Article

Peptide Binding to a PDZ Domain by Electrostatic Steering via Nonnative Salt Bridges

Nicolas Blöchliger,[1] Min Xu,[1] and Amedeo Caflisch[1,*]
[1]Department of Biochemistry, University of Zurich, Zurich, Switzerland

ABSTRACT We have captured the binding of a peptide to a PDZ domain by unbiased molecular dynamics simulations. Analysis of the trajectories reveals on-pathway encounter complex formation, which is driven by electrostatic interactions between negatively charged carboxylate groups in the peptide and positively charged side chains surrounding the binding site. In contrast, the final stereospecific complex, which matches the crystal structure, features completely different interactions, namely the burial of the hydrophobic side chain of the peptide C-terminal residue and backbone hydrogen bonds. The simulations show that nonnative salt bridges stabilize kinetically the encounter complex during binding. Unbinding follows the inverse sequence of events with the same nonnative salt bridges in the encounter complex. Thus, in contrast to protein folding, which is driven by native interactions, the binding of charged peptides can be steered by nonnative interactions, which might be a general mechanism, e.g., in the recognition of histone tails by bromodomains.

INTRODUCTION

The fundamental process of protein-protein binding can be conceptualized as diffusional association followed by formation of the stereospecific complex (1–3). Long-range electrostatic forces can significantly accelerate and guide diffusional association, a phenomenon termed electrostatic steering (1,3–9). Association results in a relatively weak encounter complex, which is stabilized mainly by nonspecific interactions and whose binding interface is not yet fully desolvated (1,2,7,10,11). Crossing the transition state, potentially through multiple pathways (12), specific short-range hydrogen bonds and hydrophobic interactions form and the stereospecific complex is reached (1).

PDZ (PSD-95/Discs large/ZO-1) domains, which are found in scaffold proteins involved in signaling (13–15), have been used as model systems to study peptide binding (16,17). They share a common fold with six β-strands and two α-helices (Fig. 1 *A*) and mainly interact with target proteins by binding their C-termini (13,18), although binding to internal protein segments has been reported as well (19). In the stereospecific complex the side chain of a hydrophobic residue at the C-terminus of the target is buried and its carboxylate group interacts with the carboxylate-binding loop. In addition, backbone hydrogen bonds create an intermolecular β-sheet, and specificity is achieved by side-chain interactions (13,20,21).

In a previous molecular dynamics (MD) study on the third PDZ domain of the postsynaptic density protein 95, we focused the analysis on the binding site of the PDZ domain by comparing MD runs of the apo structure of the PDZ domain with MD runs started from the bound state (22). This comparison suggested that the peptide binds by conformational selection. No peptide dissociation event was observed because the length of each of the four trajectories started from the bound state was <0.2 μs. Several other MD simulations of PDZ domains have been performed during the last few years (23–27). However, we are not aware of any MD simulations of the binding of peptides to PDZ domains. Although unbiased MD simulations of sub-μs length have been used already to study the (reversible) binding of small and mainly rigid molecules to proteins (28–32), it is much more challenging to simulate the binding of flexible (oligo)peptides to proteins because the larger conformational space requires significantly longer trajectories (9,12,33).

Here, we report on unbiased, multiple MD simulations of 2.1–3.6 μs each, which were carried out to characterize the binding and unbinding of the C-terminal hexapeptide segment Acetyl-EQVSAV of the Ras-associating guanine nucleotide exchange factor 2 (RA-GEF2, also known as PDZ-GEF2 or RapGEF6) (34) to the second PDZ domain of protein tyrosine phosphatase 1E (PTP1E, also known as PTPL1, FAP-1, or PTP-Bas) (35,36). This study focuses on the intermolecular interactions during the (un)binding process and was motivated by the following questions. Is it possible to capture the spontaneous binding of a flexible hexapeptide to the PDZ domain by MD simulations on a conventional compute cluster? What is the role of the electrostatic interactions in the initial association and final binding? Does the binding proceed through native interactions, i.e., via the intermolecular contacts of the stereospecific complex as observed in the crystal structure? Is unbinding the reverse of binding?

Submitted February 20, 2015, and accepted for publication March 17, 2015.

*Correspondence: caflisch@bioc.uzh.ch

Editor: Rohit Pappu.

0006-3495/15/05/2362/9 $2.00

http://dx.doi.org/10.1016/j.bpj.2015.03.038

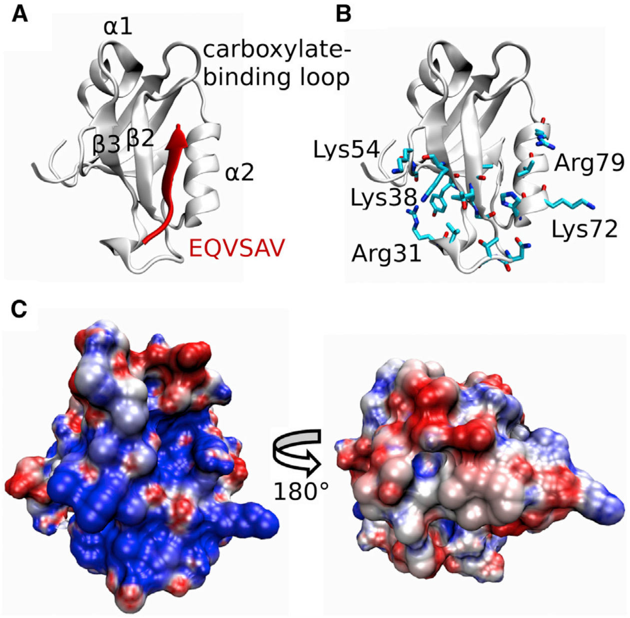

FIGURE 1 Initial association of the peptide and the PDZ domain by electrostatic steering. (*A*) Crystal structure of the stereospecific complex of PTP1E PDZ2 and the C-terminal RA-GEF2 peptide (PDB code 3LNY). The PDZ domain is shown in white with some secondary structure elements labeled. The peptide and its sequence are in red. (*B*) Surface of initial association. PDZ residues having an average contact frequency with the peptide residues of at least 0.1 during association are shown in a stick-like representation (see Materials and Methods). The structure and the orientation are the same as in (*A*). The contact frequency values are shown in Fig. S1. (*C*) Electrostatic surface potential. The color scale ranges from −5 kT/e (*red*) to 5 kT/e (*blue*). The orientation is the same as in (*A*) for the left panel. All illustrations were rendered with VMD (85). To see this figure in color, go online.

We observed multiple events of spontaneous binding in 10 MD runs of ~2.3 μs each started from fully unbound, and several rebinding events in 10 MD runs of ~3.4 μs each started from the crystal structure of the complex. Fast initial association is always driven and stabilized by long-range electrostatic interactions between negatively charged carboxylate groups in the peptide and positively charged side chains in the vicinity of the binding site. These salt bridges are not present in the final stereospecific complex.

MATERIALS AND METHODS

MD simulations

We carried out 10 independent simulations starting with the peptide placed randomly in the simulation box (called binding runs in the following) and 10 simulations started from the bound state (called unbinding runs). The total simulation length amounted to 57 μs.

The coordinates of PTP1E PDZ2 in complex with the C-terminal RA-GEF2 peptide were downloaded from the protein database (Protein Data Bank (PDB) code 3LNY, URL www.rcsb.org) (37). The sequence of the C-terminal RE-GEF2 peptide used here is EQVSAV, and its N-terminus was capped with acetyl. To reproduce neutral pH conditions, the side chains of aspartates and glutamates were negatively charged, those of lysines and arginines were positively charged, and the histidine side chains were neutral. The structure was solvated in a cubic water box. For the binding runs the peptide was placed randomly in the box with a resulting mean distance to the PDZ domain of 12 Å. The size of the box was 73 Å for the binding runs and 63 Å for the unbinding runs. The simulation system contained sodium and chloride ions to approximate an ionic strength of 150 mM and to compensate for the total charge of the two molecules. The simulations were carried out with GROMACS 4.5.5 (38) using the CHARMM27 force field (39,40) and the TIP3P water model (41). Periodic boundary conditions were applied, and electrostatic interactions were evaluated using the particle-mesh Ewald summation method (42). The van der Waals interactions were truncated at a cutoff of 10 Å. The temperature of 310 K was kept constant by an external bath with velocity rescaling (43), and the pressure was kept close to 1 atm by the Berendsen barostat (44). The LINCS algorithm was used to fix the covalent bonds involving hydrogen atoms (45). The integration time step was 2 fs, and snapshots were saved every 10 ps. Each MD run was carried out on 16 cores (i.e., four Xeon5560 CPUs) of the Schrödinger supercomputer at the University of Zurich, which required ~1 week per μs.

SAPPHIRE plot

Recently, we have developed an algorithm for the analysis of long MD trajectories (46,47). The resulting SAPPHIRE (States And Pathways Projected with HIgh REsolution) plot is a comprehensive visualization of the thermodynamics and kinetics of the simulated system. A function measuring distance between snapshots is needed to generate SAPPHIRE plots and can be freely chosen by the user. We chose the Euclidean distance function on 29 distances between atoms of the peptide and the binding site of the PDZ domain for the present application. Table S1 in the Supporting Material contains the full list of atom pairs used.

We briefly describe the method here and refer the reader to the original publications for more details (46,47). Starting from an arbitrary snapshot, all the snapshots are sequentially ordered in a stepwise fashion. In each step, the snapshot closest to any snapshot prior in the sequence becomes the next entry. The complete sequence of snapshots is called progress index. Assuming high snapshot density within free energy basins, snapshots belonging to the same basin are grouped together and distinct states do not overlap (46). A stochastic algorithm to generate an approximate progress index has been developed. This algorithm is scalable to large data sets and was used here. It is important to note that the progress index is not a reaction coordinate. It is rather a sorting of all MD snapshots to identify basins without any a priori clustering.

We employ three types of annotation functions to highlight and interpret the states along the progress index and the pathways connecting them (Fig. 2). First, we use a kinetic annotation function to localize the individual states on the progress index. Specifically, for every snapshot i along the progress index, we plot the average of the mean first-passage times between A_i and S_i, denoted τ_{MFP}, where A_i is the set of snapshots added to the progress index before i and S_i is the set of those added after i. The value of this annotation function is low within a state and high in transition regions, and barriers are highlighted reliably (although they cannot be interpreted quantitatively) (46). Second, we plot the actual sampling time of the individual snapshots to illustrate when and in which sequence the different states were sampled. This information appears as red dots in Fig. 2 and corresponds to the trace of the temporal evolution of the system, i.e., the detailed sequence of events for each MD run. Third, we characterize the states themselves by a structural annotation. In this case we have used the distance between the peptide and the PDZ domain, the solvent accessible surface area of the peptide, the root mean-square deviation (RMSD) of the peptide with respect to a reference structure after alignment on the PDZ domain, as well as several interatomic distances.

Trajectories from the individual simulation runs were concatenated and subsampled at 20 ps to generate the SAPHHIRE plot. For the unbinding runs, the size of the simulation box was adjusted to match the binding runs after the system has been centered on the PDZ atoms. The stochastic

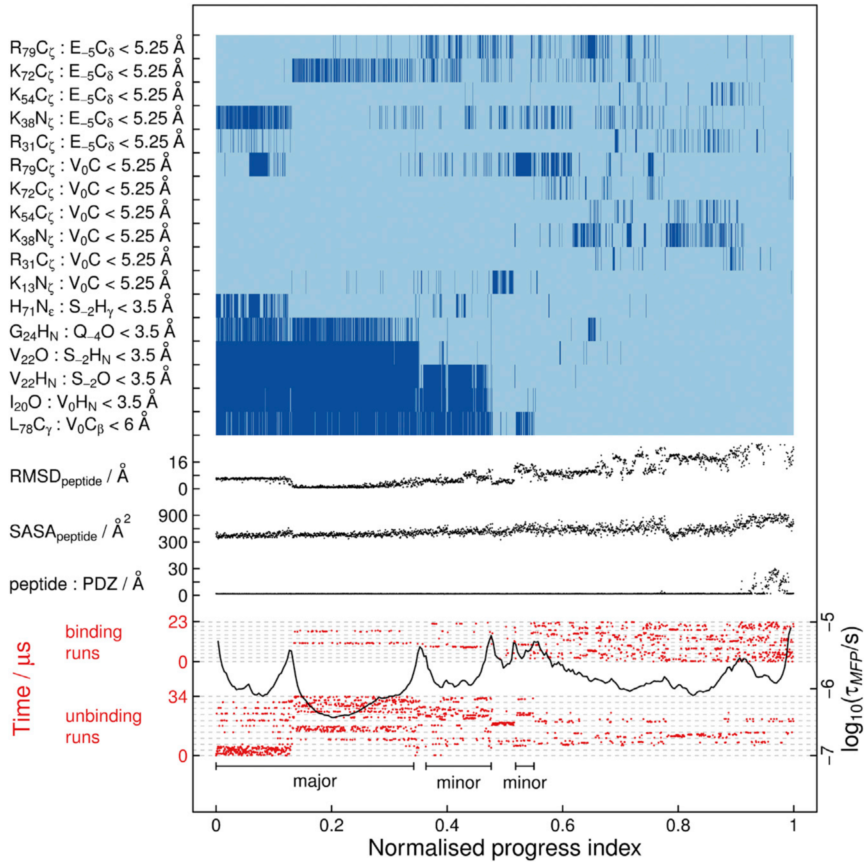

FIGURE 2 SAPPHIRE plot illustrating the sampled ensemble. The progress index (x axis) represents a reordering of the trajectory snapshots that groups similar snapshots next to each other (see Materials and Methods). The progress index is annotated with kinetic information (τ_{MFP}, a function whose value is low within states and high in transition regions, *black profile* in the *bottom*), sampling time (*red dots*), and structural information (*middle* and *top*). The annotation in the top part of the panel uses binning, with dark blue meaning that the distance given on the left side (reporting on burial of the Val0 side chain, intermolecular backbone hydrogen bonds, and salt bridges) is below the indicated threshold. $RMSD_{peptide}$ was computed on the C_α atoms of the peptide with respect to a representative structure of the major binding mode after alignment on the PDZ domain. $SASA_{peptide}$ is the solvent accessible surface area of the peptide. Peptide: PDZ denotes the minimal distance between the peptide and the PDZ domain. Gray dashed lines indicate the boundaries between individual simulation runs. The major and minor binding modes of the stereospecific complex are labeled (*black horizontal segments* in the *bottom*). To see this figure in color, go online.

algorithm mentioned previously is scalable because of the preorganization of the data via tree-based, hierarchical clustering (48). The lower and upper threshold radius and the tree height for the clustering were set to 0.6 Å, 10 Å, and 12. The first snapshot on the progress index is the starting structure of the first run, i.e., the crystal structure of the bound complex (49). The number of guesses to find nearest neighbors (46) was set to 10^4. The method is implemented in the CAMPARI simulation and analysis package (http://campari.sourceforge.net).

Contact frequencies

First, the MD trajectory segments were classified as stereospecific complex or other, where other includes fully unbound and encounter complex. For this binary classification the kinetic annotation of the SAPPHIRE plot was used (*black profile* in Fig. 2), as well as the RMSD of the peptide with respect to the major binding mode, and various distances between the peptide and the binding pocket (Figs. S2–S21). This classification is illustrated in the top of Figs. S2–S21. The 10 binding runs were then used to compute contact frequencies by employing only the segments annotated as other (i.e., fully unbound and encounter complex). A contact is considered to be formed between a residue of the peptide and a residue of the PDZ domain if two atoms are within 5 Å. The acetyl at the N-terminus of the peptide was considered as an independent residue, and CAMPARI (http://campari.sourceforge.net) was used for this analysis.

Electrostatic surface potential

The electrostatic potential on the surface of the PDZ domain was calculated with PDB2PQR (50,51) and APBS (52) using the conformation of the PDZ domain in the stereospecific complex (PDB entry 3LNY).

Binding time and k_{on}

Mean binding times were separately estimated for the binding runs and the rebinding events observed in the unbinding runs as $\tau = t_{unbound}/n$, where $t_{unbound}$ is the total time the peptide is not bound as in the stereospecific complex (defined previously and annotated in Figs. S2–S21) and n is the number of binding events. For the binding runs $n = 5$ and $\tau = 3.8$ μs. We observed $n = 3$ rebinding events in the unbinding runs, resulting in $\tau = 2.4$ μs. The rate constant k_{on} was estimated to be $1/\tau$ [peptide], where [peptide] $= 4.3$ mM and 6.7 mM for the binding and unbinding runs, respectively. The resulting values for k_{on} are 61 $\mu M^{-1}s^{-1}$ and 63 $\mu M^{-1}s^{-1}$ for the binding and unbinding runs, respectively.

Free energy profile

Cut-based (53,54) and conventional, histogram-based free energy profiles were computed using Fep1d (55).

RESULTS AND DISCUSSION

We performed 10 simulations starting from the peptide placed randomly in the simulation box. These 10 simulations (called binding runs in the following) were completely agnostic of the binding site, and no biasing force or restraint was used. In addition 10 independent runs were started from the bound state using the crystal structure of the complex (PDB code 3LNY (49)) as starting conformation and different random seeds for the initial assignment of the velocities. Total simulation time amounted to 57 μs.

Association via electrostatic steering and nonnative salt bridges

We observed fast association of the peptide and the PDZ domain in all of the binding runs. The intermolecular distance dropped below 2.5 Å within 5 ns on average. The residues most involved in complex formation are Val22, Thr23, Gly24, His71, Val75, and Arg79 located along the binding site, Asn27, Thr28, Val30, Arg31, Tyr36, and Lys38 in the β2-β3 loop and on strand β3, as well as Lys54, Gly55, and Lys72 (Figs. 1 *B* and S1 and Materials and Methods). On the other hand, contact frequencies are low for the carboxylate binding loop, for helix α1, which contains two negatively charged residues, and the β-sheet formed by the β1, β6, β4, and β5 strands, which is located on the other side of the domain with respect to the binding site. The electrostatic potential on this surface of initial association is positively charged (Fig. 1 *C*), and diffusion of the peptide, which bears two negative charges, to the vicinity of the binding site is thus mainly driven by electrostatic steering.

Various salt bridges are formed in the encounter complex, which features multiple relative orientations of the peptide and PDZ domain. In the fifth binding run, for example, the carboxylate group of Val0 (peptide residues are numbered from −5 to 0) forms salt bridges with Arg79, Lys13, and Lys72 before committing to the final binding pose (Fig. 3 and Movie S1). The detailed sequence of events and the roles played by the individual charged residues are different in the other simulation runs (Figs. S2–S21), illustrating the heterogeneity of the encounter complex and the lack of specific interactions (10,11,56–58). However, in all of the binding runs salt bridges are dynamically formed in the encounter complex.

The solvent accessible surface area of the peptide is larger in the encounter complex than in the stereospecific complex (Figs. 3 and S2–S21). This indicates that the binding interface of the encounter complex is not fully desolvated and that specific intermolecular hydrogen bonds or hydrophobic interactions are of lesser importance (10,56,57).

Final binding in antiparallel β-sheet arrangement

We have recently developed a method for the visualization of long MD trajectories (46,47). The main output of the method is the SAPPHIRE plot, which offers an intuitive illustration of the states and sequence of events encountered during the simulation (see Materials and Methods). Previously, we used SAPPHIRE plots to analyze protein folding and conformational changes in the native state of a protein (47) as well as multiple conformations of a loop of the prion protein (59). Here, we apply the method to a binding process.

The SAPPHIRE plot of the combined binding and unbinding runs (Fig. 2) shows that the major binding mode is stabilized by the canonical burial of the Val0 side chain (formed contact with side chain of Leu78 in the binding pocket). In addition, the Val0 carboxylate group interacts with the carboxylate-binding loop (Fig. 3 *A*) and intermolecular backbone hydrogen bonds are formed between the carbonyl

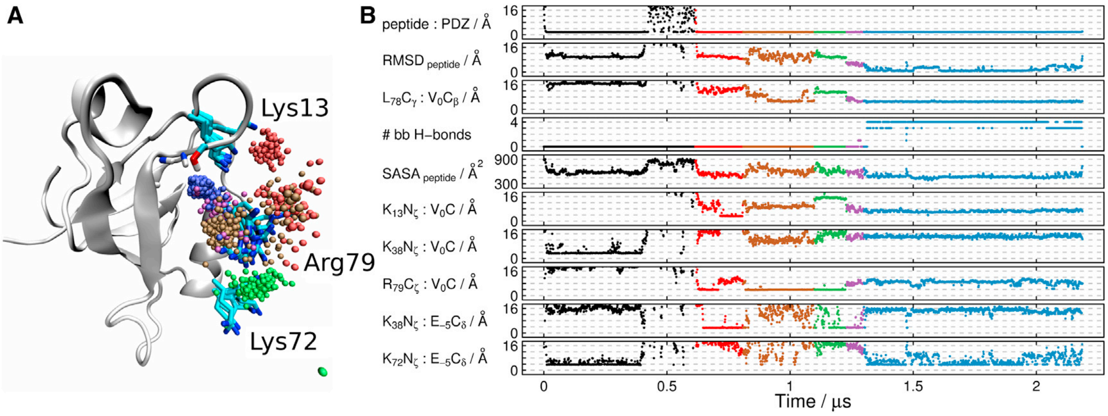

FIGURE 3 Salt bridges stabilize the encounter complex. (*A*) Nonspecific salt bridges in the encounter complex during binding in the fifth binding run (Movie S1). A representative conformation of the PDZ domain in the major binding mode is shown along with the backbone amide groups of the carboxylate-binding loop and the side chain of Ser17. A sphere is drawn every 1 ns at the positon of the carbon atom of the Val0 carboxylate group during binding and colored according to time, as indicated in (*B*). The side chains of selected basic residues involved in salt bridges with the carboxylate group of Val0 are drawn every 150 ns. The illustration was rendered with VMD (85). (*B*) Analysis of the fifth binding run. The minimal distance between the peptide and the PDZ domain (peptide: PDZ), the RMSD of the peptide C_α atoms after alignment on the PDZ domain, the number of backbone hydrogen bonds formed (i.e., distances between the carbonyl oxygen of Ile20 and the NH group of Val0, between both polar groups of Val22 and Ser-2, and between the NH group of Gly24 and the carbonyl oxygen of Gln-4 below 3.5 Å), the solvent accessible surface area of the peptide ($SASA_{peptide}$), and distances between selected atom pairs are plotted as median values in a window of 1 ns. Corresponding plots for the other simulation runs are given in Figs. S2–S21. To see this figure in color, go online.

oxygen of Ile20 and the NH group of Val0, between both polar groups of Val22 and Ser-2, and between the NH group of Gly24 and the carbonyl oxygen of Gln-4 (Fig. 2). Of importance, the most populated binding mode is essentially identical to the crystal structure. The barrier at a value of the normalized progress index of ~0.15 is due to reorientation of the Glu-5 side chain, which can either point toward the solvent or form a salt bridge with Lys72. Regarding the crystal structure, note that the atoms of the Glu-5 residue had very high B-factors and the side chain did not show any electron density (49). Furthermore, the peptide used by Zhang et al. is slightly longer than the one we simulated and has an additional charged residue (Glu-7), which is likely to affect, at least in part, the orientation of the N-terminal segment of the peptide in the bound conformation.

A minor binding mode is located between normalized progress index values of ~0.36 and ~0.48 (Fig. 2). In this binding mode the C-terminal part of the peptide is bound as in the crystal structure, whereas the N-terminal segment protrudes into the solvent. Only the two backbone hydrogen bonds toward the C-terminus of the peptide (between the carbonyl oxygen of Ile20 and the NH group of Val0 and between the NH group of Val22 and the carbonyl oxygen of Ser-2) are formed, in agreement with recent experimental results obtained by amide-to-ester mutations (57). Another alternative binding mode, which is short-lived and was repeatedly sampled, is located between normalized progress index values of ~0.52 and ~0.55. This binding mode features burial of the Val0 side chain (as in the crystal structure), whereas the Val0 carboxylate group forms a salt bridge with Arg79 instead of interacting with the carboxylate-binding loop and no intermolecular backbone hydrogen bonds are present. Finally, snapshots representing the encounter complex are found between normalized progress index values of ~0.55 and ~0.9. Fully unbound conformations accumulate at the end of the progress index.

The stereospecific complex (major or minor binding modes) was reached in five out of the 10 binding runs (Figs. 2, 3, and S2–S11). Additionally, the peptide rebound in three of the six unbinding runs in which full dissociation was observed (Figs. S12–S21 and Movie S2). Our estimate for k_{on} based on these eight binding events is ~60 $\mu M^{-1}s^{-1}$ (see Materials and Methods). We note that the TIP3P water model used here shows a self-diffusion constant higher by a factor of 2–3 than the experimentally measured value (60), which might influence k_{on}. Experimental values for k_{on} collected at lower temperatures and similar or higher ionic strength range from 2.9 to 36 $\mu M^{-1}s^{-1}$ for the same PDZ domain or its mouse ortholog PTP-BL PDZ2 and the peptide ENEQVSAV or *dansyl*-EQVSAV (49,57,61–63).

The dissociation of the encounter complex is frequent on the timescale of binding in our simulations as the average lifetime of the encounter complex is ~200 ns (see distance between peptide and PDZ domain in Figs. S2–S21). The encounter complex is thus located before the rate-limiting step (2,56,57). This observation is validated by the free energy profile along the distance between the Val0 side chain and the hydrophobic pocket of the PDZ domain (Fig. 4), which confirms that the main barrier accounts for the burial of the Val0 side chain. Comparing Fig. 3 with the corresponding figures for the other simulation runs (Figs. S2–S21) shows that the stereospecific complex can be reached from the encounter complex via various pathways. The burial of the Val0 side chain takes place before the formation of the backbone hydrogen bonds or almost simultaneously (e.g., in the binding run 8, Fig. S19). Thus, the sequence of events for binding starts with the formation of nonnative salt bridges in the encounter complex (which does not always lead to full binding) followed by burial of the Val0 side chain, and formation of the backbone hydrogen bonds between residues Val0/Ser-2 and the PDZ β2 strand in an antiparallel β-sheet arrangement.

Inverse sequence of events during unbinding

It is interesting to analyze the unbinding process and compare with binding. Peptide dissociation starts by the rupture of the backbone hydrogen bonds, which takes place before the Val0 side chain exits from the hydrophobic pocket of the PDZ domain. Thus, the initial events of unbinding are the reverse of the final events of binding. Furthermore, the peptide does not immediately diffuse away from the PDZ domain after the native interactions of the stereospecific complex break apart. Instead, the peptide remains in contact with the PDZ domain for several hundred nanoseconds (Movie S2 and Figs. S4, S5, S7–S9, and S21). Quantitatively, the residence time in the encounter complex is 650 ± 900 ns during unbinding and 200 ± 300 ns during binding. Of importance, the same nonnative salt bridges provide kinetic stabilization to the encounter complex during

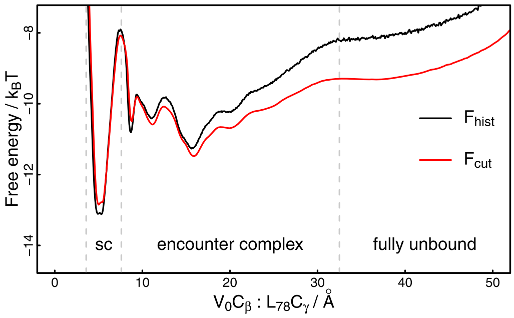

FIGURE 4 Free energy profile along a geometric order parameter. Histogram-based (*black*) and cut-based (*red*) (53,54) free energy profiles are shown as a function of the distance between the C_β atom of Val0 and the C_γ atom of Leu78, which reports on burial of the Val0 side chain. Barriers separating the stereospecific complex (sc), the encounter complex, and fully unbound conformations are indicated by gray, dashed lines. Note that this simple projection introduces overlap and hides crucial information, which, in contrast, is fully resolved by the SAPPHIRE plot (Fig. 2), e.g., the presence of minor binding modes. To see this figure in color, go online.

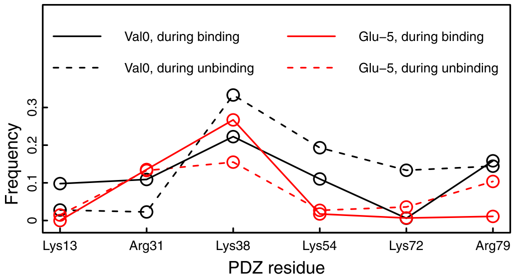

FIGURE 5 Salt bridges in the encounter complex during binding and unbinding. The trajectory segments that correspond to the encounter complex during binding and unbinding were extracted based on Figs. S2–S21, and salt bridges were considered to be formed if the N_ζ atom (for the PDZ lysines) or the C_δ atom (for the PDZ arginines) was within 6 Å of the carboxylate carbon of Val0 or Glu-5, respectively. To see this figure in color, go online.

both peptide association and dissociation (Fig. 5). Thus, the sequence of events for full dissociation is the reverse of binding.

CONCLUSIONS

We have used unbiased MD simulations to analyze the binding of the C-terminal hexapeptide segment of a natural ligand to the second PDZ domain of PTP1E. The general view of the binding process is schematically depicted in Fig. 6 and a representative binding event is shown in Movie S1. Initial association is driven by the long-range electrostatic interactions between the peptide and the PDZ domain (Fig. 1 *C*). In the resulting encounter complex the peptide is weakly bound in the vicinity of the binding site (Figs. 1 *B* and 3). The complex is maintained by nonspecific electrostatic interactions, which allows the peptide to sample multiple orientations (Fig. 3). After the rate-limiting step the side chain of Val0 is buried in a hydrophobic pocket (Figs. 2 and 3). At this point, up to four backbone hydrogen bonds between the peptide and β2 can form depending on whether the major binding mode is reached directly or via distinct minor binding modes (Fig. 2). The comparison of the sequence of events for binding and unbinding shows that the two processes are one the inverse of the other.

To further investigate the influence of the encounter complex on the rate constant for binding, we suggest to measure experimentally the salt dependence of the binding rate, e.g., by the Förster resonance energy transfer technique. These measurements have already been reported for PTP-BL PDZ2 and a dansylated peptide without any charged side chains (64). Whereas k_{off} was independent of the ionic strength, k_{on} decreased with increasing ionic strength, which was attributed to the negative charge of the C-terminal carboxylate group. A stronger influence on k_{on} is predicted, on the basis of our MD simulation results, for a similar peptide with one or two negatively charged side chains. On the other hand, electrostatic steering has been ruled out for binding of a peptide with no net charge (*dansyl*-KQTSV) to PDZ3 of postsynaptic density protein 95 (which has glutamic acids at the positions of Arg31 and Lys72) (65).

The mechanism of initial association guided by nonspecific electrostatic steering is likely to be valid for other (small, single-domain) peptide-binding proteins (4,7,9,66,67). As an example, the binding of histone tails to bromodomains is most probably driven by the negative electrostatic potential on the surface surrounding the acetylated lysine binding site, whereas the final stereospecific complex is stabilized by the hydrogen bond between the acetyl carbonyl and the side chain of the evolutionary conserved Asn (68). Other examples include the binding of phosphorylated peptides to SH2 domains (12) as well as intrinsically disordered proteins (69), which tend to contain more charged residues than globular proteins (70). Regarding the coupled binding and folding of intrinsically disordered proteins (71), experimental and theoretical (72,73) studies have highlighted nonnative salt bridges in the encounter complex (74), enhanced on-rates due to electrostatic interactions (75–77), nonnative steering (78), and late formation of native contacts (79).

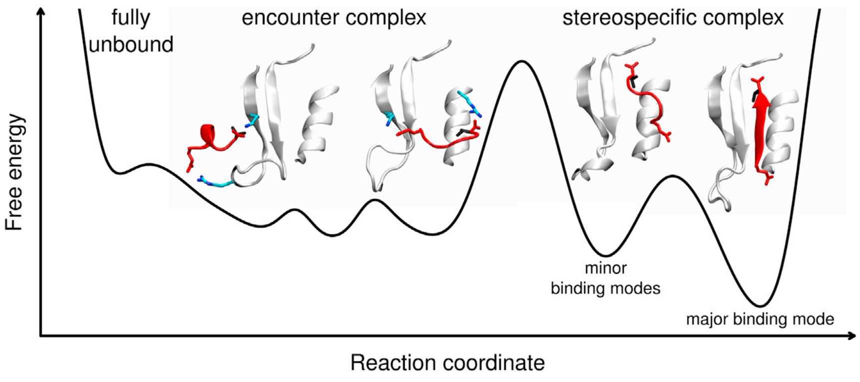

FIGURE 6 Schematic free energy profile of the binding process. After association accelerated by electrostatic steering a weak encounter complex is formed, which is stabilized by nonspecific intermolecular salt bridges. In contrast, the stereospecific complex features burial of the Val0 side chain and multiple binding modes differing among each other only in the orientation of the N-terminal part of the peptide. In this qualitative illustration, relative barrier heights roughly reflect the kinetics observed in the MD simulations, i.e., fast formation, reconfiguration and dissociation of the encounter complex, interconversions among major and minor binding modes on an intermediate timescale, and slow transitions between encounter complex and stereospecific complex. (*Insets*) The ribbon illustrations focus on the binding site, i.e., only the following structural elements are shown for clarity: Carboxylate-binding loop, β2-β3 strands and loop, and helix α2 of the PDZ domain (*gray*), backbone of the peptide, C-terminal carboxylate group, and the Glu-5 side (*red*), Val0 side chain (*black*), and the side chains of basic residues of the PDZ domain involved in salt bridges with the peptide (*cyan*). The ribbon illustrations were prepared with VMD (85). To see this figure in color, go online.

Finally, it is interesting to compare protein folding with peptide binding as they differ in the number of molecules involved but they are both governed by noncovalent interactions. Protein folding is driven by progressive formation of native interactions, which are in general more favorable than nonnative contacts (80–84). In contrast, our simulation results provide evidence that the binding of a charged peptide to a protein surface with opposite charge can be steered by long-range polar interactions that are not present in the final bound state.

SUPPORTING MATERIAL

Twenty-one figures, one table, and two movies are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(15)00299-4.

AUTHOR CONTRIBUTIONS

All authors contributed to the study design. M.X. performed the MD simulations and prepared the Supplementary Movies, N.B. and A.C. analyzed the data and wrote the article.

ACKNOWLEDGMENTS

This work was supported by a grant from the Swiss National Science Foundation to A.C.

REFERENCES

1. Schreiber, G., G. Haran, and H. X. Zhou. 2009. Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* 109:839–860.
2. Schreiber, G. 2002. Kinetic studies of protein-protein interactions. *Curr. Opin. Struct. Biol.* 12:41–47.
3. Berg, O. G., and P. H. von Hippel. 1985. Diffusion-controlled macromolecular interactions. *Annu. Rev. Biophys. Biophys. Chem.* 14:131–160.
4. Schreiber, G., and A. R. Fersht. 1996. Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.* 3:427–431.
5. Selzer, T., S. Albeck, and G. Schreiber. 2000. Rational design of faster associating and tighter binding protein complexes. *Nat. Struct. Biol.* 7:537–541.
6. Hemsath, L., R. Dvorsky, …, M. R. Ahmadian. 2005. An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. *Mol. Cell.* 20:313–324.
7. Northrup, S. H., J. O. Boles, and J. C. L. Reynolds. 1988. Brownian dynamics of cytochrome *c* and cytochrome *c* peroxidase association. *Science.* 241:67–70.
8. Gabdoulline, R. R., and R. C. Wade. 1997. Simulation of the diffusional association of barnase and barstar. *Biophys. J.* 72:1917–1929.
9. Ahmad, M., W. Gu, and V. Helms. 2008. Mechanism of fast peptide recognition by SH3 domains. *Angew. Chem. Int.Ed. Engl.* 47:7626–7630.
10. Tang, C., J. Iwahara, and G. M. Clore. 2006. Visualization of transient encounter complexes in protein-protein association. *Nature.* 444:383–386.
11. Suh, J. Y., C. Tang, and G. M. Clore. 2007. Role of electrostatic interactions in transient encounter complexes in protein-protein association investigated by paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* 129:12954–12955.
12. Giorgino, T., I. Buch, and G. De Fabritiis. 2012. Visualizing the induced binding of SH2-phosphopeptide. *J. Chem. Theory Comput.* 8:1171–1175.
13. Sheng, M., and C. Sala. 2001. PDZ domains and the organization of supramolecular complexes. *Annu. Rev. Neurosci.* 24:1–29.
14. Harris, B. Z., and W. A. Lim. 2001. Mechanism and role of PDZ domains in signaling complex assembly. *J. Cell Sci.* 114:3219–3231.
15. Feng, W., and M. Zhang. 2009. Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat. Rev. Neurosci.* 10:87–99.
16. Jemth, P., and S. Gianni. 2007. PDZ domains: folding and binding. *Biochemistry.* 46:8701–8708.
17. Chi, C. N., A. Bach, …, P. Jemth. 2012. Ligand binding by PDZ domains. *Biofactors.* 38:338–348.
18. Doyle, D. A., A. Lee, …, R. MacKinnon. 1996. Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell.* 85:1067–1076.
19. Hillier, B. J., K. S. Christopherson, …, W. A. Lim. 1999. Unexpected modes of PDZ domain scaffolding revealed by structure of nNOS-syntrophin complex. *Science.* 284:812–815.
20. Songyang, Z., A. S. Fanning, …, L. C. Cantley. 1997. Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science.* 275:73–77.
21. Stiffler, M. A., J. R. Chen, …, G. MacBeath. 2007. PDZ domain binding selectivity is optimized across the mouse proteome. *Science.* 317:364–369.
22. Steiner, S., and A. Caflisch. 2012. Peptide binding to the PDZ3 domain by conformational selection. *Proteins: Struct., Funct. Bioinf.* 80:2562–2572.
23. Kong, Y., and M. Karplus. 2009. Signaling pathways of PDZ2 domain: a molecular dynamics interaction correlation analysis. *Proteins: Struct., Funct. Bioinf.* 74:145–154.
24. Mostarda, S., D. Gfeller, and F. Rao. 2012. Beyond the binding site: the role of the β_2-β_3 loop and extra-domain structures in PDZ domains. *PLOS Comput. Biol.* 8:e1002429.
25. Buchli, B., S. A. Waldauer, …, P. Hamm. 2013. Kinetic response of a photoperturbed allosteric protein. *Proc. Natl. Acad. Sci. USA.* 110:11725–11730.
26. Buchenberg, S., V. Knecht, …, G. Stock. 2014. Long-range conformational transition of a photoswitchable allosteric protein: molecular dynamics simulation study. *J. Phys. Chem. B.* 118:13468–13476.
27. Dhulesia, A., J. Gsponer, and M. Vendruscolo. 2008. Mapping of two networks of residues that exhibit structural and dynamical changes upon binding in a PDZ domain protein. *J. Am. Chem. Soc.* 130:8931–8939.
28. Shan, Y., E. T. Kim, …, D. E. Shaw. 2011. How does a drug molecule find its target binding site? *J. Am. Chem. Soc.* 133:9181–9183.
29. Dror, R. O., A. C. Pan, …, D. E. Shaw. 2011. Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proc. Natl. Acad. Sci. USA.* 108:13118–13123.
30. Huang, D., and A. Caflisch. 2011. Small molecule binding to proteins: affinity and binding/unbinding dynamics from atomistic simulations. *ChemMedChem.* 6:1578–1580.
31. Gohlke, H., U. Hergert, …, L. Schmitt. 2013. Binding region of alanopine dehydrogenase predicted by unbiased molecular dynamics simulations of ligand diffusion. *J. Chem. Inf. Model.* 53:2493–2498.
32. Buch, I., T. Giorgino, and G. De Fabritiis. 2011. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA.* 108:10184–10189.
33. Magno, A., S. Steiner, and A. Caflisch. 2013. Mechanism and kinetics of acetyl-lysine binding to bromodomains. *J. Chem. Theory Comput.* 9:4225–4232.
34. Gao, X., T. Satoh, …, T. Kataoka. 2001. Identification and characterization of RA-GEF-2, a Rap guanine nucleotide exchange factor that serves as a downstream target of M-Ras. *J. Biol. Chem.* 276:42219–42225.

35. Erdmann, K. S. 2003. The protein tyrosine phosphatase PTP-Basophil/Basophil-like. Interacting proteins and molecular functions. *Eur. J. Biochem.* 270:4789–4798.

36. Abaan, O. D., and J. A. Toretsky. 2008. PTPL1: a large phosphatase with a split personality. *Cancer Metastasis Rev.* 27:205–214.

37. Berman, H. M., J. Westbrook, …, P. E. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.

38. Van Der Spoel, D., E. Lindahl, …, H. J. C. Berendsen. 2005. GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26:1701–1718.

39. MacKerell, Jr., A. D., D. Bashford, …, M. Karplus. 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B.* 102:3586–3616.

40. Mackerell, Jr., A. D., M. Feig, and C. L. Brooks, 3rd. 2004. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* 25:1400–1415.

41. Jorgensen, W. L., J. Chandrasekhar, …, M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

42. Darden, T., D. York, and L. Pedersen. 1993. Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* 98:10089–10092.

43. Bussi, G., D. Donadio, and M. Parrinello. 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126:014101.

44. Berendsen, H. J. C., J. P. M. Postma, …, J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.

45. Hess, B., H. Bekker, …, J. G. E. M. Fraaije. 1997. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* 18:1463–1472.

46. Blöchliger, N., A. Vitalis, and A. Caflisch. 2013. A scalable algorithm to order and annotate continuous observations reveals the metastable states visited by dynamical systems. *Comput. Phys. Commun.* 184:2446–2453.

47. Blöchliger, N., A. Vitalis, and A. Caflisch. 2014. High-resolution visualization of the states and pathways sampled in molecular dynamics simulations. *Sci. Rep.* 4:6264.

48. Vitalis, A., and A. Caflisch. 2012. Efficient construction of mesostate networks from molecular dynamics trajectories. *J. Chem. Theory Comput.* 8:1108–1120.

49. Zhang, J., P. J. Sapienza, …, A. L. Lee. 2010. Crystallographic and nuclear magnetic resonance evaluation of the impact of peptide binding to the second PDZ domain of protein tyrosine phosphatase 1E. *Biochemistry.* 49:9280–9291.

50. Dolinsky, T. J., J. E. Nielsen, …, N. A. Baker. 2004. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* 32:W665–W667.

51. Dolinsky, T. J., P. Czodrowski, …, N. A. Baker. 2007. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* 35:W522–W525.

52. Baker, N. A., D. Sept, …, J. A. McCammon. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. USA.* 98:10037–10041.

53. Krivov, S. V., and M. Karplus. 2006. One-dimensional free-energy profiles of complex systems: progress variables that preserve the barriers. *J. Phys. Chem. B.* 110:12689–12698.

54. Krivov, S. V., and M. Karplus. 2008. Diffusive reaction dynamics on invariant free energy profiles. *Proc. Natl. Acad. Sci. USA.* 105:13841–13846.

55. Banushkina, P. V., and S. V. Krivov. 2015. Fep1d: a script for the analysis of reaction coordinates. *J. Comput. Chem.* Published online February 25, 2015. http://dx.doi.org/10.1002/jcc.23868.

56. Haq, S. R., C. N. Chi, …, P. Jemth. 2012. Side-chain interactions form late and cooperatively in the binding reaction between disordered peptides and PDZ domains. *J. Am. Chem. Soc.* 134:599–605.

57. Eildal, J. N. N., G. Hultqvist, …, P. Jemth. 2013. Probing the role of backbone hydrogen bonds in protein-peptide interactions by amide-to-ester mutations. *J. Am. Chem. Soc.* 135:12998–13007.

58. Volkov, A. N., J. A. R. Worrall, …, M. Ubbink. 2006. Solution structure and dynamics of the complex between cytochrome *c* and cytochrome *c* peroxidase determined by paramagnetic NMR. *Proc. Natl. Acad. Sci. USA.* 103:18945–18950.

59. Huang, D., and A. Caflisch. 2015. Evolutionary conserved Tyr-169 stabilizes the $\beta 2$-$\alpha 2$ loop of the prion protein. *J. Am. Chem. Soc.* 137:2948–2957.

60. Mahoney, M. W., and W. L. Jorgensen. 2001. Diffusion constant of the TIP5P model of liquid water. *J. Chem. Phys.* 114:363–366.

61. Gianni, S., A. Engström, …, P. Jemth. 2005. The kinetics of PDZ domain-ligand interactions and implications for the binding mechanism. *J. Biol. Chem.* 280:34805–34812.

62. Gianni, S., T. Walma, …, G. W. Vuister. 2006. Demonstration of long-range interactions in a PDZ domain by NMR, kinetics, and protein engineering. *Structure.* 14:1801–1809.

63. Gianni, S., S. R. Haq, …, P. Jemth. 2011. Sequence-specific long range networks in PSD-95/discs large/ZO-1 (PDZ) domains tune their binding selectivity. *J. Biol. Chem.* 286:27167–27175.

64. Di Silvio, E., D. Bonetti, …, S. Gianni. 2014. The mechanism of binding of the second PDZ domain from the Protein Tyrosine Phosphatase-BL to the Adenomatous Polyposis Coli tumor suppressor. *Protein Eng. Des. Sel.* 27:249–253.

65. Chi, C. N., A. Engström, …, P. Jemth. 2006. Two conserved residues govern the salt and pH dependencies of the binding reaction of a PDZ domain. *J. Biol. Chem.* 281:36811–36818.

66. Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science.* 268:1144–1149.

67. Sheinerman, F. B., R. Norel, and B. Honig. 2000. Electrostatic aspects of protein-protein interactions. *Curr. Opin. Struct. Biol.* 10:153–159.

68. Filippakopoulos, P., S. Picaud, …, S. Knapp. 2012. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell.* 149:214–231.

69. Dyson, H. J., and P. E. Wright. 2005. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6:197–208.

70. Uversky, V. N. 2002. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* 11:739–756.

71. Dyson, H. J., and P. E. Wright. 2002. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* 12:54–60.

72. Baker, C. M., and R. B. Best. 2014. Insights into the binding of intrinsically disordered proteins from molecular dynamics simulation. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 4:182–198.

73. Chen, T., J. Song, and H. S. Chan. 2014. Theoretical perspectives on nonnative interactions and intrinsic disorder in protein folding and binding. *Curr. Opin. Struct. Biol.* 30:32–42.

74. Zhang, W., D. Ganguly, and J. Chen. 2012. Residual structures, conformational fluctuations, and electrostatic interactions in the synergistic folding of two intrinsically disordered proteins. *PLOS Comput. Biol.* 8:e1002353.

75. Rogers, J. M., A. Steward, and J. Clarke. 2013. Folding and binding of an intrinsically disordered protein: fast, but not 'diffusion-limited'. *J. Am. Chem. Soc.* 135:1415–1422.

76. Ganguly, D., S. Otieno, …, J. Chen. 2012. Electrostatically accelerated coupled binding and folding of intrinsically disordered proteins. *J. Mol. Biol.* 422:674–684.

77. Ganguly, D., W. Zhang, and J. Chen. 2013. Electrostatically accelerated encounter and folding for facile recognition of intrinsically disordered proteins. *PLOS Comput. Biol.* 9:e1003363.

78. De Sancho, D., and R. B. Best. 2012. Modulation of an IDP binding mechanism and rates by helix propensity and non-native interactions: association of HIF1α with CBP. *Mol. Biosyst.* 8:256–267.
79. Dogan, J., X. Mu, …, P. Jemth. 2013. The transition state structure for coupled binding and folding of disordered protein domains. *Sci. Rep.* 3:2076.
80. Dobson, C. M., A. Šali, and M. Karplus. 1998. Protein folding: a perspective from theory and experiment. *Angew. Chem. Int.Ed. Engl.* 37:868–893.
81. Dill, K. A., and H. S. Chan. 1997. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* 4:10–19.
82. Onuchic, J. N., Z. Luthey-Schulten, and P. G. Wolynes. 1997. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* 48:545–600.
83. Onuchic, J. N., and P. G. Wolynes. 2004. Theory of protein folding. *Curr. Opin. Struct. Biol.* 14:70–75.
84. Best, R. B., G. Hummer, and W. A. Eaton. 2013. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. USA.* 110:17874–17879.
85. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28.