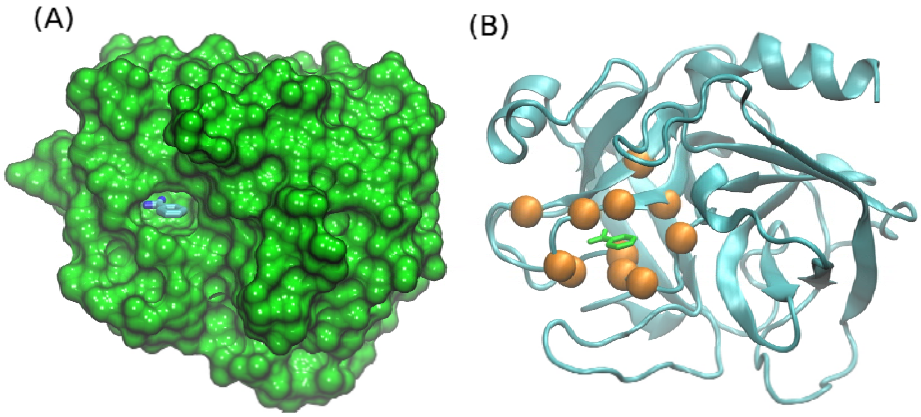


Fig. 1 – The structure of the trypsin-benzamidine complex. (A) The protein is shown in surface representation to highlight the binding pocket. (B) The protein is shown in ribbon and the benzamidine ligand is shown in green licorice. The alpha carbons of the residues in the binding pocket used to define the collective variable are shown in orange sphere.

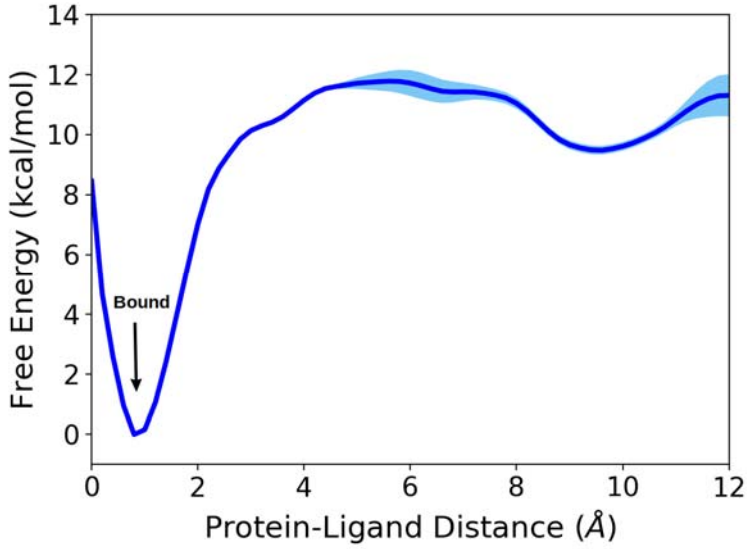


Fig. 2 – The 1D-PMF of the dissociation of the trypsin-benzamidine complex obtained from the well-tempered meta-eABF simulation (WTM-eABF).

From the barrier height of 11.8 ± 0.3 kcal/mol, a simple Arrhenius approximation gives the unbinding rate constant (k_{off}) value of (1.8 ± 0.66) × 10^4 s^{-1} which is different from the experimental unbinding rate constant (600 ± 300 s^{-1}) by only a factor of 30. As a result, our predicted residence time of benzamidine ligand in trypsin is 55.6 μs, as opposed to the experimental value of 1700 μs. The accuracy of our estimated dissociation kinetics is comparable to the results obtained by Buch et al. (k_{off} = (9.5 ± 3.3) × 10^4 s^{-1})[25], Plattner *et al*. (k_{off} = (1.31 ±1.09) ×10^4 s^{-1})[26], Tiwary *et al*. (k_{off} = 9.1 ±2.5 s^{-1})[8] and Brotzakis *et al*. (k_{off} = 4176 ± 324 s^{-1})[9], although the first three studies required total simulation times beyond multiple microseconds. The computational cost of the funnel metadynamics study by Brotzakis *et al*. (500 ns) is similar to the current work (~550 ns).

To elucidate the binding and unbinding mechanism, we monitored the protein-ligand distance, non-bonded interaction energy, and the formation and the dissolution of the hydrogen bonds between the protein and the ligand (Figure 3). During the course of our enhanced sampling simulation, we observed two distinct binding and unbinding events as monitored by the collective variable of protein-ligand distance, as discussed in the methods section. The simulation begins from an unbound state and quickly reaches in a bound state from 1-10 ns (Figure 3A). Then the ligand dissociates and explores the unbound and intermediate states before gradually getting inside the binding pocket at ~130 ns. It remains in the binding pocket until ~310 ns after which it dissociates into the unbound conformation and does not bind any further until the end of the trajectory (~550 ns). Interestingly, two intermediate states are explored by the benzamidine ligand at a protein-ligand distance of 5 Å and 7 Å where it does not occupy the binding site but does remain in the vicinity of the trypsin protein. In this state, there are a reduced number of hydrogen bonds (1-2) between the protein and the ligand compared to the true bound conformation (3). This intermediate state is further explored by the ligand after the second unbinding event, although it does not reach to the true bound state anymore. We discuss these intermediate states in further detail later in this section.

To get a better understanding of the thermodynamics of the association and dissociation of the trypsin-benzamidine complex, we computed the electrostatic, van der Waals, and the total non-bonded interaction energy between the protein and the ligand throughout the trajectory (Figure 3B). It is clear that both the van der Waals and the electrostatic energy have a stabilizing effect when the ligand is approaching the protein, As we mentioned previously, we observed two binding and unbinding events in the trajectory. The unbound state is better characterized by the absence of any hydrogen bonds between the protein and the ligand, particularly between the 340-400 ns region. This region has an electrostatic as well as total non-bonded interaction energy in the range of 20-30 kcal/mol. In the bound state, contrarily, these numbers are approximately -15 kcal/mol and -30 kcal/mol respectively. The contribution of the vdW energy in stabilizing the bound state is significantly lower than the

electrostatic effect because the benzamidine ligand is positively charged and shows strong electrostatic interaction with polar protein residues. These results show that non-bonded energy contribution in the binding energy of trypsin-benzamidine complex can be as large as 50-60 kcal/mol. But the actual binding free energy is less than 10 kcal/mol because our non-bonded energy analysis does not include the effect of entropy as well as the stabilization of the bound and unbound systems through interaction with water and ions.

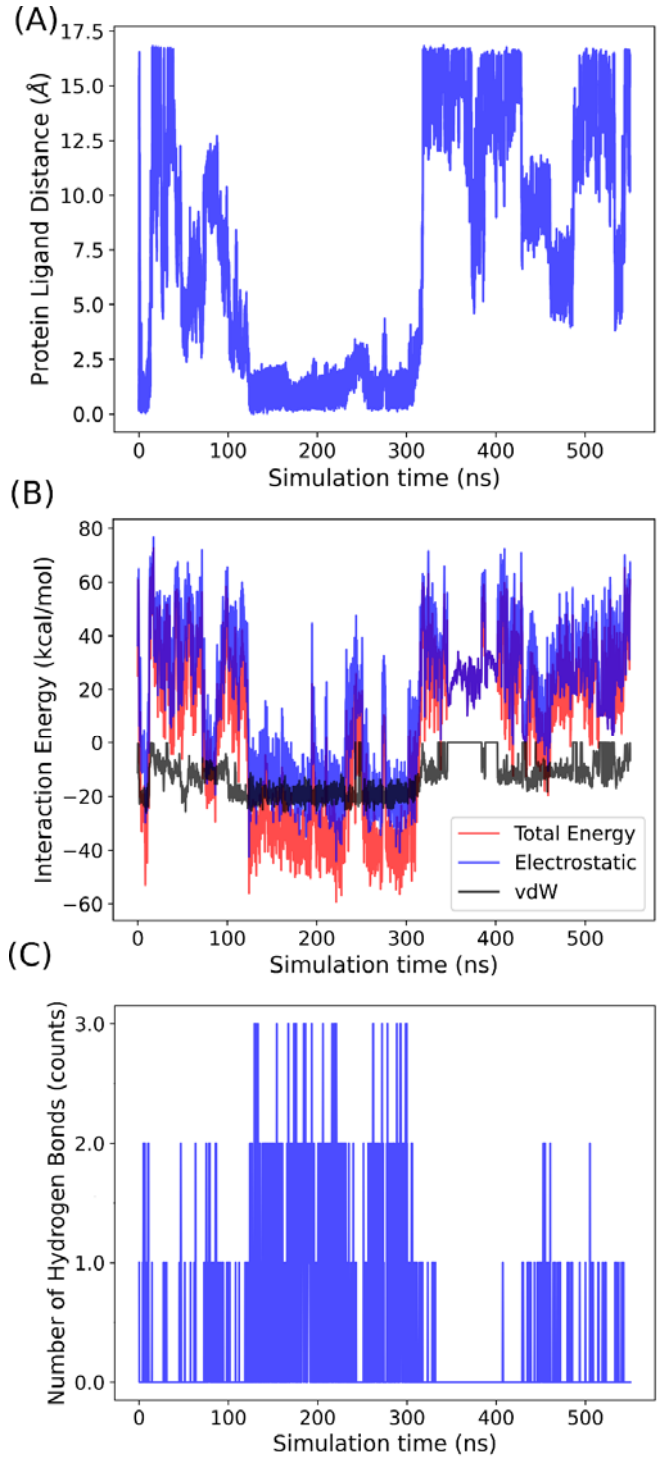


Fig. 3 – (A) The protein-ligand distance from well-tempered meta-eABF method simulation reveals several binding and unbinding events. (B) The vdW and electrostatic interaction energies were also calculated, and reveal the electrostatic forces that influence the binding of benzamidine. (C) The absolute number of hydrogen bonds formed between trypsin and benzamidine as a function of simulation time confirms the presence of binding and unbinding events in the enhanced simulation.

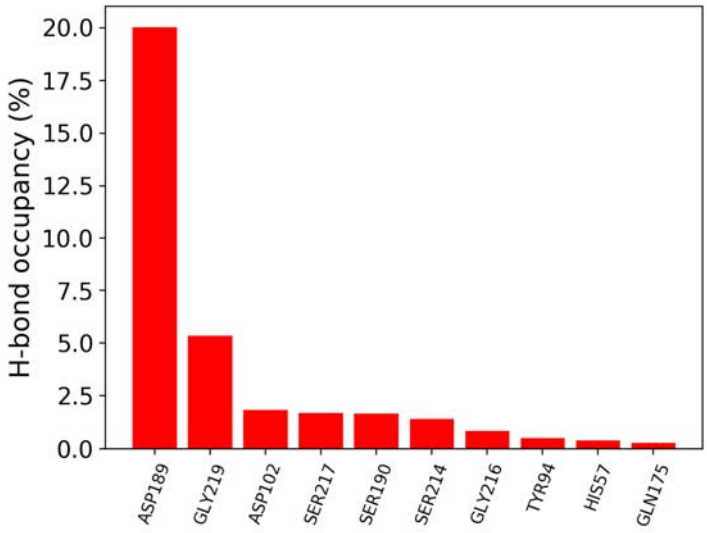


Fig. 4 – The hydrogen bond occupancy formed between trypsin and benzamidine.

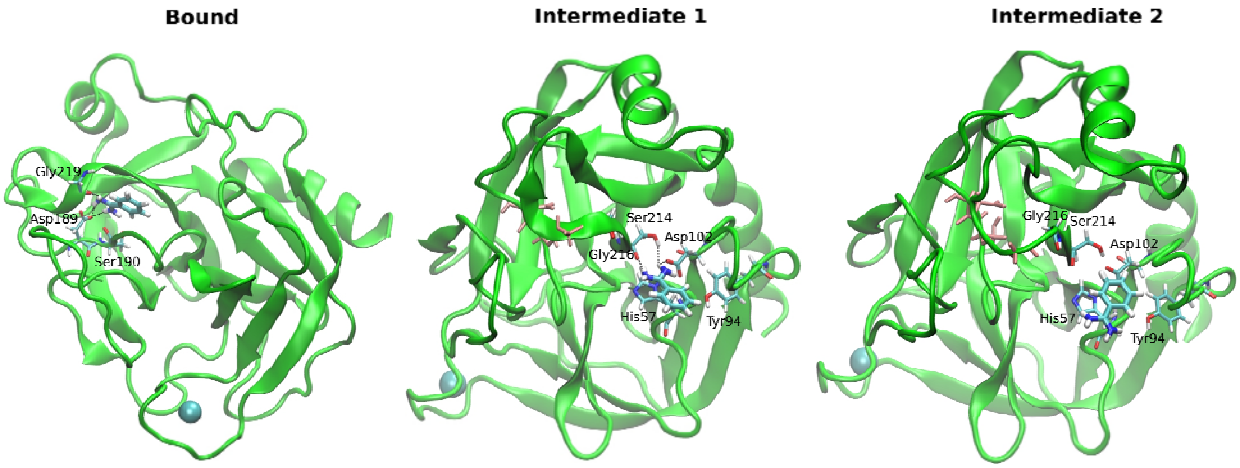


Fig. 5 – The bound state and the two intermediate states were observed in the WTM-eABF trajectory of the trypsin-benzamidine complex.

To get a mechanistic insight of the binding and unbinding process, we tracked the protein residues that form hydrogen bonds with the benzamidine in the entire trajectory. The percentage occupancy of all hydrogen bonds with >0.5% occupancy are shown in Figure 4. Apart from the two hydrogen bonds present in the bound state with Asp189, Ser190 and Gly219,[9] we observed the appearance of many new hydrogen bonds with other residues such as Gly216, Ser217, Ser214 and even some residues, distant from the binding pocket in the protein sequence like Asp102, Tyr94, His57. These residues interact with the benzamidine when it is in the process of transition between the bound and the unbound states. Previous studies by Tiwary *et al.* [8] and Brotzakis *et al.*[9] have also identified these residues to be interacting with benzamidine ligand at various intermediate states during the dissociation of the complex.

There are two intermediate states, which we captured through the WTM-eABF simulation, indicated by the protein-ligand distance of 5-6 Å (between 50-80 ns and 460-490 ns of the trajectory), and 7-8 Å (between 430-460 ns). We refer to them as Intermediate 1 and Intermediate 2 respectively (Figure 5). Intermediate 1 has predominantly hydrogen bonding interaction between the benzamidine and the residues His57, Tyr94 and Asp102, alongside residues near the binding pocket such as Ser214 and Gly216, indicating that the intermediate 1 is similar in configuration with the P state observed by Brotzakis *et al*. The Intermediate 2, on the contrary, has the polar group of the ligand facing outward and stabilized primarily by the solvation effect of the water and counterions, while the interaction with the protein is via hydrophobic interactions, much like the B state in Brotzakis *et al*.[9]

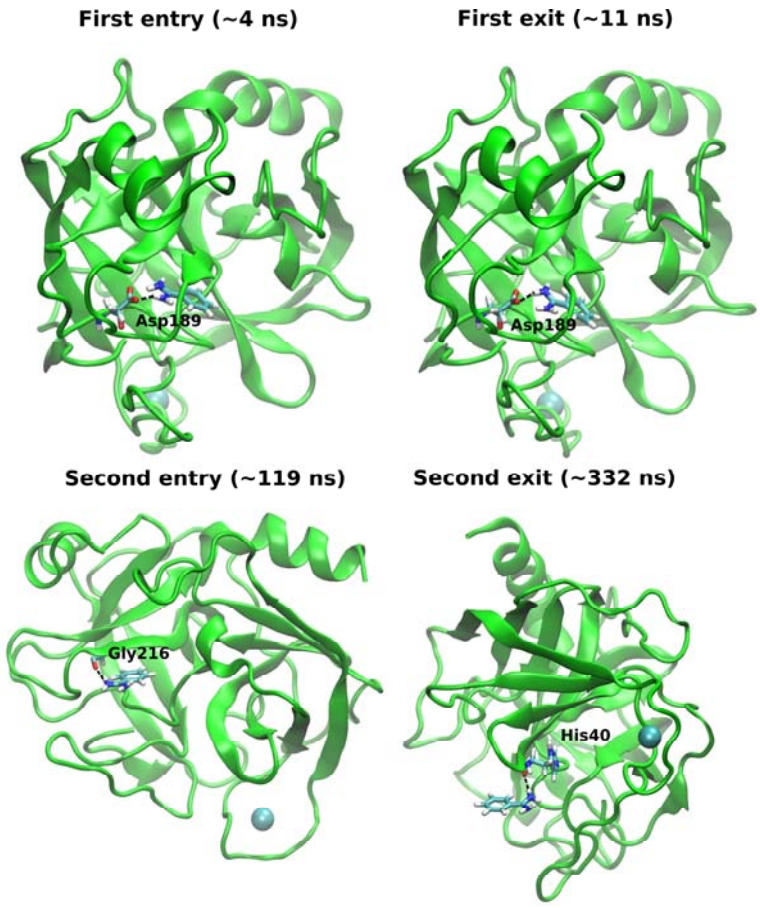


Fig. 6 – The structures of the entry points and the exit points for the two binding events observed in our simulation. The entry and exit points are defined by the frames where the first hydrogen bond forms and the last hydrogen bond breaks, respectively, between the protein and ligand. The times mentioned in the figure are from the biased simulation and do not correspond to real-time.

We also monitored the hydrogen bond which forms first when the ligand is approaching the protein and the one which breaks last during the dissociation, and the results are depicted in Figure 6. During the first binding event between ~4-11 ns, the first hydrogen bond forms with residue Asp189, and it is also the last residue to interact with the benzamidine ligand via hydrogen bonding. This is indicative of a direct binding and unbinding event, where both the entry and exit happen without much assistance from the neighboring residues. The ligand has to approach from a very specific solid angle to follow such a binding mechanism, making it less likely to occur by random chance. The second binding event, which takes place between ~119-332 ns, makes the ligand travel across the surface of the protein both during association and dissociation. The first residue that forms a hydrogen bond to the incoming ligand is Gly216, a residue that does not form a hydrogen bond in the final bound state. Similarly, during unbinding the last hydrogen bond to be broken is with His40, which rarely forms any hydrogen bond with the ligand in the rest of the trajectory. This pathway is much more general and does not require the ligand to be positioned along any specific orientation. We do observe the frequent change of orientation of the ligand during the transition between bound and unbound state for the second binding event.

CONCLUSIONS

We performed enhanced sampling molecular dynamics simulations using the well-tempered meta-eABF (WTM-eABF) methodology to compute the free energy landscape, binding free energy, unbinding kinetics and to study the mechanism of the association and dissociation of the trypsin-benzamidine complex in atomistic detail. In about 550 ns of enhanced simulation, we

observed two independent association and dissociation events for a process that takes place in the beyond-millisecond timescale. The topology of the free energy landscape computed using this approach is in qualitative agreement with previous, much more extensive computer simulation studies on the same system. The free energy difference between the bound and the unbound state does not reflect the true binding free energy measured in the experiment, as our approach does not take into account the effect of spatial entropy of the ligand in the unbound state with sufficient rigor, and this is an area which requires further investigation in future studies. The ~3 kcal/mol difference in the binding free energy obtained in our calculation can also be due to the use of sub-optimal reaction coordinate of the protein-ligand distance. Repeating this study with advanced path collective variables (path-CV)[50] or machine-learning-based reaction coordinates like SGOOP can potentially increase the accuracy of the computed free energy difference.[51] Despite this limitation, the dissociation kinetics we estimated from our embarrassingly short MD trajectory using a very crude Arrhenius equation argument is within 2 orders of magnitude of the experimental value, an accuracy that is comparable with previous studies performed using advanced techniques such as Markov State modeling (MSM) and machine learning guided infrequent metadynamics. We also explicitly showed that the contribution of electrostatic energy difference in the binding free energy can be very large (50-60 kcal/mol) due to the interaction of the charged protonated nitrogen in the benzamidine with the polar/charged residues in the protein. But that effect is largely compensated due to the solvation and entropic stabilization due to the availability of a larger number of microstates for the unbound conformation of the ligand.

To obtain a more detailed understanding of the association and dissociation mechanisms, we looked at the hydrogen bonds formed between the protein and the ligand during the simulation. We observed that many residues outside the binding pocket also participate in hydrogen bonding with benzamidine and will likely determine the binding pathway and kinetics. Particularly, the first hydrogen bond that is formed during the binding process and the last hydrogen bond that is broken during the dissociation process can change depending on the orientation of the ligand and the solid angle at which it approaches the protein. From the time series of protein-ligand distance and the hydrogen bonding data, we identified two intermediate states between the bound and the unbound states, whose characteristics are in agreement with much more detailed and extensive molecular dynamics studies on this system.

Previously a protocol has been proposed to compute protein-ligand binding free energy using the novel well-tempered meta-eABF enhanced sampling approach, by computing the free energy profile of individual bond angle, torsion etc.[52] We, here, examined the capability of this scheme in studying the protein-ligand binding mechanism, using a very crude and simple distance-based reaction coordinate in one dimension and a single MD run. Although following complex protocols can improve the accuracy of quantitative predictions like binding free energy and kinetics, the intermediates and the pathway elucidated from our primitive approach is consistent with the more extensive computational studies which incur orders of magnitude more computational cost. Overall, we believe that the WTM-eABF approach can find widespread applications, both in academic and industrial research, for screening small molecule drugs in terms of their therapeutic efficacy.

Acknowledgments. The authors thank Rommie Amaro and Lane Votapka for sharing the structure and topology files of the trypsin-benzamidine complex. This work was supported partially by the National Science Foundation (NSF) via grant MCB 2028443. SES acknowledges support through Pfizer - La Jolla Academic Industrial Relations (AIR) Diversity Research Fellowship in Chemistry. The authors thank the UC Irvine High Performance Computing (HPC) facility for providing the computational resources. The authors declare no competing financial interest.

REFERENCES

1. M. Karplus and J. A. McCammon, *Nat. Struct. Biol.*, **2002**, *9*, 646–652. https://doi.org/10.1038/nsb0902-646.
2. L. T. Chong, A. S. Saglam and D. M. Zuckerman, *Curr. Opin. Struct. Biol.*, **2017**, *43*, 88–94. https://doi.org/10.1016/J.SBI.2016.11.019.
3. G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, **1977**, *23*, 187–199. https://doi.org/10.1016/0021-9991(77)90121-8.
4. B. Roux, *Comput. Phys. Commun.*, **1995**, *91*, 275–282. https://doi.org/10.1016/0010-4655(95)00053-I.
5. W. You, Z. Tang and C. A. Chang, *J. Chem. Theory Comput.,* **2019**, *15*, 2433–2443. https://doi.org/10.1021/acs.jctc.8b01142.
6. A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, **2002**, *99*, 12562–12566. https://doi.org/10.1073/pnas.202427399.
7. A. Barducci, G. Bussi and M. Parrinello, *Phys. Rev. Lett.*, **2008**, *100*, 020603. https://doi.org/10.1103/PhysRevLett.100.020603.

8. P. Tiwary, V. Limongelli, M. Salvalaglio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, **2015**, *112*, E386–E391. https://doi.org/10.1073/pnas.1424461112.
9. Z. F. Brotzakis, V. Limongelli and M. Parrinello, *J. Chem. Theory Comput.*, **2018**, *15*, 743–750. https://doi.org/ 10.1021/ACS.JCTC.8B00934.
10. D. Pramanik, Z. Smith, A. Kells and P. Tiwary, *J. Phys. Chem. B*, **2019**, *123*, 3672–3678. https://doi.org/ 10.1021/acs.jpcb.9b01813.
11. V. Limongelli, M. Bonomi and M. Parrinello, *Proc. Natl. Acad. Sci.*, **2013**, *110*, 6358–6363. https://doi.org/ 10.1073/pnas.1303186110.
12. P. Söderhjelm G. A. Tribello and M. Parrinello, *Proc. Natl. Acad. Sci.*, **2012**, *109*, 5170–5175. https://doi.org/ 10.1073/pnas.1201940109.
13. L. Fusani, D. S. Palmer, D. O. Somers and I. D. Wall, *J. Chem. Inf. Model.*, **2020**, *60*, 1528–1539. https://doi.org/ 10.1021/acs.jcim.9b00843.
14. C. A. Söldner, A. H. C. Horn, H. Sticht, *Int. J. Mol. Sci.*, **2019**, *20*, 1970–1978. https://doi.org/10.3390/ ijms20081970.
15. E. Darve, D. Rodríguez-Gómez and A. Pohorille, *J. Chem. Phys.*, **2008**, *128*, 144120–144126. https://doi.org/ 10.1063/1.2829861.
16. J. Comer, J. C. Gumbart, J. Hénin, T. Lelièvre, A. Pohorille and C. Chipot, *J. Phys. Chem. B*, **2015**, *119*, 1129–1151. https://doi.org/10.1021/jp506633n.
17. A. Lesage, T. Lelièvre, G. Stoltz and J. Hénin, *J. Phys. Chem. B*, **2017**, *121*, 3676–3685. https://doi.org/10.1021/ acs.jpcb.6b10055.
18. H. Fu, H. Zhang, H. Chen, X. Shao, C. Chipot and W. Cai, *J. Phys. Chem. Lett.*, **2018**, *9*, 4738–4745. https://doi.org/10.1021/acs.jpclett.8b01994.
19. D. Ray and I. Andricioaei, *Biophys. J.*, **2020**, *119*, 1568–1579. https://doi.org/10.1016/j.bpj.2020.08.031.
20. C. Yang, E. Kim, M. Lim and Y. Pak, *J. Chem. Theory Comput.*, **2019**, *15*, 751–761. https://doi.org/10.1021/ acs.jctc.8b00936.
21. C. Yang, E. Kim and Y. Pak, *Nucleic Acids Res.*, **2015**, *43*, 7769–7778. https://doi.org/10.1093/nar/gkv796.
22. H. Fu, X. Shao, W. Cai and C. Chipot, *Acc. Chem. Res.*, **2019**. https://doi.org/10.1021/acs.accounts.9b00473.
23. H. Fu, H. Chen, X. Wang, H. Chai, X. Shao, W. Cai and C. Chipot, *J. Chem. Inf. Model.*, **2020**, *60*, 5366–5374. https://doi.org/10.1021/acs.jcim.0c00279.
24. F. Guillain and D. Thusius, *J. Am. Chem. Soc.*, **1970**, *92*, 5534–5536. https://doi.org/10.1021/ja00721a051.
25. I. Buch, T. Giorgino and G. De Fabritiis, *Proc. Natl. Acad. Sci. U. S. A.*, **2011**, *108*, 10184–10189. https://doi.org/10.1073/pnas.1103547108.
26. N. Plattner and F. Noé, *Nat. Commun.*, **2015**, *6*, 1–10. https://doi.org/10.1038/ncomms8653.
27. A. Dickson and S. D. Lotz, *Biophys. J.*, **2017**, *112*, 620–629. https://doi.org/10.1016/J.BPJ.2017.01.006.
28. N. Donyapour, N. M. Roussey and A. Dickson, *J. Chem. Phys.*, **2019**, *150*, 244112–244118. https://doi.org/ 10.1063/1.5100521.
29. L. W. Votapka, B. R. Jagger, A. L. Heyneman, R. E. Amaro, *J. Phys. Chem. B*, **2017**, *121*, 3597–3606. https://doi.org/10.1021/acs.jpcb.6b09388.
30. B. R. Jagger, A. A. Ojha and R. E. Amaro, *J. Chem. Theory Comput.*, **2020**, *16*, 5348–5357. https://doi.org/ 10.1021/acs.jctc.0c00495.
31. A. K. Faradjian and R. Elber, *J. Chem. Phys.*, **2004**, *120*, 10880–10889. https://doi.org/10.1063/1.1738640.
32. R. A. Elber, *Biophys. J.*, **2007**, *92*, L85–L87. https://doi.org/10.1529/BIOPHYSJ.106.101899.
33. J. M. Bello-Rivas and R. Elber, *J. Chem. Phys.*, **2015**, *142*, 094102. https://doi.org/10.1063/1.4913399.
34. E. Vanden-Eijnden and M. Venturoli, *J. Chem. Phys.*, **2009**, *130*, 194101–194108. https://doi.org/10.1063/ 1.3129843.
35. D. Ray and I. Andricioaei, *J. Chem. Phys.*, **2020**, *152*, 234114–234120. https://doi.org/10.1063/5.0008028.
36. G. A. Huber and S. Kim, *Biophys. J.*, **1996**, *70*, 97–110. https://doi.org/10.1016/S0006-3495(96)79552-8.
37. B. W. Zhang, D. Jasnow and D. M. Zuckerman, *J. Chem. Phys.*, **2010**, *132*, 054107. https://doi.org/10.1063/ 1.3306345.
38. D. M. Zuckerman and L. T. Chong, *Annu. Rev. Biophys.*, **2017**, *46*, 43–57. https://doi.org/10.1146/annurev-biophys-070816-033834.
39. D. Ray, T. Gokey, D. L. Mobley and I. Andricioaei, *J. Chem. Phys.*, **2020**, *153*, 154117. https://doi.org/10.1063/ 5.0021953.
40. L. Maragliano, E. Vanden-Eijnden and B. Roux, *J. Chem. Theory Comput.*, **2009**, *5*, 2589–2594. https://doi.org/10.1021/ct900279z.
41. D. Ray, S. E. Stone and I. Andricioaei, *J. Chem. Theory Comput*.**, 2022,** *18*, 79-95. DOI: 10.1021/ acs.jctc.1c00803
42. M. Marquart, J. Walter, J. Deisenhofer, W. Bode and R. Huber, *Acta Crystallogr. Sect. B*, **1983**, *39*, 480–490. https://doi.org/10.1107/S010876818300275X.
43. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, **2015**, *11*, 3696–3713. https://doi.org/10.1021/ acs.jctc.5b00255.
44. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, **2004**, *25*, 1157–1174. https://doi.org/10.1002/jcc.20035.
45. H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J. Chem. Phys.*, **2004**, *120*, 9665–9678. https://doi.org/10.1063/ 1.1683075.
46. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, **2005**, *26*, 1781–1802. https://doi.org/10.1002/jcc.20289.
47. G. Fiorin, M. L. Klein and J. Hénin, *Mol. Phys.*, **2013**, *111*, 3345–3362. https://doi.org/10.1080/ 00268976.2013.813594.
48. NAMD Energy Plugin, Version 1.4. https://www.ks.uiuc.edu/ Research/vmd/plugins/namdenergy/
49. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, **1996**, *14*, 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.
50. D. Branduardi, F. L. Gervasio M. Parrinello, *J. Chem. Phys.*, **2007**, *126*, 054103. https://doi.org/10.1063/ 1.2432340.
51. P. Tiwary and B. J. Berne, *Proc. Natl. Acad. Sci. U. S. A.*, **2016**, *113*, 2839–2844. https://doi.org/10.1073/ pnas.1600917113.
52. H. Fu, H. Chen, W. Cai, X. Shao and C. Chipot, *J. Chem. Inf. Model.*, **2021**. https://doi.org/10.1021/ acs.jcim.1c00269.