Identifying Key Residues for Protein Allostery through Rigid Residue Scan
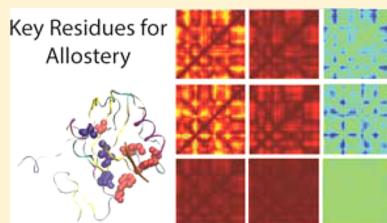
Robert Kalescky,[†] Jin Liu,*[,†,‡] and Peng Tao*[,†]

[†]Department of Chemistry, Center for Drug Discovery, Design, and Delivery (CD4), and Center for Scientific Computation, Southern Methodist University, 3215 Daniel Avenue, Dallas, Texas 75275, United States
[‡]Department of Biophysics, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390, United States

ⓢ *Supporting Information*

ABSTRACT: Allostery is a ubiquitous process for protein regulatory activity in which a binding event can change a protein's function carried out at a distal site. Despite intensive theoretical and experimental investigation of protein allostery in the past five decades, effective methods have yet to be developed that can systematically identify key residues involved in allosteric mechanisms. In this study, we propose the rigid residue scan as a systematic approach to identify important allosteric residues. The third PDZ domain (PDZ3) in the postsynaptic density 95 protein (PSD-95) is used as a model system, and each amino acid residue is treated as a single rigid body during independent molecular dynamics simulations. Various indices based on cross-correlation matrices are used, which allow for two groups of residues with different functions to be identified. The first group is proposed as "switches" that are needed to "turn on" the binding effect of protein allostery. The second group is proposed as "wire residues" that are needed to propagate energy or information from the binding site to distal locations within the same protein. Among the nine residues suggested as important for PDZ3 intramolecular communication in this study, eight have been reported as critical for allostery in PDZ3. Therefore, the rigid residue scan approach is demonstrated to be an effective method for systemically identifying key residues in protein intramolecular communication and allosteric mechanisms.

1. INTRODUCTION

Allostery is the remote-control capacity of biomolecules. Protein allostery has been under investigation for more than five decades since the Monod–Wyman–Changeux (MWC)[1] and Koshland–Némethy–Filmer (KNF)[2] models were published in the 1960s. In the past decade, allosteric regulation has been observed in many proteins, leading to the idea that allostery is an intrinsic feature of all proteins.[3] Allosteric regulation has also been discovered to be a part of most processes of a living cell, including cell signaling, phosphorylation, and protein degradation.[4] Thus, understanding allosteric mechanisms is essential to understanding cell regulation. Perhaps more significantly, this understanding builds the foundation upon which to develop allosteric drugs. This class of drugs may be successful in the treatment of many diseases that are difficult to treat with traditional drugs because of the lack of selectivity and specificity, which is particularly true for cancer and neurodegenerative diseases.[5] Therefore, investigation on the working mechanism of allosteric regulation, considered as the "second secret of life",[6] is vital to understanding many complex biological systems and will aid the pharmaceutical industry in the development of more selective, potent, and effective allosteric drugs.

The original definition of allostery, which was based on crystallographic structures of hemoglobin, is the conformational change of one site of a protein caused by an effector at another site called the allosteric site.[4] Allostery has become well-accepted as a dynamical property of proteins since this initial definition, as is evident via NMR studies of various proteins, including imidazole glycerol phosphate synthase (IGPS), catabolite activator protein (CAP), the third PDZ domain (PDZ3) of postsynaptic density 95 protein (PSD-95), the catalytic subunit of protein kinase A (PKA-C), and 70 kDa heat shock proteins (HSP70).[7−10]

Numerous experimental and computational approaches have been used to identify intramolecular allosteric pathways. Experimentally, NMR spectroscopy has been widely used to identify allosteric communication pathways.[11,12] Chemical shift covariance analysis (CHESCA) was recently developed to quantitatively describe the relative contribution of each residue to allostery.[13] Double-mutant cycle analysis has been used to measure the energetic couplings between residues.[14] Computationally, statistical coupling analysis (SCA),[15] a sequence-based method using multiple sequence alignments to identify the coevolutionary residues that form protein communication pathways, has been successfully applied to PDZ domains and HSP70 proteins to identify allosteric pathways.[16] Normal mode
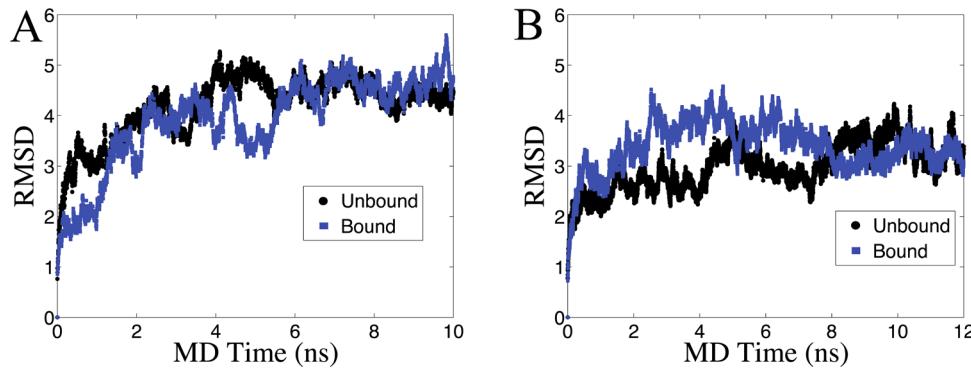
Figure 1. RMSD plots for unbound and bound PDZ3 in MD simulations: (A) 10 ns isothermal−isobaric (*NPT*) ensemble; (B) 12 ns canonical (*NVT*) ensemble.

analysis (NMA) has been used to describe allosteric pathways using a single protein structure by examining the response of the protein to local structural perturbations.[17−20] Molecular dynamics (MD) simulations have been widely used as the basis for protein residue coupling analyses. Cross-correlation analysis has been used as a general strategy to identify allosteric sites by searching for large correlated motion changes of distinct sites of a protein structure upon perturbation events.[21] Anisotropic thermal diffusion (ATD) heats a single residue to observe the heat propagation.[22,23] Pump−probe MD (PPMD) has atoms excited by oscillating forces to monitor the transmission of oscillations.[24]

Here we present a new approach, termed rigid residue scan, in which systematic rigid-body molecular dynamics is carried out to quantify the contribution of each residue to the overall protein dynamics. Modification of individual residues may significantly affect the overall protein dynamics through the collective influence of residues on the protein's structure. Furthermore, changing a residue's dynamics may significantly disrupt allosteric pathways, as protein allostery is considered to be a dynamical process and proposed major allosteric pathways often contain several residues. A direct way to change single-residue dynamics behavior is to treat the residue as a single rigid body, i.e., the residue has no internal degrees of freedom. The influence of a single residue on allostery can be investigated directly through rigid-body simulations. In the amino acid mutagenesis study, all of the bonded and nonbonded interactions related to the mutated residue are completely changed.

Although rigid-body methods are effective tools in MD simulations, it has been a technical challenge to implement these methods for a large number of atoms. Recently, an efficient rigid-body integrator, known as SHAPE, was reported to apply rigid-body constraints on an arbitrary selection of atoms.[25] SHAPE, used in this work, has an accuracy comparable to those of the widely applied SHAKE[26] and RATTLE[27] rigid-body methods. However, it is much more efficient for large numbers of atoms than both the SHAKE and RATTLE algorithms.

In this work, we present a systematic way to identify key residues that are significant to protein intramolecular communication and that may also be important in allosteric pathways by quantifying the contribution of each residue to the whole-protein dynamics.

2. MATERIALS AND METHODS

2.1. Molecular Dynamics Simulations. The protein was subjected to normal MD simulations (herein called unperturbed simulations) as a reference for the rigid residue scans. Systematic rigid-body MD simulations (herein called perturbed simulations) were carried out to probe each residue's influence on the whole-protein dynamics. A single amino acid residue was treated as a single rigid body in each perturbed MD simulation. Therefore, the total number of rigid-body (perturbed) MD simulations was equal to the total number of residues in the protein, as each residue regardless of its type or position was subjected to rigid-body simulation.

The model protein used in this study was the PDZ3 domain of PSD-95, as its allosteric mechanisms have been extensively studied.[15,18,28−38] A crystal structure of PDZ3 bound with the peptide CRIPT was obtained from the Protein Data Bank (PDB ID 1BE9).[39] CRIPT is known to bind to PDZ3 preferentially over the first and second PDZ domains in PSD-95.[34]

Hydrogen atoms were added to the crystal structure of PDZ3 using the hydrogen position construction facility (HBUILD) of the CHARMM (38b1) molecular simulation package.[40] The CRIPT peptide was removed to create an unbound PDZ3 structure. Rigid residue scans were performed using both the unbound and bound states. Each structure was solvated in a cubic box of water using the TIP3P explicit water model.[41] The systems were neutralized by adding sodium cations, and additional sodium and chloride ions were added to reproduce typical physiological ion concentrations. The neutralized systems were subjected to 200 steps of energy minimization using steepest decent (SD) and 12000 steps of adopted-basis Newton−Raphson (ABNR) energy minimization in CHARMM. The minimized structures were heated to 300 K in 12 ps. After heating, both unbound and bound PDZ3 simulation boxes were subjected to 10 ns MD simulations in the isothermal−isobaric (*NPT*) ensemble controlled through extended system constant-pressure and constant-temperature algorithms.[42−44] A 2 fs simulation time step was used in the MD simulations. The temperature and pressure were 300 K and 1 atm, respectively. The root-mean-square deviation (RMSD) plots of PDZ3 from unperturbed *NPT* MD simulations for both the unbound and bound states show that the systems are equilibrated roughly after 2 ns of simulation (Figure 1A). After the *NPT* MD simulations, the trajectory frame with the box size closest to the average box size was selected from the trajectory of the last 8 ns of MD simulation, with the first 2 ns considered as equilibration. The

simulation snapshots were saved to trajectory files for every picosecond throughout the simulations.

The selected trajectory frames for both unbound and bound PDZ3 were used for 12 ns canonical (*NVT*) ensemble Langevin dynamics simulations using a 2 fs time step. The RMSD plots of PDZ3 for both states show that the systems are equilibrated roughly after 2 ns (Figure 1B). Therefore, the first 2 ns was considered as equilibration, and only the last 10 ns was considered in the final analysis. The selected trajectory frames for both PDZ3 states were used for the rigid residue scans. A single residue was treated as a single rigid body in each perturbed simulation. Each simulation was carried out for a total of 12 ns, with the first 2 ns discarded as equilibration and the last 10 ns subjected to analysis.

Two additional 12 ns unperturbed trajectories were carried out for both unbound and bound states to investigate the simulation consistency. The initial structures for these simulations were the same ones selected from the *NPT* simulations, but different random numbers were used to generate the initial velocities. Similarly, the first 2 ns was considered as equilibration, and only the last 10 ns was considered in the final analysis. A few additional rigid residue scan simulations (10 ns) were done to further examine the simulation consistency. However, these were not additional independent trajectories, but instead were continuations of the first 10 ns of the rigid residue scan simulations.

Overall, more than 2900 ns of total simulation time was carried out in this study. All of the simulations were carried out using the CHARMM force field (version 27) and the CHARMM molecular dynamics package (38b1).[40] Periodic boundary conditions (PBCs) were used in all of the simulations. The particle mesh Ewald algorithm[45] was applied to calculate long-range electrostatic interactions. All of the chemical bonds with hydrogen were fixed using SHAKE to remove the high oscillation frequency of these bonds.

The sequence numbers of the 115 residues start at 301 and end at 415 in the original PDB file. Residue sequence numbers starting at 1 and ending at 115 are used throughout this work for clarity. The corresponding residue sequence numbers in referenced works should be 300 higher than those used in this work, unless mentioned otherwise.

2.2. Analysis of MD Trajectories. *2.2.1. RMSD.* The RMSD can be used as a measure of the average change of a protein's structure during the course of an MD simulation. A protein's tertiary structure is represented by N atomic Cartesian coordinate vectors $\mathbf{r}_i$ ($i = 1$ to N). The RMSD (R) between a particular frame along the trajectory and a reference structure is calculated as

$$R = \sqrt{\frac{\sum_{i=1}^{N}(\mathbf{r}_i^0 - \mathbf{U}\mathbf{r}_i)^2}{N}} \tag{1}$$

In eq 1, $\mathbf{r}_i^0$ is the Cartesian coordinate vector of the ith atom in the reference structure and $\mathbf{U}$ is the transformation matrix for the best-fit alignment between the given protein structure and the reference structure. The RMSD can be calculated using all or a subset of atoms for given protein.

2.2.2. Cross-Correlation (Normalized Covariance) Matrix. The cross-correlation matrix $\mathbf{C}$ is the main tool used in this work to examine the intrinsic properties of internal protein dynamics. This matrix is based on correlated atomic motions, and each matrix element, C_{ij}, is calculated through the following equation:

$$C_{ij} = \frac{c_{ij}}{c_{ii}^{1/2} c_{jj}^{1/2}} = \frac{\langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle - \langle \mathbf{r}_i \rangle \cdot \langle \mathbf{r}_j \rangle}{[(\langle r_i^2 \rangle - \langle r_i \rangle^2)(\langle r_j^2 \rangle - \langle r_j \rangle^2)]^{1/2}} \tag{2}$$

In eq 2, i and j are the atomic indices of the protein structure, c_{ij}, c_{ii}, and c_{jj} are the covariance matrix elements, and $\mathbf{r}_i$ and $\mathbf{r}_j$ are the Cartesian coordinate vectors (with magnitudes r_i and r_j) from the least-squares-fitting structures. The cross-correlation matrix can be calculated using all of the atoms or a subset of the atoms (e.g., the C_α carbon of each amino acid residue) in a protein.

2.2.3. Average Residue Correlation Index. We define the average residue correlation (ARC) index, which describes a protein's intrinsic dynamics and its rigidity, to facilitate the direct comparison between rigid residue scan results. The ARC index, C, for each simulation system is calculated through the following equation:

$$C = \frac{1}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} C_{ij} \tag{3}$$

This is the average of the matrix elements in the upper triangle of the cross-correlation matrix excluding all of the diagonal elements, which refer to self-correlation and are equal to unity. The standard deviation of the ARC index, σ_C, calculated as

$$\sigma_C = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (C_{ij} - C)^2} \tag{4}$$

is also used in this study.

The ARC index reflects the overall rigidity of the protein, with higher values indicating increased protein rigidity. In the limit of a completely rigid protein, the ARC index will have a value of unity, as all of the atoms are perfectly correlated.

2.2.4. Residue Correlation Similarity Index. We use the residue correlation similarity (RCS) index, R^{ab}, to compare simulations a and b using their cross-correlation matrices $\mathbf{C}^a$ and $\mathbf{C}^b$:

$$R^{ab} = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (C_{ij}^a - C_{ij}^b)^2} \tag{5}$$

2.2.5. Comparisons of Different Simulations. One should be aware that when any index is used alone, the risk of a false positive or false negative is rather high. Therefore, we used the difference in the ARC indices (ΔARC) and the RCS index to compare the unperturbed/perturbed simulations of unbound/bound PDZ3 in multiple ways.

Perturbed Simulations of Unbound PDZ3 against Perturbed Simulations of Bound PDZ3. When a given residue is being treated as a rigid body, its effect on the overall protein dynamics in either the unbound or bound state could provide some information about the given residue's role in protein allostery. For example, if perturbed simulations of the unbound and bound states display similar cross-correlation patterns under such conditions, this suggests that treating that specific residue as a rigid body eliminates the difference between the unbound and bound states, therefore removing the allosteric effect upon ligand binding. The smaller the ΔARC value and the RCS index are, the stronger is the effect of removing the allostery for the given residue. We propose that those residues with small ΔARC values and RCS indices may act as "switches" for protein allostery because the internal dynamics of these

residues are important in the differentiation of the unbound and bound states. However, larger ΔARC values and RCS indices may not necessarily indicate a stronger allosteric effect.

Unperturbed Simulations of Bound PDZ3 against Perturbed Simulations of Unbound PDZ3. These indices are calculated to answer the following question: when treating a particular residue as a rigid body in the unbound state, would the protein in the unbound state behave similarly to the bound-state protein in the unperturbed simulation? In other words, can we make the unbound protein behave similarly to the bound protein by treating some residues as rigid bodies? The smaller the ΔARC value and the RCS index are for this comparison, the more similar the perturbed simulation of the unbound protein is to the unperturbed simulation of the bound protein.

Unperturbed Simulations of Unbound PDZ3 against Perturbed Simulations of Bound PDZ3. Similar to the comparison above, here one would ask whether the protein in the bound state could behave similarly to the unperturbed protein in the unbound state when a particular residue in the bound protein is treated as a rigid body. This comparison may identify those residues that could reverse the allosteric effect upon binding when they are treated as rigid bodies.

2.2.6. Unbound and Bound Difference Index. It is also interesting to ask the following question: for which residues is the difference between perturbed simulations of the unbound and bound proteins similar to the difference between unperturbed simulations of the unbound and bound proteins. In other words, for which residues, when being treated as rigid bodies, does the protein still exhibit allosteric effects similar to those observed in the unperturbed simulations? To identify such residues, we define the unbound and bound difference (UBD) index for residue i, D_i, to select residues in the rigid body scan where differences between unbound and bound PDZ3 resemble those from the unperturbed MD simulations:

$$D_i = |(C_i^{bound} - C_i^{unbound}) - (C_{unpert}^{bound} - C_{unpert}^{unbound})|$$
$$+ |R_i^{unbound\_bound} - R_{unpert}^{unbound\_bound}| \qquad (6)$$

In eq 6, $C_{unpert}^{unbound}$ and C_{unpert}^{bound} are the ARC indices from unperturbed simulations of unbound and bound PDZ3, respectively; $C_i^{unbound}$ and C_i^{bound} are the ARC indices from perturbed simulations of unbound and bound PDZ3, respectively, when residue i is treated as a rigid body; $R_{unpert}^{unbound\_bound}$ is the RCS index for the unperturbed simulations of unbound and bound PDZ3; and $R_i^{unbound\_bound}$ is the RCS index for the perturbed simulations of unbound and bound PDZ3 when residue i is treated as a rigid body. The first absolute-value term in eq 6 is designed to measure the similarity of the differences between the unbound and bound states obtained from the unperturbed and perturbed simulations using the ARC index, and the second term measures this similarity using the RCS index. Because the relative importance of the ARC and RCS indices for the purpose of this comparison is not yet well understood, these two terms have equal weight in eq 6.

One could understand the UBD index in the following way: The smaller the UBD index is for a given residue, the better the difference between unbound and bound states in the rigid residue scan resembles the difference between the unbound and bound states in the unperturbed simulation. In other words, making a residue rigid or dynamic would not alter the intramolecular communication changes upon binding events if the UBD index is very small for that residue. The degree of dynamics of this residue would not disrupt the whole-protein communication changes upon binding. It is noteworthy that the energy flow along allosteric pathways may cause significant dynamics changes of residues along the pathways; the least disruptive residues for the effects of binding on intramolecular communication are mostly suitable to carry out the propagation of energy. Therefore, we propose that the residues with small UBD indices are the key residues that may function as "wire residues" to propagate information/energy from a certain part of the protein to distal places in the protein. However, there are certainly many other residues that could serve as key residues to transmit energy/information but are not selected using the UBD index. For example, residues with large UBD indices may serve as part of different allosteric pathways depending on the environment of the protein. In this sense, the term "wire" may be considered mainly to differentiate the residues selected in the previous section.

3. RESULTS

3.1. Rigid Residue Scan for the Unbound State of PDZ3. Unbound PDZ3 was subjected to unperturbed and rigid residue scan MD simulations. The RMSD analysis of each trajectory (Table S1 in the Supporting Information) indicates that the PDZ3 structure is stable throughout the simulation. The unperturbed MD simulation of unbound PDZ3 has an ARC index of 0.742 and a standard deviation of 0.199 (Table 1) when all of the protein atoms are included in the analysis.

Table 1. Average Residue Correlation (ARC) Index and Its Standard Deviation σ_C

	unperturbed unbound		unperturbed bound	
	all-atom	C_α	all-atom	C_α
ARC	0.742	0.741	0.930	0.934
σ_C	0.199	0.189	0.055	0.051

The cross-correlation matrix for all of the protein atoms is depicted in a heat map (Figure 2A), and the histogram shows the distribution of correlation among the atomic pairs (Figure 2B). The ARC index is 0.741 with a standard deviation of 0.189 (Table 1) when only the C_α carbons along the peptide backbone are used. The heat map and histogram of the C_α cross-correlation matrix are shown in Figure 2C,D, respectively. It is clear that the analysis of only C_α is sufficient to represent the overall protein because of its similarity to the all-atom analysis.

The Supporting Information contains the RMSD plots (Table S1), ARC indices and their standard deviations (Table S3), and cross-correlation matrix heat maps and histogram plots for all 115 unbound rigid residue scan simulations (Table S2). The ARC indices are plotted against the residue numbers in Figure 3A. Intuitively one may expect that making part of the protein rigid would increase the overall rigidity. However, this is not the case. The ARC indices from the rigid-body simulations are rather evenly distributed from 0.2 to 0.9. The cross-correlation matrices of the rigid residue scan differ significantly. Sorted ARC indices using the standard deviations as error bars are plotted in Figure 3B, where red and blue points indicate hydrophobic and hydrophilic residues, respectively. The plot shows a uniform distribution of hydrophobic and hydrophilic residues, indicating that there is no correlation between the ARC indices and hydrophobicity. The ARC
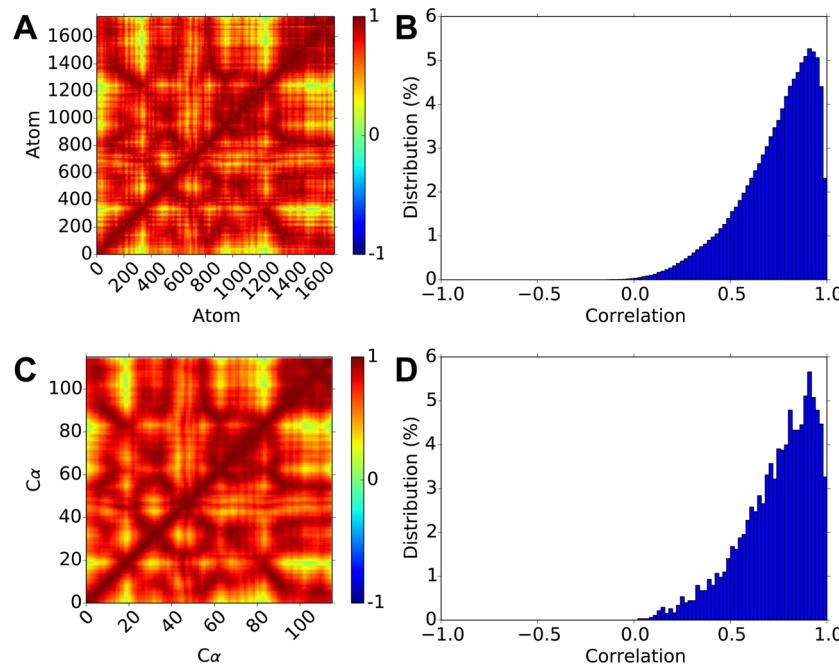
Figure 2. Cross-correlation matrices for the unperturbed MD simulation of unbound PDZ3: (A) heat map and (B) histogram using the all-atom cross-correlation matrix; (C) heat map and (D) histogram using the C_α cross-correlation matrix. In the heat maps, red, green, and blue indicate positive, none, and negative correlation, respectively. The histogram plots have a bin size of 0.2. The heat map color scheme and histogram bin size are consistent throughout the work.
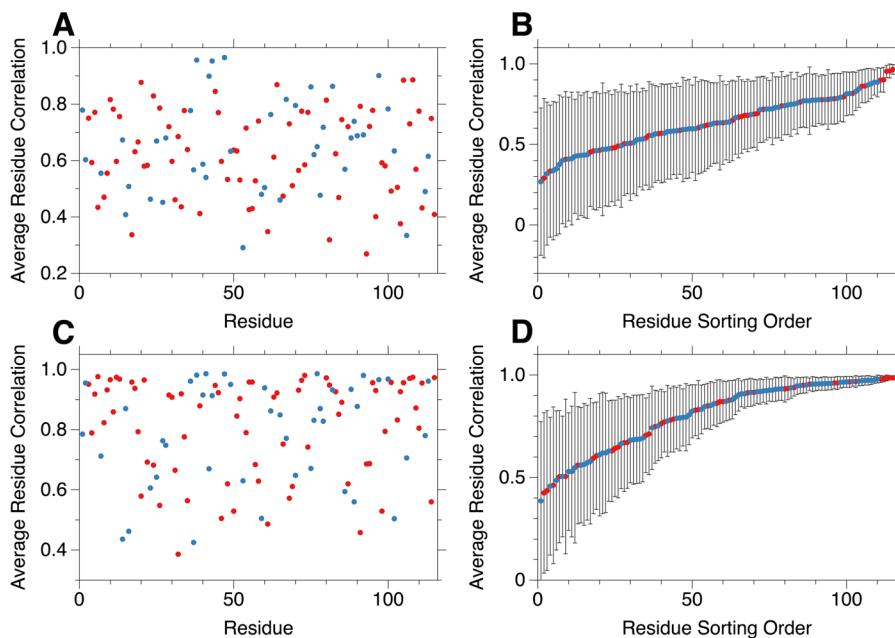
Figure 3. Average residue correlation (ARC) index for rigid residue scans of (A, B) unbound and (C, D) bound PDZ3: (A, C) rigid-body simulation ARC indices as functions of residue number; (B, D) sorted rigid-body simulation ARC indices as functions of sorting index, using the standard deviations as error bars. Red and blue points indicate hydrophobic and hydrophilic residues, respectively.

indices from 80 rigid residue scan simulations are smaller than the ARC index from the unperturbed simulation of unbound PDZ3 (see Table S3 in the Supporting Information). This indicates that making a residue rigid generally does not increase the overall positive correlation. The ARC indices for 12 residues from the perturbed simulations are at least 0.1 larger than that from the unperturbed simulation of unbound PDZ3. Ala47 has the greatest ARC index increase (0.224), while Ile38

and Ala43 are the other two residues with index increases larger than 0.2.

3.2. Rigid Residue Scan for the CRIPT-Bound State of PDZ3. The CRIPT-bound state of PDZ3 was also subjected to unperturbed MD simulations and 115 rigid residue scan MD simulations. The RMSD analysis of each trajectory (Table S1 in the Supporting Information) indicates that the complex is stable throughout the simulations. The all-atom ARC index from the unperturbed simulation of bound PDZ3 is 0.930 and
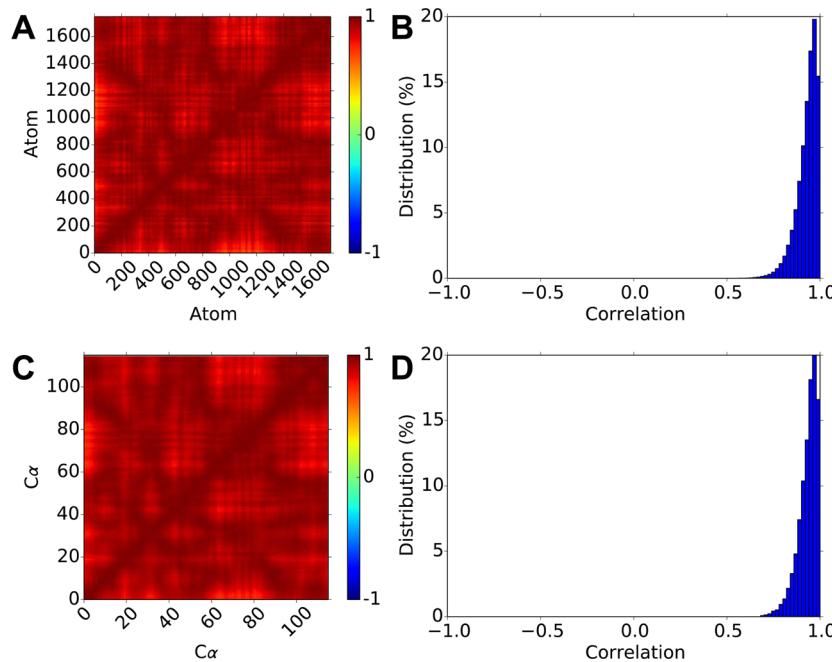
Figure 4. Cross-correlation matrices for the unperturbed MD simulation of bound PDZ3: (A) heat map and (C) histogram using the all-atom cross-correlation matrix; (C) heat map and (D) histogram using the C_α cross-correlation matrix. In the heat maps, red, green, and blue indicate positive, none, and negative correlation, respectively. The histogram plots have a bin size of 0.2.

Table 2. Residues Sorted by Either Residue Correlation Similarity (RCS) or Unbound and Bound Difference (UBD) Indices Based on Comparison of Rigid Residue Scans of Unbound and Bound PDZ3[a]

| | 1: perturbed unbound−perturbed bound[b] | | | 2: unperturbed bound−perturbed unbound[b] | | | 3: unperturbed unbound−perturbed bound[b] | | | 4 | |
no.	residue[c]	ΔARC[d]	RCS[e]	residue	ΔARC[d]	RCS[f]	residue	ΔARC[d]	RCS[g]	residue	UBD[h]
1	**47**[i]	−0.020	0.020	**38**	−0.022	0.026	66	−0.012	0.063	**29**	0.016
2	**38**	−0.026	0.033	**47**	−0.031	0.027	24	0.059	0.078	**90**	0.017
3	43	0.039	0.050	43	−0.019	0.029	4	−0.049	0.080	**72**	0.017
4	97	−0.065	0.052	42	0.035	0.043	8	−0.082	0.081	**100**	0.019
5	82	−0.069	0.055	97	0.033	0.046	110	−0.065	0.082	**36**	0.020

[a]C_α cross-correlation matrices are used. Only the top five rigid residue scan simulations are listed. See Table S5 in the Supporting Information for the complete list. [b]The first three lists are sorted using the RCS index. [c]In all four lists, the "residue" column indicates which residue is treated as a rigid body in the perturbed simulation. [d]In the first three lists, the average residue correlation index difference (ΔARC) is calculated as indicated in the list title. For example, for each residue in list 1 (perturbed unbound−perturbed bound), ΔARC is calculated as ARC(perturbed unbound) − ARC(perturbed bound). [e]For list 1, the RCS index is calculated using the cross-correlation matrices of unbound and CRIPT-bound PDZ3 when the indicated residue is treated as a rigid body in both perturbed simulations. [f]For list 2, the RCS index is calculated using the cross-correlation matrix of unbound PDZ3 when the indicated residue is treated as rigid body and the cross-correlation matrix from the unperturbed simulation of CRIPT-bound PDZ3. [g]For list 3, the RCS index is calculated using the cross-correlation matrix of CRIPT-bound PDZ3 when the indicated residue is treated as rigid body, and the cross-correlation matrix from the unperturbed simulation of unbound PDZ3. [h]In list 4, the UBD index is used to sort the residues. Please refer to section 2.2.6 for the definition of the UBD index. [i]Residues listed in bold type have been reported as important for PDZ3 allostery. See Table 5 for details.

its standard deviation is 0.055, while the C_α ARC index and standard deviation from the unperturbed simulation of bound PDZ3 are 0.934 and 0.051, respectively (Table 1). The all-atom cross-correlation matrix heat map and histogram plot are given in Figure 4A,B, respectively, and the C_α cross-correlation matrix heat map and histogram plot are given in Figure 4C,D, respectively. The small difference between the all-atom and C_α analyses indicates that C_α is sufficient to represent the overall protein dynamics. Therefore, all subsequent analyses were carried out using C_α cross-correlation matrices only.

The correlation matrix heat map and histogram plots show that there are significantly more positive correlations among residues in the bound state than in the unbound state. The bound ARC index (0.934; Table 1) is 0.193 greater than that of

the unbound state, indicating that binding with CRIPT makes PDZ3 more rigid. This can be understood as CRIPT restraining the movement of PDZ3 near the binding interface, which then increases the rigidity of the entire protein.

The Supporting Information provides RMSD plots (Table S1), C_α ARC indices and their standard deviations (Table S3), and C_α cross-correlation matrix heat maps and histogram plots for all 115 CRIPT-bound rigid residue scan simulations (Table S2). The ARC indices are plotted against the residue numbers in Figure 3C, while the sorted ARC indices using the standard deviations as error bars are shown in Figure 3D. As for the unbound state, the different rigid-body constraints significantly affect the overall protein dynamics. The bound-protein rigid residue scan ARC indices have a wide distribution of 0.386 to

Table 3. ARC Indices and Their Standard Deviations, σ_C, for Three Independent Simulations of Unperturbed Unbound and Unperturbed CRIPT-Bound PDZ3

state	1 ARC	1 σ_C	2 ARC	2 σ_C	3 ARC	3 σ_C	average ARC	average σ_C
bound	0.934	0.051	0.970	0.023	0.987	0.012	0.964 ± 0.027	0.029 ± 0.020
unbound	0.741	0.189	0.751	0.190	0.803	0.124	0.765 ± 0.033	0.168 ± 0.038

0.986 and overall are larger than the corresponding unbound indices, indicating a general increase in rigidity through complexation.

As in the unbound case, 80 rigid residue scan ARC indices are less than the ARC index from the unperturbed simulation of the bound protein, indicating that making a residue rigid does not increase the overall positive correlation of the bound state. Residues Ile41, Ala47, and Ile77 have ARC indices that are at least 0.05 larger than that from the unperturbed simulation of the bound state. It is noteworthy that when Ala47 is rigid, both the unbound and bound states are significantly more rigid than corresponding unperturbed states.

3.3. Comparison between Unbound and CRIPT-Bound States of PDZ3. *3.3.1. Unbound PDZ3 (in Rigid Residue Scans) Similar to Unperturbed Bound PDZ3.* It is interesting to examine whether any unbound rigid residue scan system behaves similarly to the unperturbed bound state. To this end, we use the RCS index instead of the ARC index. The RCS indices are below 0.1 for 13 rigid residue scan simulations of unbound protein relative to the unperturbed simulation of the bound state (see section 2 in Table 2 and Table S5 in the Supporting Information). Ala47 also stands out in this analysis with the second-smallest RCS index difference from the unperturbed bound state. The unbound rigid residue scan ARC indices for 12 of the top 13 residues differ from that of the unperturbed bound state by no more than 0.1, demonstrating good correlation between the RCS and ARC indices.

3.3.2. Bound PDZ3 (in Rigid Residue Scans) Similar to Unperturbed Unbound PDZ3. The RCS indices are less than 0.1 for 13 rigid residue scan simulations of bound protein relative to the unperturbed simulation of the unbound state (see section 3 in Table 2 and Table S5 in the Supporting Information). The ARC index differences for these 13 residues relative to that of the unperturbed unbound state are all less than 0.1.

3.3.3. Unbound PDZ3 Similar to Bound PDZ3 (Both in Rigid Residue Scans). The scenario in which the perturbed unbound and bound states with the same rigid residue behave similarly is interesting to examine. The RCS indices for the perturbed unbound and bound states are less than 0.1 for 14 rigid residue scan simulations (see section 1 in Table 2 and Table S5 in the Supporting Information). The ARC index differences between the perturbed unbound and bound rigid residue scan simulations for 12 of the top 14 residues are less than 0.1. The perturbed unbound and bound states where Ala47 is rigid have the smallest RCS indices. The top three residues (Ala47, Ile38, and Ala43) coincide with the top three residues in the comparison in section 3.3.1.

Another comparison of the unbound and bound rigid residue scan systems is to determine which residues cause the difference between the unbound and bound states in the perturbed simulations to resemble the difference between the unbound and bound states in the unperturbed MD simulations. The UBD index is used for this comparison. The top five residues with the smallest UBD indices are listed in section 4 in

Table 2. Gly29 has the smallest UBD index (0.016), and His72 and Ala90 have the second-smallest ones.

3.4. Consistency of the Simulations. Only one trajectory could be carried out for each rigid residue scan simulation because of limited computational resources. Two additional unperturbed simulations were carried out for the unbound and bound systems to estimate the consistency of the simulations. The ARC indices and corresponding standard deviations of all six trajectories are listed in Table 3. All three trajectories show consistent ARC indices and standard deviations for both the unbound and bound states. Only one trajectory from each unperturbed unbound and bound simulation was used for analysis for compatibility with the rigid residue scan. It is expected that the analysis results will remain similar when multiple trajectories can be carried out for all of the simulations and used for analysis.

Additional 10 ns MD simulations were carried out after those in the original scan for nine residues in the rigid residue scan simulations of unbound PDZ3. These residues were selected because their unbound ARC indices were significantly higher than that of the unperturbed MD simulation. The ARC indices and standard deviations are listed in Table 4. For three out of

Table 4. ARC Indices and Their Standard Deviations, σ_C, for Selected Simulations with an Additional 10 ns Simulation from Unbound PDZ3 Rigid Residue Scan

residue	first 10 ns ARC	first 10 ns σ_C	second 10 ns ARC	second 10 ns σ_C	first − second[a] ΔARC	first − second[a] $\Delta\sigma_C$
42	0.899	0.074	0.791	0.197	0.108	−0.123
43	0.953	0.038	0.614	0.247	0.339	−0.209
47	0.965	0.030	0.970	0.023	−0.005	0.007
64	0.869	0.103	0.763	0.172	0.106	−0.069
75	0.861	0.102	0.963	0.027	−0.102	0.075
82	0.863	0.105	0.980	0.017	−0.117	0.088
97	0.901	0.077	0.969	0.026	−0.068	0.051
105	0.885	0.079	0.649	0.264	0.236	−0.185
108	0.886	0.091	0.962	0.029	−0.076	0.062

[a]ΔARC and $\Delta\sigma_C$ are the difference between two simulations: (value for the first 10 ns) − (value for the second 10 ns).

these nine simulations (Ala47, Tyr97, and Ser108), both the ARC index difference and standard deviation difference for the two trajectories are less than 0.1. This demonstrates that a certain noise level is expected from the analysis results. The main purpose of the current study is to serve as a proof of concept. These results indicate that the proposed method can provide some useful information for individual residues with regard to their role in protein allostery.

To test whether saving snapshots every 1 ps is too frequent and leads to highly correlated snapshots, cross-correlation matrices were also generated using snapshots saved every 100 ps. The heat maps and histograms of the cross-correlation matrices generated from snapshots saved every 1 and 100 ps are plotted for the following simulations: unperturbed unbound,

unperturbed bound, perturbed unbound, and perturbed bound, the latter two with residue 1 being treated as a rigid body (Figures S1−S4 in Supporting Information). For all four simulations, the correlation heat maps and histograms based on snapshots saved every 1 and 100 ps are very similar to each other, indicating that saving snapshots every 1 ps is adequate for the purpose of this study.

4. DISCUSSION

4.1. Key Residues May Function as "Switches" for the Effect of Binding on Intramolecular Communication. The top five residues are Ala47, Ile38, Ala43, Tyr97, and Ala82 from section 1 in Table 2 (perturbed unbound−perturbed bound). The cross-correlation pattern of perturbed unbound PDZ3 is similar to that of the perturbed bound state when any of these residues is made rigid. Therefore, the perturbed simulations of the unbound and bound states display similar cross-correlation patterns when one of these five residues is treated as a rigid body, thereby eliminating the difference between the unbound and bound states, suggesting that the internal dynamics of these residues are important in the differentiation of the unbound and bound states. As proposed in section 2.2.5, we predict these residues to be key residues that function as "switches" for the effects of binding on intramolecular communication. Interestingly, four out of these five residues (Ala47, Ile38, Ala43, and Tyr97) are also among the top four residues in section 2 in Table 2 (unperturbed bound−perturbed unbound). This suggests that these "switch" residues may switch the unbound state (more flexible with less positive correlation) to the bound state (more rigid with more positive correlation).

Consistent with our results, these residues have been reported by various studies to be important for PDZ3 allosteric mechanisms (Table 5). Ala47 is a core residue buried inside the protein, and it is on the opposite side of the key residue His72 to which it is coupled. His72 is located at the binding interface

Table 5. Selected Residues That Have Also Been Reported in the Literature as Important for PDZ3 Allostery

References

ref 15 (Lockless and Ranganathan, *Science* **1999**, *286*, 295): multiple sequence alignment (MSA) of 274 PDZ domains

ref 33 (McLaughlin et al, *Nature* **2012**, *491*, 138): statistical coupling analysis (SCA) of 1578 mutations

ref 37 (Petit et al., *PNAS* **2009**, *106*, 18249): NMR relaxation of side-chain dynamics

ref 35 (Kaya et al., *Nucleic Acids Res.* **2013**, *41*, W249): Monte Carlo (MC) path generation calculations

Residue	Reported Significance
Gly29	ref 33: shows the largest average mutational effect in the whole protein
Ile36	ref 33: shows significant sensitivity to mutation
Ile38	ref 35: occurs in two populated allosteric pathways, also identified as plausible functional residue
Ala47	ref 15: coupled to key residue His72 ref 33: shows significant sensitivity to mutation.
His72	ref 15: determines the ligand specificity ref 33: most conserved residue in PDZ3
Ala90	ref 33: buried in the protein core, highly conserved, but shows little effect on mutation.
Tyr97	ref 37: enhances CRIPT binding through an entropy-driven mechanism ref 35: identified as a plausible functional residue
Phe100	ref 37: enhances CRIPT binding through an entropy-driven mechanism.

with CRIPT and determines the ligand's specificity.[15] (Ala47 and His72 are denoted as Ala51 and His76, respectively, in ref 15.) The distance between the C_α's of Ala47 and His72 is about 19 Å. Ala47 has been reported to show significant sensitivity to mutation.[33] Ile38 was reported as a key residue for two commonly reported PDZ3 allosteric pathways.[35] Tyr97 was reported as a plausible functional residue for allosteric pathways,[35] and it is also one of a group of potential phosphorylation target positions in PDZ3.[46] Phosphorylation is one of the most important biological switch mechanisms for protein signaling processes, and it is largely involved in protein allostery.

Neither Leu42 nor Ala43 has been reported as important for PDZ3 allosteric mechanisms. Interestingly, the additional 10 ns of rigid-body simulation for these residues led to lower ARC indices (Table 4), suggesting that the selection of these two residues may be noise. In future studies, longer simulation times will certainly help to reduce noise and produce more reliable predictions. Furthermore, Leu42 could be ruled out without additional simulations by combining perturbed unbound−perturbed bound and unperturbed bound−perturbed unbound ARC indices.

The rigid residue scan results for the key residues discussed, including the unbound- and bound-state cross-correlation matrix heat maps and histogram plots, are listed in Table 6. The heat map of the difference between the two cross-correlation matrices for each scan is also listed in Table 6.

None of the top residues listed in section 3 in Table 2 (unperturbed unbound−perturbed bound) were reported as important for PDZ3 allostery, indicating that these indices should not be used for selecting residues. The purpose of this column is to show which rigid residues may make the perturbed bound protein behave similarly to the unperturbed unbound protein. The rigid-body constraints need to be disruptive to make the perturbed bound protein behave similarly to the unperturbed unbound protein, as the unperturbed bound PDZ3 is more rigid with more positive correlation than the unperturbed unbound state. However, the rigid-body constraints on most residues are disruptive and lower the positive correlation within PDZ3. It is understandable that important residues could not stand out in this analysis. Therefore, it is not recommended to use the indices listed in section 3 in Table 2 (unperturbed unbound−perturbed bound) to select important residues for protein allostery.
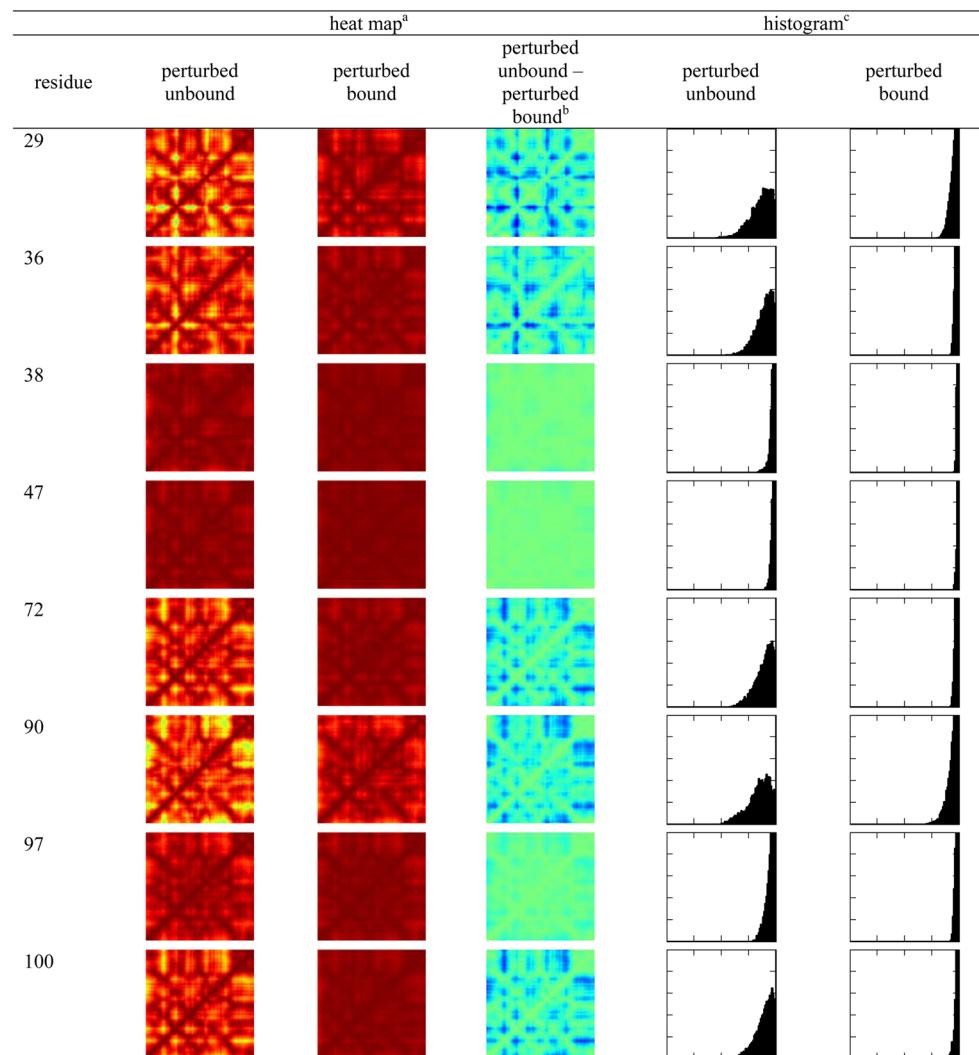
4.2. Key Residues May Function as "Wire Residues" for Intramolecular Communication. All of the top five residues using the UBD index (section 4 in Table 2) have been reported as important for PDZ3 allosteric mechanisms. Gly29 is at the top of the table with the smallest UBD index among all of the residues. Significantly, this residue has been reported to show the largest average mutational effect in the entire protein.[33] Gly29 is immediately adjacent to His72 with a $C_\alpha−C_\alpha$ distance of 6 Å. The residues His72 and Ala90 have the same UBD value. Both residues were reported in the previous studies of PDZ3 allosteric function (Table 5). His72 displays no tolerance to any mutations.[33] As mentioned in the previous section, His72 is located at the binding interface with CRIPT and determines the ligand's specificity.[15] (His72 is denoted as His76 in ref 15.) On the other hand, Ala90 is buried in the protein core and is highly conserved but shows little effect upon mutation.[33] Essentially all of the amino acid substitutions, including those with dramatic chemical differences, can be tolerated at this position. This observation strongly suggests

Table 6. Heat Maps and Histograms of C_α Cross-Correlation Matrices for Selected Residues in PDZ3 Treated as a Rigid Body in Both the Unbound and CRIPT-Bound States

residue	heat map[a]			histogram[c]	
	perturbed unbound	perturbed bound	perturbed unbound − perturbed bound[b]	perturbed unbound	perturbed bound
29					
36					
38					
47					
72					
90					
97					
100					

[a]In these heat maps, red, blue, and green denote positive, negative, and no correlation, respectively. [b]The "perturbed unbound − perturbed bound" heat map is plotted using the difference matrix between the cross-correlation matrices of unbound and CRIPT-bound PDZ3 when the target residue is treated as rigid body. [c]In the histogram plots, the normalized distributions are plotted using a bin width of 0.2.

that the spatial position of this residue, not the characteristics of its side chain, is important for the function of the protein. The fourth position in this section is Phe100, which is located on the C-terminal α-helix ($\alpha3$) of PDZ3. $\alpha3$ was reported to enhance CRIPT binding through an entropy-driven mechanism proposed on the basis of an NMR study.[37] Two aromatic rings on $\alpha3$ are suspected to be important for the entropy-driven binding enhancement (Phe100 and Tyr97). The fifth position in this section is Ile36, which is buried inside the protein and located on the one β-sheet. This residue was reported to show significant sensitivity to mutation.[33] The significance of the top residues ranked using the UBD index indicates that this index could be used as a reliable criterion to select important residues for protein allostery.

4.3. Overall View of Important Residues in the PDZ3 3D Structure. The key residues discussed in the previous two sections are illustrated in the three-dimensional (3D) structure of PDZ3 given in Figure 5. Three residues, Ile38, Ala47, and Tyr97, proposed to be important as "switch residues" in this study, are shown in blue. The spatial arrangement of these

residues suggests that they could participate in some pathways, although this has not been previously reported. Ile38 is located in the central part of a β-strand, Ala47 is located in the central part of an α-helix, and Tyr97 is located in the central part of another α-helix ($\alpha3$) in PDZ3. The analysis of the 3D structure of PDZ3 does not provide any obvious indication (e.g., disulfide bonds, salt bridges, etc.) of the allosteric importance of these residues. This supports the hypothesis that internal dynamics determines allosteric propagation. In addition, these residues may play an important role in dynamic communication along these secondary structures.

The "wire residues" Gly29, Ala90, His72, Phe100, and Ile36 are illustrated in red in Figure 5. It is clear that Gly29 is in direct contact with His72 with a C_α–C_α distance of 6 Å and that both are located near the CRIPT binding interface. Interestingly, Ala90, which is buried in the protein, is located at the opposite side of the protein and is far from the CRIPT binding interface. The distance between the C_α's of Ala90 and the central threonine residue of CRIPT is 17 Å. Also, the side chain of Phe100 is directly behind the backbone of Gly29 with
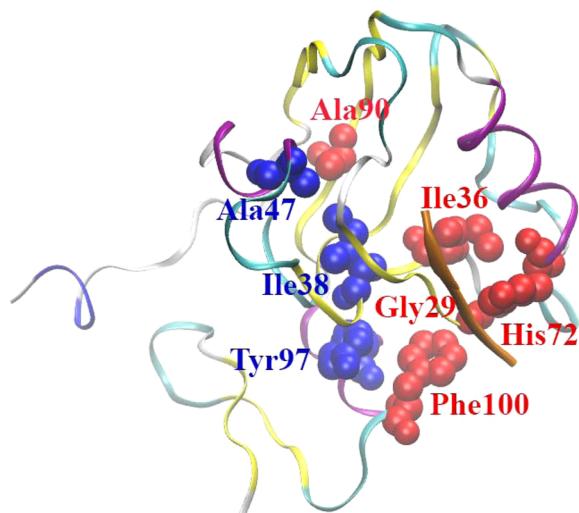
Figure 5. Structure of PDZ3 highlighting selected residues from the rigid residue scan analysis. The identified "switch residues" are illustrated in blue. The identified "wired residues" are illustrated in red.

a $C_\alpha-C_\alpha$ distance of around 8 Å. The residues Phe100, Gly29, and His72 are arranged linearly with a Gly29−His72 $C_\alpha-C_\alpha$ distance of around 6 Å. It is reasonable to suggest that these "wire residues" could participate in some allosteric pathways that are important for PDZ3 functions on the basis of this analysis and the spatial arrangement of these residues. The reported significance of these illustrated residues is summarized in Table 5.

4.4. Further Discussion of the Rigid Residue Scan Method. Overall it has been demonstrated that the rigid residue scan method is a robust tool to select potentially important residues for protein allostery without any a priori knowledge of protein allosteric mechanisms. For PDZ3, only one residue (Ala43) out of four residues selected using a combination of sections 1 and 2 in Table 2 (Ala47, Ile38, Ala43, and Tyr97) could be considered as a false positive. The top five residues selected using the UBD index have been reported in several experimental studies as highly important for PDZ3 allostery. This novel method has a success rate close to 90% and should certainly be considered in any mechanistic study of protein allostery.

In addition, the rigid residue scan method has the following advantages compared with other methods to probe protein dynamics. Compared with other scanning methods such as alanine scan, the rigid residue scan focuses on and can test the hypothesis that internal dynamics determine the allosteric propagation of information in a protein. All of the internal motions of the targeted residue are completely suppressed, but all physical interactions between the residue and its environment are conserved during the simulations. If one would mutate the target residue to a different type for simulation, one could argue that the different interactions between the mutated residue and its protein environment would lead to overall changes including those of protein dynamics and allostery. This would obscure the effect on the whole-protein dynamics by the internal motion of certain types of residues.

In the anisotropic thermal diffusion (ATD) method,[22] some knowledge about protein allostery is required a priori in order to select a hot spot for the simulation. In the original ATD study, the highly conserved His76 in PDZ3 (denoted as His72 in the present study) was selected on the basis of a previous study.[15] This method works only for the proteins with known effector or ligand binding sites. It does not provide a systematic way to probe all of the possible allosteric sites of a given protein. Residues that are potentially important for multiple allosteric pathways may not stand out through ATD analysis. In the ATD method, a substantial portion of the protein needs to be cooled to an extremely low temperature, with 10 K used in the original study. No large-scale and low-frequency protein motion important for overall protein allostery can be detected during the simulation at such low temperatures.

Another simulation method, pump−probe molecular dynamics (PPMD),[24] is similar to the rigid residue scan method in the sense that different atoms and residues can be selected for a given simulation. However, as stated in the article, the pumping motion of atoms circulates the Z axis with the same phase and direction of force solely because of its simplicity. In other words, a different setup of the pumping motion may lead to different results. In comparison, the application of rigid-body constraints to each residue in the rigid residue scan method, although unphysical as well, is unambiguous and unique and removes predetermined and arbitrary factors in the simulation.

It is challenging to predict the effect on protein dynamics from the rigid-body constraints of individual residues because of the limited knowledge of the relationships between individual residues and the overall protein dynamics. One can only find these out through the actual simulation. In the case of PDZ3, the constraints on most residues decrease the overall positive correlation inside PDZ3, making the protein more flexible. On the other hand, the rigid-body constraints on some residues increase the overall protein positive correlation, indicating that these residues are special in the sense that they may carry additional functionality for intramolecular communication through their internal motions. One could observe the change of overall protein dynamics and deduce information about potential protein allostery by switching off the internal motions of these residues. This situation is similar to protein mutagenesis experiments, where most mutations have unfavorable effects on protein structure, some are neutral, and some can stabilize the protein structure. One can only find out the results by carrying out an actual experiment.

5. CONCLUSIONS

In this study, we developed a systematic approach, rigid residue scan, to identify key residues for intramolecular communication under the influence of a binding event. In the rigid residue scan method, multiple MD simulations (one for each residue in the protein) were carried out for both unbound and bound proteins. In each of these simulations, one and only one residue was treated as a single rigid body. Cross-correlation analyses were carried out for each simulation to represent the intramolecular communication patterns. Several indices, including the average residue correlation (ARC), residue correlation similarity (RCS), and unbound and bound difference (UBD) indices, were proposed and applied to measure and compare cross-correlation matrices from different simulations. Key residues for intramolecular communication under the influence of ligand binding can be selected using these indices. It was hypothesized in this study that residues selected using different indices play two different roles in protein intramolecular communication. One role is "switches" for the effect of binding on intramolecular communication. When these "switches" are turned off (i.e., treated as rigid bodies to remove their internal dynamics), the unbound and

1698

dx.doi.org/10.1021/jp5083455 | J. Phys. Chem. A 2015, 119, 1689−1700

bound proteins will behave similarly. In other words, the binding effect can be "turned off" by these "switches". The other role is "wire residues" that propagate energy from one part of the protein to distal parts of the same protein. Eight out of the nine residues selected in this study were reported previously as important for PDZ3 allostery. Therefore, the rigid residue scan method is demonstrated to be an effective approach to identify key residues for protein intramolecular communication and potentially allosteric mechanisms.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Correlation heat maps and histograms generated from snapshots saved every 1 and 100 ps for selected simulations; RMSD plots, ARC indices and their standard deviations, RCS indices, UBD indices, and cross-correlation matrix heat maps and corresponding histogram plots for all of the simulations in rigid residue scan of unbound and CRIPT-bound PDZ3; sorted unbound and CRIPT-bound PDZ3 ARC, RCS, and UBD indices from rigid residue scan; complete lists for Table 2; and complete refs 36 and 40. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors
*E-mail: liuj@smu.edu.
*E-mail: ptao@smu.edu.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Monod, J.; Wyman, J.; Changeux, J.-P. On the Nature of Allosteric Transitions: A Plausible Model. *J. Mol. Biol.* **1965**, *12*, 88–118.

(2) Koshland, D. E.; Némethy, G.; Filmer, D. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits. *Biochemistry* **1966**, *5*, 365–385.

(3) Gunasekaran, K.; Ma, B. Y.; Nussinov, R. Is Allostery an Intrinsic Property of All Dynamic Proteins? *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 433–443.

(4) Tsai, C. J.; Nussinov, R. A Unified View of "How Allostery Works". *PLoS Comput. Biol.* **2014**, *10*, No. e1003394.

(5) Nussinov, R.; Tsai, C. J. Allostery in Disease and in Drug Discovery. *Cell* **2013**, *153*, 293–305.

(6) Fenton, A. W. Allostery: An Illustrated Definition for the "Second Secret of Life". *Trends Biochem. Sci.* **2008**, *33*, 420–425.

(7) Amaro, R. E.; Sethi, A.; Myers, R. S.; Davisson, V. J.; Luthey-Schulten, Z. A. A Network of Conserved Interactions Regulates the Allosteric Signal in a Glutamine Amidotransferase. *Biochemistry* **2007**, *46*, 2156–2173.

(8) Jault, J.-M.; Fieulaine, S.; Nessler, S.; Gonzalo, P.; Di Pietro, A.; Deutscher, J.; Galinier, A. The HPR Kinase from *Bacillus subtilis* Is a Homo-oligomeric Enzyme Which Exhibits Strong Positive Cooperativity for Nucleotide and Fructose 1,6-Bisphosphate Binding. *J. Biol. Chem.* **2000**, *275*, 1773–1780.

(9) Collier, G.; Ortiz, V. Emerging Computational Approaches for the Study of Protein Allostery. *Arch. Biochem. Biophys.* **2013**, *538*, 6–15.

(10) Zhuravleva, A.; Gierasch, L. M. Allosteric Signal Transmission in the Nucleotide-Binding Domain of 70-kDa Heat Shock Protein (Hsp70) Molecular Chaperones. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6987–6992.

(11) Tzeng, S.-R.; Kalodimos, C. G. Protein Dynamics and Allostery: An NMR View. *Curr. Opin. Struct. Biol.* **2011**, *21*, 62–67.

(12) Manley, G.; Loria, J. P. NMR Insights into Protein Allostery. *Arch. Biochem. Biophys.* **2012**, *519*, 223–231.

(13) Selvaratnam, R.; Chowdhury, S.; VanSchouwen, B.; Melacini, G. Mapping Allostery through the Covariance Analysis of NMR Chemical Shifts. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6133–6138.

(14) Zandany, N.; Ovadia, M.; Orr, I.; Yifrach, O. Direct Analysis of Cooperativity in Multisubunit Allosteric Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 11697–11702.

(15) Lockless, S. W.; Ranganathan, R. Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families. *Science* **1999**, *286*, 295–299.

(16) Smock, R. G.; Rivoire, O.; Russ, W. P.; Swain, J. F.; Leibler, S.; Ranganathan, R.; Gierasch, L. M. An Interdomain Sector Mediating Allostery in Hsp70 Molecular Chaperones. *Mol. Syst. Biol.* **2010**, *6*, 414.

(17) Tehver, R.; Chen, J.; Thirumalai, D. Allostery Wiring Diagrams in the Transitions That Drive the Groel Reaction Cycle. *J. Mol. Biol.* **2009**, *387*, 390–406.

(18) De Los Rios, P.; Cecconi, F.; Pretre, A.; Dietler, G.; Michielin, O.; Piazza, F.; Juanico, B. Functional Dynamics of PDZ Binding Domains: A Normal-Mode Analysis. *Biophys. J.* **2005**, *89*, 14–21.

(19) Gerek, Z. N.; Keskin, O.; Ozkan, S. B. Identification of Specificity and Promiscuity of PDZ Domain Interactions through Their Dynamic Behavior. *Proteins: Struct., Funct., Bioinf.* **2009**, *77*, 796–811.

(20) Gerek, Z. N.; Ozkan, S. B. Change in Allosteric Network Affects Binding Affinities of PDZ Domains: Analysis through Perturbation Response Scanning. *PLoS Comput. Biol.* **2011**, *7*, No. e1002154.

(21) Liu, J.; Nussinov, R. Allosteric Effects in the Marginally Stable Von Hippel–Lindau Tumor Suppressor Protein and Allostery-Based Rescue Mutant Design. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 901–906.

(22) Ota, N.; Agard, D. A. Intramolecular Signaling Pathways Revealed by Modeling Anisotropic Thermal Diffusion. *J. Mol. Biol.* **2005**, *351*, 345–354.

(23) Liu, J.; Tawa, G. J.; Wallqvist, A. Identifying Cytochrome P450 Functional Networks and Their Allosteric Regulatory Elements. *PLoS One* **2013**, *8*, No. e81980.

(24) Sharp, K.; Skinner, J. J. Pump–Probe Molecular Dynamics as a Tool for Studying Protein Motion and Long Range Coupling. *Proteins* **2006**, *65*, 347–361.

(25) Tao, P.; Wu, X.; Brooks, B. R. Maintain Rigid Structures in Verlet Based Cartesian Molecular Dynamics Simulations. *J. Chem. Phys.* **2012**, *137*, No. 134110.

(26) Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics Of *n*-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(27) Andersen, H. C. RATTLE: A "Velocity" Version of the SHAKE Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52*, 24–34.

(28) Vijayabaskar, M. S.; Vishveshwara, S. Interaction Energy Based Protein Structure Networks. *Biophys. J.* **2010**, *99*, 3704–3715.

(29) Lee, H. J.; Zheng, J. J. PDZ Domains and Their Binding Partners: Structure, Specificity, and Modification. *Cell Commun. Signaling* **2010**, *8*, 8.

(30) Basdevant, N.; Weinstein, H.; Ceruso, M. Thermodynamic Basis for Promiscuity and Selectivity in Protein–Protein Interactions: PDZ Domains, a Case Study. *J. Am. Chem. Soc.* **2006**, *128*, 12766–12777.

(31) Fuentes, E. J.; Der, C. J.; Lee, A. L. Ligand-Dependent Dynamics and Intramolecular Signaling in a PDZ Domain. *J. Mol. Biol.* **2004**, *335*, 1105–1115.

(32) Tochio, H.; Hung, F.; Li, M.; Bredt, D. S.; Zhang, M. Solution Structure and Backbone Dynamics of the Second PDZ Domain of Postsynaptic Density-95. *J. Mol. Biol.* **2000**, *295*, 225−237.

(33) McLaughlin, R. N., Jr.; Poelwijk, F. J.; Raman, A.; Gosal, W. S.; Ranganathan, R. The Spatial Architecture of Protein Function and Adaptation. *Nature* **2012**, *491*, 138−142.

(34) Niethammer, M.; Valtschanoff, J. G.; Kapoor, T. M.; Allison, D. W.; Weinberg, R. J.; Craig, A. M.; Sheng, M. CRIPT, a Novel Postsynaptic Protein That Binds to the Third PDZ Domain of PSD-95/SAP90. *Neuron* **1998**, *20*, 693−707.

(35) Kaya, C.; Armutlulu, A.; Ekesan, S.; Haliloglu, T. MCPATH: Monte Carlo Path Generation Approach To Predict Likely Allosteric Pathways and Functional Residues. *Nucleic Acids Res.* **2013**, *41*, W249−W255.

(36) Thorsen, T. S.; Madsen, K. L.; Rebola, N.; Rathje, M.; Anggono, V.; Bach, A.; Moreira, I. S.; Stuhr-Hansen, N.; Dyhring, T.; Peters, D.; et al. Identification of a Small-Molecule Inhibitor of the Pick1 PDZ Domain That Inhibits Hippocampal LTP and LTD. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 413−418.

(37) Petit, C. M.; Zhang, J.; Sapienza, P. J.; Fuentes, E. J.; Lee, A. L. Hidden Dynamic Allostery in a PDZ Domain. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18249−18254.

(38) Kong, Y.; Karplus, M. Signaling Pathways of PDZ 2 Domain: A Molecular Dynamics Interaction Correlation Analysis. *Proteins* **2009**, *74*, 145−154.

(39) Doyle, D. Crystal Structures of a Complexed and Peptide-Free Membrane Protein Binding Domain: Molecular Basis of Peptide Recognition by PDZ. *Cell* **1996**, *85*, 1067−1076.

(40) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(42) Andersen, H. C. Molecular Dynamics Simulations at Constant Pressure and/or Temperature. *J. Chem. Phys.* **1980**, *72*, 2384−2393.

(43) Nosé, S.; Klein, M. L. Constant Pressure Molecular Dynamics for Molecular Systems. *Mol. Phys.* **1983**, *50*, 1055−1076.

(44) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31*, 1695−1697.

(45) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(46) Ballif, B. A.; Carey, G. R.; Sunyaev, S. R.; Gygi, S. P. Large-Scale Identification and Evolution Indexing of Tyrosine Phosphorylation Sites from Murine Brain. *J. Proteome Res.* **2008**, *7*, 311−318.