

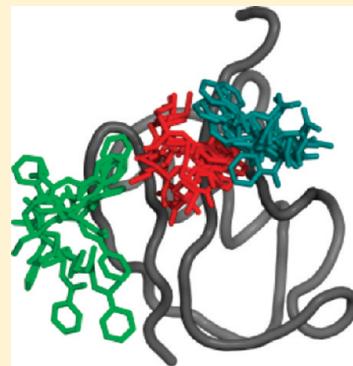
Molecular Simulations of Multimodal Ligand–Protein Binding: Elucidation of Binding Sites and Correlation with Experiments

Alexander S. Freed, Shekhar Garde, and Steven M. Cramer*

The Howard P. Isermann Department of Chemical and Biological Engineering, and Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, New York

 Supporting Information

ABSTRACT: Multimodal chromatography, which employs more than one mode of interaction between ligands and proteins, has been shown to have unique selectivity and high efficacy for protein purification. To test the ability of free solution molecular dynamics (MD) simulations in explicit water to identify binding regions on the protein surface and to shed light on the “pseudo affinity” nature of multimodal interactions, we performed MD simulations of a model protein ubiquitin in aqueous solution of free ligands. Comparisons of MD with NMR spectroscopy of ubiquitin mutants in solutions of free ligands show a good agreement between the two with regard to the preferred binding region on the surface of the protein and several binding sites. MD simulations also identify additional binding sites that were not observed in the NMR experiments. “Bound” ligands were found to be sufficiently flexible and to access a number of favorable conformations, suggesting only a moderate loss of ligand entropy in the “pseudo affinity” binding of these multimodal ligands. Analysis of locations of chemical subunits of the ligand on the protein surface indicated that electrostatic interaction units were located on the periphery of the preferred binding region on the protein. The analysis of the electrostatic potential, the hydrophobicity maps, and the binding of both acetate and benzene probes were used to further study the localization of individual ligand moieties. These results suggest that water-mediated electrostatic interactions help the localization and orientation of the MM ligand to the binding region with additional stability provided by nonspecific hydrophobic interactions.



I. INTRODUCTION

Recent advances in MM chromatography have generated significant interest for biopharmaceutical applications. Multimodal (or mixed-mode) (MM) chromatography methods simultaneously employ more than one mode of interaction between chromatographic ligands and solutes (proteins) (e.g., hydrophobic interactions and hydrogen bonding,¹ cation exchange and metal affinity,² etc.), thereby providing alternative and unique selectivities. Johansson et al. have developed libraries of MM ligands that can capture charged proteins even under high salt solution conditions,^{3–5} or development of resin materials capable of weak anion-exchange and reversed phase interactions for the simultaneous separation of acidic, basic, and neutral pharmaceutical compounds.⁶ Applications of MM chromatography have ranged from screening of biological mixtures and preparative protein purifications⁷ to front-end separations for mass spectrometry analysis^{8,9} and more.

Although significant effort has been made to enhance our fundamental understanding of protein interactions in ion-exchange^{10–17} and hydrophobic interaction chromatography,^{18–21} the current lack of fundamental molecular level understanding of the nature of interactions in MM systems has hindered their translation into becoming a mainstream tool.² The low affinity or “pseudo affinity” nature of multiple interactions involved in MM chromatography presents special challenges and questions to molecular level interpretations: For example, from a structural perspective, what

spatial regions of the protein and ligands are important in binding interactions? What are the major contributors to selectivity vs strength of binding? Do different interactions that contribute to binding act synergistically? These questions not only are fundamental to progress toward better design of ligands for specialized separations using MM chromatography, but they are also of interest in other problems in which protein–surface interactions play an important role.^{22,23}

We have recently employed homologous libraries of mutants of cold shock protein B (CspB)²⁴ and ubiquitin²⁵ to examine differences in their binding on ion-exchange (SP Sepharose FF) and MM chromatographic surfaces (Capto MMC). A majority of the mutant proteins showed stronger retention on Capto MMC relative to that on SP Sepharose FF, indicating stronger interactions of both proteins with multimodal ligands. More importantly, the retention time of ubiquitin on Capto MMC changed significantly in response to specific mutations, suggesting regions on the protein surface that were putatively involved in interactions with multimodal ligands. We also performed a complementary analysis of these interactions using NMR spectroscopy experiments on (labeled) wild type ubiquitin in aqueous solutions containing free multimodal ligands. Although the conformational

Received: April 24, 2011

Revised: September 1, 2011

Published: September 26, 2011

space (and geometric context) available for protein–ligand interactions is different when the ligand is free vs when it is tethered to a chromatographic surface, on the basis of previous work,^{3,26} we expected the free solution NMR study to provide qualitative, yet important, structural insights into protein–ligand binding. Indeed, the broader binding region with a concentration of sites on a single face of ubiquitin identified by NMR was consistent with that obtained using chromatography experiments.²⁵ Despite this overlap between chromatography and NMR spectroscopy experiments on ubiquitin, the picture of protein–multimodal-ligand interactions remains coarse-grained.

Identifying regions on protein surfaces where a ligand of specific chemistry binds is a challenging problem. Even when the ligand is simple (e.g., hydrophobic in nature), the context dependence of the hydrophobicity of the protein surfaces makes it especially challenging to predict the regions for specific proteins using simple (context-independent) hydropathy scales. When the ligand is multimodal, the problem is even harder. In addition, information about weak interactions such as those involved in protein–ligand binding in multimodal chromatography is difficult to come by experimentally.

The goal of this work was to explore whether MD-based calculations with explicit solvent provides a reasonable description of structural aspects of binding of multimodal ligands to proteins. In this work, we present an approach that complements NMR data for protein–ligand interactions with atomically detailed molecular dynamics (MD) simulations of ubiquitin and multimodal ligands in aqueous solution to obtain insights into the molecular structure and dynamics of these interactions. Analysis of MD simulations provides a clearer picture of the structural details of specific binding regions on the ubiquitin surface and dominant conformations of the interacting ligands, as well as flexibility and dynamics of those conformations. Spatially resolved binding of different chemical moieties of the MM ligand provides a picture of the importance of different interactions (e.g., electrostatic vs hydrophobic) in determining specificity of binding and its strength. By performing additional simulations of binding of separate chemical groups on the MM ligand (e.g., carboxylic, aromatic) to the protein surface, we also comment on the synergy of these interactions in overall binding. Collectively, this analysis presents a clearer picture of the ligand–protein binding process and the nature of pseudo affinity interactions in multimodal chromatographic systems. These results will have implications for improved *in silico* design of multimodal ligands with unique selectivities, and the combined approach of NMR with MD simulations will have significant potential for elevating the state of the art of understanding protein interactions in a wide range of areas, such as biomaterials and biosensors.

II. DETAILS OF MOLECULAR DYNAMICS SIMULATIONS

Atomically detailed simulations were performed using GROMACS²⁷ to characterize protein–MM-ligand binding. Simulations included one copy of the wild-type protein ubiquitin (pdb ID: 1d3z) solvated in a solution containing over 9000 water molecules, 6 copies of the multimodal ligand (based on the Capto MMC material), and counterions added to make the system electrically neutral (see Figure 1a). The protonation states of ionizable residues of the protein and ligands corresponded to a pH of 5 (the same as employed in the NMR experiments). At this pH, lysines (7 total), arginines (4), and histidine (1) are positively charged, whereas glutamic (6) and aspartic (5) acid groups are

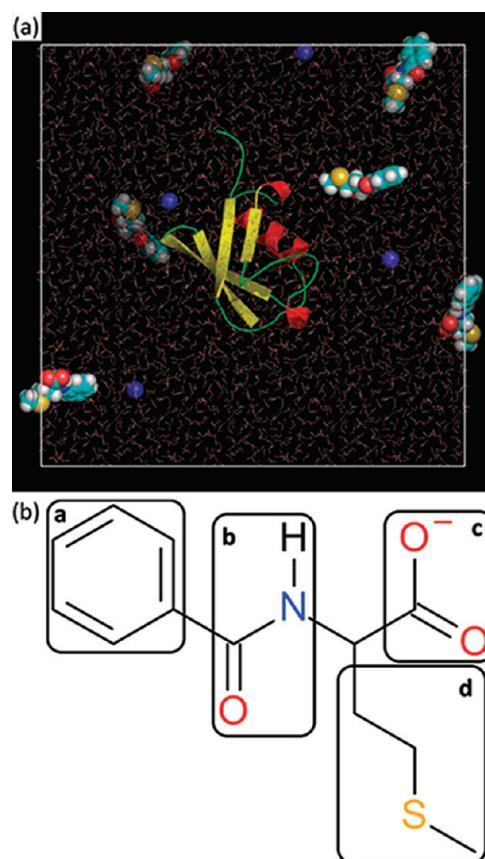


Figure 1. (a) A snapshot of the MD simulation system shows the protein ubiquitin (cartoon representation), water molecules (red and white wireframe), ligand molecules (spacefill), and sodium counterions (blue spheres). (b) Chemical structure of the multimodal ligand based on the Capto MMC resin. The ligand comprises four chemical groups as shown by boxes: (a) aromatic, (b) hydrogen-bonding, (c) electrostatic, and (d) aliphatic.

negatively charged, giving rise to a total charge on the protein of +1e. The carboxylic group of the MM ligand is negatively charged; therefore, five Na⁺ counterions were added to neutralize the system.

Our main objective was to follow the dynamics of the protein and ligands over a sufficiently long time during which ligands can effectively sample the space surrounding the protein, to observe binding and unbinding events, and to determine the locations of binding sites. The starting protein coordinates were taken from one of the NMR structures of ubiquitin (1d3z), and a truncated form of the multimodal ligand was used without the linker (see Figure 1b). Parameters for covalent (bond, angle, dihedral) and noncovalent (van der Waals, electrostatic partial charges) interactions for the protein were taken from the Amber94 force field.²⁸ Although parameters for the MM ligand are not available in the literature, those for its constituent groups are available in the Amber94 force field. Table 1 of the Supporting Information lists these parameters for completeness. Water molecules were modeled with the extended simple point charge (SPC/E) model, which represents water with three partial charge sites (one located on oxygen and two on the hydrogen atoms) and a Lennard-Jones interaction size located at the oxygen center.²⁹ (Note: although other classical force fields for protein and water are available, we expect the broad structural features of binding of binding, which

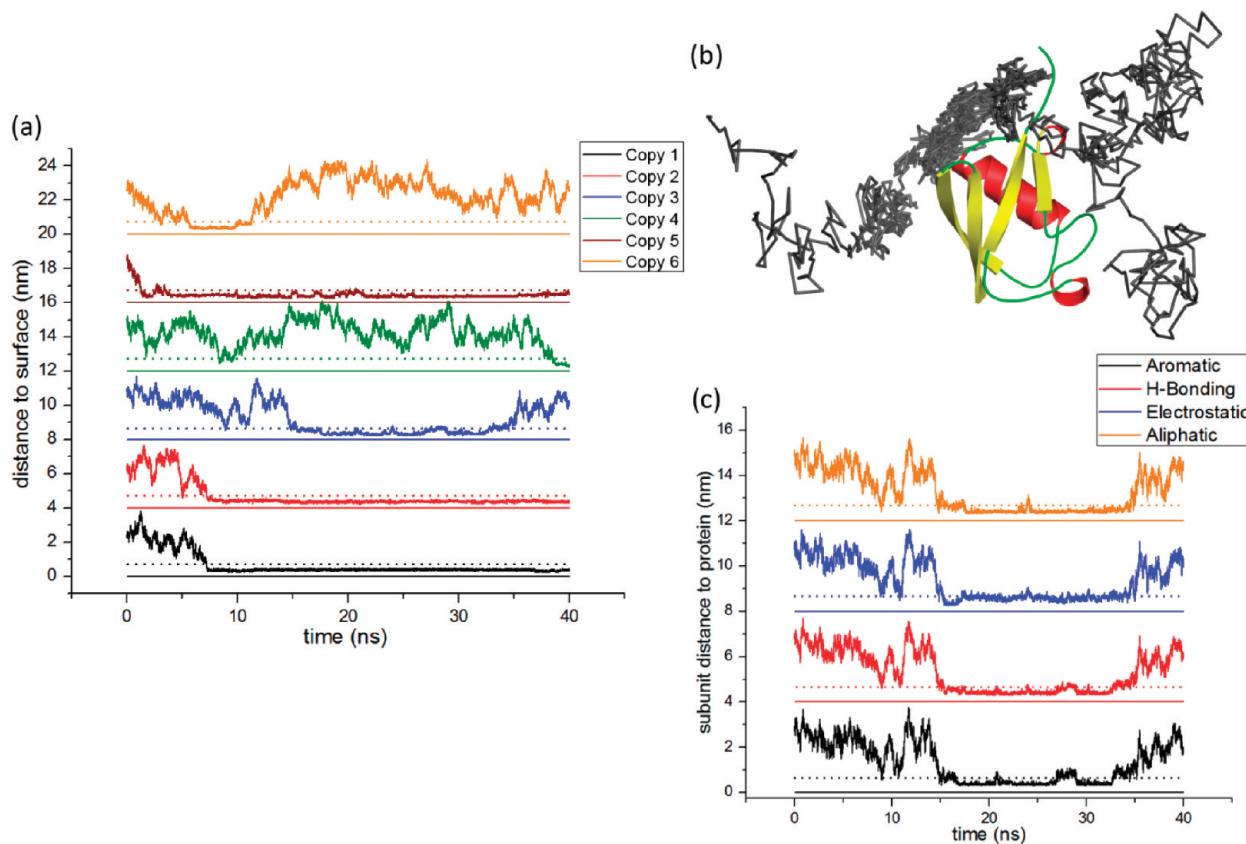


Figure 2. (a) Time dependence of the distance between ligand center of mass and the nearest heavy atom on the protein surface plotted for six copies of the MM ligand in one of the six 40 ns long MD trajectories. For clarity, profiles are shifted vertically by 4 nm. For reference, horizontal dotted lines show a threshold distance of 0.65 nm for each copy. (b) Trajectory of the COM of ligand 3 (i.e., copy no. 3 in A) shown pictorially relative to the ubiquitin molecule. The colors indicate progression in time going from blue to red. (c) Distances between COMs of four different ligand subunits and the nearest protein heavy atom for copy no. 3 in the trajectory shown in panel A. Horizontal dotted lines show a threshold distance of 0.55 nm for each subunit.

depend on direct and water-mediated protein–ligand interactions, to be similar in those models).

Simulations were conducted in the isothermal–isobaric ensemble at a temperature of 298 K and a pressure of 1 atm. The particle mesh Ewald (PME) method was used to calculate long-range electrostatic interactions. Covalent bonds containing hydrogen atoms were treated using the SHAKE algorithm.³⁰ An initial well-dispersed configuration was created for the system in which each of the six ligands was placed at least 2 nm from each other as well as from the protein, distributed randomly in the system. We performed six separate simulations with different initial distributions of the starting ligand placement and configurations. Each of the simulations was run for 40 ns, leading to a total simulation time of 240 ns. The first 1 ns of each run was treated as an equilibration run, leading to about 234 ns of production time for further analysis. Configurations were stored every 1 ps.

Analysis. To determine the distribution of ligands over the protein surface, distances between the ligand and protein atoms were calculated. The ligand was divided into four subunits (see Figure 1b), and various distances between the center of mass (COM) of each subunit and the nearest protein atoms were calculated. The entire ligand was considered to be “bound” if the distance between subunit COM and the nearest protein heavy atom was less than 5.5 Å for at least three subunits simultaneously. Protein heavy atoms were determined to be “interacting”

with the bound ligand if they were within 4.5 Å of any heavy atom of the ligand. This process was performed over the entire simulation, and analysis of the average number of ligand heavy atoms, $\langle n_i \rangle$, near different protein atoms was performed. Because the possibility of having two heavy atoms of two different ligands to be within 4.5 Å of a given protein atom is negligible, the maximum value of the average number described above is 1. All protein surface atoms showed only a fractional occupancy of their vicinity by the MM ligand. Finally, by tracking the movements of individual ligands, binding sites were extracted and compared with those determined by the NMR experiments.

III. RESULTS AND DISCUSSION

Characterizing protein–ligand binding events: The goal of our MD simulations is to allow sufficient sampling of the space surrounding the protein by ligands such that water-mediated protein–ligand interactions can be characterized. Figure 2a shows the distance between ligand COM and the nearest heavy atom on the protein surface for six different ligands in a 40 ns MD trajectory (which is one of six different trajectories). Figure 2b correspondingly shows the spatial trajectory of copy no. 3 of the ligand relative to the protein. It is clear that the ligand samples the region near the protein, coming within the threshold distance and staying within that distance for a period of time before leaving the protein surface. Other ligands also show a similar behavior (some binding

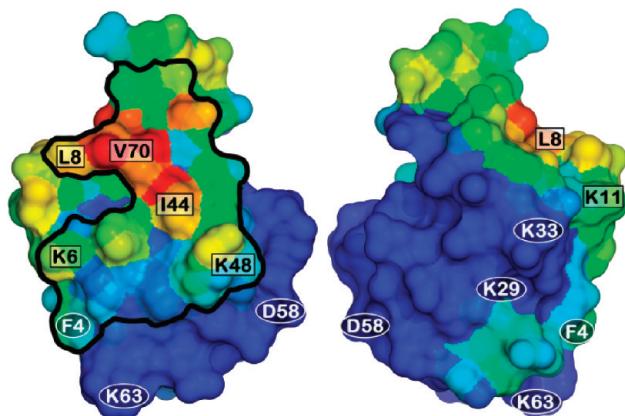


Figure 3. Analysis of ligand binding averaged over 234-ns-long MD simulations overlapped onto results of the chromatography and NMR experiments. Average ligand subunit counts in the vicinity of different protein atoms are color-coded using a rainbow scheme from red (high, 0.4) to blue (low, ~ 0). (Left image, front view; right image, back view). Residues that affect retention time in chromatography are shown in boxes; those that do not are shown with ovals around them. Binding region identified by NMR is within the solid black contour.

for a longer time and others for a much shorter time). Figure 2c shows a typical time dependence of the distance between four different subunit COMs and the protein surface for copy no. 3 of the ligand. When COMs of three or more subunits of a given ligand are simultaneously within 0.55 nm of the protein surface, the ligand is considered to be “bound”. In such bound states, we identified protein heavy atoms (those within 0.45 nm of any heavy atom of the ligand) that participated in binding to the ligand. We note that the choice of 0.55 and 0.45 nm for the above two cutoff distances was used on the basis of van der Waals radii of various heavy atoms, such that direct protein–ligand contacts are correctly identified, but solvent-separated contacts are excluded.

Figure 3 shows the results from MD simulations averaged over a 234-ns-long trajectory (six parallel simulations) overlapped onto results from the chromatography and NMR experiments. Average ligand counts in the vicinity of different protein atoms are color-coded using a rainbow scheme from red (high) to blue (low). Although this representation of results lacks details of ligand orientation or that of a binding site (see the next subsection), it highlights regions on the protein surface near which there is a larger probability of finding (some) part of a bound ligand. Surface residues that were found to affect retention time in chromatographic experiments (and therefore considered to be part of the binding region)—L8, K6, K11, I44, K48, and V70—are shown with rectangular boxes around them, whereas those that do not affect retention time—F4, K29, K33, D58, and K63—are shown with ovals.²⁵ The contiguous region identified by NMR experiments as potentially important in binding is enveloped by a black line.²⁵ The overall agreement between the three methods is reasonably good; the distribution of ligands obtained from the MD simulations also indicates that there is a clear preferred binding region on the protein surface.

Although both MD and NMR focus on interactions in free ligand systems, it is difficult to compare their results quantitatively with each other (e.g., average counts in MD with KD values in NMR). First, ligand and protein concentrations in MD simulations are 29 and 5 mM, respectively, which are higher than in the NMR study. Second, the color map of the protein surface represents

an average over data from six ligands. This allows for overlapping binding site regions, resulting in higher probabilities for specific sites at the intersection of such sites. In contrast, the results obtained from NMR are determined exclusively from noncompeting sites. This leads to some differences in the results, as discussed in the next section. Despite these differences, the agreement in Figure 3 is noteworthy.

Determination of Binding Sites on the Protein Surface. By examining the trajectory of each ligand in the simulations, it is possible to determine specific binding regions or sites on the protein surface. We note that the protein–ligand interaction is not a high affinity one, and the term “binding site” here simply refers to the region or collection of residues on the protein surface that most closely interact with the ligand when at least three of the four ligand subunits are in contact with the protein, that is, when the ligand is “bound”. Because ligands explore different parts of the protein surface, several such binding sites are observed, with some residues that are common between different binding sites. Below, we discuss residues involved in these sites, compare them with results from the NMR spectroscopy, and highlight the underlying interactions that appear to be common to most of the binding sites.

The major binding sites determined from the MD analysis are shown in Figure 4b along with the corresponding NMR data in Figure 4a. The NMR data include residues that were associated with specific binding sites (colored similarly) as well as residues that exhibited multiphasic behavior; that is, they were involved in more than one binding site. Site 1 obtained using MD simulations (Figure 4b) includes residues 6, 8, 68, 44, and 70, which matches well with the high-energy NMR binding site composed of residues 6 and 44 as well as the adjoining multiphasic residues 68, 8, and 70. Site 2 contained residues 44, 45, 66, and 68, which corresponded to the NMR binding site containing 66 as well as the adjoining multiphasic residues 45 and 68. Site 3, which included residues 44, 46, 47, 48, and 49, corresponded to the NMR site centered on residues 46 and 47 as well as adjoining multiphasic residues. The MD results for this site also included residues 44 and 49, which were found to be associated with other sites in the NMR experiment. Site 4 included residues 42, 72, 73, 75, and 76, which corresponded to the NMR site defined by residues 42 and 73. Note that although the NMR results did not include residues 72, 75, and 76, binding site 4 identified from the MD makes logical sense on the basis of the size and shape of the ligand.

Although site 5 (residues 8, 10, 70, 71, 74) did not correspond to a specifically defined site from the NMR, it was proximal to two strong binding sites (1 and 4) that were identified and included residues from these sites. Furthermore, although site 6 (residues 4, 6, 10, 11, 12, 14) did not correspond to a clearly defined site from the NMR, it involved residues 6 and 4, which were identified from the NMR along with adjacent residues.

We note that MD/NMR comparison is qualitative for several reasons. Both methods contain methodological issues: NMR (signal-to-noise ratio) may not lead to sufficient signal to pick up low affinity interactions, whereas methodological issues in MD are related to sufficient sampling, accuracy of the force field, and solution conditions. To obtain sufficient data, we included a higher concentration of ligands in MD relative to that employed in the NMR studies, which will lead to differences, as well. Nevertheless, the agreement above is encouraging.

There were two sites observed in MD (sites 7 and 8) that were not detected by NMR (Figure 5a). One of these (site 7) had a somewhat low population density relative to the other (site 8),

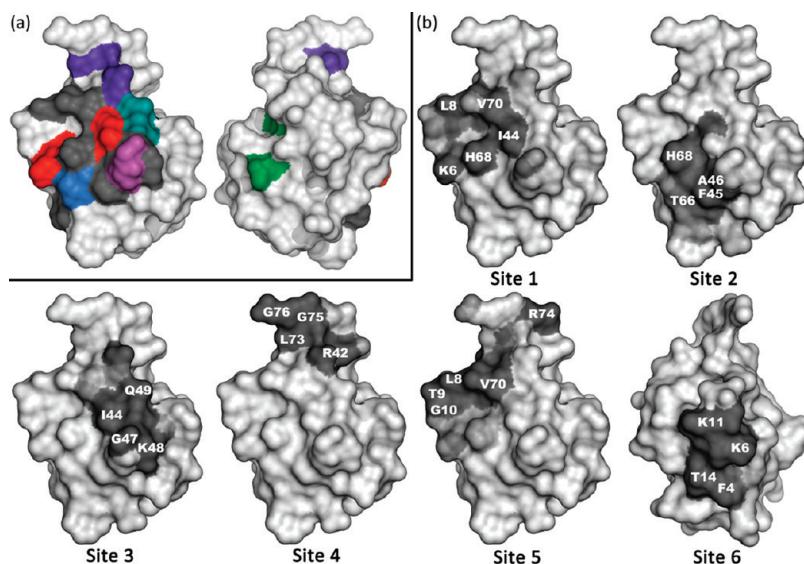


Figure 4. Binding sites obtained from (a) NMR experiments (reported elsewhere²⁵) (note: residues that participated in the same binding site are colored similarly); and (a) individual ligand trajectories extracted from the MD simulations.

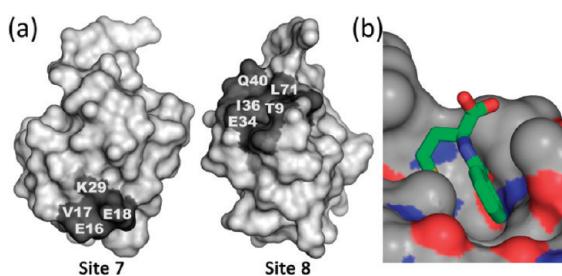


Figure 5. (a) Two additional binding sites determined from the MD simulations that did not match the NMR sites. (b) The position of the multimodal ligand at binding site 8. The MM ligand is shown in stick representation and the vicinal protein is shown using a surface representation.

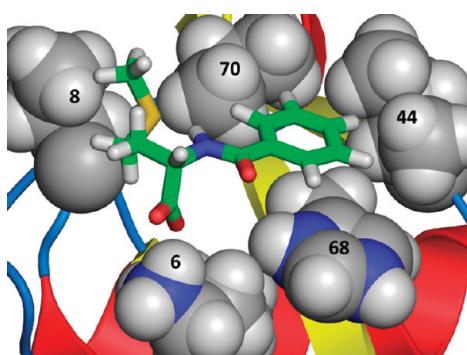


Figure 6. The multimodal ligand adsorbed at the binding site formed by residues 6, 8, 44, 68, and 70. The MM ligand is shown with sticks (white, hydrogen; green, carbon; and blue, nitrogen), and the interacting residues are shown in spacefill representation (gray, carbons; blue, nitrogen; and white, hydrogens).

indicative of a weaker binding. Site 8 involved primarily hydrophobic interactions, as can be seen in Figure 5b. In addition, this site on the protein surface was made up entirely of hydrophobic residues. This is in sharp contrast with the other sites, which

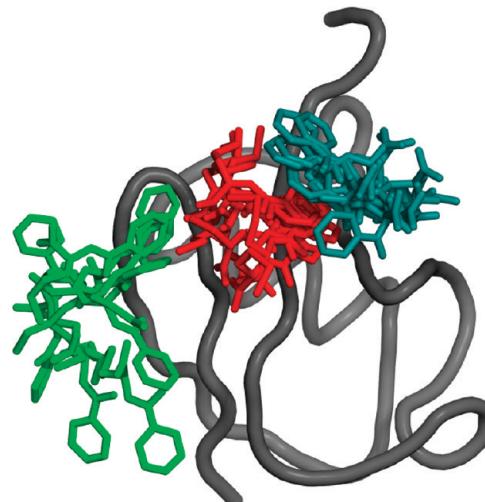


Figure 7. Conformational flexibility of ligands in the vicinity of binding sites 6, 1, and 4 (from Figure 4b), indicated in green, red, and cyan, respectively.

involved a combination of electrostatic and hydrophobic interactions. Identification of site 8 in MD is likely the result of the high ligand concentration employed in the simulations. Interestingly, when NMR experiments were carried out at high salt (1 M) concentrations (results not shown), these two sites were identified. This result is consistent with the increased strength of hydrophobic interactions at elevated salt concentrations.³¹

It is of interest to examine the specific interactions involved between the multimodal ligand and the protein surface. Figure 6 shows a snapshot of the multimodal ligand adsorbed to binding site 1 on the surface of ubiquitin. Several protein side chains (spheres) are highlighted to show their interactions with the multimodal ligand (sticks). The multimodal ligand has a direct electrostatic interaction with Lys 6 and a hydrogen bond with His 68. This adsorption is further aided by interactions with a large hydrophobic surface area on the protein. The aromatic part of the

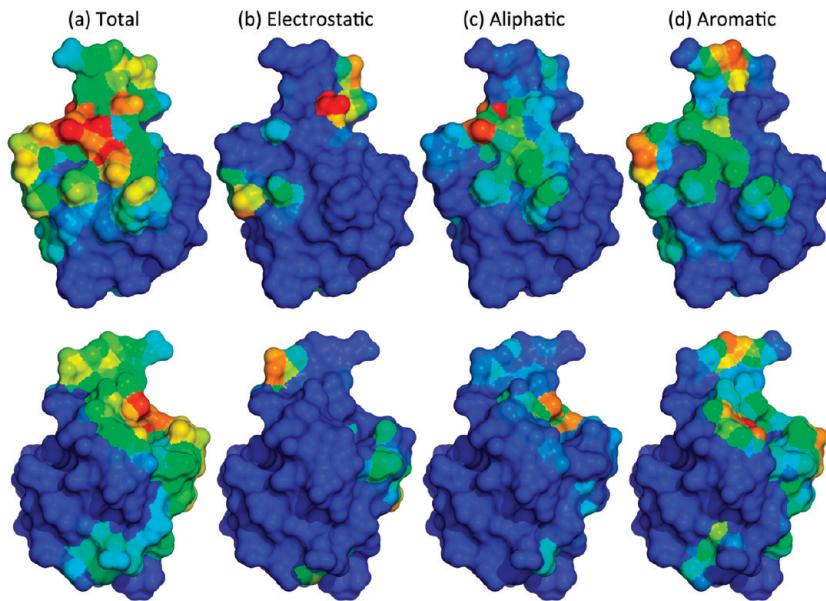


Figure 8. MD results for different contributions to protein–ligand binding. The locations of binding of the (b) electrostatic, (c) aromatic, and (d) aliphatic subunits of the MM ligand are indicated along with the total. The protein surface color is determined by $\langle n_i \rangle$ value for each surface atom with blue = low and red = high. The color scheme for panels b, c, and d are normalized to enable visual comparison.

ligand is able to interact with a hydrophobic pocket created among residues 70, 44, and 68. In addition, the aliphatic chain of the ligand is able to thread itself between residues 8 and 70. This hydrophobic interaction added to the charge–charge interaction and hydrogen bonding acts synergistically, which leads to the high affinity for this site determined in the NMR experiments.

This comparison between the MD and NMR results showed relatively good agreement for the determination of binding site locations. Importantly, both methods identified the same general interaction region of the protein surface. This is important both for proving the model's effectiveness and for gaining a better understanding of the NMR results.

Ligand Binding Dynamics. Figure 7 shows the configurations of the ligands (at 1 ns intervals) in the vicinity of three distinct binding sites (sites 6, 1, and 4) taken from a 10 ns piece of a MD trajectory. The figure highlights the conformational flexibility of the ligands near these binding sites. For the ligand near site 6 (shown in green), there is significant movement of the ligand in the vicinity of this binding site. In contrast, for sites 1 and 4, the ligand conformations appear more constrained. The relative conformational homogeneity of ligands near site 1 may be driven by the aromatic stacking with the ligand bound to site 4. This conformational behavior is consistent with the enhanced KD value seen in the NMR experiment for this binding site. Although ligand conformations near sites 1 and 4 appear homogeneous, they still exhibit some dynamics near the binding sites. Conformational flexibility near the binding site is expected to be important for stabilizing the protein–resin complex when the ligand is immobilized on a surface, since it would allow for the protein to have interactions with the surface, even if the immobilized ligands were sterically hindered.

Subunit Analysis. We further analyzed the simulation trajectories to understand how different subunits of the ligand shown in Figure 1b interact with the protein. Protein surface atoms are color-coded in Figure 8a, b, c, and d using $\langle n_i \rangle$ counts for total (i.e., all), charged, aliphatic, and aromatic subunits of the ligand,

respectively. The figure thus depicts the locations where the particular subunits of the ligand are bound to the protein surface.

We observe distinct preferences between different ligand subunits. As expected, Figure 8b, which shows the binding sites for the carboxylate subunit, displays a high binding density near the arginine, lysine, and histidine residues. In particular, this subunit was found to be located at distinct regions on the periphery of the preferred binding face of the protein. Although there is significant overlap between binding regions of the aromatic and aliphatic subunits (Figure 8c and d), which would be expected, they also seem to sample different regions on the protein surface. The aliphatic subunit clusters in a single highly hydrophobic pocket while the aromatic part of the ligand is able to sample a wider area. This could be a result of the difference between the aromatic and aliphatic molecular properties or steric limitations. The inability to separate these multiple effects at present makes it difficult to draw quantitative conclusions from the different binding distributions.

To understand ligand binding preferences quantitatively, in future work, we plan to perform systematic calculations of the free energy of binding of each ligand subunit separately to different regions of the protein surface. Such an analysis will also address the questions about the synergy of the interactions and many-body effects quantitatively. In the absence of such an analysis, qualitative insights can be obtained into the physics of subunit interactions by examining the nature of the protein surface from electrostatic and hydrophobic perspectives.

Analysis of Protein Surface Properties. To examine binding of the negatively charged carboxylate subunit, we calculated the electrostatic potential (EP) near the protein surface using APBS (adaptive Poisson–Boltzmann solver). We also performed additional simulations of the protein in water containing acetate ($\text{CH}_3\text{—COO}^-$) solutes and monitored the binding of this moiety to the protein. Figure 9a shows the EP surface for ubiquitin, with positive EP denoted in blue and negative EP, in red. There is a high concentration of positive EP on the front face of the protein among residues 6, 72, and 48, whereas the back face shows mixed

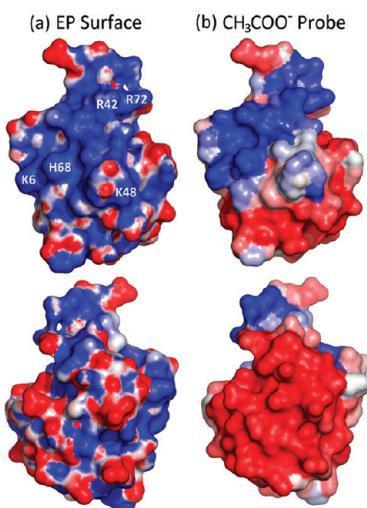


Figure 9. The (a) electrostatic potential surface of ubiquitin (blue = positive, red = negative) and (b) $\langle n_i \rangle$ values for binding of an acetate probe molecule (blue = strong binding, red = weak). Note that the binding strength color convention for acetate is reversed to keep it consistent with standard electrostatic potential maps.

EP with positive and negative regions interspersed among each other. Figure 9b shows the results of the MD simulation using the acetate probes. The probes interact favorably with residues 6, 68, 42, and 72, which is consistent with the high positive electrostatic potential on the front surface of the protein. In contrast, very little binding is observed on the back face with the mixed EP color map. Thus, the contiguous region of positive EP on the front top half of the protein appears to be an important contributor to the binding of the acetate moiety. Interestingly, this is the same general region favored by the MM ligand, as indicated by NMR and chromatography experiments.

We also characterized the hydrophobicity map of the protein surface. Granick and Bae have articulated the difficulty of characterizing hydrophobicity of nanoscopic objects for which macroscopic contact angle measurements are not feasible.³² We have recently shown that water density fluctuations or, alternatively, binding of probe hydrophobic solutes provide robust molecular level characterization of hydrophobicity.^{33,34} Further, we have shown that such a characterization captures the context dependence of hydrophobicity, especially for chemically heterogeneous surfaces, which the traditional (context independent) hydropathy scales are likely to miss. To this end, we performed MD simulations of ubiquitin in an aqueous solution containing benzene (subunit a in Figure 1c) as a probe of hydrophobicity. The probability of observing a benzene center of mass in the vicinity of the protein characterizes the “hydrophobicity” of the protein surface to the extent that it is relevant to the binding of subunit a.

Figure 10a and b compares the hydrophobicity maps of the ubiquitin surface obtained by applying the Kyte–Doolittle hydrophobicity scale³⁵ with that from the benzene probe MD method. The residues of high hydrophobicity (per Kyte–Doolittle scale) are shown in red, and those with low hydrophobicity, in blue. In the benzene probe density results, areas of high probe density (hydrophobic) are shown in red, and the low levels of probability density are in blue. It is clear that although there is not an exact agreement between the two methods, there is similarity in the hydrophobic binding areas. The differences highlight the role

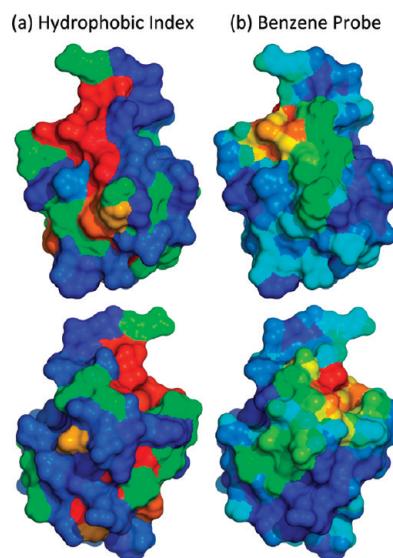


Figure 10. Analysis of the surface hydrophobicity of ubiquitin. (a) Protein surface color-coded using the Kyte–Doolittle scale,³⁵ and (b) using benzene probe binding. Red indicates more hydrophobic residues in part a and strong binding of benzene in part b.

that local context (topography and chemistry) plays in the binding of hydrophobic probe solutes. The aromatic probe binds weakly to a large part of the protein surface, with specific binding hot spots in a few selected areas (red). Our studies of binding aliphatic probes to other proteins also show a similar general behavior.

These results indicate that the hydrophobic parts of the MM ligand provide an additional thermodynamic driving force for ligand binding. The primarily nonspecific nature of hydrophobic interactions allow the MM ligand to sample several conformations that satisfy both electrostatic and hydrophobic interactions. In addition, the MM ligand's ability to span between the charged sites and the hydrophobic hot spots can lead to the high binding in specific areas, such as would be expected in synergistic binding events.³⁶ These results suggest that although the hydrophobicity of the protein is sufficient in several surface locations, it is the electrostatic interactions that determine the ligand localization.

IV. CONCLUSIONS

In this paper, we employed MD simulations to examine the binding of a multimodal ligand in free solution to the protein ubiquitin by calculating the probability distribution of ligand binding on the protein surface. Importantly, a comparison between the MD and previous NMR results indicated that both methods selected the same preferred binding region on the protein surface. Further, by following the trajectories of individual ligands over the course of the simulation, binding sites identified by the MD simulations could be compared with those from NMR. This use of MD simulations in concert with NMR offers an important new approach for studying protein–ligand and protein–surface interactions.

An analysis of the ligand binding dynamics indicated that even for a strong interaction site, the ligand was sufficiently flexible to access a number of favorable configurations. This behavior is expected to play an important role in solid phase multimodal chromatographic systems in which multiligand and avidity effects

often occur. A comparison was also performed to study the localization of the individual ligand moieties and the protein surface properties (electrostatic potential and hydrophobicity). These results in concert with data obtained with benzene and acetate probes indicated that water-mediated electrostatic interactions are more important for the localization of the MM ligand to the binding region, with additional stability provided by a combination of local interactions, such as hydrophobicity and hydrogen bonding. This analysis gives a clearer picture of the ligand–protein binding process and which protein surface factors govern the localization of the ligand to a specific binding site. Although the present manuscript focuses on binding phenomena in multimodal chromatography, water-mediated interactions are broadly relevant in many biological systems of interest to physical chemists, and thus, this work is expected to have implications for a range of systems in which protein–ligand interactions can play an important role (e.g., biomaterials, drug delivery, drug design, etc.).

Future work will examine more quantitatively aspects of synergy of multimodal interactions in protein–ligand binding and will study protein binding to surfaces that present a range of immobilized MM ligands. We will also examine the effects of various fluid phase additives that have been shown to dramatically enhance selectivity in these systems.

■ ASSOCIATED CONTENT

Supporting Information. The Supporting Information provides details about the force field parameters (atom type and partial charge) used to describe the ligand atoms in the simulation. This information is available free of charge via the Internet at <http://pubs.acs.org/>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: crames@rpi.edu.

■ ACKNOWLEDGMENT

This work was supported by National Science Foundation Grant CBET 0933169.

■ REFERENCES

- (1) Shen, Y.; Shao, X.; O'Neill, K.; Bradshaw, J. S.; Lee, M. L. *J. Chromatogr. A* **2000**, *866*, 1.
- (2) Gagnon, P.; Cheung, C. W.; Yazaki, P. *J. J. Sep. Sci.* **2009**, *32*, 3857.
- (3) Liu, F. F.; Dong, X. Y.; Wang, T.; Sun, Y. *J. Chromatogr. A* **2007**, *1175*, 249.
- (4) Johansson, B. L.; Belew, M.; Eriksson, S.; Glad, G.; Lind, O.; Maloisel, J. L.; Norrman, N. *J. Chromatogr. A* **2003**, *1016*, 35.
- (5) Johansson, B. L.; Belew, M.; Eriksson, S.; Glad, G.; Lind, O.; Maloisel, J. L.; Norrman, N. *J. Chromatogr. A* **2003**, *1016*, 21.
- (6) Liu, X. D.; Pohl, C. *Am. Lab.* **2007**, *39*, 22.
- (7) Gao, D.; Yao, S. J.; Lin, D. Q. *J. Appl. Polym. Sci.* **2008**, *107*, 674.
- (8) Bicker, W.; Lammerhofer, M.; Lindner, W. *Anal. Bioanal. Chem.* **2008**, *390*, 263.
- (9) Boersema, P. J.; Divecha, N.; Heck, A. J. R.; Mohammed, S. *J. Proteome Res.* **2007**, *6*, 937.
- (10) Brooks, C. A.; Cramer, S. M. *AIChE J.* **1992**, *38*, 1969.
- (11) Weaver, L. E., Jr.; Carta, G. *Biotechnol. Prog.* **1996**, *12*, 342.
- (12) Chung, W. K.; Hou, Y.; Freed, A.; Holstein, M.; Makhadze, G. I.; Cramer, S. M. *Biotechnol. Bioeng.* **2009**, *102*, 869.
- (13) Freed, A. S.; Cramer, S. M. *Langmuir* **2011**, *27*, 3561.
- (14) Yao, Y.; Lenhoff, A. M. *Anal. Chem.* **2004**, *76*, 6743.
- (15) Yao, Y.; Lenhoff, A. M. *Anal. Chem.* **2005**, *77*, 2157.
- (16) Fu, J. Y.; Balan, S.; Potty, A.; Nguyen, V.; Willson, R. C. *Anal. Chem.* **2007**, *79*, 9060.
- (17) Dismer, F.; Hubbuch, J. *J. Chromatogr. A* **2010**, *1217*, 1343.
- (18) Chen, J.; Yang, T.; Cramer, S. M. *J. Chromatogr. A* **2008**, *1177*, 207.
- (19) Chen, J.; Yang, T.; Luo, Q.; Breneman, C. M.; Cramer, S. M. *React. Funct. Polym.* **2007**, *67*, 1561.
- (20) Makrodimiris, K.; Fernandez, E. J.; Woolf, T. B.; O'Connell, J. P. *Anal. Chem.* **2005**, *77*, 1243.
- (21) Jones, T. T.; Fernandez, E. J. *Adv. Colloid Interface Sci.* **2003**, *259*, 27.
- (22) Karajanagi, S. S.; Vertegel, A. A.; Kane, R. S.; Dordick, J. S. *Langmuir* **2004**, *20*, 11594.
- (23) Sethuraman, A.; Han, M.; Kane, R. S.; Belfort, G. *Langmuir* **2004**, *20*, 7779.
- (24) Chung, W. K.; Hou, Y.; Holstein, M.; Freed, A.; Makhadze, G. I.; Cramer, S. M. *J. Chromatogr. A* **2010**, *1217*, 191.
- (25) Chung, W. K.; Freed, A. S.; Holstein, M. A.; McCallum, S. A.; Cramer, S. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 16811.
- (26) Lv, Y.; Lin, Z.; Tan, T.; Feng, W.; Qin, P.; Li, C. *Sens. Actuators, B* **2008**, *133*, 15.
- (27) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701.
- (28) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- (29) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
- (30) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
- (31) Ghosh, T.; Kalra, A.; Garde, S. *J. Phys. Chem. B* **2005**, *109*, 642.
- (32) Granick, S.; Bae, S. C. *Science* **2008**, *322*, 1477.
- (33) Godawat, R.; Jamadagni, S. N.; Garde, S. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15119.
- (34) Sarupria, S.; Garde, S. *Phys. Rev. Lett.* **2009**, *103*, 037803.
- (35) Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105.
- (36) Kane, R. S. *Langmuir* **2010**, *26*, 8636.