THE PROTEIN SOCIETY

Molecular dynamic study of MlaC protein in Gram-negative bacteria: conformational flexibility, solvent effect and protein-phospholipid binding

Yu-ming M. Huang,[1]* Yinglong Miao,[2] Jason Munguia,[3] Leo Lin,[3] Victor Nizet,[3,4] and J. Andrew McCammon[1,2,5]

[1]Department of Pharmacology, University of California, San Diego, La Jolla, California 92093
[2]Howard Hughes Medical Institute, University of California, San Diego, La Jolla, California 92093
[3]Department of Pediatrics, University of California, San Diego, La Jolla, California 92093
[4]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093
[5]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093

Received 29 February 2016; Accepted 20 April 2016
DOI: 10.1002/pro.2939
Published online 00 Month 2016 proteinscience.org

Abstract: The composition of the outer membrane in Gram-negative bacteria is asymmetric, with the lipopolysaccharides found in the outer leaflet and phospholipids in the inner leaflet. The MlaC protein transfers phospholipids from the outer to inner membrane to maintain such lipid asymmetry in the Mla pathway. In this work, we have performed molecular dynamics simulations on apo and phospholipid-bound systems to study the dynamical properties of MlaC. Our simulations show that the phospholipid forms hydrophobic interactions with the protein. Residues surrounding the entrance of the binding site exhibit correlated motions to control the site opening and closing. Lipid binding leads to increase of the binding pocket volume and precludes entry of the water molecules. However, in the absence of the phospholipid, water molecules can freely move in and out of the binding site when the pocket is open. Dehydration occurs when the pocket closes. This study provides dynamic information of the MlaC protein and may facilitate the design of antibiotics against the Mla pathway of Gram-negative bacteria.

Keywords: Mla pathway; ABC transporter; phospholipid; molecular dynamics simulation

Introduction

The cell envelope of Gram-negative bacteria consists of two different membranes, the outer membrane (OM) and the inner membrane (IM), which are separated by the periplasm[1] (Fig. 1). The OM in Gram-negative bacteria has an asymmetric lipid distribution, with lipopolysaccharides at the outer leaflet and phospholipids at the inner leaflet. To maintain this lipid asymmetry, an ATP-binding cassette (ABC) transport system prevents phospholipid accumulation in the outer leaflet by the action of proteins in the Mla pathway that remove the misplaced phospholipids from the OM.[2,3] Inactivating any of the key components from the Mla pathway increases permeability of OM, impairs bacterial functions, and increases sensitivity to antibiotic treatment.

The MlaC protein is one of the key elements in the Mla pathway in Gram-negative bacteria. It functions as a lipid-transfer protein, delivering the phospholipid from OMs to IMs to maintain membrane integrity.[3] The crystal structure of MlaC–phospholipid complex from *Ralstonia solanacearum* shows that MlaC protein is folded into nine alpha helices and five beta stands (Fig. 2). The hydrophobic fatty

*Correspondence to: Yu-ming M. Huang; E-mail: yuh155@ucsd.edu

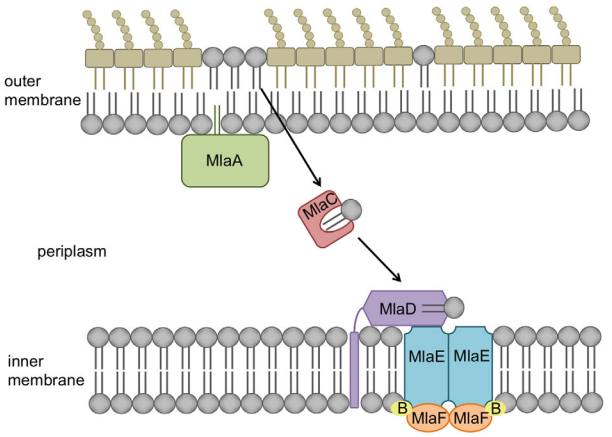

Figure 1. Model of Mla pathway in Gram-negative bacteria.

acids of the phospholipid are located inside the protein and make direct contacts with the α1, α6, α7, β3, and β4 regions.

As a novel antibiotic sensitizing concept, one can design small molecules that bind to the Mla proteins to interrupt the phospholipid transport on the membranes and limit the viability of bacteria during infection and therapy. As a first potential adjunctive antibiotic design, understanding the dynamical properties of the Mla proteins can provide an instructive platform to identify latent drug-binding pockets and further implement the design. To this end, *in silico* molecular dynamic modeling has served as a powerful tool to generate a series of conformations that show the physical movements of a protein. A number of existing studies through molecular dynamics (MD) investigation have successfully provided useful information for drug discovery.[4–6]

In this work, we sought to obtain dynamic insights of the MlaC protein through computational modeling. To better understand dynamic motions of the MlaC protein from different organisms, we performed MD simulations of MlaC from both *Ralstonia solanacearum* and *Acinetobacter baumannii*. The mechanisms of phospholipid binding, correlated protein motions and dynamics of water molecules were investigated in detail.

Results and Discussion

Phospholipid binding

The BLAST calculations show that the sequence identity of MlaC protein between *Ralstonia solanacearum* and *Acinetobacter baumannii* is 29%. Albeit with a low sequence identity, most residues surrounding the lipid-binding site, such as α3, α7, β3, and β4, are either the same or share similar amino acid properties (Fig. 2).

The lipid-binding site of MlaC protein is composed of a large number of hydrophobic residues (Fig. 3). Furthermore, the nonpolar side chains of these residues, Leu, Ile, Val, Lys, and Phe, can move in unison with hydrophobic fatty acids of the phospholipid. Thus, the phospholipid tails are mobile within the binding site. In contrast to the hydrophobic tails that form close interactions with the protein, the hydrophilic phosphate head group of the lipid is exposed to solvent; and it is highly flexible shown from the RMSD calculations (Fig. 4).

Flexibility and rigidity of MlaC protein

The RMSD is a measure of average distance between atoms of two structures, which indicates

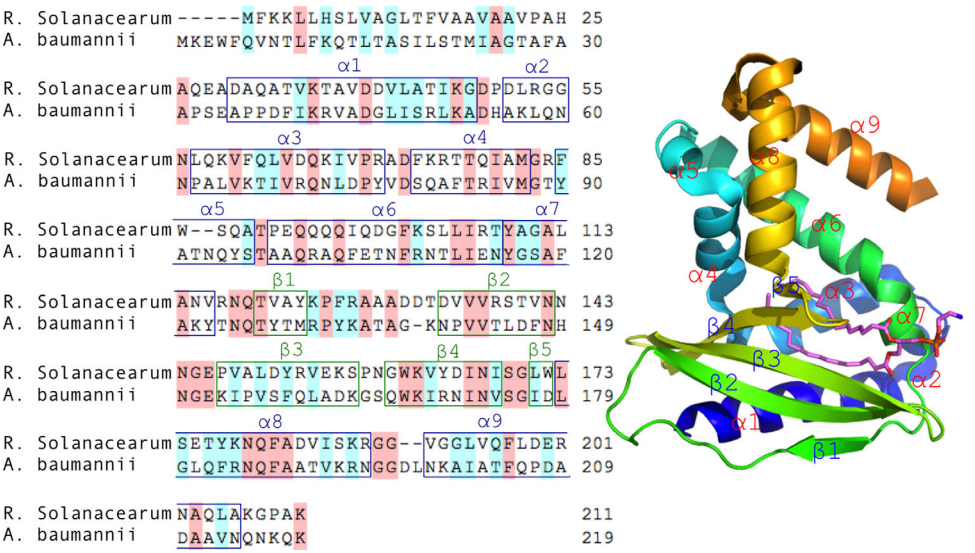

Figure 2. Sequence alignment and structure of MlaC protein. The MlaC sequences from *Ralstonia solanacearum* and *Acinetobacter baumannii* are aligned using BLAST program. Red and blue highlights indicate identical and chemically similar amino acids, respectively. The structure of MlaC protein is composed of nine alpha helices and five beta stands. The bound-phospholipid is colored in pink.

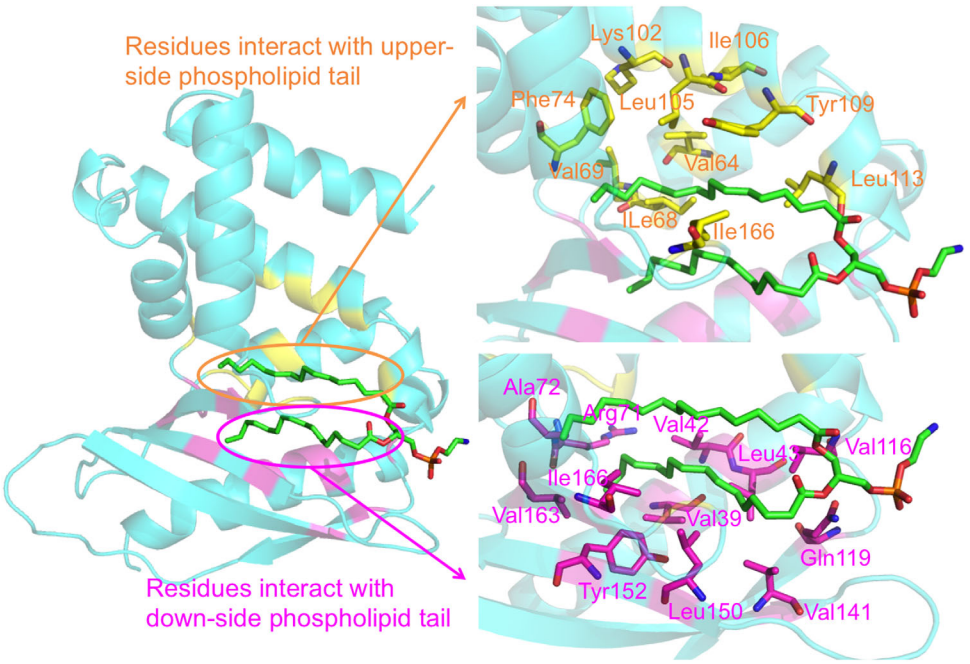

Figure 3. MlaC protein–phospholipid interactions. Residues interacting with the phospholipid tails on top and bottom side are shown in yellow and magenta, respectively.

the structural changes of a protein during MD simulations. Figure 5(A) shows that all systems are reasonably equilibrated after 10-ns simulations. The RMSD of the entire protein in Apo and Holo MlaC (*R. S.*) system are generally within 1.5–2.0 Å. However, Apo and Holo MlaC (*A. B.*) system exhibit larger RMSD values. This can be attributed to the fact that the initial coordinates of the protein in Apo and Holo MlaC (*A. B.*) system are from homology modeling, thus somewhat larger conformational adjustments occur during the simulations.

To better understand structural variations surrounding the lipid-binding site, we examined RMSD of MlaC protein within 5 Å of the phospholipid. As shown in Figure 5(B), generally, the free MlaC proteins display larger RMSD than the complexes. This indicates that the binding of the phospholipid restricts motions of MlaC protein and reduces the plasticity of the binding site.

In addition to the study of global and local conformational flexibility, we measured RMSF of Cα on the protein backbone to illustrate the fluctuations of each residue over the simulation time. Figure 6 shows that the loop between β2 and β3 and the loop between β4 and β5 are highly flexible. Notably, these two loops are located near the entrance of the lipid-binding site, implying that the plasticity of the loops could allow the protein to alter the shape

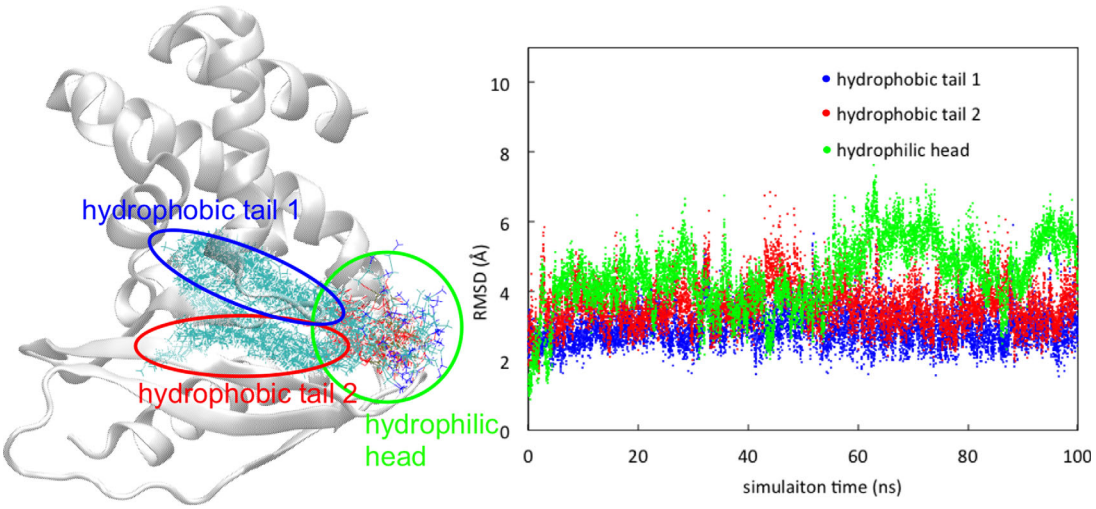

Figure 4. RMSD calculations of phospholipid. The RMSD of hydrophobic tail 1, hydrophobic tail 2, and hydrophilic head of the phospholipid are shown in blue, red, and green, respectively.

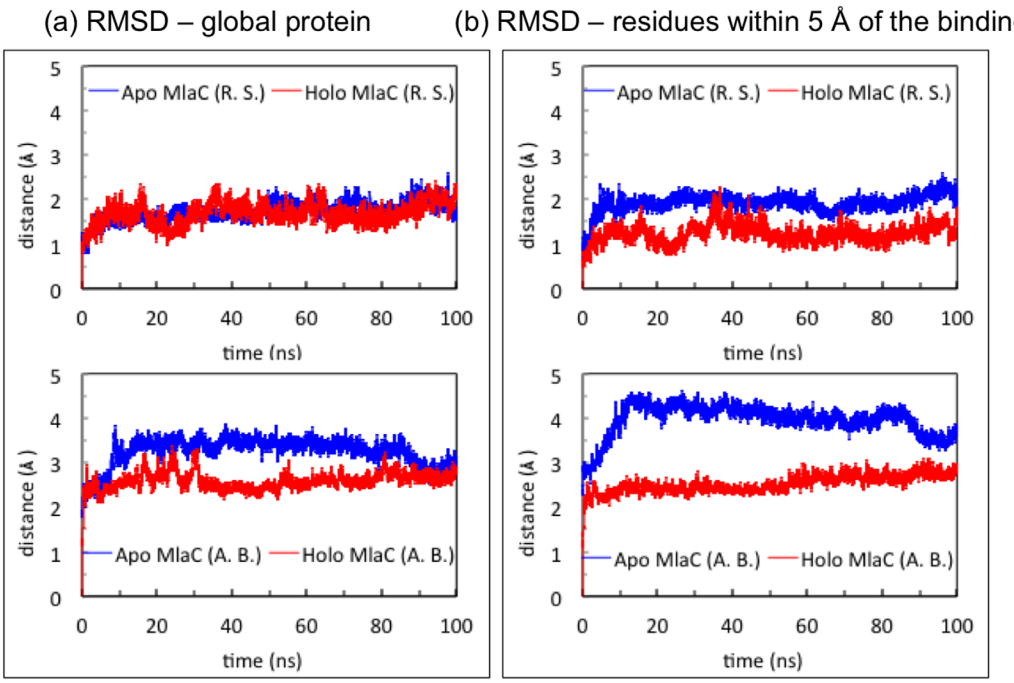

Figure 5. RMSDs of MlaC protein relative to the simulation starting structure. We calculated RMSD of the entire protein (a) and residues within 5 Å around the lipid-binding site (b). The free protein and complex structures are shown in blue and red, respectively.

and size of the binding site to fit the lipid conformations.

Opening and closing of the phospholipid-binding site

We next explored the fluctuations in volume of the lipid-binding site during the MD simulations. Figure 7 shows that the volume of the binding site changes over the simulation time. In the lipid-bound states [Holo MlaC (*R. S.* and *A. B.*)], an opening of the lipid-binding site was observed, by which the volume of the pocket could open up to ~600 Å^3. However, the size of the binding site shrinks to 50–200 Å^3 in the absence of the lipid [Apo MlaC (*R. S.* and *A. B.*)]. The release of the phospholipid induces the closing of the binding pocket.

We further examined the distance changes of several residues surrounding the lipid-binding site. Figure 8 shows that the Ala112 (Ala119)–Ser169 (Ser176) and Val116 (Tyr123)–Ser169 (Ser176) distances are increased when the phospholipid is bound to the MlaC protein. Notably, all these six residues are located at the entrance of the binding site, Ala112 (Ala119) and Val116 (Tyr123) are in α7, and the Ser169 (Ser176) are in β5. This indicates that the movement of α7 and β5 can alter the size of the lipid-binding site and change the binding pocket from the open to closed form.

Correlations

Because the binding pocket size changes between the lipid-free and -bound states, we inspected if any correlated motions of the protein that could help to adjust the protein conformations and adapt for lipid binding. Calculations of generalized correlation show that two highly correlated motions occur near the lipid-binding site (see Fig. 9). First, α7 correlates with β2 and β3. Second, α7 correlates

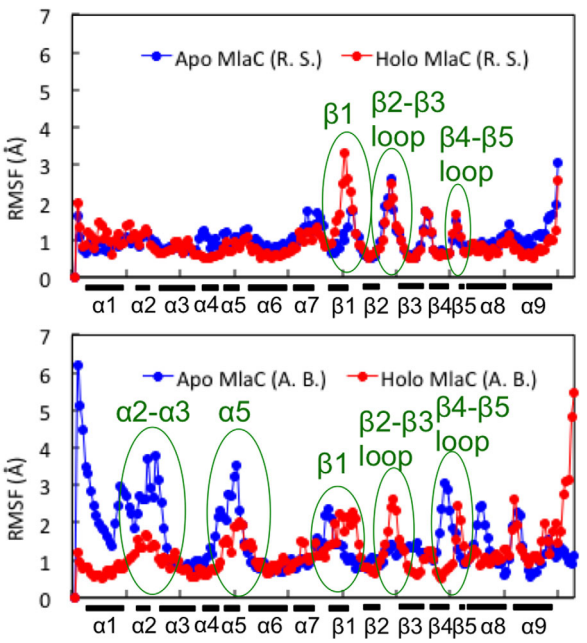

Figure 6. RMSFs of MlaC protein. We calculated RMSF on Systems 1–4. The green circles highlight the flexible regions. The free protein and complex structures are shown in blue and red, respectively.

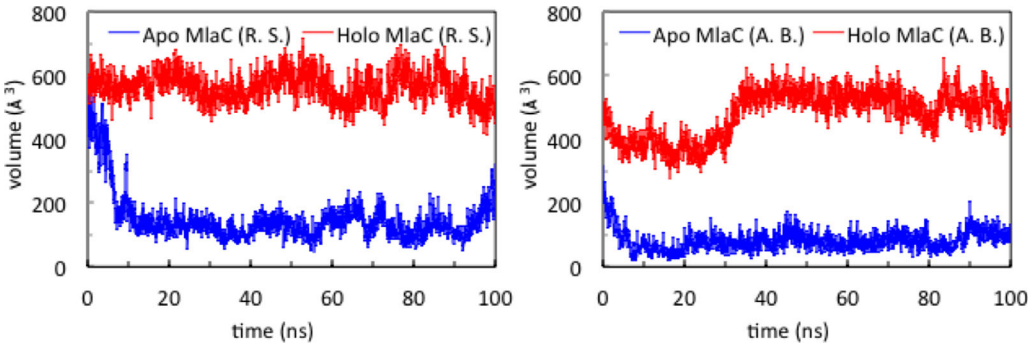

Figure 7. Volume of the lipid-binding site of MlaC protein. The binding site volume fluctuations of the MlaC protein from *Ralstonia solanacearum* and *Acinetobacter baumannii* are shown on the left and right, respectively.

	System 1 and 2	System 3 and 4
Distance 1	Ala112Cα---Ser169Cα	Ala119Cα---Ser176Cα
Distance 2	Val116Cα---Ser169Cα	Tyr123Cα---Ser176Cα

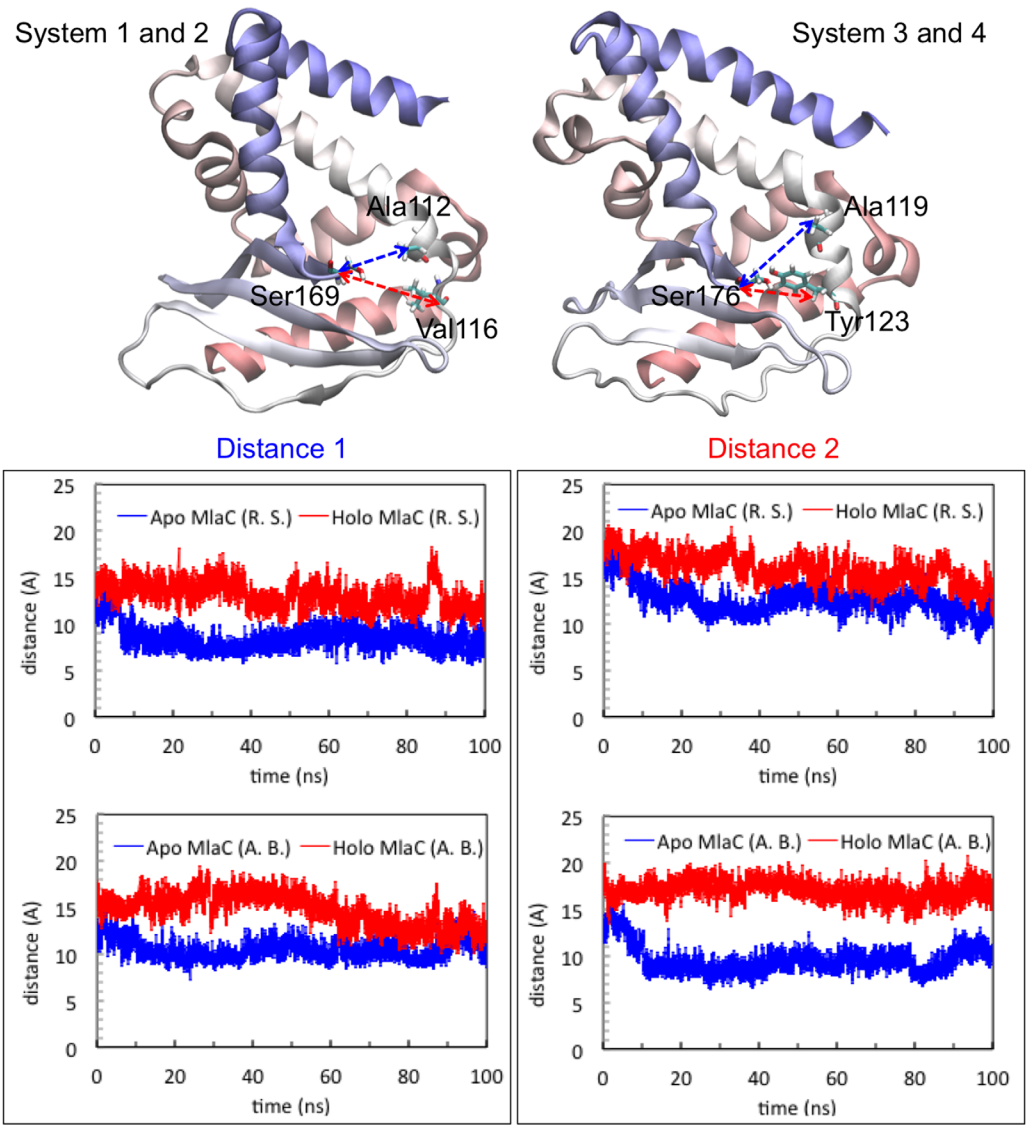

Figure 8. Distance changes of the residues around the binding site. We presented the distance fluctuations between few residues around the entrance of the lipid-binding site.

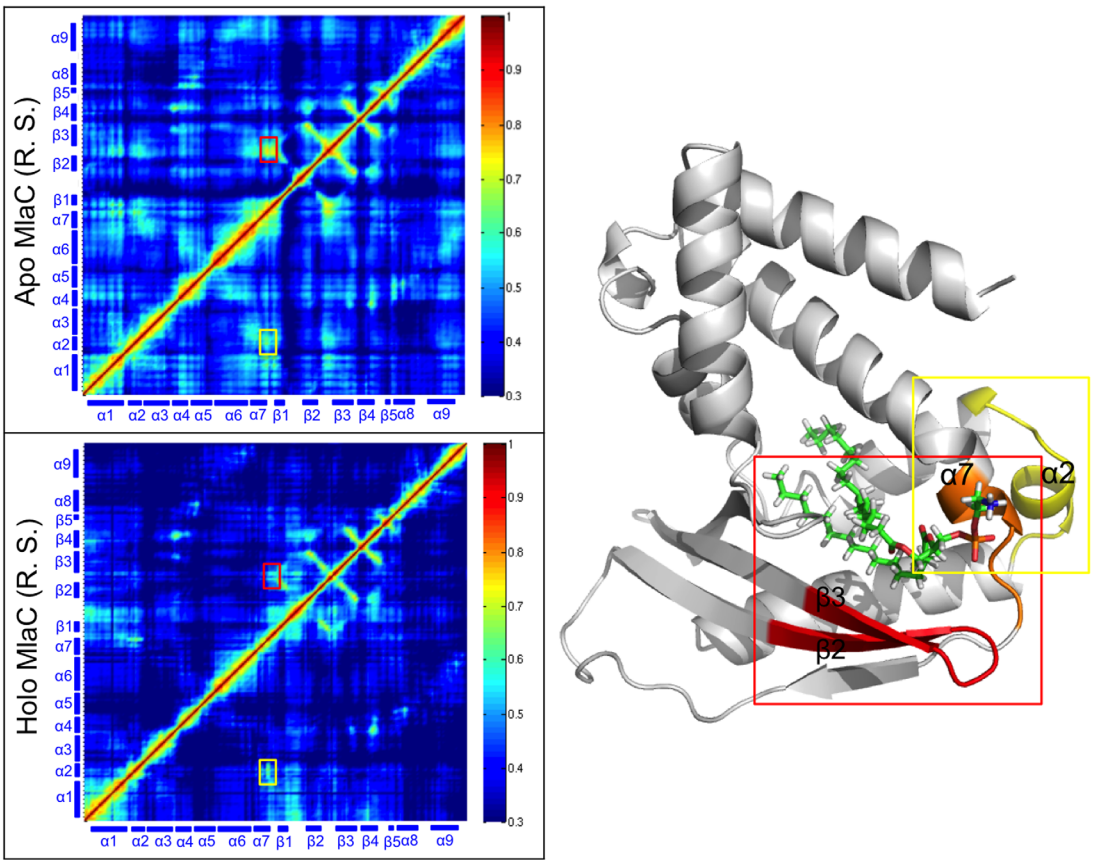

Figure 9. Correlation map of MlaC protein. The correlation map of the MlaC protein from *Ralstonia solanacearum* are on the left. The MlaC protein (gray) and phospholipid (colored bonds) are shown on the right. The red box shows that α7 appears correlated motions with β2 and β3; and the yellow box indicates the correlations between α7 and α2.

with α2. Accordingly, with motion of α7, loops near α2, β2 and β3 can co-evolve to change the shape and size of the binding site. These correlated motions are also shown in Apo and Holo MlaC (*A. B.*).

Water molecules in the phospholipid-binding site

Water molecules could play an important role in ligand binding, and understanding of the water dynamics may facilitate further drug design.[7] Hence, we calculated the number of water molecules in the binding site in both the absence and presence of the phospholipid (Fig. 10). From our calculations, both Holo MlaC (*R. S.* and *A. B.*) show less than 20 waters in the binding pocket upon lipid binding. In the lipid-free state, Apo MlaC (*R. S.*) displays that water molecules can freely move in and out from the binding pocket, thus, the number of waters fluctuates significantly. However, in Apo MlaC (*A. B.*), dehydration occurs because the hydrophobic residues around the binding site form contacts when the

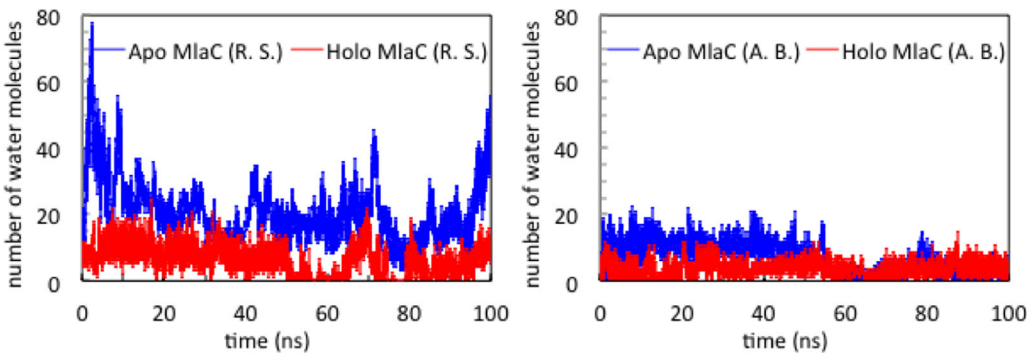

Figure 10. Fluctuations of number of water molecules in the lipid-binding site. We calculated number of water molecules in the binding site of Apo and Holo MlaC (*R. S.* and *A. B.*) over the simulation time.

Table I. *List of MlaC Systems Studied in the Present MD Simulations*

No.	Abbreviation	System	Details	Source bacteria
System 1	Holo MlaC (*R. S.*)	MlaC + phospholipid	Coordinate from PDB code 2QGU	*Ralstonia solanacearum*
System 2	Apo MlaC (*R. S.*)	MlaC protein	Remove phospholipid from System 1	*Ralstonia solanacearum*
System 3	Apo MlaC (*A. B.*)	MlaC protein	Build by homology modeling	*Acinetobacter baumannii*
System 4	Holo MlaC (*A. B.*)	MlaC + phospholipid	Re-dock phospholipid to System 3	*Acinetobacter baumannii*

protein is in apo form. Moreover, the binding pocket tends to close at the end of simulation. These results are consistent with those obtained from calculations of the pocket volume. Although both Apo MlaC (*R. S.* and *A. B.*) are lipid-free system, the pocket volume of Apo MlaC (*A. B.*) is smaller than Apo MlaC (*R. S.*). A dewetting process occurs at the binding interface in Apo MlaC (*A. B.*). The difference between Apo MlaC (*R. S.*) and Apo MlaC (*A. B.*) can be attributed to the composition of residues in the binding pocket. In the Apo MlaC (*A. B.*) system, the side chains of two Phe residues, Phe120 and Phe158, point into the pocket, which exclude waters. However, in the Apo MlaC (*R. S.*) system, Leu113 and Tyr152 replace the position of Phe120 and Phe158, respectively. The hydroxyl group of Tyr152 allows some water molecules to move around and the side chain of Leu113 is smaller than Phe120, which explains why more waters are in the Apo MlaC (*R. S.*) binding site.

Conclusions

In this study, we investigated dynamic properties of an important bacterial lipid-transfer protein—MlaC. MD simulations were carried out on both lipid-free and lipid-bound states of MlaC protein from *Ralstonia solanacearum* and *Acinetobacter baumannii*. Our simulations showed that the hydrophobic tails of the phospholipid form nonpolar contacts with the protein and that the hydrophilic head of the phospholipid can freely move in the solvent. The binding of the phospholipid restricts the motions of the protein, especially the residues around the binding site. In addition, loops near β2, β3, β4, and β5 are flexible; and α7 can co-move with α2, β2, and β3. These protein motions surrounding the entrance of the lipid-binding site tune the shape and size of the binding pocket to accommodate the substrate binding. The open and closed forms of the binding site were thus observed. The calculations of pocket volume showed that the size of the binding site decreases markedly once the phospholipid is released from the protein. Finally, analysis of water motions suggested that, in the lipid-free state, the opening of the pocket allows water molecules to stay in the lipid-binding site; however, the waters are released from the pocket during closing of the binding site.

Methods

Molecular systems

In this study, we considered the MlaC proteins from *Ralstonia solanacearum* and *Acinetobacter baumannii*. For comparison, sequence alignment was performed on the MlaC proteins from the two source bacteria using Basic Local Alignment Search Tool (BLAST).[8]

Four systems of MlaC protein were studied in this work (Table I): (1) System 1: The MlaC protein in complex of a phospholipid [named, 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine (DPPE)] from *Ralstonia solanacearum* was taken from Protein Data Bank (PDB) with the code 2QGU, (2) System 2: Creation of an apo MlaC protein from *Ralstonia solanacearum* by removing the phospholipid in System 1, (3) System 3: Construction of a MlaC protein structure in *Acinetobacter baumannii* by performing homology modeling on the i-Tasser server,[9] using 2QGU coordinates as a template, and (4) System 4: Building of a MlaC protein–phospholipid complex through re-docking the phospholipid back to System 3.

MD simulations

Before starting MD simulations, we applied the software Propka version 3.1 to assign static protonation states of the MlaC protein at physiological pH 7.0.[10] The standard Amber 14 package with graphics processing unit (GPU) acceleration was applied for MD simulations.[11–13] The protein and phospholipid were modeled using the Amber 14SB and Amber lipid14 force field, respectively.[14,15] First, we minimized the MlaC systems regarding the hydrogen atoms, side chains, and the entire protein complex for 500, 5000, and 50,000 steps, respectively. The protein structures were solvated in a rectangular box of the TIP3P water molecules with a 12 Å buffer region.[16] The counterion, Cl^-, based on the Coulombic potential, was added to neutralize the systems. Then, the water molecules and entire system were minimized for 1000 and 5000 steps with periodic boundary condition in a constant volume, respectively. After water equilibration for 100 ps, all molecules, including protein, substrate, ions and solvent, were gradually relaxed by heating the system at 50, 100, 150, 200, 250, and 300 K for 10 ps, respectively. We turned on the particle mesh Ewald to consider long-range electrostatic interactions;[17,18] and the Langevin

thermostat with a damping constant of 2 ps^{-1} was also applied to maintain a temperature of 300 K. The SHAKE algorithm was used to constrain hydrogen atoms during the MD simulations.[19] The resulting trajectories were collected every 1 ps with a time step of 2 fs at the isotropic pressure in the isothermic–isobaric (NPT) ensemble, and the total simulation time of each system was 100 ns. No restriction of the protein in all simulations here.

Analysis of simulation data

The root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), distance of atoms and number of water molecules located in the lipid-binding site were calculated using VMD.[20] The volumes of the lipid-binding site were predicted using the POVME program, version 2.0, with snapshots extracted every 50 ps from the MD simulations.[21] We measured generalized correlations that include mutual information of nonlinear contributions of the protein by applying the g_correlation program in Gromacs 4.5.5 package developed by Lange and Grubmüller.[22] The MD trajectories and snapshots were examined using both VMD and PyMol packages.[20,23]

Acknowledgments

We are grateful to Clarisse Ricci for helpful suggestions about the settings of MlaC–phospholipid structure. This work was supported by NIH, NSF, HHMI, NBCR, and SDSC.

REFERENCES

1. Kamio Y, Nikaido H (1976) Outer membrane of salmonella-typhimurium—accessibility of phospholipid head groups to phospholipase-c and cyanogen-bromide activated dextran in external medium. Biochemistry 15:2561–2570.
2. Narita S-i (2011) ABC transporters involved in the biogenesis of the outer membrane in Gram-negative bacteria. Biosci Biotechnol Biochem 75:1044–1054.
3. Malinverni JC, Silhavy TJ (2009) An ABC transport system that maintains lipid asymmetry in the Gram-negative outer membrane. Proc Natl Acad Sci USA 106:8009–8014.
4. Kim MO, Feng X, Feixas F, Zhu W, Lindert S, Bogue S, Sinko W, de Oliveira C, Rao G, Oldfield E, McCammon JA (2015) A molecular dynamics investigation of mycobacterium tuberculosis prenyl synthases: conformational flexibility and implications for computer-aided drug discovery. Chem Biol Drug Des 85:756–769.
5. Tautermann CS, Seeliger D, Kriegl JM (2015) What can we learn from molecular dynamics simulations for GPCR drug design? Comput Struct Biotechnol J 13: 111–121.
6. Nair PC, Miners JO (2014) Molecular dynamics simulations: from structure function relationships to drug discovery. In Silico Pharmacol 2:1–4.
7. de Beer SBA, Vermeulen NPE, Oostenbrink C (2010) The role of water molecules in computational drug design. Curr Top Med Chem 10:55–66.
8. Zhang JH, Madden TL (1997) PowerBLAST: a new network BLAST application for interactive or automated sequence analysis and annotation. Genome Res 7:649–656.
9. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. Nat Methods 12:7–8.
10. Li H, Robertson AD, Jensen JH (2005) Very fast empirical prediction and rationalization of protein p*K*a values. Proteins 61:704–721.
11. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688.
12. Goetz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC (2012) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. J Chem Theory Comput 8:1542–1555.
13. Salomon-Ferrer R, Case DA, Walker RC (2013) An overview of the Amber biomolecular simulation package. WIREs Comput Mol Sci 3:198–210.
14. Case DA, Berryman JT, Betz RM, Cerutti DS, Cheatham TE, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Homeyer N, Izadi S, Janowski P, Kaus j, Kovalenko A, Lee TS, LeGrand S, Li P, Luchko TL, R., Madej B, Merz KM, Monard G, Needham P, Nguyen H, Nguyen HT, Omelyan I, Onufriev A, Roe DR, Roitberg A, Salomon-Ferrer R, Simmerling CL, Smith W, Swails J, Walker RC, Wang J, Wolf RM, Wu X, York DM, Kollman PA (2015) AMBER. University of California, San Francisco.
15. Dickson CJ, Madej BD, Skjevik AA, Betz RM, Teigen K, Gould IR, Walker RC (2014) Lipid14: The Amber lipid force field. J Chem Theory Comput 10:865–879.
16. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79: 926–935.
17. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593.
18. Salomon-Ferrer R, Goetz AW, Poole D, Le Grand S, Walker RC (2013) Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. J Chem Theory Comput 9:3878–3888.
19. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-interaction of Cartesian equations of motion of a system with constraints—molecular-dynamics of N-alkanes. J Comput Phys 23:327–341.
20. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14:33–38. Model
21. Durrant JD, Votapka L, Sorensen J, Amaro RE (2014) POVME 2.0: an enhanced tool for determining pocket shape and volume characteristics. J Chem Theory Comput 10:5047–5056.
22. Lange OF, Grubmuller H (2006) Generalized correlation for biomolecular dynamics. Proteins Struct, Funct, Bioinform 62:1053–1061.
23. The PyMOL molecular graphics system, version 1.7.4 Schrödinger, LLC.