Changes in Protonation States of In-Pathway Residues can Alter Ligand Binding Pathways Obtained From Spontaneous Binding Molecular Dynamics Simulations

Helena Girame, Marc Garcia-Borràs and Ferran Feixas *

Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, Girona, Spain

Protein-ligand binding processes often involve changes in protonation states that can be key to recognize and orient the ligand in the binding site. The pathways through which (bio) molecules interplay to attain productively bound complexes are intricate and involve a series of interconnected intermediate and transition states. Molecular dynamics (MD) simulations and enhanced sampling techniques are commonly used to characterize the spontaneous binding of a ligand to its receptor. However, the effect of protonation state changes of in-pathway residues in spontaneous binding MD simulations remained mostly unexplored. Here, we used molecular dynamics simulations to reconstruct the trypsin-benzamidine binding pathway considering different protonation states of His57. This residue is part of the trypsin catalytic triad and is located more than 10 Å away from Asp189, which is responsible for benzamidine binding in the trypsin S1 pocket. Our MD simulations showed that the binding pathways that benzamidine follow to target the S1 binding site are critically dependent on the His57 protonation state. Binding of benzamidine frequently occurs when His57 is protonated in the delta nitrogen while the binding process is significantly less frequent when His57 is positively charged. Constant-pH MD simulations retrieved the equilibrium populations of His57 protonation states at trypsin active pH offering a clearer picture of benzamidine recognition and binding. These results indicate that properly accounting for protonation states of distal residues can be important in spontaneous binding MD simulations.

Keywords: ligand binding pathways, protonation states, spontaneous binding simulations, constant-pH molecular dynamics, trypsin-benzamidine complex

INTRODUCTION

Characterizing the mechanisms of ligand binding and unbinding to a biomolecule is crucial to elucidate the molecular basis of biological processes and improve the potency of drugs (Bernetti et al., 2019). The pathways through which (bio)molecules interplay to attain stable (and often transient) bound complexes are intricate and involve a series of interconnected intermediate, misbound, and transition states. Molecular dynamics (MD) simulations and enhanced sampling techniques are frequently used to characterize the spontaneous binding pathways of drugs, substrates, or peptides to its biological receptors (Dror et al., 2011; Shan et al., 2011; Decherchi and Cavalli, 2020). In these

FIGURE 1 | Overview of trypsin structure and S1 pocket. **(A)** Trypsin (in purple) structure corresponding to PDB ID 3PTB with benzamidine inhibitor (in grey) bound to the S1 pocket (orange surface). His57 and Asp189 residues are highlighted in yellow and in orange sticks, respectively. Calcium ion is depicted as a green sphere. Overview of the S1 pocket with catalytic residues (Asp102, His57, and Ser195) shown in yellow sticks. The distance between the carbon of the side-chain of Asp189 and the side-chain of His57 is 13 Å.

simulations, one or more ligands are commonly placed in the solvent and are allowed to freely diffuse without biasing the MD simulation toward a particular protein region. Providing sufficient simulation time and an accurate description of the system, the ligand freely explores the dynamic protein surface until it spontaneously finds its presumed binding site (Betz and Dror, 2019). Markov-State Models were used to completely reconstruct ligand binding and unbinding pathways and the associated kinetics of enzyme-inhibitor complexes (Buch et al., 2011; Plattner and Noé, 2015). Unconstrained enhanced sampling methods were used to simulate binding of allosteric modulators into G-protein coupled receptors (Miao and McCammon, 2016) or substrate binding in allosterically regulated enzymes (Calvó-Tusell et al., 2022). The predictive power of spontaneous binding simulations relies on being able to access the timescale required to sample the binding event and critically depends on the accurate description of the simulated system.

Around 60% of protein-ligand binding events involve changes in protonation states (Aguilar et al., 2010; Onufriev and Alexov, 2013). Properly accounting for protonation states of protein residues is crucial to characterize ligand binding with MD simulations. The prediction of protonation states from rigid X-ray structures can lead to their incorrect assignment as even subtle structural fluctuations can affect each residue environment. With constant-pH molecular dynamics (CpH-MD) it is possible to model pH effects retrieving the protonation equilibriums of titratable residues coupled to protein conformational dynamics (Mongan et al., 2004; Khandogin and Brooks, 2005; Chen et al., 2014; Huang et al., 2016). Recently, Vo and co-workers reconstructed how fentanyl binds μ-opiod receptor with CpH-MD showing that the protonation of His257 at the binding pocket plays a crucial role to properly orient fentanyl (Vo et al., 2021). These results point out the importance of correctly accounting for protonation states of residues in the binding pocket to characterize the thermodynamics and kinetics of ligand-binding. However, as

captured by spontaneous binding MD simulations, the ligand can establish contact with different protein residues in its pathway toward the binding site. The nature of these interactions will also be determinant for the kinetics of the ligand binding process. Despite the number of studies of protein-ligand pathways, the effect of protonation state changes of in-pathway residues in spontaneous binding simulations remains mostly unexplored.

The binding of benzamidine to trypsin has been commonly used as an enzyme-inhibitor model system for studying spontaneous binding and benchmarking enhanced sampling techniques due to the rapid formation of the trypsin-benzamidine complex (Betz and Dror, 2019). Trypsin is a serine protease responsible of hydrolyzing proteins through a catalytic triad formed by Ser195, His57, and Asp102 (*see* **Figure 1**). The positively charged inhibitor benzamidine is recognized in the specific S1 pocket which contains a negatively charged Asp189 located more than 10 Å away from the catalytic triad. In a landmark publication, Buch et al. reconstructed the free-energy landscape of benzamidine binding from a total of 495 MD simulations of 100 ns, observing productive binding in 38% of the simulations (Buch et al., 2011). By analyzing the independent trajectories, they observed that catalytic His57 and Ser195 residues were commonly found in the binding pathway of benzamidine in its way toward the S1 pocket. The binding of benzamidine have also been studied using unconstrained enhanced sampling methods by Miao and co-workers who reconstructed binding and unbinding pathways using Gaussian accelerated molecular dynamics (GaMD) (Miao et al., 2020). Interestingly, GaMD unbinding pathways showed that benzamidine passes next to His57 in its dissociation from the S1 pocket to the solvent. Therefore, His57 play a prominent role in both catalysis and binding.

Enzymes are sensitive to pH changes and catalytic residues commonly change their protonation states at different stages of the catalytic cycle. Trypsin is active in a pH range between 7.0 and

9.0 as His57 is required to alter between two protonation states along binding, acylation and deacylation steps of the hydrolysis reaction (Sipos and Merkel, 1970; Malthouse, 2020). Czodrowski and co-workers studied the protonation changes in ligand binding in trypsin concluding that His57 is responsible for the most relevant pKa shifts during binding and catalysis (Czodrowski et al., 2007). Most common software to assign protonation states from X-ray structures predict a positively charged His57 (both delta and epsilon nitrogens protonated, HIP) at pH = 7.0 while a less clear picture arises at pH = 8.0, where both HIP and neutral His57 with the delta nitrogen protonated (HID) are possible protonation states. Short-time scale MD simulations revealed that both HIP and HID were possible protonation states of His57 (Uranga et al., 2012). Spontaneous binding simulations of the benzamidine-trypsin system have been commonly performed with His57 in the HID state, which is the assumed protonation state when the Michaelis complex is formed (Wahlgren et al., 2011). The question is whether the protonation state of His57 can influence benzamidine binding.

Here, we use spontaneous binding MD simulations to reconstruct the trypsin-benzamidine binding pathway considering different protonation states of His57. This histidine is part of the trypsin catalytic triad and is located more than 10 Å away from Asp189 responsible for benzamidine binding in the S1 pocket (**Figure 1**). Our MD simulations show that the spontaneous binding pathways are critically dependent on the His57 protonation state. Binding of benzamidine frequently occurs in a few hundreds of nanoseconds when histidine is protonated in delta (HID) while productive binding is scarcely observed when His57 it is positively charged (HIP). CpH-MD simulations reflect that both HID and HIP forms are significantly populated at the pH range between 7.0 and 8.0 showing the displacement of the equilibrium toward the HID protonation state upon pH increase. These results indicate that properly accounting for protonation states of distal residues can be key to obtain reliable pathways in spontaneous binding simulations.

METHODS

System Preparation
We used the crystal structure of benzamidine-bound *Bos taurus* trypsin (PDB ID 3PTB) as starting point for our molecular dynamics (MD) simulations. First, benzamidine was removed from the S1 pocket to protonate the system. Second, protonation states of all protein residues were assigned based on 3.0 H++ webserver (http://biophysics.cs.vt.edu/H++) at pH 7.0 (Anandakrishnan et al., 2012). To explore the role of His57 protonation in benzamidine spontaneous binding, we manually assigned the protonation of His57 residue to either HID, HIE, or HIP. Once protonated, four benzamidine molecules were arbitrarily placed in the solvent, 30 Å away from Asp189 binding pocket, as described by Miao (Miao et al., 2020). Benzamidine parameters for MD simulations were obtained from the generalized AMBER force field (GAFF) (Wang et al.,

2004), with partial charges set to fit the electrostatic potential generated at HF/6-31G* level of theory by restrained electrostatic potential model (Bayly et al., 1993). The atomic charges were calculated according to the Merz–Singh–Kollman (Singh and Kollman, 1984; Besler et al., 1990) scheme using Gaussian 09 (Frisch et al., 2016).

Conventional Molecular Dynamics Simulations
Spontaneous MD simulations starting from three different protonation states of His57 were performed in explicit water using AMBER18 package (Case et al., 2018). AMBER-ff14SB force field (Maier et al., 2015) was used to describe the protein, GAFF for benzamidine, and TIP3P for water molecules (Jorgensen et al., 1983). Each system was solvated in a cubic box with a 12 Å buffer of TIP3P water molecules and was neutralized by adding chloride counterions (Cl⁻). Subsequently, a two-stage geometry optimization approach was performed: *1)* a short minimization of water molecules, with positional restraints on solute molecules; *2)* an unrestrained minimization of all the atoms in the simulation cell. Then, the systems were heated using six steps of 50 ps, incrementing the temperature 50 K each step (0–300 K) under constant-volume, periodic-boundary conditions, and the particle-mesh Ewald approach to introduce long-range electrostatic effects (Darden et al., 1993). A 10 Å cut-off was applied to Lennard-Jones and electrostatic interactions. Bonds involving hydrogen were constrained with the SHAKE algorithm (Ryckaert et al., 1977). The Langevin equilibration scheme is used to control and equalize the temperature (Wu and Brooks, 2003). The time step was kept at 2 fs during the heating stages. Each system was then equilibrated for 2 ns with a 2 fs timestep at a constant pressure of 1 atm to relax the density of the system. After the systems were equilibrated in the NPT ensemble, 50 replicas of 200 ns MD simulations for each protonation state of His57 (HID, HIE, HIP) were performed under the NVT ensemble and periodic-boundary conditions.

Constant-pH Molecular Dynamics Simulations
A total of 30 replicas of 200 ns of spontaneous binding constant-pH MD simulations (CpH-MD) were run at pH 7.0 and 8.0 considering all His residues as titratable (His40, His57, and His91). Here, discrete CpH-MD simulations have been carried out following the protocol described by Swails and co-workers as implemented in Amber (Swails et al., 2014): the MD is propagated in explicit solvent following the previously described protocol while the protonation state changes are carried out using a Generalized-Born (GB) implicit solvent model. A salt concentration of 0.1 M was also introduced to reproduce the same GB conditions the original algorithm was parametrized for (Mongan et al., 2004). Every 50 MD steps (100 fs) in explicit solvent, the protonation state of selected titratable residues in random order could change *via* Metropolis Monte Carlo attempts in an implicit solvent
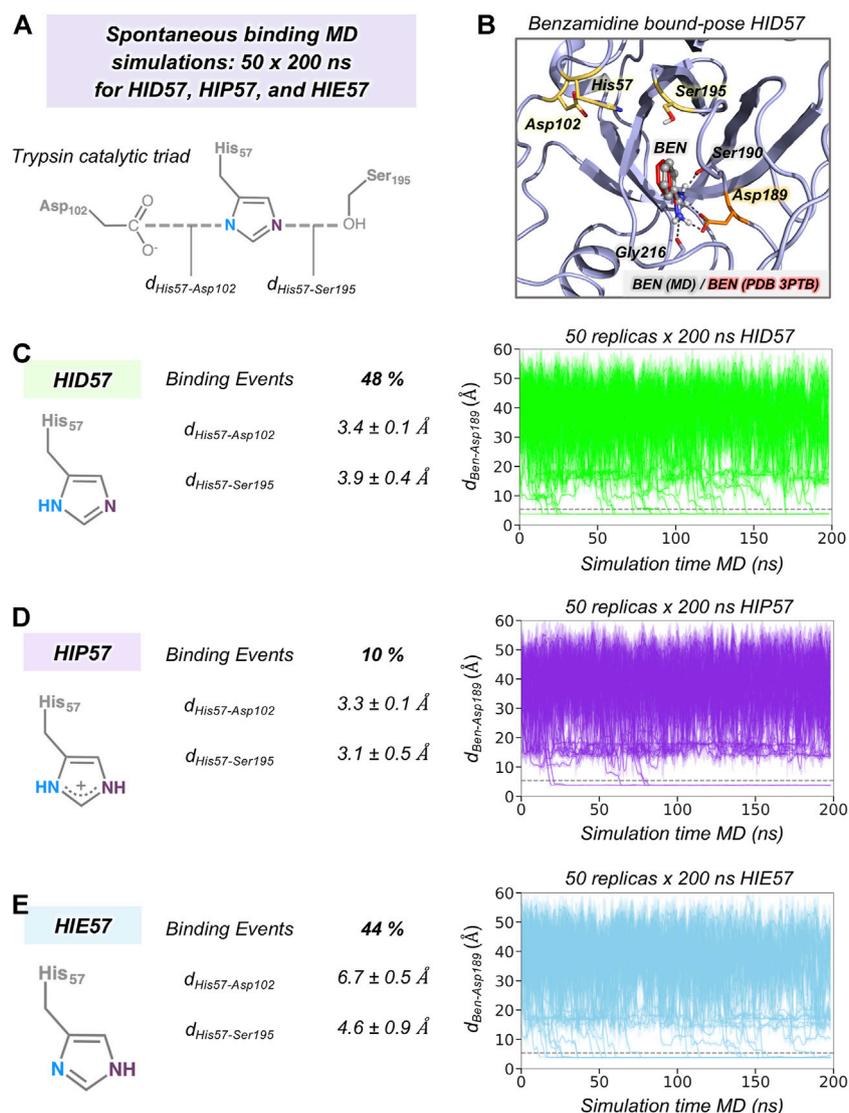
FIGURE 2 | The effect of His57 protonation in benzamidine binding. **(A)** Representation of the trypsin catalytic triad formed by Asp102, His57, and Ser195. The distance between Asp102 and His57 residues is calculated between the carbon of the carboxylate group of Asp102 and the delta nitrogen of His57. The distance between His57 and Ser195 is calculated between the epsilon nitrogen of His57 and the oxygen of the hydroxyl group of Ser195. **(B)** Representative conformation of the trypsin-benzamidine bound complex predicted from spontaneous binding HID57 MD simulations. The binding pose of benzamidine obtained from molecular dynamics (MD) simulations is shown in grey while the X-ray orientation (PDB 3PTB) is depicted in red. Catalytic residues and Asp189 are coloured in yellow and in orange, respectively. **(C)** Analysis of 50 replicas of 200 ns of HID57 MD simulations with the delta nitrogen of His57 protonated. **(D)** Analysis of 50 replicas of 200 ns of HIP57 MD simulations with His57 positively charged. **(E)** Analysis of 200 ns of HIE57 MD simulations with the epsilon nitrogen of His57 protonated. The percentage of binding events and the average distances (in Å) between catalytic residues is provided for each protonation state. Plot of the distance between the carbon atom of the amidine group of benzamidine and the carbon of the carboxylate group of Asp189 side chain along the 50 replicas of 200 ns MD simulations for each protonation state. The grey horizontal dashed line indicates when productive benzamidine binding takes place (distance below 5.4 Å).

framework. When the protonation state changes, a total of 100 relaxation steps (200 fs) were performed to relax the explicit solvent dynamics.

RESULTS

To evaluate the impact of His57 protonation state in the binding of benzamidine, we performed spontaneous binding molecular dynamics (MD) simulations placing four benzamidine molecules at least 30 Å away from the trypsin S1 pocket. From this starting point, benzamidine molecules are allowed to freely diffuse from the solvent and explore the trypsin surface in the predefined simulation time. As shown by Buch et al., spontaneous binding can readily occur in the nanosecond time scale (Buch et al., 2011). Spontaneous binding MD simulations were performed using two different strategies. First, a total of 50 replicas of 200 ns MD simulations were run for the three possible protonation states of

His57: *1*) delta nitrogen protonated (HID57); *2*) epsilon nitrogen protonated (HIE57); and *3*) positively charged histidine (HIP57). Second, we carried out a total of 30 replicas of 200 ns of constant-pH MD simulations at pH 7.0 and 8.0. From these simulations, the percentage of binding events and ligand pathways of benzamidine into trypsin were extracted. To evaluate the binding of benzamidine along the MD simulations, we monitored the distance between the amidine carbon of benzamidine and the carboxylate carbon of Asp189 side chain. We considered that productive binding is attained when this distance is below 5.4 Å, which is the minimum distance to retain the salt bridge interaction between the inhibitor and Asp189 in the S1 pocket.

The Effect of His57 Protonation States in Benzamidine Binding

Our MD simulations showed that spontaneous binding of benzamidine to the S1 pocket can occur in all protonation states of His57 (*see* **Figure 2**). The preferred binding mode obtained from MD simulations in the three protonation states is equivalent to the one observed in the PDB 3PTB (trypsin-benzamidine complex), which is also the principal binding pose predicted in previous computational studies (Buch et al., 2011; Plattner and Noé, 2015; Miao et al., 2020). However, a different number of binding events was observed for the three protonation states of His57. In particular, benzamidine binds the S1 pocket in 48% of HID57 simulations (24 out of 50 replicas of 200 ns), 44% of HIE57 (22 out of 50 replicas), and 10% of HIP57 (5 out of 50 replicas), as shown in **Figure 2**. Therefore, binding frequently occurs when His57 is found in its neutral form (either delta or epsilon nitrogen protonated) and becomes a less frequent event in its positively charged state (HIP57). Buch et al. described a total of 38% of productive binding events protonating His57 as HID but using shorter simulation time and only one molecule of benzamidine (Buch et al., 2011). Considering only the replicas that captured productive binding events, the average binding times are: 88, 80, and 55 ns for HID57, HIE57, and HIP57 respectively. These results point out that binding can readily occur in all cases irrespective of the lower probability of binding events observed for HIP57. Therefore, the open question is how the different protonation states of His57 alter the probability of binding of benzamidine.

In terms of global conformational dynamics, trypsin showed similar flexibility in the three protonation states of His57. Root-mean square fluctuations (RMSF) indicated that the flexibility of the S1 pocket is not significantly altered (*see* **Supplementary Figure S1**). The main differences were observed when analysing the stability of the catalytic triad formed by Asp102, His57, and Ser195 (*see* **Figure 2**). When His57 is positively charged (HIP57), the catalytic triad remains significantly stable, being 3.2 ± 0.5 Å and 3.3 ± 0.1 Å the Ser195-His57 and Asp102-His57 distances respectively. Additional flexibility is gained in the HID57 state with distances of 3.9 ± 0.4 Å (Ser195-His57) and 3.4 ± 0.1 Å (Asp102-His57). The intrinsic dynamism of the Ser195-His57 interaction in HID57 can be key to accommodate trypsin substrates triggering the formation of the Michaelis complex.

Finally, the HIE57 protonation state is the least probable in trypsin because as shown in the MD simulations, when protonated in epsilon, His57 destabilizes the catalytic triad (*see* **Figure 2**). For this reason, all subsequent analyses will be focused on HID57 and HIP57 states.

Characterization of Benzamidine Binding Pathways

To gain insight into the molecular basis of the effect of His57 protonation changes in the binding process, we explored the binding pathways that benzamidine followed to get into trypsin S1 pocket. First, we collectively represented all spontaneous binding MD simulations corresponding to each protonation state using two coordinates (*see* **Figure 3** and **Supplementary Methods**): *1*) the binding distance ($d_{Ben-Asp189}$, x axis) between the amidine carbon of benzamidine and the carboxylate carbon of Asp189; and *2*) the distance between the amidine carbon of benzamidine and the epsilon nitrogen of His57, for either HID57 or HIP57 ($d_{Ben-His57}$, y axis). We selected the epsilon nitrogen because it is directly interacting with the hydroxyl group of catalytic Ser195. In both HID57 and HIP57, the most populated state of this 3D free-energy landscape (FEL) is the benzamidine-bound state (i.e. $d_{Ben-Asp189}$ below 5 Å and $d_{Ben-His57}$ around 10 Å, *see* **B** in **Figure 3**). Benzamidine also accumulates in both cases in a region of the FEL defined by a range of $d_{Ben-Asp189}$ [15,20] Å and $d_{Ben-His57}$ [10,15] Å, which corresponds to trypsin hydrophobic S3 pocket composed by Trp210 and surrounding residues. Interestingly, the FEL displays significant differences in the binding patterns of HID57 and HIP57 when benzamidine approaches the S1 pocket ($d_{Ben-Asp189}$ within [5,15] Å). For HID57, the FEL shows a metastable state where benzamidine directly interacts with His57 ($d_{Ben-His57}$ below 5 Å while $d_{Ben-Asp189}$ is still found between 10 and 15 Å, *see* **I₁** and **I₂** states in **Figure 3A**). This state is not visited in the FEL of HIP57 indicating that the interaction between His57 and benzamidine is not established in these MD simulations (*see* **Figure 3B**). These results are not surprising considering that both benzamidine and His57 are positively charged in HIP57 simulations resulting in repulsive interactions that prevent the interaction. From these simulations, we estimated the free-energy difference between the unbound conformation and the transition state that leads to productive binding (*see* **Supplementary Figure S2**). The free-energy difference is lower for HID57 (around 2.5 kcal/mol) than for HIP57 (around 4 kcal/mol), pointing out that the binding of benzamidine is globally slowed down when His57 is positively charged. Despite the free energy differences are not significant, the different distribution of binding events and the reshape of the FEL indicates that benzamidine binding is modulated by the protonation state of His57 which is located more than 10 Å away from Asp189.

To characterize the molecular basis of the ligand binding processes, we independently analyzed the MD trajectories projecting them into the corresponding FEL (*see* **Figure 3** and **Supplementary Figure S3**). The analysis of independent HID57 MD trajectories showed that benzamidine binding occurred mainly through two different pathways. In the major binding
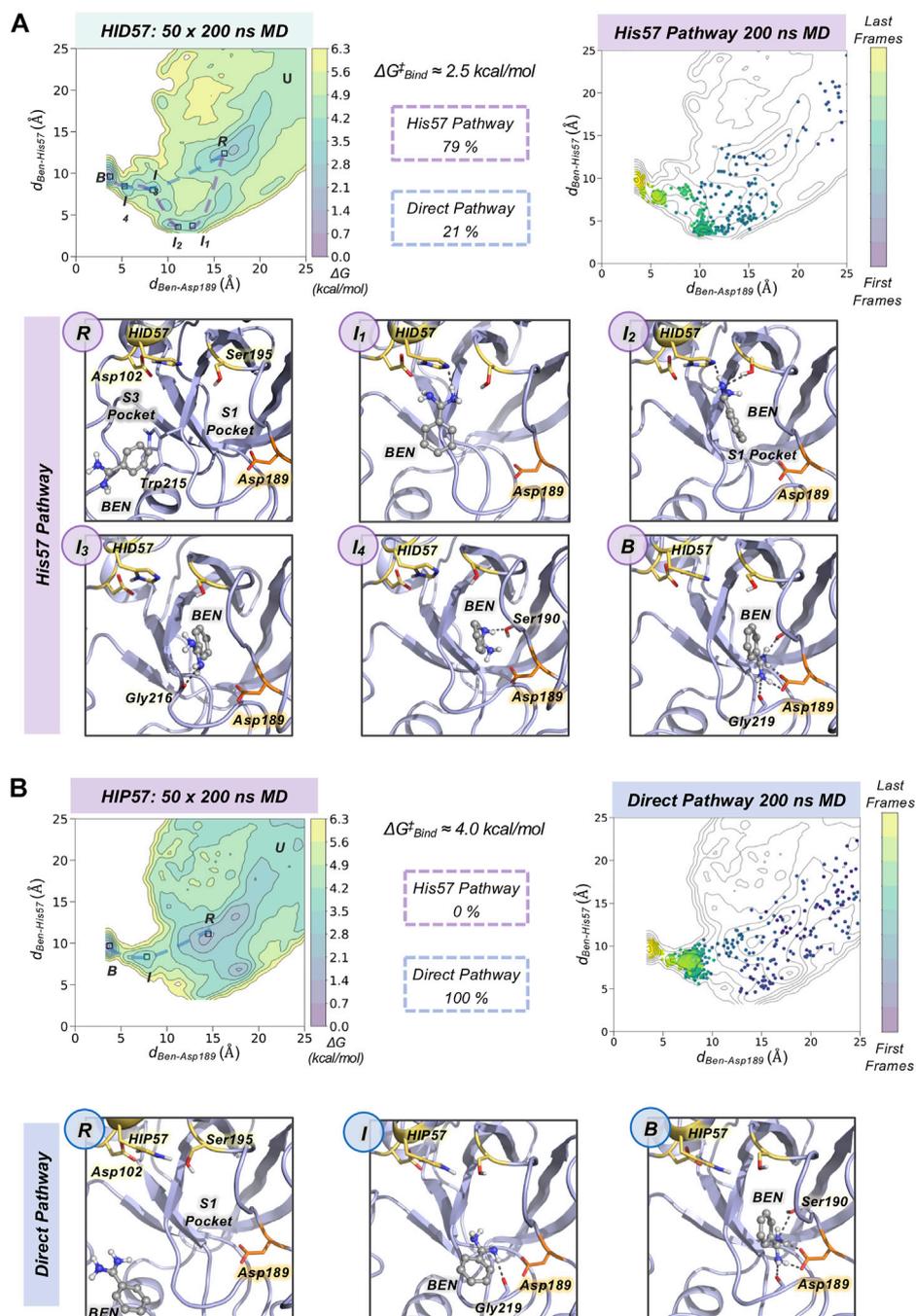
FIGURE 3 | Characterization of benzamidine binding pathways. Free energy landscape (FEL) reconstructed from 50 replicas of 200 ns of spontaneous binding MD simulations of HID57 **(A)** and HIP57 **(B)** using the binding distance ($d_{Ben-Asp189}$, x axis) between the carbon atom of the amidine group of benzamidine and the carbon of the carboxylate group of Asp189 and the distance between the carbon atom of the amidine group of benzamidine and the epsilon nitrogen of His57 ($d_{Ben-His57}$, y axis). The most relevant states of the FEL are highlighted in black boxes: U (unbound), R (recognition), B (bound) and I (intermediate) states. The free-energy difference between the unbound conformation and the transition state that leads to productive binding is given in kcal/mol. The purple dashed line indicates the trajectory of the His57 pathway while the blue dashed line indicates the trajectory of the direct pathway. The percentage of binding events (considering only productive binding simulations) that follow each pathway is provided. Projection of a representative spontaneous binding 200 ns MD trajectory on the FEL of each HID57 and HIP57. The time evolution of the ligand binding pathway is represented in a colour scale ranging from purple for the first frames to yellow for the last frames of the MD trajectory. Molecular representation of the most relevant states of the FEL corresponding to the His57 and direct pathways. Catalytic residues are shown in yellow, benzamidine in grey, and Asp189 in orange.

FIGURE 4 | Spontaneous Benzamidine Binding with Constant-pH Molecular Dynamics Simulations. Free energy landscape (FEL) reconstructed from 30 replicas of 200 ns of spontaneous binding constant-pH MD simulations at pH = 7.0 **(A)** and pH = 8.0 **(B)** using the binding distance ($d_{Ben-Asp189}$, x axis) between the carbon atom of the amidine group of benzamidine and the carbon of the carboxylate group of Asp189 and the distance between the carbon atom of the amidine group of benzamidine and the epsilon nitrogen of His57 ($d_{Ben-His57}$, y axis). The most relevant states of the FEL are highlighted in black boxes: U (unbound), R (recognition), B (bound) and I (intermediate) states. Projection of a representative 200 ns trajectory of the His57 pathway showing the protonation state of each frame in different colour: HID in green, HIP in purple, and HIE in blue. The equilibrium populations of each protonation state retrieved from the 30 replicas of 200 ns is provided for pH = 7.0 and 8.0.

pathway (observed in 79% of productive binding simulations), benzamidine first enters the S3 pocket being recognized by Trp210. Catalytic Asp102 may be involved in the attraction of positively charged benzamidine into the S3 pocket (*see* **Supplementary Figure S4**). Second, the inhibitor rolls along this pocket to establish a hydrogen bond interaction with His57 while keeping the phenyl moiety in the S3 pocket. Then, the aromatic ring of benzamidine repositions from the S3 to the S1 pocket maintaining the hydrogen bonding with catalytic His57 and Ser195. Finally, through a series of successive steps, benzamidine turns around to establish a salt bridge interaction with the carboxylate group of Asp189, attaining the benzamidine-bound pose. We term this predominant pathway as "His57 pathway" since establishing a hydrogen bond interaction with His57 is a requisite to access the S1 pocket. A similar binding pathway was described by Buch and co-workers using 495 100 ns MD simulations (Buch et al., 2011). A second, less frequent, pathway was observed in 21% of productive binding simulations. In this case, benzamidine directly evolves from the solvent to the S1 pocket without establishing a hydrogen bond interaction with HID57 (*see* **Figure 3**). The direct access to the S1 pocket from the

solvent can take place from different sites but the common feature is that benzamidine directly enters the pocket establishing a hydrogen bond with Ser190 and then repositions to interact with Asp189 through a salt-bridge interaction. In this particular pathway, the binding of benzamidine is practically a pure diffusion from the solvent to the S1 pocket, for this reason we term it as the "direct pathway."

In HIP57 MD simulations, benzamidine binding only occurred through the direct pathway (5 out of 50 replicas), as shown in **Figure 3B**. The repulsion between the positive charges of both protonated His57 (HIP) and benzamidine prevents their approximation and interaction. Thus, the His57 pathway is not observed in HIP57 simulations, making the number of binding events significantly less frequent than in HID57. Despite benzamidine can be recognized also in the S3 pocket by Trp210, when it approaches the positively charged HIP57 a hydrogen bond with this catalytic residue cannot be established and benzamidine returns back to the solvent (*see* **Supplementary Figure S5**). Therefore, when His57 is positively charged, binding will preferentially occur through direct diffusion from the solvent. These results demonstrate that the protonation state of His57

determines which ligand binding pathways can be populated and, thus, the path that benzamidine preferentially takes to access the S1 pocket of trypsin.

Spontaneous Benzamidine Binding with Constant-pH Molecular Dynamics Simulations

Using fixed protonation states for His57 can offer a limited picture of benzamidine binding, considering that this residue is responsible for the most relevant pKa shifts during binding and catalysis in trypsin (Czodrowski et al., 2007). To account for the pH effects in the benzamidine binding process, we performed 30 replicas of 200 ns of spontaneous binding constant-pH MD simulations at pH 7.0 and 8.0 (*see* **Figure 4** and **Supplementary Figure S6** for a complete analysis and S7 for convergence of equilibrium populations). In these simulations, we allowed the three histidines of trypsin to change their protonation state. Only His57 is found along the binding pathway and can play a key role in the benzamidine recognition process. At pH 7.0, the populations of the different protonation states obtained from CpH-MD simulations were 74, 26, and 0% for HIP, HID, and HIE, respectively. From these CpH-MD simulations at pH 7.0, we obtained a FEL of the binding process that resembles the one captured for HIP57 protonation state (*see* **Figures 3B, 4A**). However, the intermediate states corresponding to the interaction between benzamidine and His57 became moderately populated. This increases the number of binding events compared to HIP57 simulations up to 23% (7 out of 30 replicas). In this case, benzamidine accessed the S1 pocket through both direct and His57 pathways. Interestingly, CpH-MD simulations showed that the His57 pathway is activated only when His57 attains the HID protonation state (*see* **Figure 4A**). At pH 8.0, we observed equilibrium populations of protonation states for His57 of 57% for HID, 42% for HIP, while HIE is scarcely populated. The FEL obtained from CpH-MD simulations at pH 8.0 clearly highlights the stabilization of the interaction between His57 and benzamidine (*see* I_1 and I_2 states in **Figure 4B**). Again, binding of benzamidine proceeded through both the direct and His57 pathways. The number of binding events retrieved from CpH-MD simulations at both pH 7.0 and 8.0 is found between the values observed in HIP57 and HID57 MD simulations offering a clearer picture of the His57 protonation state ensemble. Therefore, properly accounting for the equilibrium of protonation states of His57 may be key to retrieve accurate kinetics for the binding of benzamidine to trypsin.

DISCUSSION

Spontaneous binding MD simulations showed that the protonation state of His57, which is located more than 10 Å away from the gorge of the S1 pocket, plays a key role in determining the binding pathway of benzamidine to trypsin. Binding is more favorable when His57 attains a neutral HID protonation state while is less probable in the positively charged HIP protonation. These results are in line with kinetic experiments that indicate that K_s of substrate N-α-benzyloxycarbonyl-l-lysine-p-nitroanilide increases more than 80 fold when His57 is protonated (Malthouse, 2020). Benzamidine

can access the S1 pocket through two main pathways: the His57 pathway and the direct pathway from the solvent. We observed that His57 is found in the way of benzamidine to the S1 pocket through the most probable binding pathway in the HID protonation state, establishing a hydrogen bond that is key to drive benzamidine toward the binding site. Therefore, subsequent kinetic analysis of the binding process will provide different outcomes depending on how the protonation states are defined at the beginning of the simulation. Constant-pH MD simulations naturally account for the protonation state ensemble of His57 offering a more accurate description of the spontaneous binding of benzamidine at a fixed pH.

Based on these results, we suggest to always consider the impact of protonation changes of residues that are found along the ligand binding pathway (even distal residues) when performing spontaneous binding MD simulations. In particular, in systems with slow binding processes that follow complex pathways through different metastable intermediate and transition states. It is important to remark that in our study we have excluded protonation changes of other residues (e.g. Asp, Glu, . . .) which may further alter the binding pathways obtained. These observations are not limited to the study of drug-binding into their biological receptors. In protein folding studies, it was reported that changes protonation states of certain residues were important to describe the intrinsic dynamics of amyloid-β peptides (Li et al., 2017). In enzyme engineering, substrate access tunnels are commonly engineered through point mutations to evolve the enzyme toward a new function or broad its substrate scope. Thus, properly accounting for coupled protonation state changes between the residues conforming the access tunnel and the substrate will be important to evaluate the substrate binding pathways in enzymes. By unravelling the details of ligand binding and unbinding it is possible to gain insight into the detailed molecular mechanisms of relevant biochemical processes and then, harness this information to rationally improve the potency of drugs and/or evolve an enzyme toward novel functions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

FF, HG, and MG-B designed research. HG performed research. HG, MG-B, and FF analysed data and wrote the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.922361/full#supplementary-material

REFERENCES

Aguilar, B., Anandakrishnan, R., Ruscio, J. Z., and Onufriev, A. V. (2010). Statistics and Physical Origins of pK and Ionization State Changes upon Protein-Ligand Binding. *Biophysical J.* 98, 872–880. doi:10.1016/J.BPJ.2009.11.016

Anandakrishnan, R., Aguilar, B., and Onufriev, A. V. (2012). H++ 3.0: Automating pK Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res.* 40, W537–W541. doi:10.1093/NAR/GKS375

Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993). A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J. Phys. Chem.* 97, 10269–10280. doi:10.1021/j100142a004

Bernetti, M., Masetti, M., Rocchia, W., and Cavalli, A. (2019). Kinetics of Drug Binding and Residence Time. *Annu. Rev. Phys. Chem.* 70, 143–171. doi:10.1146/annurev-physchem-042018-052340

Besler, B. H., Merz, K. M., and Kollman, P. A. (1990). Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.* 11, 431–439. doi:10.1002/JCC.540110404

Betz, R. M., and Dror, R. O. (2019). How Effectively Can Adaptive Sampling Methods Capture Spontaneous Ligand Binding? *J. Chem. Theory Comput.* 15, 2053–2063. doi:10.1021/acs.jctc.8b00913

Buch, I., Giorgino, T., and De Fabritiis, G. (2011). Complete Reconstruction of an Enzyme-Inhibitor Binding Process by Molecular Dynamics Simulations. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10184–10189. doi:10.1073/pnas.1103547108

Calvó-Tusell, C., Maria-Solano, M. A., Osuna, S., and Feixas, F. (2022). Time Evolution of the Millisecond Allosteric Activation of Imidazole Glycerol Phosphate Synthase. *J. Am. Chem. Soc.* 144, 7146–7159. doi:10.1021/jacs.1c12629

Case, D., Ben-Shalom, I., Brozell, S. R., Cerutti, D. S., Cheatham, T., Cruzeiro, V. W. D., et al. (2018). *AMBER 2018.* University of California, San Francisco.

Chen, W., Morrow, B. H., Shi, C., and Shen, J. K. (2014). Recent Development and Application of Constant pH Molecular Dynamics. *Mol. Simul.* 40, 830–838. doi:10.1080/08927022.2014.907492

Czodrowski, P., Sotriffer, C. A., and Klebe, G. (2007). Protonation Changes upon Ligand Binding to Trypsin and Thrombin: Structural Interpretation Based on pKa Calculations and ITC Experiments. *J. Mol. Biol.* 367, 1347–1356. doi:10.1016/J.JMB.2007.01.022

Darden, T., York, D., and Pedersen, L. (1993). Particle Mesh Ewald: AnN·Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* 98, 10089–10092. doi:10.1063/1.464397

Decherchi, S., and Cavalli, A. (2020). Thermodynamics and Kinetics of Drug-Target Binding by Molecular Simulation. *Chem. Rev.* 120, 12788–12833. doi:10.1021/acs.chemrev.0c00534

Dror, R. O., Pan, A. C., Arlow, D. H., Borhani, D. W., Maragakis, P., Shan, Y., et al. (2011). Pathway and Mechanism of Drug Binding to G-Protein-Coupled Receptors. *Proc. Natl. Acad. Sci. U.S.A.* 108, 13118–13123. doi:10.1073/pnas.1104614108

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). *Gaussian 09, Revision A.02.* Wallingford CT: Gaussian, Inc.

Girame, H., Garcia-Borràs, M., and Feixas, F. (2022). Changes in Protonation States of In-Pathway Residues Can Alter Ligand Binding Pathways Obtained from Spontaneous Binding Molecular Dynamics Simulations. *bioRxiv.* doi:10.1101/2022.04.30.490145

Huang, Y., Chen, W., Wallace, J. A., and Shen, J. (2016). All-Atom Continuous Constant pH Molecular Dynamics with Particle Mesh Ewald and Titratable Water. *J. Chem. Theory Comput.* 12, 5411–5421. doi:10.1021/acs.jctc.6b00552

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 79, 926–935. doi:10.1063/1.445869

Khandogin, J., and Brooks, C. L. (2005). Constant pH Molecular Dynamics with Proton Tautomerism. *Biophysical J.* 89, 141–157. doi:10.1529/BIOPHYSJ.105.061341

Li, J., Lyu, W., Rossetti, G., Konijnenberg, A., Natalello, A., Ippoliti, E., et al. (2017). Proton Dynamics in Protein Mass Spectrometry. *J. Phys. Chem. Lett.* 8, 1105–1112. doi:10.1021/acs.jpclett.7b00127

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 11, 3696–3713. doi:10.1021/ACS.JCTC.5B00255

Malthouse, J. P. G. (2020). Kinetic Studies of the Effect of pH on the Trypsin-Catalyzed Hydrolysis of N-α-Benzyloxycarbonyl-L-Lysine-P-Nitroanilide: Mechanism of Trypsin Catalysis. *ACS Omega* 5, 4915–4923. doi:10.1021/acsomega.9b03750

Miao, Y., Bhattarai, A., and Wang, J. (2020). Ligand Gaussian Accelerated Molecular Dynamics (LiGaMD): Characterization of Ligand Binding Thermodynamics and Kinetics. *J. Chem. Theory Comput.* 16, 5526–5547. doi:10.1021/acs.jctc.0c00395

Miao, Y., and McCammon, J. A. (2016). Graded Activation and Free Energy Landscapes of a Muscarinic G-Protein-Coupled Receptor. *Proc. Natl. Acad. Sci. U.S.A.* 113, 12162–12167. doi:10.1073/PNAS.1614538113

Mongan, J., Case, D. A., and McCammon, J. A. (2004). Constant pH Molecular Dynamics in Generalized Born Implicit Solvent. *J. Comput. Chem.* 25, 2038–2048. doi:10.1002/JCC.20139

Onufriev, A. V., and Alexov, E. (2013). Protonation and pK Changes in Protein-Ligand Binding. *Quart. Rev. Biophys.* 46, 181–209. doi:10.1017/S0033583513000024

Plattner, N., and Noé, F. (2015). Protein Conformational Plasticity and Complex Ligand-Binding Kinetics Explored by Atomistic Simulations and Markov Models. *Nat. Commun.* 6, 1–10. doi:10.1038/ncomms8653

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977). Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* 23, 327–341. doi:10.1016/0021-9991(77)90098-5

Shan, Y., Kim, E. T., Eastwood, M. P., Dror, R. O., Seeliger, M. A., and Shaw, D. E. (2011). How Does a Drug Molecule Find its Target Binding Site? *J. Am. Chem. Soc.* 133, 9181–9183. doi:10.1021/ja202726y

Singh, U. C., and Kollman, P. A. (1984). An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.* 5, 129–145. doi:10.1002/JCC.540050204

Sipos, T., and Merkel, J. R. (1970). Effect of Calcium Ions on the Activity, Heat Stability, and Structure of Trypsin. *Biochemistry* 9, 2766–2775. doi:10.1021/bi00816a003

Swails, J. M., York, D. M., and Roitberg, A. E. (2014). Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation. *J. Chem. Theory Comput.* 10, 1341–1352. doi:10.1021/ct401042b

Uranga, J., Mikulskis, P., Genheden, S., and Ryde, U. (2012). Can the Protonation State of Histidine Residues Be Determined from Molecular Dynamics Simulations? *Comput. Theor. Chem.* 1000, 75–84. doi:10.1016/j.comptc.2012.09.025

Vo, Q. N., Mahinthichaichan, P., Shen, J., and Ellis, C. R. (2021). How μ-opioid Receptor Recognizes Fentanyl. *Nat. Commun.* 12, 1–11. doi:10.1038/s41467-021-21262-9

Wahlgren, W. Y., Pál, G., Kardos, J., Porrogi, P., Szenthe, B., Patthy, A., et al. (2011). The Catalytic Aspartate Is Protonated in the Michaelis Complex Formed between Trypsin and an In Vitro Evolved Substrate-like Inhibitor. *J. Biol. Chem.* 286, 3587–3596. doi:10.1074/jbc.M110.161604

Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004). Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 25, 1157–1174. doi:10.1002/JCC.20035

Wu, X., and Brooks, B. R. (2003). Self-guided Langevin Dynamics Simulation Method. *Chem. Phys. Lett.* 381, 512–518. doi:10.1016/j.cplett.2003.10.013