Ligand Escape Pathways and (Un)Binding Free Energy Calculations for the Hexameric Insulin-Phenol Complex

Harish Vashisth and Cameron F. Abrams
Department of Chemical and Biological Engineering, Drexel University, Philadelphia, Pennsylvania

ABSTRACT Cooperative binding of phenolic species to insulin hexamers is known to stabilize pharmaceutical preparations of the hormone. Phenol dissociation is rapid on hexamer dissolution timescales, and phenol unbinding upon dilution is likely the first step in the conversion of (pharmaceutical) hexameric insulin to the active monomeric form upon injection. However, a clear understanding of the determinants of the rates of phenol unbinding remains obscure, chiefly because residues implicated in phenol exchange as determined by NMR are not all associated with likely unbinding routes suggested by the best-resolved hexamer structures. We apply random acceleration molecular dynamics simulation to identify potential escape routes of phenol from hydrophobic cavities in the hexameric insulin-phenol complex. We find three major pathways, which provide new insights into (un)binding mechanisms for phenol. We identify several residues directly participating in escape events that serve to resolve ambiguities from recent NMR experiments. Reaction coordinates for dissociation of phenol are developed based on these exit pathways. Potentials of mean force along the reaction coordinate for each pathway are resolved using multiple independent steered molecular dynamics simulations with second-order cumulant expansion of Jarzynski's equality. Our results for ΔF agree reasonably well within the range of known experimental and previous simulation magnitudes of this quantity. Based on structural analysis and energetic barriers for each pathway, we suggest a plausible preferred mechanism of phenolic exchange that differs from previous mechanisms. Several weakly-bound metastable states are also observed for the first time in the phenol dissociation reaction.

INTRODUCTION

The insulin-phenol complex is a pharmacologically important protein-ligand system. Insulin is a dual chain hormone (A-chain with 21 residues, and B-chain with 30 residues) responsible for carbohydrate metabolism and is used in the treatment of insulin-dependent (type 1) diabetes mellitus. Insulin monomers (51 residues and 5800 Da each) self-associate at physiological concentration (1 ng/ml) to form torus-shaped hexamers in the presence of zinc ions (1). Hexamers exist in three allosteric states termed $\mathbf{T}_6$, $\mathbf{T}_3\mathbf{R}_3$, and $\mathbf{R}_6$ (2–11), related by a $\mathbf{T}_6 \leftrightarrow \mathbf{T}_3\mathbf{R}_3 \leftrightarrow \mathbf{R}_6$ dynamic equilibrium, which is shifted to $\mathbf{R}_6$ only by phenolic derivatives (12,13). However, the $\mathbf{T}_3\mathbf{R}_3$ state can be achieved either by concentrated anionic medium (Cl^-, SCN^-, etc.) or phenolic species or both. Phenolic compounds act as antimicrobial agents and increase the shelf-life of industrial formulations by stabilizing the $\mathbf{R}_6$ state (14). Six hydrophobic cavities are present for the phenolic ligands in the $\mathbf{R}_6$ state and none in the $\mathbf{T}_6$ state. The conformation of the N-terminal B-chain residues (B1–B8) is extended in the **T** state and helical in the **R** state. Spectroscopic evidence (15) for preexisting **T** and **R** state hexamers suggests that binding pockets come into existence only after the $\mathbf{T} \rightarrow \mathbf{R}$ transition (16), which also means phenol binding/unbinding happens in **R**-state hexamers.

Phenolic exchange does not require hexamers to dissociate because the lifetime of **R**-state hexamers (days) is several orders-of-magnitude larger than the characteristic time of phenolic dissociation (2–5 ms) (17). Also, fast-to-intermediate (relative to the NMR timescale) exchange of phenols is surprising given that the ligand is almost completely buried in **R**-state hexamers (5). Although NMR spectroscopy predicts the existence of gatekeeper residues whose flexibility provides portals for entry/exit of phenol, the A-chain Ile^{10} being one example (18), many other such important residues remain unidentified. Additionally, unambiguous identification of residues displaying several aromatic ring flips correlated to phenol exchange remains elusive (18). The N-terminal A-chain α-helix (residues A2–A8) is speculated to facilitate entry/exit of phenol (18), but exact identities of residues whose movement in particular aids this event are also unclear. Crystallographic thermal factors (β-factors) are an indirect measure of residue mobility in proteins, but large-scale displacements observed in different residues during phenol escape are not obvious from β-factor data. Moreover, analysis of x-ray crystal structures provides limited information on whether several other entry/exit routes exist.

Given the need to understand the dissociation mechanisms of phenol from the insulin-phenol complex, molecular simulations can play an important role by providing valuable insights into dynamics of phenol release on the molecular level. The purpose of this contribution is to identify and characterize a variety of phenol binding/unbinding pathways using a judicious combination of molecular simulation techniques. The first technique, random acceleration molecular dynamics (RAMD) (19,20), is used to discover exit routes. RAMD is essentially an unbiased search for escape

Submitted June 9, 2008, and accepted for publication July 18, 2008.

Address reprint requests to Harish Vashisth, Tel.: 215-869-5978; E-mail: hl332@drexel.edu.

Editor: Gregory A. Voth.

0006-3495/08/11/4193/12 $2.00

doi: 10.1529/biophysj.108.139675

pathways. We then classify pathways into distinct reaction coordinates, and perform steered molecular dynamics (SMD) in combination with cumulant expansion of Jarzynski's equality (21) to construct free energy profiles of ligand dissociation reactions. We find three different exit pathways for the phenolic ligands and determine that a previously studied gatekeeper pathway is likely not the preferred pathway for unbinding. We also note several weakly bound (~3–4 Kcal/mol) metastable states on the surface of the insulin hexamer.

METHODS

System setup and equilibration details

All trajectories are generated using NAMDv2.5 (22) and the CHARMM force field (23). VMD Ver. 1.8.6 (24) is used for visualization purposes. Initial coordinates are taken from the x-ray crystal structure (2.0 Å) of the human $\mathbf{R}_6$-state insulin hexamer (PDB code 1ZNJ). The A- and B-type chain pairs for six monomers in the PDB file are labeled and grouped as follows: (A,B), (C,D), (E,F), (G,H), (I,J), and (K,L), respectively (Fig. 1 *b*). Three dimers in the hexamer are made up of chains (A,B,C,D), (E,F,G,H), and (I,J,K,L), respectively. The original structure has two Zn^{2+} ions, two Cl^- ions in tetrahedral coordination sites on the Zn^{2+}, seven phenols, and 331 water molecules in addition to the protein atoms. The missing threonines from C-terminus position 30 of the chains D, F, and H are added. The seventh phenol molecule situated next to one of the tetrahedrally coordinated Cl^- ions is deleted. Missing hydrogen atoms are added. The hexamer is then solvated using water molecules in a rectangular box of volume 82.46 × 82.65 × 80.90 $Å^3$ such that its dimensions exceeded those of protein by ~12 Å in all directions. Charge neutrality is ensured with random placement of 34 Na^+ and 24 Cl^- ions at 0.2 M concentration with a minimum distance of 5 Å between ions. The solvated and ionized system has 50,717 atoms. The interaction parameters for phenols are adapted directly from those of tyrosine. A schematic representation of the simulation system is given in Fig. 1 *a*.

This system is energy-minimized via 1000 cycles of conjugate-gradient optimization. To relax the solvent, a 20-ps NPT molecular dynamics simulation is run. The temperature is held at 300 K using the Langevin thermostat with damping coefficient of 5 ps^{-1}, and pressure using the Nosé-Hoover barostat. This equilibration phase is continued in the NVT ensemble for ~1 ns. In all simulations, no rigid bonds are used, periodic boundary conditions are applied, full electrostatics are computed every time step using the particle mesh Ewald method, and van der Waals interactions are cut off beyond 12 Å with the switching function taking effect at 10 Å distance. Configurations are saved every 5 ps and energylike quantities every half picosecond. The initial conditions for all RAMD and SMD trajectories are randomly sampled from the ensemble of configurations in the final ~100 ps of this equilibration simulation.

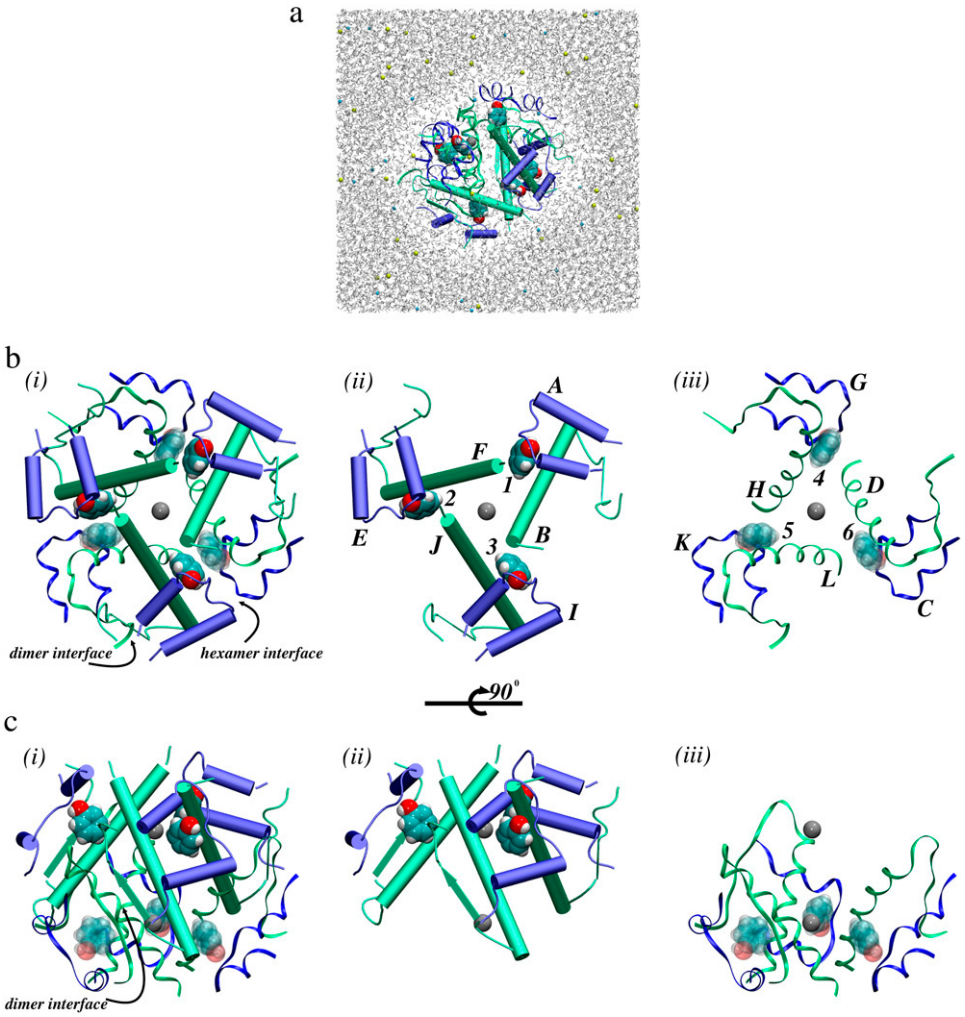

FIGURE 1 Schematic representation of the simulation system. (*a*) Simulation domain (82.46 Å × 82.65 Å × 80.90 Å) as viewed along the *z* axis: all the A- and B-type chains are colored blue and green, respectively. Six phenols are depicted with space-filling spheres. Both zincs are rendered as gray spheres. Water molecules are in wireframe representation, along with neutralizing media, sodium ions (*yellow spheres*), and chloride ions (*cyan spheres*). The system contains 50,717 atoms in all. (*b*) (*i*) Top view of R_6 insulin hexamer (PDB code 1ZNJ) is shown along threefold axis of symmetry. Hexamer and dimer interfaces are marked with arrows. Three top monomers are in cartoon representation and the remaining three, below them, are in ribbon representation. Top three phenols are darkened and lower three phenols are transparent. (*b*) (*ii*) Top view of three upper monomers. The A- and B-type chains for each are named individually as per PDB designations. Three phenols, one for each monomer, are marked as 1, 2, and 3, respectively. (*b*) (*iii*) Top view of three lower monomers. The A- and B-type chains are named accordingly. Three lower phenols are marked as 4, 5, and 6, respectively. (*c*) (*i–iii*) Side views of panel *b* show the hexamer assembly from one of the dimer interface β-sheet side (*green antiparallel arrow*).

Random acceleration molecular dynamics (RAMD) and reaction coordinates

The RAMD method of Lüdemann et al. (19) was originally applied to cytochrome P450cam to discover exit routes of camphor. The two main advantages of RAMD are 1), it speeds up dissociation kinetics, thereby making it feasible to study at nanosecond timescales; and 2), it allows unbiased pathway search through various escape routes. In RAMD, one applies randomly directed external forces to the ligand molecules following a specific protocol. First, a force direction is chosen based on a randomly oriented unit vector, $\hat{r}$, so that the external force applied to the ligand is given by

$$\vec{f}_{\text{ext}} = f_0\hat{r}, \tag{1}$$

where f_0 is the constant magnitude of the randomly chosen force. The force is maintained for a predetermined number of steps, m. The ligand is expected to maintain a certain threshold velocity $v_{\text{min}} = r_{\text{min}}/m\Delta t$ over m steps, where Δt is the time step (1 fs), and r_{min} is the specified minimum distance before a direction change. The average velocity of the ligand over the previous m steps is computed. If the ligand's velocity falls below v_{min}, a new random force vector is computed. RAMD is discontinued when the ligand has escaped. It is generally nontrivial to find suitable RAMD parameters to guarantee successful expulsions. The success probability varied between 19% and 42% depending on f_0 and m. For example, we tested m values between 10 and 100 steps and f_0 values from 5 to 20 kcal/mol·Å (Table 1). We used *tclforces* to implement RAMD in NAMD (see *tclforces* script provided as Supplementary Material in Data S1).

Each successful RAMD trajectory is approximated to a smooth curvilinear path. All such paths that lie within a cylindrical tube of radius 2.5 Å from each other are clustered into three distinct pathway classes. For each pathway class, an average path based upon all the successful trajectories of that class is constructed. A straight line having minimal standard deviation from its averaged path is determined that served as reaction coordinate in steered molecular dynamics simulations. Successful trajectories observed only with the smallest perturbing forces in RAMD simulations are used in the reaction coordinate determination.

Steered molecular dynamics (SMD) and potential of mean force (PMF) construction

Following the protocol of Jensen et al. (25), we used second-order cumulant expansion of the Jarzynski's equality (21) to compute the potential of mean force (PMF) along each reaction coordinate from work distributions obtained using constant-velocity SMD (cv-SMD) simulations. Briefly, the second-order cumulant expansion is

$$\Delta F = \langle \bar{W} \rangle - \frac{1}{2}\beta(\langle \bar{W}^2 \rangle - \langle \bar{W} \rangle^2), \tag{2}$$

where W is the work performed, $\beta = 1/k_{\text{B}}T$, k_{B} is Boltzmann's constant, T is the bulk temperature, and ΔF is the free energy difference. Bars denote averages over time windows of 20 ps, and angle brackets denote ensemble averages over independent cv-SMD trajectories. The second-order cumulant expansion is valid for Gaussian work distributions generated using cv-SMD simulations with the use of sufficiently stiff springs (26,27).

In the cv-SMD scheme, the center of mass (COM) of phenol is attached to a dummy atom via a virtual spring with spring constant k. The dummy atom is then pulled with constant velocity in a fixed direction along the reaction coordinates (RC). The distance along the pulling coordinate, λ, increases at constant velocity such that $\lambda_t = \lambda_0 + vt$, where $\lambda_0 = 0$ initially. One efficient way to sample the reaction coordinate ξ is to use high k, i.e., sufficiently stiff spring. Hence, work performed for one such trajectory is given by

$$W_{0\to t} = -kv\int_0^t [\xi - (\lambda_0 + vt)]dt. \tag{3}$$

The constant-force SMD (cf-SMD) simulations help in identifying short-lived intermediate states when the magnitude of external force is comparable to or smaller than system forces. At each such short-lived intermediate state, we then break reaction coordinates into sections S_0, S_1, etc., as given in Table 2. cv-SMD simulations are performed multiple times in each such independent section to sample efficiently that part of the RC. Overall potentials of mean force are assembled using sectional PMFs such that they match at boundaries. Performing independent pullings in each section significantly decreases nonequilibrium effects, and improves the quality of the PMFs (25).

RESULTS

Phenol escape pathways

We observed three types of escape pathways using RAMD simulations. They are denoted Pathway 1 (PW1), Pathway 2 (PW2), and Pathway 3 (PW3). PW1 and PW2 involve residues in the proximity of the hexamer interfaces as shown in Fig. 1 *b*. Three such interfaces are present in the R_6 insulin hexamer, one between each pair of dimers. PW3 mainly involves residues around the dimer interfaces visible as green antiparallel β-sheets in Fig. 1 *c*. Three dimer interfaces are also present, one between each pair of monomers forming each dimer. All pathway locations are shown in Fig. 2 *a*. Parameters and trajectory statistics for RAMD simulations are given in Table 1. The reaction coordinates along each unbinding pathway are straight lines. Three such straight lines are drawn from the phenol binding cavity to exit points on protein surface in Fig. 2 *b*. The starting coordinates for

TABLE 1 Characterization of ligand exit pathways

				PW1					PW2					PW3				
						$t_{\text{expulsion}}$ (ps)					$t_{\text{expulsion}}$ (ps)					$t_{\text{expulsion}}$ (ps)		
f_0	m	r_{min}	N_s	N_{s_1}	%	Min.	Avg.	Max.	N_{s_2}	%	Min.	Avg.	Max.	N_{s_3}	%	Min.	Avg.	Max.
20	10	0.006	84	33	39	4.80	6.00	9.30	40	48	5.70	9.20	13.1	11	13	9.30	10.9	14.0
15	10	0.005	53	18	33	11.3	13.1	14.6	30	57	9.20	12.3	15.0	5	10	9.60	12.1	14.7
10	30	0.004	45	13	29	14.4	20.3	28.4	29	64	17.6	23.3	25.0	3	7	10.8	22.2	28.5
5	100	0.030	38	6	16	60.0	113	131	32	84	54.0	89.2	147	0	0	0.00	0.00	0.00

RAMD simulation results for four different combinations of f_0 (kcal/mol·Å), m, and r_{min} (Å). N_s and % are successful trajectories and occurrence frequency, respectively. N_{s_i} denotes successful expulsions for i^{th} pathway ($i = 1$, 2, and 3 for PW1, PW2, and PW3, respectively). Minimum (Min.), average (Avg.), and maximum (Max.) expulsion times ($t_{\text{expulsion}}$) are provided. A total of 200 trajectories are carried out for each combination of f_0, m, and r_{min}.

TABLE 2 cv-SMD trajectory times in each section of the reaction coordinate for each pathway

Pathway	Section (S_i)	RC (Å)	Total simulated time in each section (ps)
PW1	S_0	0.0–2.5	1050
	S_1	2.5–4.5	900
	S_2^*	4.5–8.5	2000
	S_3	8.5–10.5	900
	S_4	10.5–11.5	600
	S_5	11.5–14.5	1500
PW2	S_0	0.0–2.0	900
	S_1	2.0–4.0	900
	S_2^*	4.0–7.0	1800
	S_3	7.0–9.0	900
	S_4	9.0–13.0	1500
	S_5	13.0–15.0	900
PW3	S_0	0.0–2.0	900
	S_1	2.0–3.5	750
	S_2	3.5–5.0	750
	S_3	5.0–7.0	900
	S_4	7.0–9.0	900
	S_5	9.0–13.0	1500
	S_6	13.0–16.0	1200

A total of three cv-SMD trajectories are carried out in each section with a velocity of 1×10^{-5} Å/fs. Asterisks indicate sections in which much slower trajectories (see article) are carried out with a velocity of 5×10^{-6} Å/fs.

three straight lines are (12.105, −12.547, and −4.271), and the unit vectors corresponding to straight lines for PW1, PW2, and PW3, are (0.645, −0.173, −0.744), (−0.338, −0.174, −0.924), and (−0.539, −0.443, 0.715), respectively. The length of reaction coordinates for PW1, PW2, and PW3, are 14.5, 15.0, and 16.0 Å, respectively. The dynamics of phenol release along the reaction coordinates for various pathways can be explicitly understood through displacements of residues lining each escape route. These root-mean-square (RMS) displacements from the crystal structure (PDB code 1ZNJ) for all the residues lining each pathway are plotted in Fig. 3. The particular phenol molecule we interrogated is marked as ''1'' and is situated in the binding pocket formed by the chains A, B, F, and H (Fig. 1 *b*).

PW1 is situated in the hexamer interfaces of the insulin-phenol complex as shown in Fig. 2 *a*. The phenol molecule in the binding pocket is surrounded by Cys^{A6}, Cys^{A7}, Thr^{A8}, Ser^{A9}, Ile^{A10}, and Cys^{A11} of the A-chain from one dimer and Val^{F2}, and His^{F5} from the adjacent dimer. While in the binding pocket, phenol forms hydrogen bonds by donating a proton to the carbonyl O atom of Cys^{A6} and accepting another from the amide N atom of Cys^{A11}. The gate of the pocket is guarded on one side by the imidazole ring of His^{F5}, and on the other side by the side chain of Ile^{A10}. A front view and a top view nearly parallel to the histidine ring of the gate are shown in Fig. 2 and Fig. 4 *a*, respectively.

During breakage of the hydrogen bonds to cysteines in the previous 4 Å of the reaction coordinate (sections S_0 and S_1 of PW1 in Table 2), the ligand significantly displaces Cys^{A6} in comparison to Cys^{A11} (Fig. 3). These hydrogen bonds are

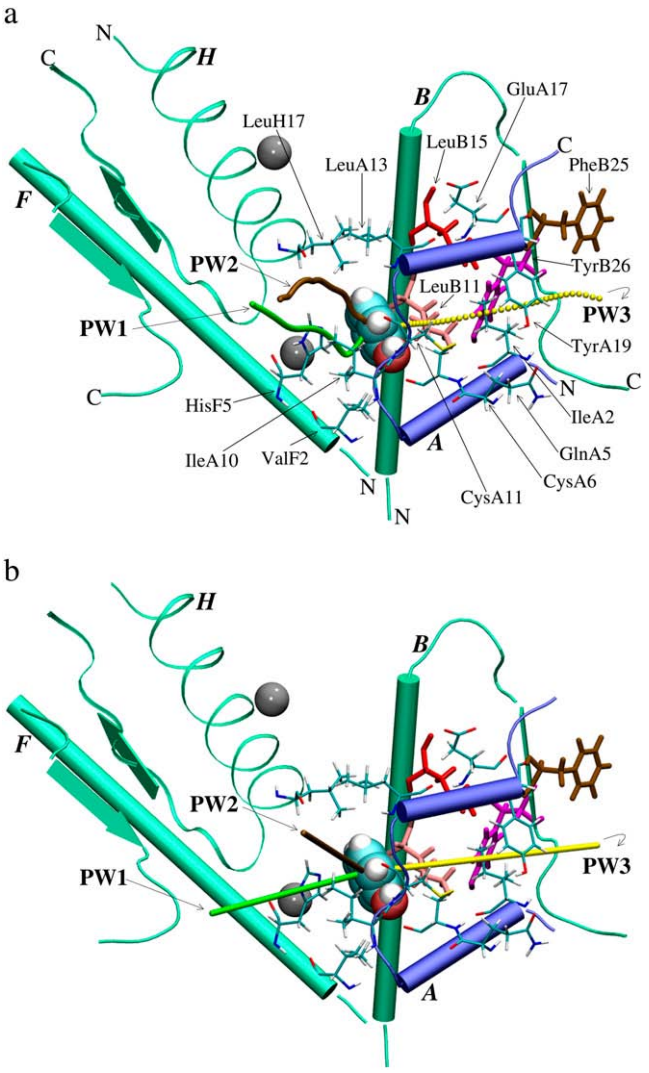

FIGURE 2 (*a*) Locations of three pathway classes determined using RAMD simulations are shown: PW1 (*green curve*), PW2 (*brown curve*), and PW3 (*yellow curve*). The A-chain is shown in blue cartoon representation, while the B-chain and F-chain are in green cartoon representations. The H-chain is in green ribbons. The terminal ends of each chain are marked by letters N and C. Phenol is rendered as space-filling. Zincs are gray spheres. The gatekeeper residues, His^{F5} and Ile^{A10}, are shown along with Val^{F2}, Leu^{H17}, Leu^{A13}, Ile^{A2}, Gln^{A5}, Cys^{A6}, Cys^{A11}, Glu^{A17}, and Tyr^{A19} (all in *licorice representation*). The B-chain residues are colored: Leu^{B11} (*pink*), Leu^{B15} (*red*), Phe^{B25} (*brown*), and Tyr^{B26} (*magenta*). (*b*) The SMD pulling directions along three pathways are rendered in same color as pathways. The length of three different reaction coordinates along three straight lines are 14.5 Å (*green*), 15.0 Å (*brown*), and 16.0 Å (*yellow*), respectively. The residue naming is the same as in panel *a*.

immediately substituted for by several water molecules (resolved crystallographically) present in the cavity, along with phenol. At this point, the aromatic ring of the ligand is in close van der Waals contact with the imidazole ring of His^{F5} and the side chain of Ile^{A10} forming the gate, where increasing deviations of these residues from their equilibrium positions are visible in Fig. 3. The imidazole ring of His^{F5} can

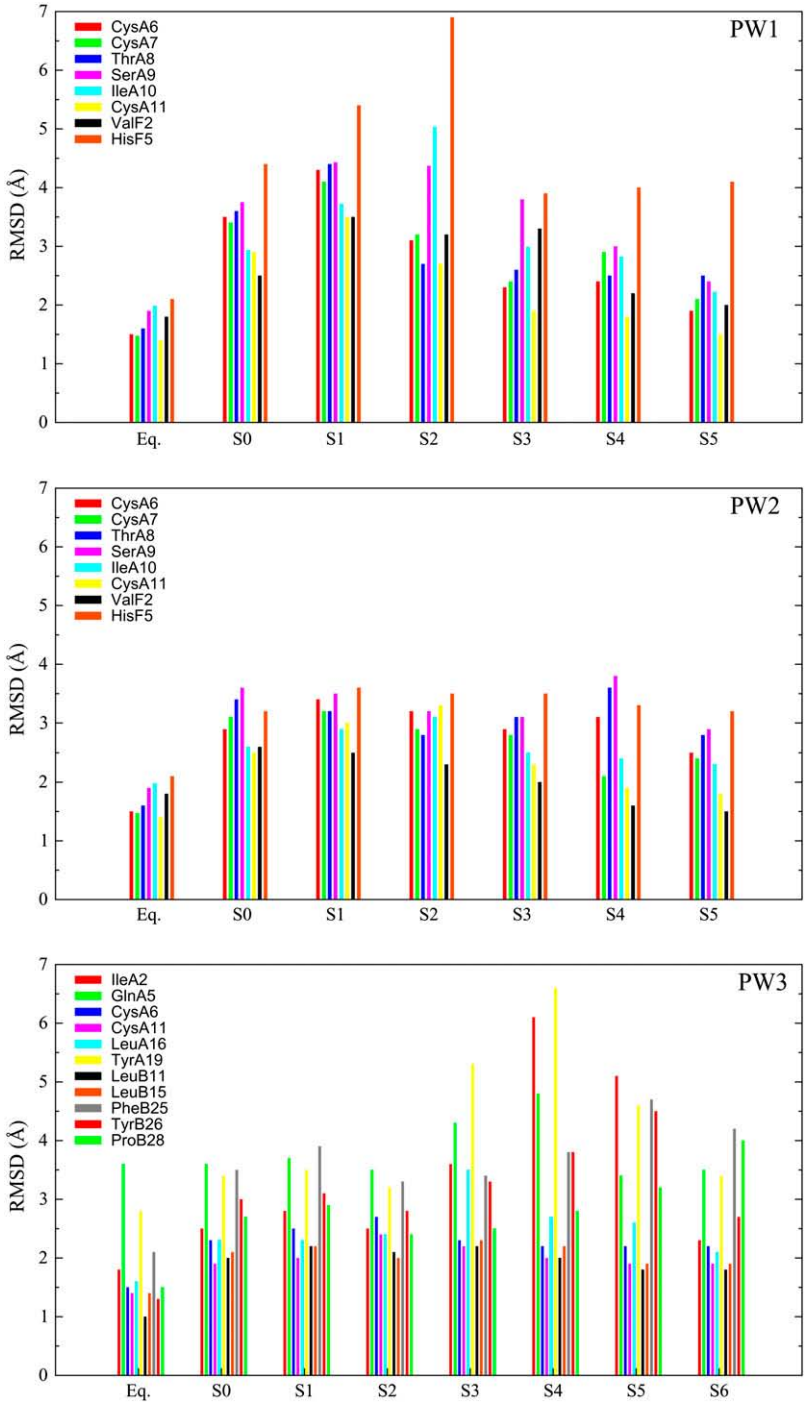

FIGURE 3 RMS displacements for all important residues (see text) for each pathway in each individual section S_i ($i = 0$–5 for PW1, PW2, and 0–6 for PW3) of the reaction coordinate (Table 2). Each residue is shown in a different color. Equilibrium fluctuations of all the residues are marked with ''Eq.'' on the abscissa. The reference x-ray structure for RMS displacement calculation is R_6 insulin hexamer (PDB code 1ZNJ).

freely rotate about C_γ of the ring, and it flips toward the solvent when it encounters an outgoing phenol, early in section S_2 (Table 2). This ring-flipping on the onset of opening the gate is captured in a snapshot, at a distance of 4.8 Å along the reaction coordinate in Fig. 4 *b*. Noticeably, the aromatic rings of phenol and His^{F5} interact parallel to each other throughout section S_2 (4.5–8.5 Å) as shown in Fig. 4, *b* and *c*. Phenol advances into the gate, pushing it wide-open by displacing the side chains of His^{F5} and Ile^{A10} significantly (Fig. 4 *c*). The pronounced RMS displacements of these two side chains, concomitant with the gate-opening mechanism, are depicted in section S_2 of PW1 in Fig. 3.

Phenol is completely out of the gate at a distance of 10.8 Å (Fig. 4 *d*). Simultaneously, concerted fluctuations in several neighboring residues of the binding pocket, located on both sides of the gate, are seen during the ligand's escape (Fig. 3). Toward the end (sections S_4 and S_5), no residue explicitly hinders the outgoing ligand although it weakly interacts hy-

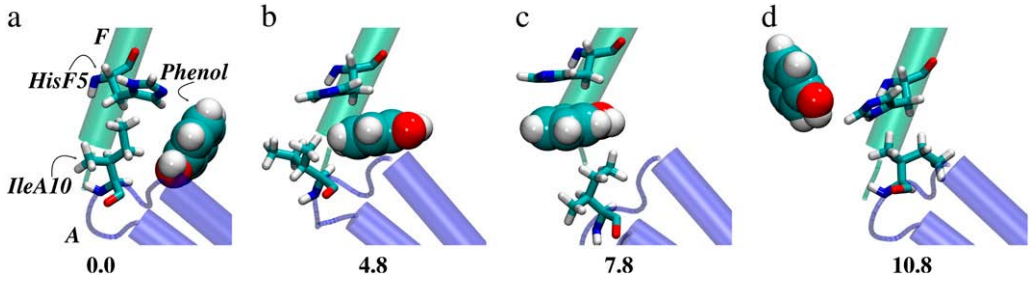

FIGURE 4 Top-view snapshots of the gate-opening mechanism (see Pathway 1 in text) at RC = 0.0, 4.8, 7.8, and 10.8 Å. The gatekeeper residues, His^{F5} and Ile^{A10}, are in licorice representation. Phenol is rendered in space-filling representation. The gate opens to the left.

drophobically with the protein. The side chain of Ile^{A10} relaxes back toward the cavity but the imidazole ring of His^{F5} remains flipped into the solvent as depicted through nearly equal RMS shifts of this residue in the last 6 Å (sections S_3, S_4, and S_5 in Fig. 3). The change in the structure of the hydrophobic pocket can also be inferred from different relaxation state of all the residues without the ligand relative to its bound state. The ligand is completely free from any interactions with the protein after 14.5 Å.

At least ~12 water molecules (seven solvent and five crystallographic) in comparison to ~2–3 molecules in the initial state, are observed within a sphere of 5 Å radius around the original COM position of the phenol in the binding pocket. The binding pockets in several $\mathbf{T}_3\mathbf{R}_3$ hexamers crystallized in the presence of concentrated anionic medium lack phenol, but possess several water molecules in its place (28,29). Similar inflow of approximately eight water molecules in the binding pocket was observed in the simulation study of phenol dissociation from the hexamer by Swegat et al. (30). ^{1}H-NMR spectroscopy experiments hint at the existence of gatekeeper residues and aromatic ring-flipping events correlated with phenolic exchange (18). Swegat et al.'s constrained molecular dynamics study of phenol dissociation suggested Ile^{A10} as one such gatekeeper residue. In PW1, we not only find involvement of Ile^{A10} but His^{F5} is also found to act as a gatekeeper residue for the ligand's escape. The unrestricted ring-flipping of His^{F5} suggested in PW1 was not observed in Swegat's study and is a new finding to be verified experimentally. However, two ring rotamers of His^{F5} in an **R**-state trimer were observed in a comparative crystal structure analysis of $\mathbf{T}_3\mathbf{R}_3$ insulin hexamers that bind *p*-hydroxybenzamide (PDB code 1ben) or methylparaben (PDB code 3mth) and are thought to have occurred during binding of these phenolic species to the **R**-state pockets (31). We argue this ring-flipping is generally important for accessing the binding pocket for any of the six ligands.

PW2 is also located in the hexamer interfaces as shown in Fig. 2 *a*. In this case, phenol encounters two new residues, Leu^{H17} and Leu^{A13}, in addition to all the residues mentioned in PW1. These leucines are involved in a close van der Waals contact in the hexamer interfaces (Fig. 2 *a*), which helps bind adjacent dimers in the insulin-phenol complex. A narrow open channel exists in the space between the gatekeeper residues (see above) and both the leucines in the hexamer interface. Front and side views of this channel are shown in Fig. 2 *a* and Fig. 5 *a* (RC = 0.0 Å), respectively.

The hydroxyl group of phenol is accompanied to the entrance of this channel by ~3–4 water molecules (also resolved in the crystal), where solvent molecules replace crystallographic water for the hydrogen bonding, in first 4 Å of escape dynamics (sections S_0 and S_1 for PW2 in Table 2). Contrary to PW1, where the gate opening is due to the aromatic ring of phenol, the ligand prefers to enter the narrow channel here with the hydroxyl group first, after considerable

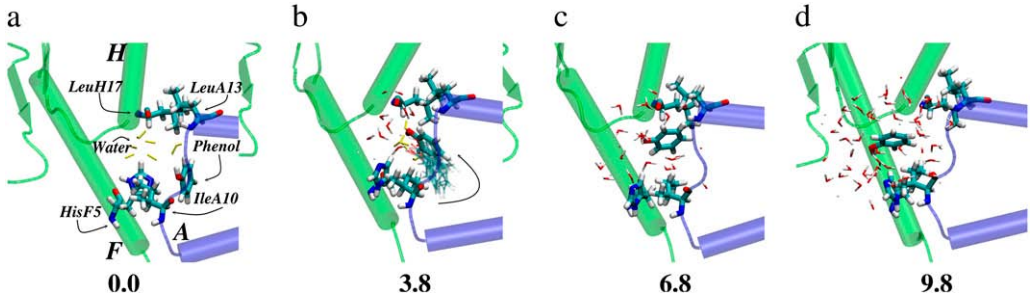

FIGURE 5 Side-view snapshots of the gate-leaping mechanism (see Pathway 2 in text) at RC = 0.0, 3.8, 6.8, and 9.8 Å. Narrow channel between gatekeeper residues and both leucines, Leu^{A13} and Leu^{H17}, is visible. Crystallographic and solvent water molecules within 5 Å of phenol are rendered in yellow and red licorice representations, respectively. All residues and phenol are also in licorice representation as indicated in panel *a*. The arrow in panel *b* shows still frames of reorienting ligand at every 15 ps in section S_1.

reorientation as represented in the snapshot in Fig. 5 *b* (RC = 3.8 Å). Phenol continues to steer through the channel making a smooth leap over the gate in the next 3 Å (section S_2 for PW2 in Table 2).

In this gate-leaping mechanism, the RMS fluctuations of His^{F5} and Ile^{A10} (the gatekeeper residues) are relatively smaller ($RC = 4.0$–7.0 Å) in comparison to the gate-opening mechanism (notice sections S_2 of PW1 and PW2 in Fig. 3). While making the leap, the phenolic ring interacts with the imidazole ring of His^{F5}, and the fluctuations of His^{F5} can be mainly attributed to the semiflipped state of this ring (Fig. 5, *c* and *d*) in this region. At RC = 9.8 Å, phenol has nearly exited this narrow opening, where it is surrounded by a cluster of solvent water molecules, shown in Fig. 5 *d*. It is not until RC = 15.0 Å, that the ligand ceases to interact with the protein. The RMS displacements for all other residues common with PW1 are depicted in Fig. 3. Only approximately six water molecules are observed in the cavity (within a sphere of 5 Å radius) after the ligand has dissociated. To a greater extent, decreased inflow of water is a result of closed-gate conformation of the gatekeeper residues, because of which the binding pocket has not significantly expanded in comparison to PW1. One such channel, similar to PW2, which extends from the binding site to the surface of **R**-state trimer, was suggested in a comparative crystal structure analysis of $\mathbf{T}_3\mathbf{R}_3$ insulin hexamers (31). Moreover, this pathway is observed in the highest frequency among all the exit routes discovered using RAMD simulations (Table 1).

PW3 is completely unrelated to the previous pathway classes because the ligand exits through the dimer interface (Fig. 1 *c*). The exact location of PW3 can be appreciated from the RAMD expulsion trajectory and the SMD pulling direction indicated in Fig. 2. The escape route passes through a sterically dense region in the core of the insulin monomer made up of the chains A and B, and is primarily lined by residues Ile^{A2}, Gln^{A5}, Cys^{A6}, Cys^{A11}, Leu^{A16}, Tyr^{A19}, Leu^{B11}, Leu^{B15}, Phe^{B25}, Tyr^{B26}, and Pro^{B28}. The first four of these residues belong to the lower N-terminal α-helix of the A-chain, the next two to the upper C-terminal α-helix of same chain, the following two to the N-terminal α-helix of the B-chain, and the last three to the dimer interface β-sheet region of the B-chain. Mostly, the A-chain residues hinder the ligand's passage in the beginning, while the B-chain residues play a role in the latter stages of escape because these residues actually participate in formation of the dimer interfaces.

At early stages of escape in section S_0 and S_1 of the reaction coordinate (Table 2), the hydroxyl group of phenol repeatedly switches hydrogen bonds between crystallographic water molecules and neighboring cysteine residues. The side chain of Leu^{B11} hinders motion of the ligand in these regions of the reaction coordinate experiencing a significant RMS displacement (Fig. 3). Comparison of crystal structures has shown previously that this side chain restricts the size and shape of phenol binding cavity in $\mathbf{R}_6$ species to be significantly smaller, and less open than $\mathbf{T}_3\mathbf{R}_3$, and is proposed to be naturally displaced in case the larger allosteric ligands such as 2,6- and 2,7-dihydroxynaphthalene have to occupy the cavity (32).

The aromatic ring of the ligand chiefly interacts through hydrophobic and van der Waals interactions with the side chains of Leu^{A16} and Leu^{B15}, and phenol moves through section S_2 with its aromatic ring oriented toward the monomer interior, possibly maintaining a weak hydrogen bond with amide proton of Cys^{A11} through its hydroxyl group. The location of these leucines (see above) does not permit them to block phenol escape directly, although small RMS shifts in these side chains are observed (Fig. 3). The phenol encounters mainly Ile^{A2} and Tyr^{A19} (Fig. 2 *a*) during its passage from sections S_3 and S_4. The tyrosine residue is exposed to the solvent but the aromatic ring is oriented toward the subunit interior, and it rotates about C_γ to flip back to the solvent region on its first collision with the escaping ligand. In the course of inducing ring flip, phenol also significantly displaces the side chain of Ile^{A2} (Fig. 3). Gln^{A5}, situated in the vicinity of Ile^{A2}, simultaneously displays a smaller displacement. Throughout section S_5, phenol weakly interacts with Ile^{A2}, Tyr^{A19}, and Tyr^{B26}, chiefly through its aromatic ring, and finally leaves with the hydroxyl group into the solvent.

At this stage, the ligand eventually finds itself in the unblocked open space between the N-terminus of the A-chain and the antiparallel β-sheet of the dimer interface. It mainly displaces the side chain of Pro^{B28} in section S_6 whereas the aromatic parts of phenol and Phe^{B25} keep interacting weakly until the end of this section. Phenol is completely dissociated beyond the reaction coordinate distances of 16 Å with the exit point located in proximity of the β-sheet region.

Aromatic ring rotation of Tyr^{A19} is related to phenol exchange, as suggested by NMR studies, and speculation exists for structural fluctuations in the N-terminal A-chain α-helix (residues A2–A8) facilitating entry/exit of phenol (18). This pathway indeed involves mainly two residues, Ile^{A2} and Gln^{A5}, which are part of this domain. More importantly, the mutation of Glu^{A17} to Gln^{A17}, which eliminates salt-bridging interactions between Arg^{B22} and Glu^{A17}, has been shown to have negligible effects on the dissociation constants for the ligand binding to phenolic pockets (33). We indeed observe that Glu^{A17} does not interact with phenol directly or indirectly in any of our pathways, which is consistent with the mutagenesis studies (33).

Thermodynamics of exit pathways

We performed both cf-SMD and cv-SMD simulations to quantify the free energy change of the ligand dissociation reaction. In the cf-SMD scheme, a constant force is applied to the COM of phenol to pull it along the respective RC for each pathway. Forces of magnitude 280 pN, 105 pN, and 350 pN were used for PW1, PW2, and PW3, respectively. The length of trajectory for each pathway was 700 ps, 1000 ps, and 400 ps, respectively. The dwelling positions for COM of phenol

along the reaction coordinate for each pathway are identified from cf-SMD simulations. Each such dwelling position serves to mark sections (e.g., S_0, S_1, etc., in Table 2) for subsequent cv-SMD simulations along the RC for each pathway.

In the subsequent cv-SMD simulations, the COM of phenol was constrained with a harmonic spring of stiffness 8 kcal/mol·Å^2 in the beginning of each section, while the system underwent 200 steps of energy minimization, and 100 ps equilibration. cv-SMD simulations were conducted in five independent sections of the RC for PW1 and PW2, and six independent sections for PW3 (Table 2). Three such trajectories in each independent section of the RC, for each pathway, were performed (Table 2). We chose a velocity of 1×10^{-5} Å/fs, and spring constant k of 5 kcal/mol·Å^2 to ensure the stiff spring approximation was valid. The velocities used in this study are much slower in comparison to several other similar studies with cv-SMD (25,34). In some sections (S_2 of PW1 and PW2), even slower velocities (5×10^{-6} Å/fs) were used to decrease nonequilibrium effects (see Discussion). The external work, W, of gradually decoupling the ligand from protein pocket was computed from several independent cv-SMD simulations in each section of the reaction coordinate (Table 2), and these microscopic work values were used to estimate ΔF using second-order cumulant expansion of Jarzynski's equality (Eq. 2). The COM position of the ligand in each pulling window, unbinding forces, and combined potential of mean force (henceforth denoted by PMF) for PW1, PW2, and PW3, are plotted in Fig. 6, Fig. 7, and Fig. 8, respectively. Overall trajectory statistics for the cv-SMD simulations are given in Table 2.

The intermediate events during the course of phenol release along three pathways can be qualitatively appreciated from their respective unbinding cv-SMD force profiles (panel *b* in Figs. 6–8). In the beginning of unbinding along each pathway, phenol is in the bound state with no external force acting on it. Early in the pull, negative cv-SMD forces indicate the larger magnitude of thermal fluctuations over the external force. Phenol escape is triggered soon after this initial drop, visible as an increasing external force dominance over system forces that try to reinstate the initial state. The steep increase in force for each pathway after this reflects the breakage of hydrogen bonds between phenol and disulfide cysteines. The continued increase in force until reaching maximum reflects displacements of side chains of residues lining each pathway (Fig. 3) that hinder unbinding of phenol, and the magnitude of this maximal force measures in some way the difficulty in unbinding for a particular pathway. It should be mentioned here that the total work done in dissociating phenol along a particular pathway mostly arises from the external force after it reaches its maximum, because in the earlier regime, the ligand has moved a considerably smaller

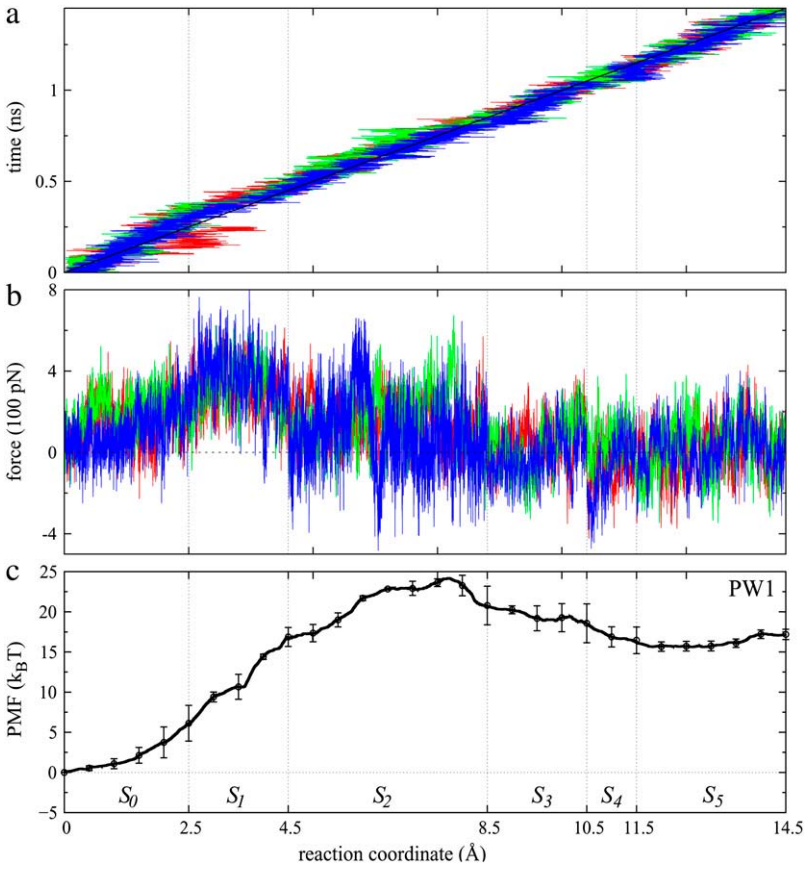

FIGURE 6 SMD simulation results for PW1 in each independent section of the reaction coordinate (*green line* in Fig. 2 *b* and Table 2). (*a*) The COM trajectory of phenol. (*b*) Unbinding forces. (*c*) Combined potential of mean forces such that they match at boundaries. Three colors (*red*, *blue*, and *green*), in panels *a* and *b*, represent data from three independent SMD simulations in each section (Table 2). Error bars for each independent section are also plotted. Vertical dotted lines represent boundaries of each independent section in which unidirectional pulling was performed.

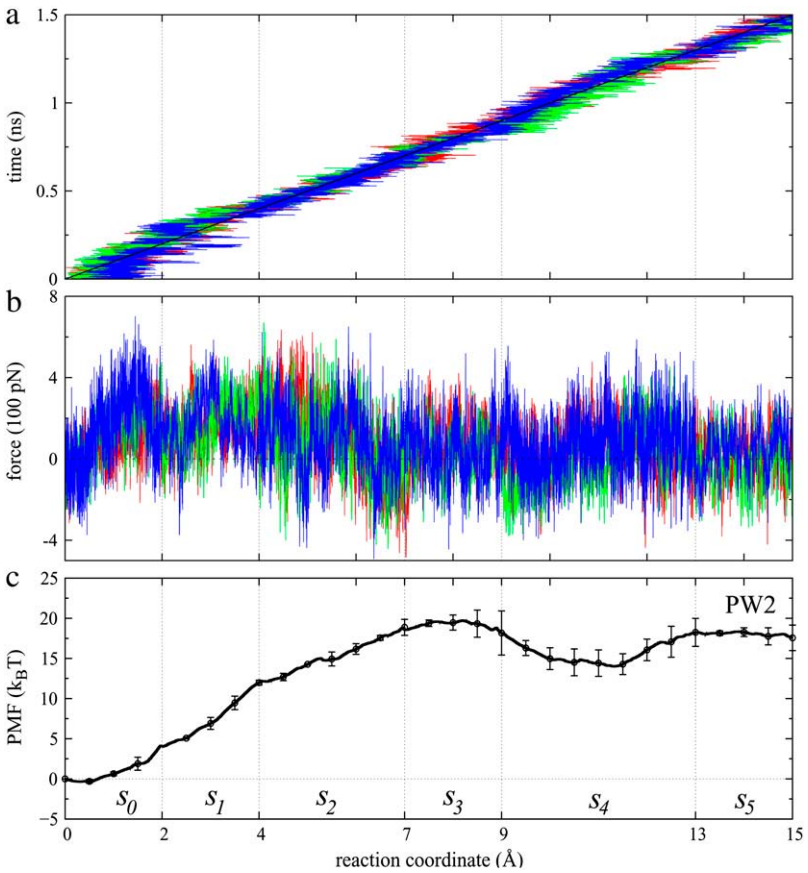

FIGURE 7 SMD simulation results for PW2 in each independent section of the reaction coordinate (*brown line* in Fig. 2 *b* and Table 2). See caption of Fig. 6 for details.

distance along the reaction coordinate for each pathway. The unbinding forces decrease further as the ligand moves away from the binding pocket. The forces fluctuate about a mean drag force when the ligand is out of the protein in the dissociated state. The dissociated state for each pathway is achieved at 14.5, 15.0, and 16.0 Å, respectively (Figs. 6–8).

RAMD simulations provide limited information on the likelihood of occurrence of each pathway, but the PMF (panel *c* in Figs. 6–8) along the chosen RC (Fig. 2 *b*) can help us gain more quantitative insights into the relative likelihoods of the several (un)binding mechanisms for phenol. It is crucial to minimize the fluctuations of the RC among different trajectories before constructing the PMF, and in this case, the COM of phenol closely follows the constraint center (λ) in all pathways, ensuring that the stiff-spring approximation is valid (panel *a* in Figs. 6–8). Minimizing fluctuations in the reaction coordinate allows us to sample efficiently the contributions to the PMF arising from the region around λ. In the respective S_0 sections, the PMFs for PW1 and PW2 are not very different from each other as all involve breakage of hydrogen bonds to Cys^{A6} and Cys^{A11} but significant deviation between both hexamer interface pathways can be seen in the latter stages of the unbinding process. The steeper PMF for PW1 in section S_1, compared to PW2, is due to close van der Waals contact of the ligand with the gatekeeper residues (see above). Section S_2 in PW1 is correlated to the gate-opening mechanism (Fig. 4), while the same section corresponds to the gate-leaping mechanism (Fig. 5) for PW2, as described earlier. The lower slope of the PMF is evident in this region for PW2 as opposed to PW1. This also makes physical sense because it is easier for the ligand to pass through the narrow channel instead of opening the gate, which in turn results in relatively higher unbinding forces for PW1. The rotational degrees of freedom of the ligand (Fig. 5 *b*) allow it to find a lower free energy path in this region by exiting the hydroxyl group first. In the dimer interface pathway (PW3 in Fig. 2), phenol meets densely packed residues in the monomer core, where steep departure of the PMF in the very early sections of the RC (S_0, S_1, and S_2) in comparison to hexamer interface pathways is seen. The unbinding forces for this pathway are the highest in magnitude (Fig. 8), and observance frequency are the least (Table 1). The maximum barrier heights with respect to initial state are measured to be 23.63 $\pm$ 0.50 k_BT (PW1), 19.33 $\pm$ 1.60 k_BT (PW2), and 26.83 $\pm$ 1.10 k_BT (PW3), at the respective reaction coordinate distances of 7.5, 8.5, and 6.0 Å. In comparison, using an equilibrium free energy method, Swegat et al. (30) calculated the barrier-height as $\sim$14.2 k_BT at 300 K.

Although the ligand has crossed the gate at RC = 10.8 Å (Fig. 4 *d*), it interacts with the flipped imidazole ring of His^{F5}

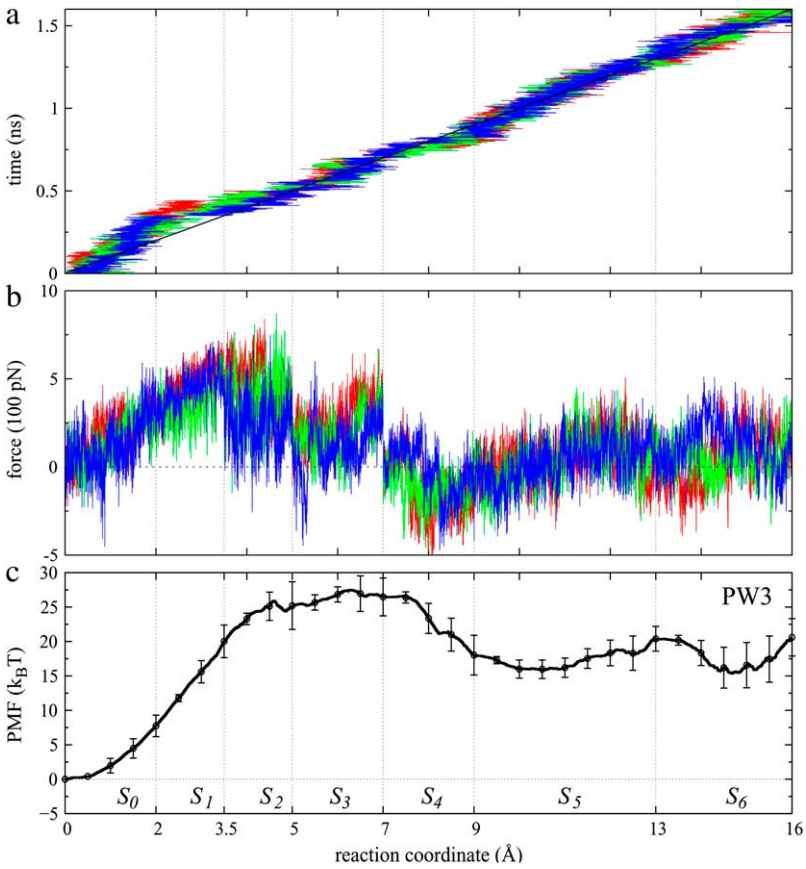

FIGURE 8 SMD simulation results for PW3 in each independent section of the reaction coordinate (*yellow line* in Fig. 2 *b* and Table 2). See caption of Fig. 6 for details.

for the next 4–5 Å, related to which a weakly bound state in this part of RC is visible (S_4, and S_5 for PW1). The stable intermediate state for PW2 corresponds to hydrophobic interactions between phenol, Leu^{H17}, Leu^{A13}, and Ile^{A10}, on the verge of crossing the narrow channel (Fig. 5 *d*). Two distinctly stable intermediate states are observed for PW3, where the first corresponds to residual interactions with Ile^{A2}, Tyr^{A19}, and Tyr^{B26}, and the second to interactions with aromatic ring of Phe^{B25}, in the β-sheet region of the dimer interface. Interestingly, several additional m-cresol binding sites are found in the crystal structures of fast-acting mutant ($Pro^{B28} \rightarrow Asp$) $\mathbf{R}_6$ insulin hexamers, in the vicinity of Tyr^{B26}, and Ile^{A2}, similar to what we observe in PW3 for native insulin hexamer (35).

The PMFs can also be used to extract the free energy change (ΔF) of this dissociation reaction. The dissociated states of the ligand are nearly identical in free energy for three pathways, as can be inferred from the individual PMF ending in the same region within error bars. The free energy difference between initial ($RC = 0.0$ Å) and the end states at 14.5 Å for PW1, 15.0 Å for PW2, and 16.0 Å for PW3, are computed to be 18.06 ± 0.60 k_BT, 17.75 ± 1.06 k_BT, and 20.60 ± 2.70 k_BT, respectively. A net positive free energy change for this unbinding reaction indicates that the ligand prefers to be in the bound state at equilibrium. The estimates of free energy change of phenol dissociation reaction show considerable variation among various experiments due to inherent limitations of each technique used in such measurements (16,36). An experimental study using isothermal titrating calorimetry reported ΔF to be ~39.24 k_BT (36), while another one with circular-dichroism spectroscopy estimated it to be ~7.22 k_BT (16). The modeling of binding curves in circular-dichroism spectroscopy (16) is suggested to be oversimplified because of a serious assumption that free ligand concentration is equal to total ligand concentration at saturation, which is found not to be the case (36). Hierarchical modeling of a complex set of phenol binding curves in isothermal titrating calorimetry experiments (36), however, is limited by the fact that the resultant heat produced is a sum of several processes (ligand binding, conformational changes, and release of water molecules) that are difficult to isolate. Different allosteric states of the insulin hexamer in solution and multiple interdependent binding cavities for phenols further complicate measurements of free energy change. It should also be noted that binding of each ligand in these experiments represents the simultaneous sum of all possible bound ligands in solution as opposed to unique atomistic picture of binding/unbinding of a single phenolic ligand presented in this work. Free energy computations in this study are used as a quantitative tool to suggest the most likely (un)binding mechanism out of three possible

escape pathways for phenolic species and should not be overinterpreted as the true free energy change of this reaction in the absence of unambiguous experimental measurements.

DISCUSSION

An important aspect of this study is the observation of different escape pathways for phenolic ligands from the insulin-phenol complex using RAMD. The knowledge of multiple exit routes provides new insights into the dissociation mechanisms of phenol and certainly complements experimental findings. As an example, the structural analysis of pathways found several aromatic ring flips involved in the ligand's release along multiple pathways, some of which are suggested by experimental studies (17,18), and thought to be related to phenol exchange. We identify such ring flips in His^{F5} (PW1) and Tyr^{A19} (PW3) during the ligand's escape. The kinetics of phenol (un)binding is not experimentally accessible yet (30), and rates of dissociation computed by using various kinetic models can depend upon the way ligand dissociates along any of multiple pathways. Due to the lack of such knowledge, a preference for a particular pathway is not given in a hierarchical modeling study of phenol binding (36), but we consider it to be important for improving and generalizing any kinetic model.

Free energy calculations were performed using piecewise cv-SMD simulations (25) which force the unbinding of phenol by pulling it gradually toward a defined point on the protein surface. Three distinct reaction coordinates for PMF construction, one for each pathway, were determined based on knowledge of exit routes from RAMD. The PMF for each pathway reveal quantitative information on the likely mechanism of phenol dissociation from the insulin-phenol complex. Given three competing dissociation mechanisms, the structural features and the energy barriers computed from the PMF's data suggest that the ligands preferentially exchange via the gate-leaping mechanism (Fig. 5) through a narrow channel along Pathway 2. Consistent with gate-leaping mechanism, we observed convergence toward PW2 on lowering the perturbing force in RAMD simulations as well (Table 1). It is likely that the choice of only one reaction coordinate limited Swegat's previous simulation of phenol dissociation (30) from finding alternate and more frequent dissociation mechanisms such as gate-leaping.

The events such as gate-opening and gate-leaping for passage across a narrow channel induce strong nonequilibrium effects on the system (25). The sensitivity of the PMF to such effects necessitates carrying them out in the quasi-equilibrium regime. Therefore, we carried out additional cv-SMD trajectories with much slower pulling velocities (5×10^{-6} Å/fs) to compensate for nonequilibrium effects in the PMF calculations (section S_2 for PW1, and PW2, in Table 2). A limited number of such slow trajectories (three in each section in our case) give a reasonably good resolution of the PMF constructed using the second-order cumulant expansion of Jarzynski's equality. The second-order cumulant expansion is suggested previously to provide an improved estimate of the PMF in comparison to exponential averaging using Jarzynski's equality given fewer slow SMD trajectories (27). We indeed observe an improved resolution of the PMF constructed using the cumulant expansion in multiple independent sections (Fig. S1). We therefore resort to the second-order cumulant expansion of Jarzynski's equality for the PMFs construction as used in similar studies (25,34).

CONCLUSIONS

Phenolic preservatives are used in pharmaceutical preparations of insulin, where the hexamer stability is ensured by noncovalent association of these ligands to six hydrophobic cavities buried inside the protein. We used random acceleration molecular dynamics (RAMD) (19) to identify escape routes of phenol from the hexameric insulin-phenol complex. Three different escape pathways, two through the hexamer interface, and one through the dimer interface, were observed. Several events relating to opening of entry/exit portals for phenol exchange were suggested by experiments, such as flipping of aromatic rings in various residues. We are the first to observe such flipping motions exclusively in two residues, His^{F5} (PW1) and Tyr^{A19} (PW3). We also observed the direct involvement of several other residues (Fig. 3) lining each pathway. Structural examination identified a gatekeeper residue Ile^{A10} in PW1 suggested previously by Swegat et al. (30), along with His^{F5} as a new gatekeeper residue. The α-helix of the A-chain (residues A2–A8) is thought to play a role in phenolic exchange (18), and we are the first to quantify important fluctuations in two residues, Ile^{A2} and Gln^{A5}, situated in this domain, that correlate with phenol unbinding. We hope future experimental studies will corroborate these observations.

Previous simulations of this complex identified only one phenol escape route (30). We argue instead for the existence of several competing dissociation mechanisms corresponding to different escape routes. More importantly, the energetics of phenol dissociation along each pathway was characterized through the PMFs computed with the help of second-order cumulant expansion of Jarzynski's equality and cv-SMD simulations. The energy barriers suggest the gate-leaping to be a preferred mechanism for phenolic exchange. Additionally, the PMFs show the existence of weakly bound metastable states along each pathway that were not observed in the previous simulation study of this complex (30) but are nevertheless supported by recent experimental evidence (35).

SUPPLEMENTARY MATERIAL

To view all of the supplemental files associated with this article, visit www.biophysj.org.

We are grateful to the National Science Foundation for support through grant No. CBET-0544933.

REFERENCES

1. Ferrari, D., J. R. Diers, D. F. Bocian, N. C. Kaarsholm, and M. F. Dunn. 2001. Raman signatures of ligand binding and allosteric conformation change in hexameric insulin. *Biopolymers.* 62:249–260.
2. Blundell, T. L., G. G. Dodson, D. C. Hodgkin, and D. Mercola. 1972. Insulin: the structure in the crystal and its reflection in chemistry and biology. *Adv. Protein Chem.* 26:279–402.
3. Bentley, G. A., E. J. Dodson, G. G. Dodson, D. C. Hodgkin, and D. Mercola. 1976. Structure of insulin in 4-zinc insulin. *Nature.* 261:166–168.
4. Baker, E. N., T. L. Blundell, J. F. Cutfield, S. M. Cutfield, G. G. Dodson, D. M. C. Hodgkin, R. E. Hubbard, M. W. Isaacs, C. D. Reynolds, K. Sakabe, N. Sakabe, and N. M. Vijayan. 1988. The structure of 2Zn pig insulin crystals at 1.5 Å resolution. *Phil. Tran. Roy. Soc. B.* 319:369–456.
5. Derewenda, U., Z. Derewenda, E. J. Dodson, G. G. Dodson, C. D. Reynolds, G. D. Smith, C. Sparks, and D. Sweson. 1989. Phenol stabilizes more helix in a new symmetrical zinc insulin hexamer. *Nature.* 338:594–596.
6. Smith, G. D., and G. G. Dodson. 1992. Structure of a rhombohedral R_6 insulin/phenol complex. *Proteins Struct. Funct. Genet.* 14:401–408.
7. Smith, G. D., E. Ciszak, and W. Pangborn. 1996. A novel complex of a phenolic derivative with insulin: Structural features related to the T↔R transition. *Protein Sci.* 5:1502–1511.
8. Smith, G. D., and R. H. Blessing. 2003. Lessons from an aged, dried crystal of T_6 human insulin. *Acta Crystallogr. D Biol. Crystallogr.* D59:1384–1394.
9. Williamson, K. L., and R. J. P. Williams. 1979. Conformational analysis by nuclear magnetic resonance: insulin. *Biochemistry.* 18:5966–5972.
10. Chang, X., A. M. M. Jorgensen, P. Bardrum, and J. J. Led. 1997. Solution structures of the R_6 human insulin hexamer. *Biochemistry.* 36:9409–9422.
11. O'Donoghue, S. I., X. Chang, R. Abseher, M. Nilges, and J. J. Led. 2000. Unraveling the symmetry ambiguity in a hexamer: calculation of the R_6 human insulin structure. *J. Biomol. NMR.* 16:93–108.
12. Kaarsholm, N. C., H. C. Ko, and M. F. Dunn. 1989. Comparison of solution structural flexibility and zinc binding domains for insulin, proinsulin, and miniproinsulin. *Biochemistry.* 28:4427–4435.
13. Roy, M., M. L. Brader, R. W.-K. Lee, N. C. Kaarsholm, J. F. Hansen, and M. F. Dunn. 1989. Spectroscopic signatures of the T to R conformational transition in the insulin hexamer. *J. Mol. Biol.* 264:19081–19085.
14. Berchtold, H., and R. Hilgenfeld. 1999. Binding of phenol to R_6 insulin hexamers. *Biopolymers.* 51:165–172.
15. Choi, W. E., D. Borchardt, N. C. Kaarsholm, P. S. Brzovic, and M. F. Dunn. 1996. Spectroscopic evidence for preexisting T- and R-state insulin hexamer conformations. *Proteins Struct. Funct. Genet.* 26:377–390.
16. Jacoby, E., P. Krüger, Y. Karatas, and A. Wollmer. 1993. Distinction of structural reorganization and ligand binding in the T↔R transition of insulin on the basis of allosteric models. *Biol. Chem. Hoppe Seyler.* 374:877–885.
17. Hassiepen, U., M. Federwisch, T. Mulders, and A. Wollmer. 1999. The lifetime of insulin hexamers. *Biophys. J.* 77:1638–1654.
18. Jacoby, E., Q. X. Hua, A. S. Stern, B. H. Frank, and M. A. Weiss. 1996. Structure and dynamics of a protein assembly. ^{1}H-NMR studies of 36 kDa R_6 insulin hexamer. *J. Mol. Biol.* 258:136–157.
19. Lüdemann, S., V. Lounnas, and R. C. Wade. 2000. How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *J. Mol. Biol.* 303:797–811.
20. Carlsson, P., S. Burendahl, and L. Nilsson. 2006. Unbinding of retinoic acid from the retinoic acid receptor by random expulsion molecular dynamics. *Biophys. J.* 91:3151–3161.
21. Jarzynski, C. 1997. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* 78:2690–2693.
22. Kalé, L., R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, and K. Schulten. 1999. NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* 151:283–312.
23. MacKerell, Jr., A. D., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.
24. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD—visual molecular dynamics. *J. Mol. Graph.* 14:33–38.
25. Jensen, M. Ø., S. Park, E. Tajkhorshid, and K. Schulten. 2002. Energetics of glycerol conduction through aquaglyceroporin GlpF. *Proc. Natl. Acad. Sci. USA.* 99:6731–6736.
26. Park, S., F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten. 2003. Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.* 119:3559–3566.
27. Park, S., and K. Schulten. 2004. Calculating potentials of mean force from steered molecular dynamics simulations. *J. Chem. Phys.* 120: 5946–5961.
28. Whittingham, J. L., S. Chaudhuri, E. J. Dodson, P. C. E. Moody, and G. G. Dodson. 1995. X-ray crystallographic studies on hexameric insulins in the presence of helix-stabilizing agents, thiocyanate, methylparaben, and phenol. *Biochemistry.* 34:15553–15563.
29. Ciszak, E., and G. D. Smith. 1994. Crystallographic evidence for dual coordination around zinc in the T_3R_3 human insulin hexamer. *Biochemistry.* 33:1512–1517.
30. Swegat, W., J. Schlitter, P. Krüger, and A. Wollmer. 2003. MD simulation of protein-ligand interaction: formation and dissociation of an insulin-phenol complex. *Biophys. J.* 84:1493–1506.
31. Smith, G. D. 1998. The phenolic binding site in $T_3R_3^f$ insulin. *J. Mol. Struct.* 470:71–80.
32. Bloom, C. R., R. Heymann, N. C. Kaarsholm, and M. F. Dunn. 1997. Binding of 2,6- and 2,7-dihydroxynaphthalene to wild-type and E-B13Q insulins: dynamic, equilibrium, and molecular modeling investigations. *Biochemistry.* 36:12746–12758.
33. Bloom, C. R., N. Wu, A. Dunn, N. C. Kaarsholm, and M. F. Dunn. 1998. Comparison of the allosteric properties of the Co(II) and Zn(II) substituted insulin hexamers. *Biochemistry.* 37:10937–10944.
34. Amaro, R., E. Tajkhorshid, and Z. Luthey-Schulten. 2003. Developing an energy landscape for the novel function of a $(\beta/\alpha)_8$ barrel: ammonia conduction through HisF. *Proc. Natl. Acad. Sci. USA.* 100:7599–7604.
35. Whittingham, J. L., D. J. Edwards, A. A. Antson, J. M. Clarkson, and G. G. Dodson. 1998. Interactions of phenol and m-cresol in the insulin hexamer, and their effect on the association properties of B28 Pro→Asp insulin analogues. *Biochemistry.* 37:11516–11523.
36. Birnbaum, D. T., S. W. Dodd, B. E. H. Saxberg, A. D. Varshavsky, and J. M. Beals. 1996. Hierarchical modeling of phenolic ligand binding to 2Zn-insulin hexamers. *Biochemistry.* 35:5366–5378.