Long-range conformational response of PDZ2 domain to ligand binding and release: a molecular dynamics study

Cheng Lu, Volker Knecht* and Gerhard Stock*

Biomolecular Dynamics, Institute of Physics, Albert Ludwigs University, 79104 Freiburg, Germany

E-mail: volker.knecht@physik.uni-freiburg.de, stock@physik.uni-freiburg.de

Abstract

The binding of a ligand to a protein may induce long-range structural or dynamical changes in the biomacromolecule even at sites physically well separated from the binding pocket. A system for which such behavior has been widely discussed is the PDZ2 domain of human tyrosine phosphatase 1E. Here we present results from equilibrium trajectories of the PDZ2 domain in the free and ligand-bound state, as well as nonequilibrium simulations of the relaxation of PDZ2 after removal of its peptide ligand. The study reveals changes in inter-residue contacts, backbone dihedral angles, and Cα positions upon ligand release. Our findings show a long-range conformational response of the PDZ2 domain to ligand release in the form of a collective shift of the secondary structure elements α2, β2, β3, α1-β4, and the C terminal loop relative to the rest of the protein away from the N-terminus, and a shift of the loops β2-β3 and β1-β2 in opposite direction. The shifts lead to conformational changes in the backbone especially in the β2-β3 loop but also in the β5-α2 and the α2-β6 loop, and are accompanied by changes of inter-residue contacts mainly within the β2-β3 loop as well as between the α2 helix and other segments. The residues showing substantial changes of inter-residue contacts, backbone conformations, or Cα positions, are considered “key residues” for the long-range conformational response of PDZ2. By comparing these residues with

*To whom correspondence should be addressed
various sets of residues highlighted by previous studies of PDZ2, we investigate the
statistical correlation of the various approaches. Interestingly, we find a considerable
correlation of our findings with several works considering structural changes, but no
significant correlations with approaches considering energy flow or networks based on
inter-residue energies.

Introduction

The binding of a ligand to a protein may change the properties at remote regions of the
molecule. This phenomenon called allostery is crucial to many biological processes.\textsuperscript{1,2} Traditionally, allostery has been attributed to changes in structure\textsuperscript{3} or shifts in the populations
of alternative conformations.\textsuperscript{4} More recently it has been proposed that allostery may be (i)
of dynamic origin with only marginal alterations in average structure,\textsuperscript{5–9} or (ii) based both
on dynamic and structural changes at the same time.\textsuperscript{10} Along with ongoing progress in experimental and theoretical approaches, tremendous effort has been devoted to understand
allosteric mechanisms on a microscopic level.\textsuperscript{10–45}

Ligand-induced allostery is observed for a subset of a class of modular domains for protein-
protein interactions denoted as PSD95/Discs large/ZO-1 (PDZ) domains.\textsuperscript{11–16,26,27,37,40} PDZ
domains mediate the clustering of membrane ion channels by binding to their C-termini.\textsuperscript{13–15}
They exhibit roughly 90 amino acids which fold into globular shape with a $\beta\beta\beta\alpha\beta\beta\alpha\beta$
topology where the second $\beta$-strand and the second $\alpha$-helix form the canonical binding
groove,\textsuperscript{11–13,16} see Figure 1 (a). The example shown in Figure 1 is the second PDZ domain
(PDZ2) of human tyrosine phosphatase 1E, a protein involved in the regulation of multiple
receptor-coupled signal transduction pathways including apoptosis.\textsuperscript{37}

Due to their abundance and small size, and owing to the influential study on evolutionarily
conserved pathways of energetic connectivity by Lockless and Ranganathan,\textsuperscript{35} PDZ domains
have become a popular model system to study allosteric communication in a protein. To this
end, a number of “communication networks” have been proposed, which aim to investigate
which residue “talks” to which.\textsuperscript{10,36–40,43–45} The construction of the network can be based on
the correlation of structural or dynamical changes of various protein residues,\textsuperscript{10,37–40} which
may report on the conformational propagation of a structural perturbation in the system.
Alternatively, the analysis has been based on the exchange of vibrational energy between
protein residues\textsuperscript{41,42} or correlations in inter-residue interaction energies.\textsuperscript{43} Finally, also cross
correlations between energetic and structural fluctuations have been considered.\textsuperscript{45} While
some general trends are similar, the concrete prediction of “key residues” (i.e., residues
that are important for allosteric communication) turn out quite different for the various
approaches.\textsuperscript{10,37–40,43–45} In a first attempt to monitor the time evolution of allosteric com-
Figure 1: Representative configurations of PDZ2 domain in the bound state as obtained from MD simulations, highlighting secondary structure segments and binding pocket. (a) Secondary structure, where α-helices are shown in purple, β-sheets in cyan, loop regions in white, and the ligand in yellow. Labels indicate the secondary structure segments $\beta_1$ (residues 6–12), $\beta_2$ (20–23), $\beta_3$ (35–40), $\alpha_1$ (45–49), $\beta_4$ (57–61), $\beta_5$ (64–65), $\alpha_2$ (73–80), and $\beta_6$ (84–90). The most important loops connecting these regions are $\beta_1$-$\beta_2$ (13–19), $\beta_2$-$\beta_3$ (24–34), $\beta_3$-$\alpha_2$ (41–44), and $\alpha_2$-$\beta_6$ (81–83). (b) Probabilities of contacts between individual residues of PDZ2 and the ligand are indicated by a color code, where green to red corresponds to the range 0 to 1.

Communication, Buchli et al. presented a time-resolved study of a conformational transition in a PDZ2 domain after photo-isomerization of an azobenzene switch attached using ultrafast vibrational spectroscopy.\textsuperscript{46} Complementing this work, we used molecular dynamics simulations to infer the changes in inter-residue contacts and local backbone conformations for the transition of this modified PDZ2 domain.\textsuperscript{47} The experiments revealed that the conformational rearrangement of the PDZ domain occurs in a highly non-exponential manner on various timescales from pico- to microseconds.

In the present work, we report microsecond MD trajectories of the wild-type PDZ2 domain in its free (PDZ2-f) and bound (PDZ2-b) form. Considering a number of observables including changes in inter-residue contacts,\textsuperscript{47} backbone conformations,\textsuperscript{47} and C\textsubscript{α} positions\textsuperscript{38,48,49}, we discuss the structural changes of PDZ2 upon ligand release. We find that PDZ2 responds via a collective shift of various secondary structure elements (particularly of $\alpha_2$, $\beta_2$, $\beta_3$ and the termini) with respect to each other, which leads to essential conformational changes of several loops of the protein (particularly of $\beta_2$-$\beta_3$). To investigate the time evolution of the structural response of PDZ2, we perform nonequilibrium MD simulations of the relaxation of PDZ2 after removal of the ligand, which show that the process occurs on a timescale of 100 ns. By considering the various sets of “key residues” identified by our and nine other studies of PDZ2, we investigate the statistical correlation of the various approaches.
Methods

Simulation set-up
Following previous work, all MD simulations of PDZ2 were performed using GROMACS with a hybrid GPU-CPU acceleration scheme. Here PDZ2 was simulated in aqueous NaCl solution. The protein was described using the all-atom Amber99sb*-ILDN force field, the water by the TIP3P model, and the ions with the model of Ref. The side chains of all four histidine residues (32, 53, 71, and 86) of PDZ2 were chosen electrostatically neutral (mono-protonated). The lengths of bonds involving hydrogens were constrained, allowing for a 2 fs time step. Long-range electrostatic interactions were evaluated in reciprocal space using Particle-Mesh Ewald with a maximum spacing for the FFT grid of 0.12 nm and an interpolation via a 6th order polynomial. The minimal cut-off distance for electrostatic and van der Waals interactions was set to 1.2 nm. The temperature of 300 K was maintained via the velocity rescaling algorithm (0.1 ps relaxation time) and the pressure \( P = 1 \text{ bar} \) was controlled using the weak coupling method of Berendsen.

Nonequilibrium simulations
MD simulations of the PDZ2 domain in the bound and free state in explicit aqueous salt solution under periodic boundary conditions were conducted as shown in Figure 2. Six 1 \( \mu \text{s} \) simulations for PDZ2 in the bound (\( b_{,eq} \)) and the free state at equilibrium (\( f_{,eq} \)) were taken from our previous study. From the ensemble \( b_{,eq} \), \( n = 20 \) snapshots were selected randomly, yielding the sub-ensemble \( b_{,sel} \). Hence, from each configuration, the ligand was removed, the resulting void filled with water molecules, and the system obtained was energy minimized using steepest descent, yielding the set of final configurations denoted as \( f^* \). From each of these configurations, a 10 ns \( NPT \) simulation with harmonic restraints on the protein atoms was conducted. The final configuration was taken with the positions and box dimensions scaled such that the volume was equal to the average volume of the second half of the \( NPT \) simulations, resulting in the set of configurations \( f^*/0 \). From each of these snapshots, a 100 ns \( NVT \) simulation was started; the set of the resulting final configurations is denoted as \( f^*/100\text{ns} \).

Analysis
Changes in contact probabilities, average backbone dihedral angles, and \( C_\alpha \) positions upon ligand removal at equilibrium and nonequilibrium conditions were determined. The equilibrium simulations of PDZ2 in the bound and in the free state were analyzed after skipping the
initial 50 ns for equilibration. Inter-residue contacts were analyzed as follows. For each pair of residues $i \neq j$, the inter-residue distance $r_{ij}$ was defined as the minimal distance between any atom of $i$ and any atom of $j$. If $r_{ij} < 0.45$ nm, the residues were considered to be in contact.

**Backbone conformation of individual residues**

In order to monitor local ligand-induced conformational transitions, changes in backbone dihedral angles $\alpha = \phi, \psi$ of PDZ2 upon ligand release were analyzed. To avoid issues due to the periodicity of the angles, the latter were transformed according to $x(\alpha) = \cos \alpha$ and $y(\alpha) = \sin \alpha$. The averages $\bar{x}_{\alpha,s} = \langle x(\alpha) \rangle_s$ and $\bar{y}_{\alpha,s} = \langle y(\alpha) \rangle_s$ were computed for the two states $s = b, f$. Defining

$$\Delta^2_\alpha = (\bar{x}_{\alpha,f} - \bar{x}_{\alpha,b})^2 + (\bar{y}_{\alpha,f} - \bar{y}_{\alpha,b})^2,$$

the overall change in the $(\phi, \psi)$ angles for a given residue was quantified as

$$\Delta^2(\phi, \psi) = \Delta^2_\phi + \Delta^2_\psi.$$

An alternative measure for the change in backbone conformation for a given residue was obtained via computing the root mean square deviation (rmsd) between the average conformation of the main chain (NH and CO group as well as $C_\alpha$ and C atoms) of this residue in the bound and the free state.
Cα shifts

Ligand-induced changes in the average positions of individual residues of PDZ2 were monitored by computing the Cα shifts

$$\Delta r_i \equiv \langle r_i \rangle_b - \langle r_i \rangle_f, \quad i = 1, \ldots, N_{\text{res}}.$$  

Here, $r_i$ denotes the position of the Cα atom of residue $i$ after fitting the corresponding configuration to the initial structure such as to minimize the rmsd of the Cα atoms of the protein from the reference structure. Furthermore, $N_{\text{res}} = 94$ is the number of residues, and $\langle \ldots \rangle_b$ or $\langle \ldots \rangle_f$ denote averages over the bound or the free state, respectively. Standard errors for the average positions at equilibrium were obtained from the averages over individual trajectories. Hence, the statistical errors of the shifts were determined via error propagation. When the size of the shift in one component was below its statistical error, the shift was considered insignificant (standard error criterion) and set to zero. This approach yielded results similar to those from a more rigorous analysis using hypothesis tests as described in the SI (Table S4).

Rotational fit

As the rotation applied for the fitting to the reference structure depends on the configuration of the molecule via the moment of inertia, the separation of internal and rigid body motions is only straightforward for relatively rigid systems. In the case of the PDZ2 domain, issues might thus arise due to flexibility of the loops. In order to address effects due to loop flexibility, we compared Cα shifts obtained employing two different sets of atoms for the fit, using either (i) all the Cα atoms or (ii) only those of two β-strands shown to be rigid ($β_1$ and $β_6$).47 The Cα shifts obtained in conjunction with these two fitting schemes are shown in Figure 3 (a). The error bars of the two curves overlap with each other, indicating that within the statistical accuracy the shifts obtained using the two fitting schemes agree with each other.

A second check was performed by determining the ligand-induced change in the vectors $V_{6,12}$ along the $β_1$ and $V_{90,84}$ along the $β_6$ strand. Here the vector $V_{6,12}$ was chosen to point from Ile-6 to Ala-12 and $V_{90,84}$ from Glu-90 to Val-84. The angle between the corresponding average vector in the bound and in the free state obtained using the two fitting schemes was calculated. As evident from Table S1, each scheme yielded a rotation angle of zero. Also the shifts of the corresponding Cα atoms were small (0.004 nm to 0.011 nm). Hence, subsequent analyses of Cα shifts were performed by fitting on the β strands, $β_1$ and $β_6$.  

6
Figure 3: (a) Rotational fit: Sizes of Cα shifts upon ligand release at equilibrium obtained from two different sets of Cα atoms for the least squares fit to remove overall translation and rotation (either all Cα atoms or only Cα atoms in the β1 and β6 strands). (b,c) Nonequilibrium simulation conditions: Sizes of Cα shifts between configurations selected from the bound state at equilibrium and (b) the initial or (c) final configurations of the nonequilibrium simulations after ligand removal. In (b,c), the sizes of Cα shifts upon ligand release at equilibrium are shown for comparison. In (a – c), the widths of the shaded regions indicate standard errors.

Initial conditions of nonequilibrium simulations

Figure 3 (b) shows the sizes of Cα shifts between the set of configurations selected from the bound state at equilibrium and the initial configurations of the nonequilibrium simulations.
after ligand removal, showing effects from the preparation of the nonequilibrium simulations (especially from the energy minimization after replacing the ligand by water). Large shifts are observed in the $\beta_2$-$\beta_3$ loop and moderate shifts in the $\beta_5$-$\alpha_2$ loop as well as in the $\alpha_2$ helix. Figure S1 displays the corresponding individual components of the shifts. The large size of the shift in the $\beta_2$-$\beta_3$ loop arises from a strong shift in $y$-direction (see coordinate system in Figure 1) which is opposite to that observed at equilibrium.

Figure 3 (c) shows the sizes of $C_\alpha$ shifts between the set of configurations selected from the bound state at equilibrium and the final configurations of the nonequilibrium simulations, indicating that the nonequilibrium simulations ultimately using solely the experimental structure of the bound state largely reproduce the transition from PDZ2-bound to PDZ2-free at equilibrium (observed from simulations using the experimental structures of both the bound and the free state).

**Results**

In order to understand the long-range conformational response of the PDZ2 domain to ligand release or binding, MD simulations of the PDZ2 domain in the bound and free state at equilibrium, as well as nonequilibrium simulations of PDZ2 after ligand release, were conducted. In the following, first the equilibrium and then the nonequilibrium simulations are discussed.

**Comparison to experiment**

The average rmsd of the simulated $C_\alpha$ atom positions from the experimental (x-ray) structure $0.16 \pm 0.01$ nm for PDZ2-f and $0.14 \pm 0.01$ nm for PDZ2-b (the standard errors were obtained from averages over individual trajectories). These rmsds are comparable to the resolution of the corresponding x-ray experiments of 0.164 nm for the free state and 0.130 nm for the bound state.\textsuperscript{15}

**Contacts**

The probabilities of contacts of individual residues of PDZ2 (residues 1 – 94) with the ligand (residues -5 – 0) when the ligand is bound to PDZ2 are given in Table S2 and indicated in Figure 1. It is found that all six residues of the peptide interact with the PDZ2 domain. The $\alpha_2$ helix and the $\beta_2$ strand forming the so called binding groove (or binding pocket) are most frequently in contact with the ligand (red color). Whereas the $\beta_2$ strand contacts all peptide residues, the $\alpha_2$ helix only contacts the C-terminal half of the peptide. Besides
the canonical binding groove, also several loops contact the peptide. In particular, the $\beta_2-\beta_3$ loop contacts the N-terminus (Glu(-5) and Gln(-4)) and the $\beta_1-\beta_2$ loop the C-terminus of the peptide (Val(0)).

Table 1 shows the inter-residue contacts whose probabilities change more than 0.2 upon ligand removal. Significant changes mainly involve residues or secondary structure segments contacting the ligand in the bound state. An exception is an increase in contact probability within the $\beta_4$ strand.

Table 1: Contacts between residues $i$ in segments $S_i$ and residues $j$ in segments $S_j$ whose probabilities change significantly upon ligand removal at equilibrium ($\Delta P_{\text{eq}} > 0.2$). The symbols $\Delta P_{\text{eq}}$ and $\Delta P_{\text{neq}}$ denote the changes in contact probabilities at equilibrium and nonequilibrium, respectively. The latter are the changes in contact probabilities in the final 20 ns of the nonequilibrium simulations and the contact probabilities in the bound state at equilibrium. The superscripts "**" or "***" indicate that, in the bound state, the respective residue is sometimes or always respectively in contact with the ligand. Contact changes mainly occur within the $\beta_2-\beta_3$ loop (residues 24 – 34) and between the $\alpha_2$ helix (residues 73 – 80) and other segments.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$j$</th>
<th>$\Delta P_{\text{eq}}$</th>
<th>$\Delta P_{\text{neq}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala-74*</td>
<td>Ala-69</td>
<td>-0.25</td>
<td>-0.32</td>
</tr>
<tr>
<td>Leu-78**</td>
<td>Leu-66</td>
<td>-0.43</td>
<td>-0.45</td>
</tr>
<tr>
<td>Leu-78**</td>
<td>Ile-20**</td>
<td>0.23</td>
<td>0.34</td>
</tr>
<tr>
<td>Leu-78**</td>
<td>Val-22**</td>
<td>0.24</td>
<td>0.26</td>
</tr>
<tr>
<td>Arg-79**</td>
<td>Lys-13*</td>
<td>-0.39</td>
<td>-0.25</td>
</tr>
<tr>
<td>Leu-78**</td>
<td>Leu-87</td>
<td>-0.40</td>
<td>-0.29</td>
</tr>
<tr>
<td>Ile-20**</td>
<td>Val-22**</td>
<td>-0.26</td>
<td>-0.27</td>
</tr>
<tr>
<td>Gly-24*</td>
<td>Asn-27*</td>
<td>0.24</td>
<td>–</td>
</tr>
<tr>
<td>Gly-24*</td>
<td>Val-30*</td>
<td>-0.27</td>
<td>-0.23</td>
</tr>
<tr>
<td>Gly-25</td>
<td>Ser-29*</td>
<td>0.25</td>
<td>–</td>
</tr>
<tr>
<td>Val-26</td>
<td>Ser-29*</td>
<td>0.21</td>
<td>–</td>
</tr>
<tr>
<td>Asn-27*</td>
<td>Ser-29*</td>
<td>0.21</td>
<td>0.35</td>
</tr>
<tr>
<td>Gly-25</td>
<td>Thr-70</td>
<td>-0.24</td>
<td>–</td>
</tr>
<tr>
<td>Val-58</td>
<td>Val-61</td>
<td>0.29</td>
<td>0.30</td>
</tr>
</tbody>
</table>

**Backbone conformational change**

The changes in backbone dihedral angles $\Delta(\phi, \psi)$ along the amino acid sequence are shown in Figure 4 (a) and mapped to the three-dimensional structure in Figure 4 (b). The residues showing changes in backbone dihedral angles $\Delta(\phi, \psi) \geq 0.16$ are listed in Table S3. Significant conformational changes occur in the loops $\beta_2-\beta_3$, $\beta_5-\alpha_2$, and $\alpha_2-\beta_6$. Especially in the former two loops which are in contact with residues -5 to -2 of the ligand as shown in Table 1, ligand release leads to a significant configurational rearrangement. Table S3 and Figure 4 and S2 show that ten out of the 11 residues of the $\beta_2-\beta_3$ loop and two out of the seven residues of the $\beta_5-\alpha_2$ loop change their average backbone conformation by more than
the above given threshold.

![Graph](image)

**Figure 4:** Changes in backbone conformation of individual residues of PDZ2 upon ligand release at equilibrium, in terms of the corresponding rmsds between the average main chain structures in the bound and the free state, represented via (a) scatter plot and (b) by mapping them onto a representative three-dimensional structure of the domain in the bound state using a color code (green to red) in the range 0 to 0.0052 nm.

**Cα Shifts**

Changes in backbone conformations upon ligand release are expected to lead to shifts in residue positions. The latter were monitored in terms of shifts in the average positions of Cα atoms within an internal coordinate system of the protein defined in Figure 1 (a). The Cα shifts of the individual residues along the three directions are shown in Figure 5. The figure and Table 2 indicate that ligand release leads to a shift of the α2 helix, the β2 and β3 strands, as well as the α1-β4 and the C-terminal loop in positive x-direction, and of residues 14 – 17 of the β1-β2 loop as well as residues 25 – 28 of the β2-β3 loop in negative x-direction. In particular, the α1-β4 loop which acts as a steric bridge (see Table S6) transduces the
shift of the $\beta_3$ strand to the C-terminal loop. Furthermore, Figure 5 shows that the C-terminus of the $\alpha_2$ helix moves in negative $y$- and negative $z$-direction, and the N-terminus of the helix in positive $y$-direction. This suggests a rotation of the helix. This rotation is presumably related to the contact changes of the $\alpha_2$ helix with other secondary structure segments described above.

Figure 5: $C_\alpha$ shifts in (a) $x$-, (b) $y$-, and (c) $z$-direction in PDZ2 upon ligand release at equilibrium. (Top panel) Estimated shifts (black curves) with upper and lower bounds (shaded areas). Both the shifts at equilibrium (solid lines and orange areas) as well as those in the nonequilibrium simulations ($C_\alpha$ positions averaged over the final 20 ns relative to the set of structures from the bound state selected to prepare the initial configurations of the nonequilibrium simulations; dotted lines and green areas) are shown. (Bottom panel) Representative configuration of PDZ2 indicating shifts at equilibrium in the range -0.05 nm to 0.05 nm, showing residues with negative, no, or positive shifts in blue, green, or red, respectively. The shifts are given with respect to the internal coordinate system defined in Figure 1 (a).

Table 2: $C_\alpha$ shifts in PDZ2 upon ligand release at equilibrium conditions. The shifts are given with respect to the internal coordinate system defined in Figure 1 (a). Only residues showing shifts with sizes above 0.025 nm are listed. The symbol “–” indicates shifts which are not significant according to Welch’s test or the standard error criterion.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>$\Delta x_{eq}$ (nm)</th>
<th>$\Delta y_{eq}$ (nm)</th>
<th>$\Delta z_{eq}$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>2</td>
<td>-0.01(1)</td>
<td>–</td>
<td>-0.03(2)</td>
</tr>
<tr>
<td>$\beta_1$-$\beta_2$</td>
<td>14</td>
<td>-0.04(1)</td>
<td>0.01(0)</td>
<td>0.03(1)</td>
</tr>
<tr>
<td>$\beta_1$-$\beta_2$</td>
<td>15</td>
<td>-0.06(1)</td>
<td>–</td>
<td>0.04(1)</td>
</tr>
</tbody>
</table>

Continued on next page
Table 2 – continued from previous page

<table>
<thead>
<tr>
<th>S</th>
<th>R</th>
<th>$\Delta x_{eq}$ (nm)</th>
<th>$\Delta y_{eq}$ (nm)</th>
<th>$\Delta z_{eq}$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1-\beta_2$</td>
<td>16</td>
<td>-0.04(1)</td>
<td>-0.03(0)</td>
<td>0.04(0)</td>
</tr>
<tr>
<td>$\beta_1-\beta_2$</td>
<td>17</td>
<td>-0.02(0)</td>
<td>-0.01(0)</td>
<td>0.03(0)</td>
</tr>
<tr>
<td>$\beta_1-\beta_2$</td>
<td>18</td>
<td>-</td>
<td>-</td>
<td>0.03(0)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>20</td>
<td>0.03(0)</td>
<td>-</td>
<td>0.02(0)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>21</td>
<td>0.06(1)</td>
<td>-</td>
<td>0.02(0)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>22</td>
<td>0.05(1)</td>
<td>-</td>
<td>0.01(0)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>23</td>
<td>0.04(1)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_2-\beta_3$</td>
<td>24</td>
<td>0.02(1)</td>
<td>0.01(1)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_2-\beta_3$</td>
<td>25</td>
<td>-0.07(5)</td>
<td>0.05(1)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_2-\beta_3$</td>
<td>26</td>
<td>-0.13(8)</td>
<td>-</td>
<td>-0.05(4)</td>
</tr>
<tr>
<td>$\beta_2-\beta_3$</td>
<td>27</td>
<td>-0.19(11)</td>
<td>-</td>
<td>0.03(2)</td>
</tr>
<tr>
<td>$\beta_2-\beta_3$</td>
<td>28</td>
<td>-0.16(11)</td>
<td>0.05(4)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_2-\beta_3$</td>
<td>29</td>
<td>-</td>
<td>0.08(6)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_2-\beta_3$</td>
<td>30</td>
<td>-</td>
<td>0.12(7)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_2-\beta_3$</td>
<td>31</td>
<td>-</td>
<td>0.07(7)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>32</td>
<td>0.03(0)</td>
<td>0.01(1)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>33</td>
<td>0.03(1)</td>
<td>-</td>
<td>0.01(0)</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>34</td>
<td>0.03(1)</td>
<td>-0.01(0)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>35</td>
<td>0.06(1)</td>
<td>-0.02(0)</td>
<td>-0.01(0)</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>36</td>
<td>0.05(1)</td>
<td>-0.02(0)</td>
<td>-0.02(0)</td>
</tr>
<tr>
<td>$\beta_3-\alpha_1$</td>
<td>37</td>
<td>0.03(1)</td>
<td>0.02(0)</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha_1-\beta_4$</td>
<td>38</td>
<td>0.03(0)</td>
<td>-0.02(0)</td>
<td>-0.01(0)</td>
</tr>
<tr>
<td>$\beta_3-\beta_5$</td>
<td>39</td>
<td>-0.02(0)</td>
<td>-</td>
<td>-0.01(0)</td>
</tr>
<tr>
<td>$\beta_3-\alpha_2$</td>
<td>40</td>
<td>-0.01(1)</td>
<td>0.03(1)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_5-\alpha_2$</td>
<td>41</td>
<td>-0.05(2)</td>
<td>0.06(2)</td>
<td>0.04(1)</td>
</tr>
<tr>
<td>$\beta_5-\alpha_2$</td>
<td>42</td>
<td>-</td>
<td>0.09(3)</td>
<td>0.06(2)</td>
</tr>
<tr>
<td>$\beta_5-\alpha_2$</td>
<td>43</td>
<td>0.03(2)</td>
<td>0.06(1)</td>
<td>0.06(2)</td>
</tr>
<tr>
<td>$\beta_5-\alpha_2$</td>
<td>44</td>
<td>0.06(3)</td>
<td>0.06(1)</td>
<td>0.03(1)</td>
</tr>
<tr>
<td>$\beta_5-\alpha_2$</td>
<td>45</td>
<td>0.06(2)</td>
<td>0.04(2)</td>
<td>-</td>
</tr>
<tr>
<td>$\beta_5-\alpha_2$</td>
<td>46</td>
<td>0.09(2)</td>
<td>0.06(1)</td>
<td>-</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>47</td>
<td>0.04(2)</td>
<td>0.06(1)</td>
<td>0.03(1)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>48</td>
<td>0.03(1)</td>
<td>-</td>
<td>-0.03(2)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>49</td>
<td>0.04(1)</td>
<td>0.02(1)</td>
<td>-0.02(1)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>50</td>
<td>-</td>
<td>-0.05(1)</td>
<td>-0.02(1)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>51</td>
<td>0.06(0)</td>
<td>-0.03(1)</td>
<td>-0.04(1)</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>52</td>
<td>0.03(0)</td>
<td>0.01(0)</td>
<td>-</td>
</tr>
<tr>
<td>C</td>
<td>53</td>
<td>0.02(0)</td>
<td>-</td>
<td>-0.02(0)</td>
</tr>
<tr>
<td>C</td>
<td>54</td>
<td>0.02(1)</td>
<td>-</td>
<td>-0.03(1)</td>
</tr>
</tbody>
</table>

Nonequilibrium simulations

The transient behavior during the relaxation of PDZ2 after ligand removal was investigated using nonequilibrium simulations. The simulation set-up is shown in Figure 2. Table 1 shows
that the changes in inter-residue contacts in the nonequilibrium simulations are (in terms of the size and sign of the alterations in corresponding contact probabilities) very similar to the changes at equilibrium. Figure 5 indicates that also the sizes of the final shifts averaged over the final 20 ns of the 100 ns nonequilibrium simulations ($|\Delta r_{\text{sel},f^{*}/100}|$) are similar to the shifts at equilibrium. Especially the shifts of the $\alpha_2$ helix, the $\beta_2$ and $\beta_3$ strands, as well as the $\alpha_1$-$\beta_4$ and the C-terminal loop in positive $x$-direction (see also Table S5), and of the $\beta_1$-$\beta_2$ and $\beta_2$-$\beta_3$ loop in negative $x$-direction, are largely reproduced.

As noted in the Methods section and displayed in Figure 3 (b), the preparation of the initial configurations of the nonequilibrium simulations lead to large shifts, especially in the $\beta_2$-$\beta_3$ loop and in the $\alpha_2$ helix. Hence, the initial configurations of the nonequilibrium simulations do not fully mimic the bound state. The $\beta_2$ and $\beta_3$ strands, as well as the $\alpha_1$-$\beta_4$ and the C-terminal loop, though, are not affected and their positions are similar to those in the bound state. Hence, the transient shifts in these regions from $f^*/0$ to $f^*/100$ might reflect features of the time-dependent relaxation behavior after ligand removal under more natural conditions.

The time dependent $C_\alpha$ shifts in $x$-direction, relative to the set of structures selected from the bound state used to derive the initial configurations of the nonequilibrium simulations, $\Delta r_{\text{sel},f^{*}/t}$, are shown in Figure 6. The shifts of the $\beta_2$ and $\beta_3$ strands, as well as the $\alpha_1$-$\beta_4$ and the C-terminal loop in positive $x$-direction are seen to built up gradually.

Figure 6: Time evolution of $C_\alpha$ shifts in PDZ2 after ligand removal. Shown are representative snapshots of PDZ2 indicating $C_\alpha$ shifts relative to set of configurations of bound state ($\text{sel,b}$) used to derive initial configurations of nonequilibrium simulations in $x$-direction during the time intervals (50 ps) 0 – 100 ps, (10 ns) 5 ns – 20 ns, and (90 ns) 80 ns – 100 ns of the nonequilibrium simulations. Colors from green to blue indicate negative and from green to red positive shifts in the overall range -0.1 nm to 0.1 nm.
Discussion and Conclusions

We first want to mention several methodological issues raised by our study. As a general point, we note that distal conformational changes due to ligand binding or release are typically quite small (Table 2). An appropriate investigation therefore requires sufficient statistical sampling (i.e., several long trajectories) as well as a sound analysis of the statistical significance of small structural changes (see SI). Employing Cartesian coordinates (e.g., to calculate $C_\alpha$ shifts), we found it important to carefully check for possible uncertainties associated with the rotational fit of the trajectory (Figure 3a), which may easily obscure small structural changes. Concerning the nonequilibrium simulations of PDZ2, we note the ambiguous choice of initial conditions reflecting the ligand release (Figure 3c), which may affect the interpretation of the resulting structural dynamics.

Dealing with the propagation of small conformational changes, it appears not surprising that relatively small modifications of the molecular system may result in quite different behavior. For example, a recent study monitoring cross correlations between inter-residue interaction energies and structural fluctuations showed that ligand binding affects PDZ3-PSD95 and PDZ2 in a different manner. Moreover, it is interesting to compare our results for wild-type PDZ2 to a previous computational study of a photoswitchable PDZ2 domain (PDZ2S), which mimics the free-bound transition by $cis \rightarrow trans$ photo-isomerization of an azobenzene photoswitch. Generally speaking, the initial conformation change due to the photoswitch in PDZ2S is larger than the rearrangement in PDZ2 due to ligand binding, which affects larger conformational heterogeneity and higher residue fluctuations of PDZ2S. As a consequence, the subsequent propagation of the structural perturbation of the two systems agrees only in the general trend but differs in atomistic detail. For example, large changes of the backbone conformation occur for 14 residues in PDZ2 and 24 residues in PDZ2S, where the two sets have nine residues in common (residues 25, 27 – 29, and 32 – 34 in the $\beta_2$-$\alpha_2$ loop as well as 67 – 68 in the $\beta_5$-$\alpha_2$ loop). However, we only find a single common residue pair (Ala-74, Ala-69), when we compare inter-residue contacts of PDZ2 and PDZ2S (Table 1 and Table 1 in Ref. 47).

To summarize our findings of the MD simulation study of PDZ2, the above results reveal that the release of the ligand affects a complex and concerted response of the PDZ2 domain, which originates from direct interactions between ligand and protein. Apart from obvious contacts with the binding-pocket segments $\alpha_2$ and $\beta_2$, the ligand directly interacts via its N- and C-termini with the $\beta_1$-$\beta_2$ and $\beta_2$-$\beta_3$ loops of PDZ2, respectively (Figure 1). These interactions are important, because they affect a concerted conformational change of the binding-pocket with respect to these loops, i.e., $\alpha_2$ and $\beta_2$ move in $+x$ direction and $\beta_1$-$\beta_2$ and $\beta_2$-$\beta_3$ in $-x$ direction in Figure 5(a). This conformational shift of $\alpha_2$ and $\beta_2$ extends via several relatively rigidly connected secondary structure elements all the way
to the C-terminus of PDZ2. That is, \( \beta_2 \) is hydrogen bonded to \( \beta_3 \), which through residue 54 of the \( \alpha_1-\beta_4 \) loop is linked to the C-terminus. As a consequence, we find an overall conformational transition of interconnected secondary-structure segments (\( \alpha_2, \beta_2, \beta_3, \alpha_1-\beta_4 \) and the C-terminal loop) with respect to the \( \beta_1-\beta_2 \) and \( \beta_2-\beta_3 \) loops of PDZ2. Introducing considerable strain in the protein, this structural change subsequently affects a significant conformational rearrangement of the \( \beta_2-\beta_3 \) loop (Figure 4). Due to the finite rigidity of the inter-residue connections, the sizes of the conformational shifts tend to decrease with distance from the binding pocket (Figure 5 and Table 2). Taken together, these findings suggest that the interconnected secondary-structure segments \( \alpha_2, \beta_2, \beta_3, \alpha_1-\beta_4 \) and the C-terminal loop may provide a mechanism of long-range conformational propagation in PDZ2. Moreover, our study identifies a list of “key residues” (i.e., residues that are important for allosteric communication), which were shown to undergo significant changes of inter-residue contacts (Table 1), backbone dihedral angles (Table S3) or \( C_\alpha \) positions (Table 2).

It is interesting to compare our results to related studies of the allosteric communication in PDZ2 domain. As a simple (yet not unambiguously defined) observable, we follow Ref.\(^{39} \) and consider the key residues predicted by each method. (A full list of all residues and methods is found in Table S7). By using Fisher’s exact test (see SI), we calculate corresponding \( p \)-values to assess the statistical correlation of the various approaches with our study (Table 3). On the experimental side, Fuentes \textit{et al.}\(^{37} \) used side-chain (\(^2\)H-methyl) and backbone (\(^{15}\)N) NMR spin relaxation to identify residues that show a significant change of pico- or nanosecond fluctuations upon ligand binding of PDZ2. Interestingly, the resulting \( p \)-value of 0.0008 (line “NMR” in Table 3) reveals a highly significant correlation of the experiment with our study. On the theoretical side, various computational studies have reported communication networks of the allosteric response in PDZ2, which are based either on the correlation of structural changes of various protein residues\(^\text{10,38–40} \) or on the analysis of energy-based correlations.\(^\text{41–43,45} \)

Table 3: Comparison of the present work to previous studies in terms of key residues identified (see text), as given from the \( p \)-values obtained from Fisher’s exact test.

<table>
<thead>
<tr>
<th>Method</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMR(^{37} )</td>
<td>0.0008</td>
</tr>
<tr>
<td>PRS(^{38} )</td>
<td>0.0009</td>
</tr>
<tr>
<td>RestrMD(^{10} )</td>
<td>0.0025</td>
</tr>
<tr>
<td>PSN-ENM(^{39} )</td>
<td>0.0555</td>
</tr>
<tr>
<td>MC(^{40} )</td>
<td>0.2012</td>
</tr>
<tr>
<td>E-RMSF(^{45} )</td>
<td>0.4806</td>
</tr>
<tr>
<td>IEcorr(^{43} )</td>
<td>0.8231</td>
</tr>
<tr>
<td>PEN(^{44} )</td>
<td>0.8938</td>
</tr>
<tr>
<td>ET(^{42} )</td>
<td>0.9595</td>
</tr>
</tbody>
</table>

The structure-based network approaches includes the study of Dhulesia \textit{et al.}\(^\text{10} \) who conducted MD simulations of PDZ2 in explicit solvent using ensemble-averaged NMR restraints.
("RestrMD"), as well as the perturbation response scanning ("PRS") approach by Gerek and Ozkan,\textsuperscript{38} who probed the average response of the protein structure (C\textsubscript{α} positions of an elastic network model) to random perturbations of the binding pocket. The obtained p-values of \(\approx 10^{-3}\) (Table 3) report on significant correlation of these methods of our study with experiment. (We note in passing, though, that a closer analysis of the PRS results revealed significant differences to our findings, see the SI). Another method was proposed by Raimondi et al.,\textsuperscript{39} who combined 20 ns MD simulations with an elastic network model to construct a protein structure network ("PSN-ENM") by computing average numbers of interatomic contacts between residues. Finally Cilia et al.\textsuperscript{40} applied Monte Carlo sampling ("MC") of the side chain conformational space to estimate a mutual information matrix, that reports on ligand-induced changes of the individual residue pairs. The predicted key residues of the latter two methods still correlate well with our results, however, with comparatively low significance (Table 3).

Alternatively, communication networks of the allosteric response in PDZ2 have been constructed via the analysis of energy-based correlations. This includes nonequilibrium MD studies of the transport of vibrational energy between protein residues,\textsuperscript{41,42} ("ET") a MD-based network of the correlations of inter-residue interaction energies,\textsuperscript{43} ("IEcorr") a protein energy network ("PEN") based inter-residue interaction energies,\textsuperscript{44} as well as a recent study of the cross correlations between energetic and structural fluctuations.\textsuperscript{45} Interestingly, none of these methods show significant correlation with our work or with the other structure-based methods, including the NMR experiments. This finding may be caused by the considerable timescale separation between local energy flow and long-range structural change. That is, vibrational energy in disordered condensed phase systems typically dissipates within picoseconds, while the propagation of structural change was found occur on a 100 ns timescale in our nonequilibrium simulations of PDZ2 (and will be certainly much longer for larger proteins). Nonetheless, it has been suggested that residues that provide a preferred pathway of vibrational energy at the same time may indicate possible allosteric connectivity in proteins.\textsuperscript{32,41,42} The present findings for PDZ2 do not support this idea, though.

To conclude, we have presented an allosteric mechanism for the response of PDZ2 to ligand binding based on collective shifts of secondary structure segments and changes in backbone conformations as well as inter-residue contacts. Our results show significant correlations with previous studies based on structural changes but not with energy-based analyses which can be explained from the timescales involved in these two classes of phenomena.

Acknowledgments

The authors thank Sebastian Buchenberg and Markus Schumacher for inspiring and helpful discussions. This work has been supported by the Deutsche Forschungsgemeinschaft and the
China Scholarship Council.

References


Figure 7: Table of Contents Graphic.