

2023 INCITE Proposal Submission

Proposal

Title: Characterizing Energy Landscape of Substrate Translocation in Bacterial Membrane

Principal Investigator: Wonpil Im

Organization: Lehigh University

Date/Time Generated: 6/16/2022 8:13:47 AM

Section 1: PI and Co-PI Information

Question #1

Principal Investigator: The PI is responsible for the project and managing any resources awarded to the project. If your project has multiple investigators, list the PI in this section and add any Co-PIs in the following section.

Principal Investigator

First Name

Wonpil

Last Name

Im

Organization

Lehigh University

Email

wonpil@lehigh.edu

Work Phone

6107584524

Address Line 1

locacco Hall, 111 Research Dr

Address Line 2

(No answer given.)

City

Bethlehem

State

PA

Zip Code

18015

Question #2

Co-PI (s)

Question #3

Institutional Contact: For the PI's institution on the proposal, identify the agent who has the authority to review, negotiate, and sign the user agreement on behalf of that institution. The person who can commit an organization may be someone in the contracts or procurement department, legal, or if a university, the department head or Sponsored Research Office or Grants Department.

Institutional Contact

Institutional Contact Name

Diane M. Mason

Institutional Contact Phone

610-758-4402

Institutional Contact Email

e067@lehigh.edu

Section 2: Project Information

Question #1

Select the category that best describes your project.

Research Category

Biological Sciences: Biophysics

Question #2

Please provide a project summary in two sentences that can be used to describe the impact of your project to the public (50 words maximum)

Project Summary

A computational biophysics study to characterize the free energy landscape of substrate translocation in the bacteria. The results will determine nutrient and drug uptake mechanism in bacteria and hold a promise in discovering a new target process for antibiotic development.

Section 3: Early Career Track

Question #1

Early Career

Starting in the INCITE 2022 year, INCITE is committing 10% of allocatable time to an [Early Career Track](#) in INCITE. The goal of the early career track is to encourage the next generation of high-performance computing researchers. Researchers within 10 years from earning their PhD (after December 31st 2012) may choose to apply. Projects will go through the regular INCITE Computational Readiness and Peer Review process, but the INCITE Management Committee will consider meritorious projects in the Early Career Track separately.

Who Can Apply: Researchers less than 10 years out from their PhD that need LCF-level capabilities to advance their overall research plan and who have not been a previous INCITE PI.

How to Apply:

In the regular application process, there will be a check-box to self-identify as early career.

- *The required CV should make eligibility clear.*
- *If awarded, how will this allocation fit into your overall research plan for the next 5 years?*

Projects will go through the regular INCITE review process. The INCITE Program is targeting at least 10% of allocatable time. When selecting the INCITE Career Track, PIs are not restricted to just competing in that track.

- *What is the Early Career Track?*
 - *The INCITE Program created the Early Career Track to encourage researchers establishing their research careers. INCITE will award at least 10% of allocatable time to meritorious projects.*
 - *Will this increase my chances of receiving an award?*
 - *Potentially, this could increase chances of an award. Projects must still be deemed scientifically meritorious through the review process INCITE uses each year.*
 - *What do I need to do to be considered on the Early Career Track?*
 - *In the application process, select 'Yes' at 'If you are within 10 years of your PhD, would you like to be considered in the Early Career Track?' You will need to write a paragraph about how the INCITE proposal fits into your 5-year research and career goals.*
 - *What review criteria will be used for the Early Career Track?*
 - *The same criteria for computational readiness and scientific merit will be applied to projects in the Early Career Track as will be applied to projects in the traditional track. The different will be manifest in awards decisions by the INCITE management committee.*
-

Early Career Track

If you are within 10 years of your PhD, would you like to be considered in the Early Career Track? Choosing this does not reduce your chances of receiving an award.

No

If 'yes', what year was your PhD? If 'no' enter N/A

N/A

If 'yes', how will this allocation fit into your overall research plan for the next 5 years? If 'no' enter N/A.

N/A

Section 4: INCITE Allocation Request & Other Project Funding/Computing Resources

Question #1

OLCF Summit (IBM / AC922) Resource Request - 2023

Question #2

OLCF Frontier (Cray Shasta) Resource Request – 2023

Question #3

OLCF Frontier (Cray Shasta) Resource Request – 2024

Question #4

OLCF Frontier (Cray Shasta) Resource Request – 2025

Question #5

ALCF Theta (Cray XC40) Resource Request - 2023

Node Hours

2,484,600

Storage (TB)

13

Off-Line Storage (TB)

13

Question #6

ALCF Polaris Resource Request - 2023

Question #7

ALCF Polaris Resource Request - 2024

Question #8

ALCF Polaris Resource Request - 2025

Question #9

ALCF Aurora (Intel X^e) Resource Request – 2023

Question #10

ALCF Aurora (Intel X^e) Resource Request – 2024

Question #11

ALCF Aurora (Intel X^e) Resource Request – 2025

Question #12

List any funding this project receives from other funding agencies.

Funding Sources

Funding Source

NSF

Grant Number

MCB-1810695

Funding Source

NIH

Grant Number

R01GM138472

Question #13

List any other high-performance computing allocations being received in support of this project.

Other High Performance Computing Resource Allocations

Section 5: Project Narrative and Supplemental Materials

Question #1

Using the templates provided here, please follow the [INCITE Proposal Preparation Instructions](#) to prepare your proposal. Elements needed include (1) Project Executive Summary, (2) Project Narrative, (3) Personnel Justification and Management Plan, (4) Milestone Table, (5) Publications Resulting from prior INCITE Awards (if appropriate), and (6) Biographical Sketches for the PI and all co-PI's. Concatenate all materials into a single PDF file. Prior to submission, it is strongly recommended that proposers review their proposals to ensure they comply with the proposal preparation instructions.

Concatenate all materials below into a single PDF file.

- 1. Project Executive Summary (One Page Max)**
- 2. Project Narrative (15 Pages Max)**
- 3. Personnel Justification and Management Plan (1 Page Max)**
- 4. Milestone Table**
- 5. Publications resulting from prior INCITE Awards (if appropriate)**
- 6. Biographical Sketches for the PI and all co-PI's.**

Concatenate-2022.pdf

The attachment is on the following page.

Project Executive Summary

Title: Characterizing Energy Landscape of Substrate Translocation in Bacterial Membrane

PI and Co-PI(s): Wonpil Im, Soohyung Park, Shasha Feng

Applying Institution/Organization: Department of Biological Sciences, Lehigh University

Number of Node Hours Requested: 2,484,600

Amount of Storage Requested: 13 TB

Executive Summary:

The overarching goal of this proposal is to characterize the free energy landscapes of the transbilayer or intra-bilayer pathways for small molecules and phospholipid substrates in the bacterial membranes. By offering a quantitative estimation of the substrate translocation process and elucidating the translocation path, the proposed study will greatly advance the investigation of drug uptake, drug resistance in bacteria, and discover a new target process for antibiotic development.

Selective porins like OccK5 in Gram-negative bacterial outer membranes impose restrictions on which substrates, such as nutrients and antibiotics, to enter the cell. Yet, the transport properties have never been elucidated in realistic outer membranes containing lipopolysaccharides (LPS), via both experimental and computational approaches. The LPS 'bushes' form a dense barrier lowering the permeability of substrates and also affecting the porin conformational dynamics. It is extremely challenging to reach a converged free energy profile for translocation in LPS-containing membranes. In contrast to trans-bilayer transport by β -barrel porins, Mla, a phospholipid transporter complex, imposes a different challenge. The MlaFEDB complex translocates lipids into the bilayer. While many groups have raced to solve the cryo-EM structure of Mla complex in the past two years and did observe phospholipids trapped inside the transmembrane MlaE, the lipid diffusion pathway remains elusive. The path was difficult to characterize as the protein itself is embedded in the lipid bilayer and crosslinking experiment or functional assays were unable to elucidate the pathway.

This proposal aims to determine the following aspects of the substrate translocation via large-scale ensemble simulations and augment the sampling through enhanced sampling technique REST2: *i*) free energy profiles and energy barriers of translocating a nutrient through negatively-charged substrate channel OccK5; *ii*) the influence of LPS polysaccharide "bushes" on the OccK5 transport properties; *iii*) phospholipid diffusion pathways in Mla complex and whether a counterion is required for negatively charged PG lipid translocation.

The proposed simulations are a pivotal study that combines our decades-long expertise of LPS research, membrane properties modeling, transporter investigation with enhanced sampling techniques and peta-scale computation facilities. This proposed study is also supported by NSF grant MCB-1810695 and NIH grant R01GM138472 on membrane protein investigation and advanced simulation method development in CHARMM-GUI.

Outcomes: This proposal will determine the minimum-energy translocation paths for small molecules through OccK5 and for phospholipids through MlaE. The free energy barriers and key binding sites will be identified from the trajectories to guide mutagenesis and crosslinking efforts. The overcoming of sampling challenge of LPS and outer membrane will pave the road for quantitatively understanding other outer membrane translocation processes. By studying phospholipid diffusion pathway, this proposal holds a promise in discovering a new target process for antibiotic development.

PROJECT NARRATIVE

1. Significance of Research

1.1 Background and significance

Gram-negative bacteria cause diseases including food poisoning, meningitis, pneumonia, wound infections, lungs, urinary tract, and stomach infections in healthcare settings.^{1, 2} Due to their high resistance to antibiotics, Gram-negative bacteria are among the most significant public health problems in the world.³ They put patients in the intensive care unit (ICU) at high risk and lead to high mortality and morbidity. Combating bacterial infection in hospitals while also keeping an eye on the rising of superbacteria has been an urgent issue for governmental response to medical crisis.⁴ In the past two years, the coronavirus plagued the world with quick spreading and high-death rate variants.^{5, 6} It is of great importance to develop medicines that can effectively treat infection, curb the virus or bacteria from developing more mutants and infecting a larger population, therefore bring down the risk of future pandemic.

Bacteria have various mechanisms to protect themselves from lethal reagents like detergents and antibiotics.⁷ One of the tools is the membrane that acts as an prominent shield from harsh environments and a gate to control the in-and-out of drugs and nutrients. Gram-negative bacteria have two unique membranes, an external outer membrane and an internal inner membrane.^{8, 9} The outer membrane is asymmetric, with the outer leaflet composed of lipopolysaccharides (LPS) and the inner leaflet composed of phospholipids. An LPS molecule consists of a hydrophobic bilayer domain called lipid A, negatively charged oligosaccharide core, and hydrophilic O-antigen polysaccharide chain (depicted as chains outside the bilayer membrane) (**Figure 1**).¹⁰ The long, thick core and O-antigen polysaccharide chain greatly reduce the permeability of bacterial membranes, and bury the outer membrane proteins underneath the “bushes”.¹¹

In this study, we focus on two aspects of the bacterial self-protection machinery (**Figure 1**): *i*) the substance translocation through OccK5 and the influence of LPS on the transport gate control; *ii*) lipid translocation in MlaE protein, a transmembrane component of outer membrane asymmetry maintenance pathway, Mla (Maintenance of lipid asymmetry). This study has been supported by NSF grant MCB-1810695 and NIH grant R01GM138472 on membrane protein investigation and advanced simulation method development in CHARMM-GUI.

LPS is the first gate keeper in small molecule translocation. Due to the negatively charged core oligosaccharide, LPS membranes form a knitted network via divalent cation bridges. This leads to a tightly packed lipid tails and dense polysaccharide chain layer. Such a unique feature makes the outer membrane a highly effective barrier to translocation.¹² Thus, to facilitate the uptake of small molecules like nutrients or antibiotics, outer membrane proteins (OMPs, β-barrel porins) are necessary.¹³ OMPs could be either non-specific (e.g., OmpF and OmpC) or specific to substrates (e.g., OccK5 and OprD in Occ family).^{11, 14-16} Since LPS chains hinder both the diffusion of small molecules to OMP and the dynamics of pore entry of OMP, the permeability of small molecules depends on the polysaccharide chain length. It has been reported that mutants lacking O-antigen tend to exhibit higher sensitivity to detergents and antibiotics compared to those with longer polysaccharide chains.¹⁷⁻¹⁹ It is important to provide atomistic detailed understanding of the influence of LPS chains on the small molecule translocation through OMPs.

Maintenance of outer membrane barrier requires delicate lipid transportation control. Phospholipids appearing at the outer leaflet of outer membranes is deleterious to the bacteria.⁸ A delicate mechanism to maintain the asymmetry is the Mla system.²⁰ MlaA is located on the outer membrane, MlaC is in the periplasmic space to shuttle phospholipids, and MlaFEDB is in the inner membrane (**Figure 1**).²¹ MlaE and MlaD are transmembrane protein, where phospholipids are found inside the pocket formed by the two proteins. In the past few years, vast structural biology

efforts were made to decipher the structure of the Mla system.²¹⁻²⁶ Yet, the solved lipids inside Mla do not have good electron map densities and the atomistic coordinates modelled by different research groups diverged.^{21, 22, 24, 26} Meanwhile, all the structural efforts failed to capture the lipids in the translocation process. Currently, it is not understood how the lipids diffuse out of the pocket of the transmembrane component MlaE into the bulk membrane.

Biomedical relevance to antibacterial drug research.

Occk5 transports the negatively charged substrates with the gating residues exerting selectivity. Understanding this gating mechanism will shed light on the translocation of charged drug molecules through Occk5, which would be important for optimizing drug absorption, distribution, metabolism, elimination and toxicity (ADMET) properties.²⁷ Studying the translocation process of the channel is also key to pave the road for studying many other β -barrel proteins in the bacterial membrane.²⁸⁻³⁰ Meanwhile, Mla is a vital system to keep the asymmetry of the outer membrane. The knockdown of Mla was shown to cause lower bacterial cell viability and

render them more susceptible to detergent treatment.²¹ The lipid diffusion pathway in Mla can be utilized as a target for developing drugs that covalently bind to certain residues. Both protein categories have significant meaningful connections with drug discovery and lead optimization in antibacterial drug research.

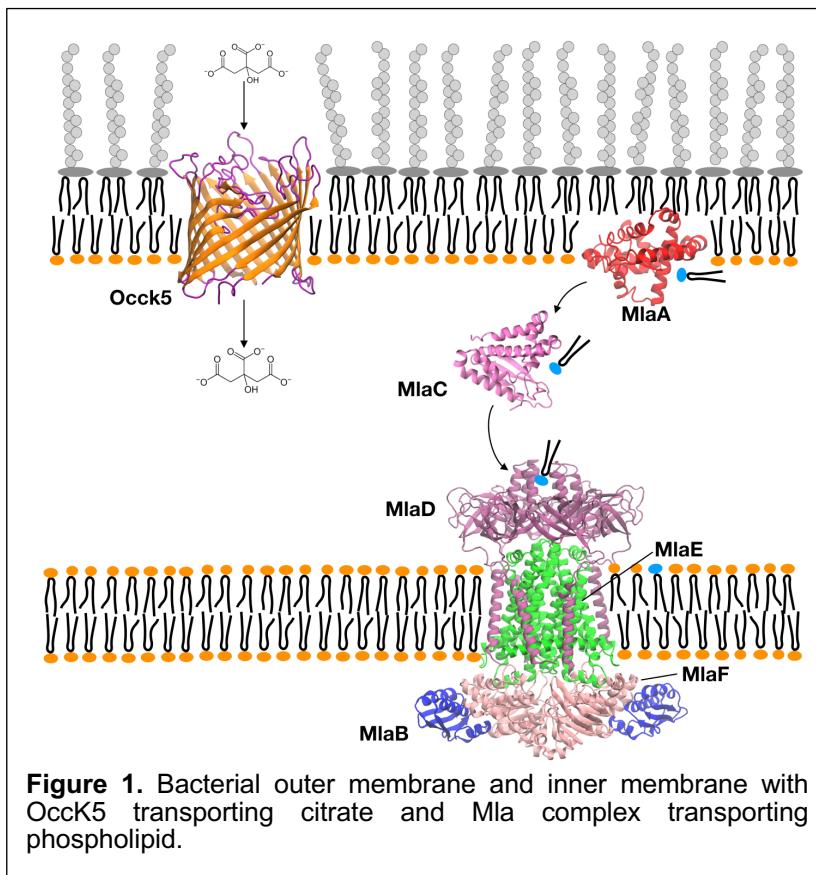


Figure 1. Bacterial outer membrane and inner membrane with Occk5 transporting citrate and Mla complex transporting phospholipid.

Advantage of high-performance computational free energy characterization. Umbrella sampling with replica exchange solute tempering (US/REST2)³¹ can describe the translocation process at an atomic level. The translocation process is usually difficult to capture in the experiments. For example, lipid diffusion in Mla has so far been a very challenging process to characterize, with crosslinking experiments unable to show the pathway.²² The energetic landscape is even beyond the reach. Employing free energy calculation by US/REST2 with high-performance computation has the benefits of elucidating the quantitative energetic profile as well as laying foundations for experiments, e.g., point mutation and cross-linking, to validate the translocation pathway. Due to the size of membrane systems, such simulations require numerous windows for sampling along the translocation path and several Hamiltonian tempering replicas to accelerate the sampling, making it only possible with leadership computing resources.

Expected outcomes and plans for federal grant submissions. The outcome of this proposal would be used to lay foundation for experimental mutagenesis study and for current federal grant renewal. The simulations will also guide the module development of CHARMM-GUI Enhanced

Sampler, which helps to make the advanced multi-replica simulation techniques available to the computational community.

1.2 Innovation

Conceptual Novelty. To date, there are limited free energy studies of substrate transport through β -barrel membrane proteins, despite their apparent abundance in bacterial membranes.³²⁻³⁹ Majority of these free energy studies used homogeneous phospholipid bilayers, e.g., DMPC or POPC, to calculate the free energy profile, which did not include LPS in the membrane.¹¹ Only recently, a study on the OprD transport of positively charged arginine in the rough LPS membrane of *E. coli* was reported. It showed that the R core has little effect on the transport of arginine through OprD,³⁹ consistent with LPS monolayer experiments.⁴⁰ However, a lack of studies on substrate translocation in the realistic bacterial membrane both in experiments and in simulations impedes the understanding of the real scenarios in the bacterial outer membrane.⁴¹⁻⁴⁴ In this proposal, we aim to elucidate the free energy profile of substrate translocation in membranes containing LPS with both R core and O-antigens.

Mla is a novel transporter distinguished from the traditional transporters that have distinct outward-open, inward-open, and intermediate states.²² A remotely similar structural fold was only recently found in the LPS transporter system.⁴⁵ While the structures share visual similarity, the specific structure fold and mechanism vary significantly, as the Lpt system transports LPS unidirectionally.⁴⁶ Experimentally, lipid transporters are challenging to study, as it is difficult to label the lipid and distinguish the substrate lipid from the lipids in the bulk bilayer. Therefore, computational study offers a unique approach to probe the transportation pathway.

The proposed simulations of Occk5 and Mla are significant in that they combine our decades-long experience of LPS research,⁴⁷⁻⁵⁵ membrane modeling,⁵⁶⁻⁶⁰ transporter investigation^{22, 61-64} with enhanced sampling technique expertise and peta-scale computation facilities.

Technical Novelty. The key challenge in sampling the translocation process is the substrate orientation and the surrounding protein residue configurations. In REST2 algorithm,⁶⁵ a collection of atoms, so-called ‘solute’, and the remaining as ‘solvent’ are defined, and the sampling of solutes are selectively accelerated by scaling solute-solvent interactions. In the replicas, solutes are in a higher effective temperature than solvents for their better conformational exploration in the phase space. So far, REST2 has not been applied to substrate translocation. It is in principle a similar approach to accelerate the degrees of freedom in terms of residues near the translocation, but offers a much simpler approach without the need of additional collective variables such as orientation of the substrate.

2. Research Objectives and Milestones

2.1 Overarching goals

The overall goal is to explore the free energy landscape along the trans-bilayer transport pathway of a nutrient and the intra-bilayer translocation of lipids. Through characterizing these processes using the enhanced free energy sampling technique, we aim to understand how environments affect the uptake of small molecules and to refine the elusive mechanism of bacterial phospholipids transport. These obscure mechanisms of transportation across the membrane or inside the membrane are largely unreachable by conventional biophysical experiments for the same spatial resolution and timescale. The two milestones in this proposal serve to understand two aspects of the bacterial membrane, i) how the LPS influences transportation of small molecules that are vital to bacterial survival and growth, ii) how phospholipids are transported to

maintain asymmetry in the outer membrane. Carefully built and executed enhanced sampling simulations can provide insights into how the key gating residues along the translocation pathway affects the transportation and what changes are needed to clear the transport pathway, which would be key information for experimentalists.

2.2 Proposed Research Plans and Milestones

Milestone 1: Quantitative characterization of effects of LPS polysaccharide length on substrate transport

Rationale. Unlike phospholipid bilayers (**Figure 2A**), the polysaccharides of LPS molecules in the outer membrane of Gram-negative bacteria create a crowded (**Figure 2B-C**). They not only interact with substrates but also hinder the dynamics of the pore-entry especially the loops of β -barrel channels. As the length of LPS polysaccharide chain increases from G2 core to G2 core-O10 antigen (**Figure 2B-C**), the sampling of substrate conformations and its entry into the pore becomes significantly more challenging. Our preliminary adaptive replica-exchange umbrella sampling (REUS) simulations⁶⁶ of citrate transport through OccK5 clearly show such a challenge: the free energy profile in a DMPC phospholipid bilayer is symmetric (**Figure 2D**) but becomes highly asymmetric in the LPS with O-antigen polysaccharide (Pa.G2.O10 membrane, **Figure 2E**). The result indicates that a simplistic one-dimensional (1D) enhanced sampling approach is not sufficient for quantitative free energy profile of substrate transport in a complex outer membrane of Gram-negative bacteria.

In addition to LPS polysaccharide chain length, the sampling of substrate itself becomes more challenging with increasing chemical complexity,^{37, 67} e.g., imipenem >> citrate > succinate (**Figure 3**). While both sampling issues (the LPS dimension and the ligand dimension) are equally important, there has been a lack of free energy estimation of small molecule transport in a complex realistic outer membrane with long O-antigen polysaccharide chains. Therefore, we use citrate as a model compound and quantitatively characterize the free energy profile of its transport in varying LPS membranes. The results will be a benchmark to understand how different parts of LPS modify the free energy landscape of transportation, while also lay a basis for studying other selective β -barrel channels such as OprD, which transports positively charged basic amino acids and carbapenem antibiotics.¹⁴

Research plan. In this milestone, we will study two LPS membranes of bacterium *P. aeruginosa*, where LPS includes either G2 core oligosaccharides (Pa.G2) or up to 10 O-antigen repeating units (Pa.G2.O10) (**Figure 2B-C**). The Pa.G2 is also called a rough LPS membrane while Pa.G2.O10 is called a smooth LPS membrane.⁶⁸

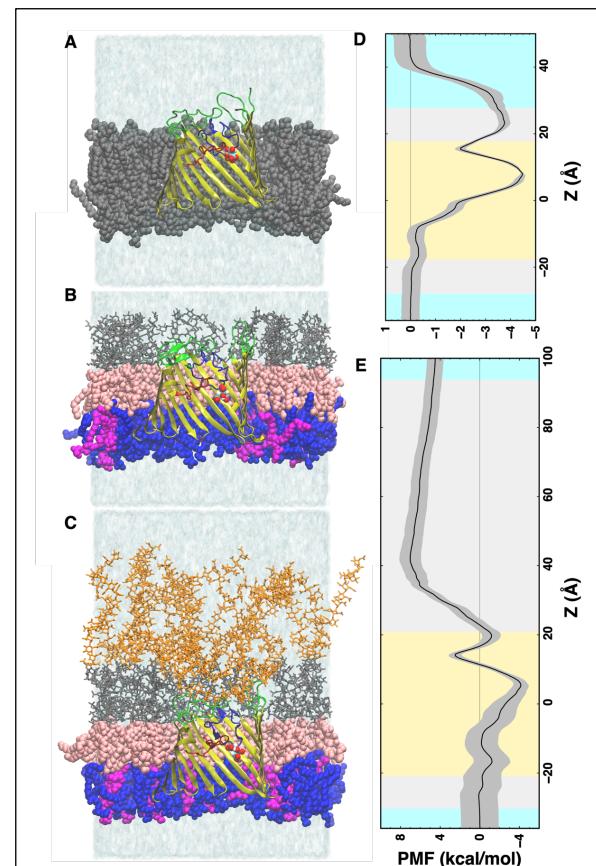


Figure 2. (A-C) Snapshots of OccK5 with a citrate embedded in a DMPC bilayer (A) and *P. aeruginosa* outer membranes, where LPS is up to G2 core (Pa.G2) (B) and with additional 10 O-antigen units (Pa.G2.O10) (C). (D-E) Free energy profiles of citrate transport along the membrane normal in (D) DMPC bilayer and (E) Pa.G2.O10 outer membrane.

The latter has a significantly increased length of extracellular LPS chains with the channels buried underneath the LPS brushes.

We embed OccK5 in the outer membrane systems (**Figure 2B-C**) and use citrate as the model substrate. We will use REST2, a Hamiltonian replica-exchange solute tempering method, combined with umbrella sampling (US) to calculate the free energy profiles (**Table 1**). The collective variable (CV) for US will be the Z-position of the substrate (Z_s), along which windows will be uniformly distributed at every 2 Å. Near the narrowest pore region (over ~20 Å), the windows will be distributed at every 1 Å. These windows will be then coupled to REST2 for better sampling of substrate orientations. The potential of mean force (PMF) as a function of Z_s will be calculated by the weighted histogram analysis method⁶⁹ (or equivalent methods such as MBAR⁷⁰ and/or DHAMed⁷¹). To ensure sufficient sampling, we will also calculate two-dimensional (2D) PMF as a function of Z_s and the orientation of the substrate. The lowest free energy path for the transport will be obtained by a string method from the calculated 2D-PMF.⁷²

We will investigate the influence of polysaccharides on the free energy profiles of small molecule transport by comparing the results from Pa.G2 and Pa.G2.O10 membranes with a DMPC membrane. We will examine how LPS polysaccharide chains affect the substrate conformation and orientation. We will also analyze OccK5 protein conformations to examine the propensity of two open states of OccK5, which have been reported in the single-channel electrical conductance measurement,⁴² and how the population is affected by LPS polysaccharide chains.

Table 1. System information of Occk5 simulations.

Bilayer	System Size (Å ³)	# Atoms	# of Windows			Theta Nodes per Window	Total Theta Nodes
			US	REST2	Total (USxREST2)		
Pa.G2	~100×100×110	~104,000	74	4	296	8	2,368
Pa.G2.O10	~100×100×150	~154,000	114	4	456	8	3,648

Windows for US along the membrane normal span across the membrane to include bulk water regions at both sides.. The estimated number of windows are lower bounds for production runs. Total number of Theta nodes is calculated by Total_number_windows x Theta_Nodes_per_Window.

Pitfalls, caveat, alternative strategies. The narrowest site along the translocation path usually yields largest errors. In that case, we will further add more windows in US and/or increase the number of replicas in REST2. If the sampling of substrate orientation is limited, we may increase the number of replicas in REST2 or add an additional collective variable along the orientation angle.

Expected outcomes and significance. We expect to determine the free energy profiles of citrate transport through OccK5 in two outer membranes, Pa.G2 and Pa.G2.O10. Through comparison among DMPC, Pa.G2, and Pa.G2.O10 systems, we aim to elucidate the influence of the LPS core and the length of O10 antigens on channel conformations and associated barrier in the substrate's entry. Such a characterization with a realistic outer membrane including LPS (with O-antigen units) has not been presented before, which will enable our understanding of the real

scenario of translocation energetics. We also expect to collect a conformational ensemble of substrates in OccK5 and identify representative conformations along the translocation pathway that could be connected to electrical conductance data. We aim to identify the adjustment of residues along the pathway to accommodate the substrate and the key-binding residues. These are important to understand other charged substrate translocation through the β -barrel channels. Upon successful completion of this milestone, the methodology can be applied to the transport of larger substrates such as antibiotics imipenem.

Milestone 2: Elucidating energetic landscapes of lipid diffusion pathways in Mla

Rationale. The recent multiple cryo-EM structures of MlaFEDB universally revealed the lipid densities in the MlaE pocket.^{21, 22, 24-26} Since the resolutions of these structures (3-10 Å) are not ideal for solving the lipid densities, different research groups proposed contrasting models. We have performed REST2 and Gaussian-accelerated MD (GaMD) simulations to sample the lipid conformations and to explore its potential diffusion pathways in the pocket (**Figure 3A**). Both simulation schemes showed that the substrate phospholipid has restricted movement inside the pocket and cannot diffuse out (**Figure 3C**). Moreover, unlike the LPS transporter, the cryo-EM structures of Mla complex have not captured an intermediate state of phospholipid diffusing halfway out. So it is not clear how the phospholipid gets out of the pocket to the bulk bilayer.

Structurally, the only possible route for the phospholipid to diffuse out to the upper leaflet of the membrane is through the vestibule between the MlaE dimer interface (**Figure 3B**). The diffusion pathway to periplasmic space is clearly through the MlaD crown region. A shuttle protein

MlaC will then dock to the MlaD crown to transfer the lipid (**Figure 1**). By performing adaptive biasing force (ABF) simulations where the free energy surface was adaptively modified by sampling and adding a biasing potential to make the energy surface flat,⁷³ we studied how the lipid may get out of the MlaE pocket. The distance between the lipid and the center of the MlaE pocket was used as a collective variable. The results revealed two major pathways as we have hypothesized, one through the crown region and the other through the sideway of the MlaE. The ABF simulations use relatively coarse bins along the collective variable, which was good for sampling the potential routes but yielded coarse/unreliable free energy profile. This is also aggravated by the fact that, upon lipid diffusion, rearrangement of the helices is necessary, while in ABF, these degrees of freedom are not considered in the collective variable. In this Milestone 2, we aim to elucidate the free energy landscape of the two diffusion pathways.

Milestone 2.a Studying the vertical diffusion pathway inside the MlaD crown

The vertical diffusion pathway through the MlaD hexamer crown is a path for phospholipids to shuttle between the inner and outer membranes of the Gram-negative bacteria. In the cryo-EM

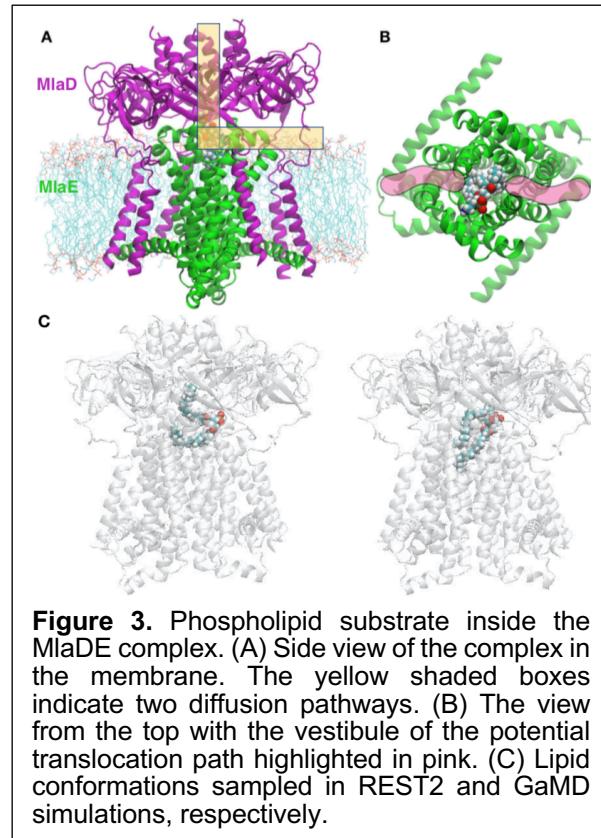


Figure 3. Phospholipid substrate inside the MlaDE complex. (A) Side view of the complex in the membrane. The yellow shaded boxes indicate two diffusion pathways. (B) The view from the top with the vestibule of the potential translocation path highlighted in pink. (C) Lipid conformations sampled in REST2 and GaMD simulations, respectively.

structures, a lipid density is commonly observed in the central pore of MlaD.^{22, 24, 25} We have performed long-scale MD simulation (3 μ s) for the whole MlaFEDB complex on Anton2 where no lipid was modelled in the MlaD crown. The MlaD hexamer collapsed and closed the tunnel after the simulation, indicating the importance of lipid molecule to maintain the structure of MlaD. This tunnel is highly hydrophobic, which can serve as a quick elevator for transportation of the lipid. The ABF simulation results showed that the phospholipid passes MlaD crown in a fast manner and spends little time in this region compared to the both ends of the tunnel. Meanwhile, the two lipid acyl chains are flexible during the transportation as they adopt either tail-together or tail-split conformations.

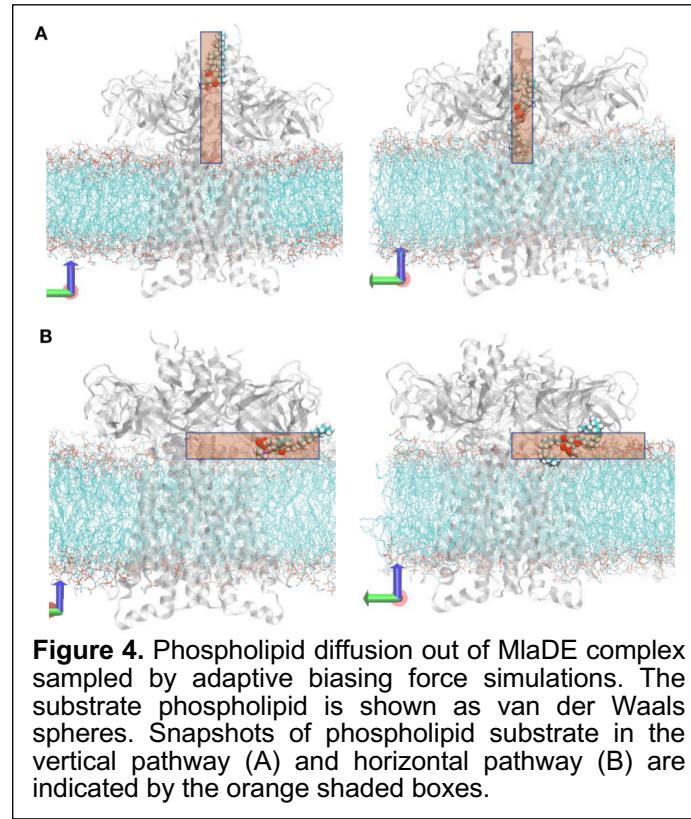
We will perform umbrella sampling and REST2 simulations to study the vertical diffusion of the phospholipid. The umbrella sampling simulations will use the distance between phospholipid headgroup and MlaE pocket along the membrane normal as a collective variable. Based on ABF simulations, we will sample a distance range of 35 Å and use a window size of 1.75 Å for US simulations (in total 20 windows). REST2 will be applied to the phospholipid substrate and protein residues that are within two shells (each shell 4-Å thickness) away from the substrate to help overcome the local energy barriers.

Mla has been shown to have no distinguishable preference on phospholipids headgroup types, indicating similar affinities for PE (0 charge) and PG (-1 charge) lipids. A neutral residue and a negatively charged residue usually make a big difference in the free energy profile in biological systems. One hypothesis is that PG might bind with a mono-valent cation to neutralize the charge. We aim to elucidate if PG without a counterion has a similar energetic landscape as neutral PE, such that a companion ion is not necessary.

In the analysis stage, to ensure sufficient sampling and to investigate the favorable transportation pose, we will calculate 2D PMF as a function of distance along the vertical distance (d_1) and the end-to-end distance of lipid acyl chains (d_2). The lowest free energy path for the transport will be obtained by a string method from the calculated 2D-PMF.

Pitfalls, caveat, alternative strategies. As the lipid gets out of the MlaD crown, the acyl chain tails will be exposed to water molecules, which will be highly unfavorable. The collective variable should therefore be limited to only the cap region of the crown not too far away. Alternatively, we may have MlaC docked onto the MlaD crown and characterize the diffusion from the MlaE pocket into the MlaC shuttle protein.

Expected outcomes and significance. We expect to determine the phospholipid diffusion and its free energy profile in the hydrophobic tunnel of MlaD crown. We expect to distinguish which tail configuration (tail-together or tail-split) is the most favorable pose for transportation. We also anticipate to elucidate how the negative lipid



headgroup impacts the free energy profile. If the energy landscape is relatively flat, that may explain why the Mla phospholipid transportation is bi-directional. And, any barriers in the free energy landscape can help identify the possible gating residues along the vertical pathway.

Milestone 2.b Studying the horizontal diffusion pathway along the MlaE vestibule

In the Milestone 2b, we aim to study the horizontal pathway between the dimer of MlaE. The ABF simulations revealed that the phospholipid substrate diffused out through this horizontal corridor fast. We hypothesize that it is likely the reason that no intermediate diffusion structure has been captured in cryo-EM studies so far. To study the free energy profile, we will again use US-REST2 to characterize the horizontal diffusion process. We will use X-Y plane distance between phospholipid headgroup and MlaE pocket as a collective variable and REST2 solute is defined as the substrate phospholipid and two shells (4 Å each shell) of protein residues around it. We will also sample a distance range of 35 Å and use a window size of 1.75 Å for US simulations (in total 20 windows) for the horizontal pathway.

We will study both PE and PG lipids to probe the influence of charge on the free energy profile. More interestingly, during the horizontal diffusion, the headgroup of phospholipid substrate is hydrated. Therefore, it offers an opportunity to observe whether a counterion will stay nearby to facilitate the lipid diffusion. It is also of interest to see how the phospholipid substrate merges into the bulk lipid bilayer. Deriving a free energy profile for the horizontal diffusion pathway will identify the barrier residues and validate the concept of a horizontal pathway. The simulation plan of Milestone 2 is summarized in **Table 2**.

Pitfalls, caveat, alternative strategies. The lipid may diffuse through either of the dimer interfaces. We will include a direction to sample only one side of the interfaces.

Expected outcomes and significance. We expect to determine the lipid conformation and free energy profile during the horizontal diffusion process. We anticipate to show the influence of lipid headgroups on the diffusion free energy. The results will elucidate a new phospholipid translocation pathway in the Mla complex that has not been found in the structural biology experiments and help interpret the functions of the crown region residues.

Table 2. System information for lipid diffusion pathways in the Mla complex.

Diffusion Pathways	Lipid	System Size (Å ³)	# Atoms	# of Windows			Theta Nodes per Window	Total Theta Nodes
				US	REST2	Total (USxREST2)		
Vertical diffusion	POPE	~107×107×148	~157,000	20	4	80	8	640
	POPG	~107×107×148	~157,000	20	4	80	8	640
Horizontal diffusion	POPE	~107×107×148	~157,000	20	4	80	8	640
	POPG	~107×107×148	~157,000	20	4	80	8	640

3. Computational Readiness

3.1. Approach

Molecular Dynamics Simulation

NAMD is a parallel MD engine that allows a high-performance simulation of large biomolecular systems.⁷⁴ It exhibits strong scaling on all DOE parallel platforms. For well-equilibrated long

simulation trajectories, the probability density distribution of a set of collective variable (CV), ξ , is given by

$$p(\xi) = \frac{\int dX \delta(\xi - \xi(X)) \exp(-U(X)/k_B T)}{\int dX \exp(-U(X)/k_B T)} \quad (1)$$

where X represents the configuration of the system, δ is the Dirac-delta function, k_B is the Boltzmann constant, T is the temperature, and $U(X)$ is the potential energy of the system. The PMF along ξ is defined by, $F(\xi) = -k_B T \ln p(\xi)$. Brute force MD simulations frequently suffer from insufficient sampling near (high) energy barriers. This can be remedied by various enhanced sampling approaches such as umbrella sampling (US)⁷⁵⁻⁷⁹ and replica exchange methods.^{65, 80-82}

Replica Exchange Umbrella Sampling with Solute Tempering (US/REST2)

The insufficient sampling in brute force MD simulations can be improved by simulations of multiple replicas in US. A typical US simulation system consists of a set of windows at a same temperature, T_0 , where the bias potential (umbrella potential) is applied to restrain $\xi(X)$ around a target value ξ_m . Typically, the bias potential is a harmonic function, $U_m^{\text{bias}}(\xi|X) = k[\xi(X) - \xi_m]^2/2$, where k is the force constant. In US, the sampling of other degrees of freedom (DOF) than the chosen CV (ξ) is not explicitly controlled, which may result in poor PMF profiles due to insufficient sampling along these DOF, e.g., the so-called Hamiltonian lagging⁸³ due to slow relaxation of environmental DOF. It has been reported that sampling along orthogonal DOF to ξ can be improved by multi-dimensional US and/or replica exchange US (REUS).⁸⁴⁻⁸⁶ Multi-dimensional US/REUS has been practically limited up to two dimensions due to exponential increase in computational costs with additional DOF. A practical approach for sampling the other DOF is a combination of REUS and replica exchange solute tempering (REST2),⁶⁵ US/REST2, which is supported in NAMD.³¹

In REST2,^{31, 65} all replicas run at the same temperature, T_0 , but the potential energy for each replica is scaled differently:

$$U_m^{\text{REST2}}(X) = \frac{\beta_m}{\beta_0} U_{ss}(X) + \sqrt{\frac{\beta_m}{\beta_0}} U_{sw}(X) + U_{ww}(X) \quad (2)$$

where U_{ss} , U_{sw} , and U_{ww} represent the solute–solute, solute–solvent, and solvent–solvent interaction energies, respectively. $\beta_m = (k_B T_m)^{-1}$. Essentially, the charge and vdw parameter of each atom of hot region is rescaled by a factor of $(\beta_m/\beta_0)^{1/2}$ and related bonded terms are rescaled correspondingly.

In US/REST2, the potential energy of the replica m is given by

$$U_m(X) = U_m^{\text{bias}}(\xi|X) + U_m^{\text{REST2}}(X) \quad (3)$$

By imposing detailed balance conditions, the exchange between replicas m and n is determined by a metropolis criterion, which is accepted with a probability, P_a

$$P_a = \begin{cases} 1 & \text{if } \Delta < 0 \\ \exp(-\Delta) & \text{otherwise} \end{cases} \quad (4.1)$$

where

$$\Delta = \frac{U_m(X_n) + U_n(X_m) - U_m(X_m) - U_n(X_n)}{k_B T_0} \quad (4.2)$$

and X_m and X_n are configurations of replicas m and n before exchange, respectively.

For uniformly distributed windows with a spacing, d , a force constant for a window potential can be optimally chosen using a simple relation, $k \approx 2k_B T(0.86/d)^2$ for $P_a \approx 0.4$.⁸⁶ For non-uniformly distributed windows, different value of k can be chosen for each window. The window distribution and associated bias force constants can be further optimized for uniform P_a between all neighboring window pairs during simulations.^{66, 87} The scaling factor of REST2 can be controlled by one of the standard keywords in the simulation control parameter class. Further, it is written into NAMD's Tcl script interface, offering flexibility for end users to change it on the fly and minimize overhead for replica exchange scheme. The nonbonded force field parameter rescaling is executed in force computing classes for direct nonbonded interaction and particle mesh Ewald (PME). For bonded terms, parameter rescaling is done in bonded tuple class template that defines a family of all bonded terms. In addition, the scaling factor can be arbitrarily adjusted if necessary for specific energy terms, instead of using original REST2 scheme. After the equilibration (typically a sequence of short simulations), US/REST2 simulations can be run without further optimization.

For each window in a US/REST2 system, all-atom simulation will be carried out using NAMD and the CHARMM C36 force field for protein,⁸⁸ lipid,⁸⁹ and carbohydrates,⁹⁰⁻⁹² and TIP3P water model.⁹³ A time step of 2 fs will be used with SHAKE algorithm⁹⁴ under constant pressure ($p = 1$ bar) and temperature ($T = 303.15$ K). Production run will be run until the PMF converges but not longer than the requested simulation time. The position and orientation of the substrate molecule will be recorded at every 0.25 ps and trajectories will be saved every 0.1 ns for the post analyses.

3.2. Parallel Programming Models

The parallel decomposition strategy used by NAMD is to treat the simulation cell (the volume of space containing the atoms) as a three-dimensional patchwork quilt, with each patch of sufficient size that only the 26 nearest-neighboring patches are involved in bonded, van der Waals, and short-range electrostatic interactions. The patches fill the simulation space in a regular grid and atoms in any pair of non-neighboring patches are separated by at least the cutoff distance at all times during the simulation. The number of patches varies from one to several hundred and is determined by the system size independently of the number of processors. NAMD is written in C++ and parallelized using Charm++, an overdecomposition-based message-driven parallel programming model. Charm++ encourages decomposition of the application domain into logical entities that resemble the data and work units of the application. The program is defined in terms of these entities, encapsulated as C++ objects, called chares. On most machines, Charm++ is implemented on top of the low-level communication layer provided by the system, for best performance (called native builds). PAMI, uGNI, and IB-Verbs are used on IBM Blue Gene/Q, Cray XE6/XK7/XC30, and InfiniBand clusters, respectively, for native builds. Charm++ can also be built on top of MPI on any system. Parallelization across threads within a process is based on the same message-driven thread-agnostic abstraction as parallelization between processes, with the option of exploiting shared memory between threads to avoid redundant copies of large unchanging data. Many Charm++ applications, including NAMD, OpenAtom, and EpiSimdemics, have been shown to scale to hundreds of thousands of cores using the native builds and the SMP mode. Charm++ supports splitting a single parallel job into multiple loosely coupled partitions, analogous to the result of MPI_Comm_split, with each partition providing a fully independent Charm++ runtime with a lower-level API for inter-partition messaging. Partitioning was originally developed to support NAMD multiple-copy algorithms (MCAs) with limited changes to the NAMD codebase. These algorithms are written using NAMD's high-level Tcl scripting interface. This generic version MCAs framework is the foundation of computational work in this proposal,

including free energy calculation and MD simulation accelerated with REST2. The MCAs framework has supported quite a few INCITE/ALCC projects since 2013.

3.3. Project workflow

NAMD has a default interface with the famous visualization software VMD. All MD trajectories will be stored in scratch directories and visualized with VMD on Cooley. All US/REST2 simulations are generic parallel jobs supported by NAMD, where no special workflow is required. Visualization and data postprocessing is executed after a simulation is done. No on-the-fly analysis/visualization is needed. VMD also provides user customized Tcl plugin for MD trajectory analysis.

3.3.1. Software packages and Programming Languages

The computational workflow utilizes the following software packages and programming languages at each stage:

- Pre-processing and Post-processing: shell/Tcl script and VMD
- MD simulation: NAMD (written in C/C++/charmm++)
- Reconstruction of the free energy landscape: WHAM code (or equivalent methods such as MBAR and DHAMed).

3.3.2. Software workflow solution

In-house shell scripts for job management will be used to monitor the overall progress of US/REST2 simulations. Additional shell/Tcl scripts can be combined to job management scripts for analysis after a simulation. The job management scripts will check the progress, status (running, pending, or completed), and submit jobs (both simulation and analysis) to facilitate the progress of the projects.

3.4. Job Characterization

A typical NAMD simulation in this proposal has a number of atoms ~100K or ~150K. Given the ensemble job nature of REST2 simulation, we evaluated the total node hours of US/REST2 simulations by measuring the performance of a small subset of US/REST2 windows (replicas). The required time and scalability were measured by benchmark runs with representative systems of ~150K atoms and 100K atoms with 16 REST2 replicas (**Figure 5**). Two types of simulations will be performed:

- Generation of configurations for REST2 and US/REST2 simulations (**Type 1**);
- US/REST2 simulations (**Type 2**) starting from configurations generated by Type 1 simulation;

Computing details of Types 1-2 are as follows (see also **Table 3**). For Milestone 1, the number of nodes required is 2,368 for 100K-atom system and 3,468 for 150K-atom system. For Milestone 2, the number of nodes required for POPE/POPG system in the horizontal and vertical pathways are 640 in each case. Initially, 1-ns brute force MD simulations with or without umbrella potential will be run, where the initial configurations with substrate located at different ξ for REST2 and

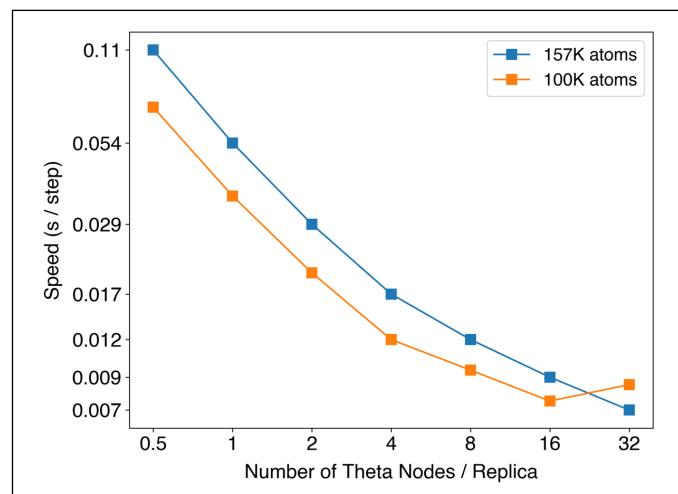


Figure 5. Scalability of simulation jobs for one replica as a function of the number of Theta nodes.

US/REST2 are generated in this stage. The simulation speed is 1.67 machine hours for 1 ns of simulation of a 150 K-atom system and 1.25 machine hours for 1 ns of simulation for a 100 K-atom system, based on the benchmark speed test with 8 Theta nodes/replica (**Figure 5**). The Type 2 simulations start from the last configurations of Type 1. The time lengths for Type 2 simulations are summarized in **Table 3**. A series of short REST2 and US/REST2 simulations will be carried out for optimization of parameters (window distribution, bias force constants, and REST2 scaling factors) for effective REST2 and US/REST2 production runs. Overall, Type 1 tasks require 14,300 node hours and for Type 2 tasks require 2,460,300 node hours. In total, we estimate that the whole project requires 2,484,600 node hours. **It needs to be stressed that NAMD CUDA (for both version NAMD2 and NAMD3) has been ported to the Polaris by ALCF early-access and exhibits similar single-node performance with general Nvidia A100 GPU cluster. With this project going during FY2023, we expect to request hour migration from Theta to Polaris.**

Table 3. Estimate of node hours for the proposed project

Milestone	System Size ^a	Number of Nodes	Job Type	Simulation Length (ns)	Speed (hour/ns) ^b	Node Hours Total ^c
1	100 K	$1 \times 2,368$	1	1	1.25	3.0 K
			2	100		275.0 K
2.a	150 K	$1 \times 3,648$	1	1	1.67	6.1 K
			2	220		1340.3 K
2.b	150 K	2×640 ^d	1	1	1.67	2.6 K
			2	200		427.5 K
Total						2484.6 K

^a An approximate number of atoms in a window/replica.

^b Speed is described by the number of machine hours needed for running 1 ns of simulation given the number of nodes in the third column.

^c The total node hours for each task is calculated by Number_of_Nodes x Simulation_Length x Speed.

^d The number of nodes is multiplied by 2, because there are two lipid systems, POPG and POPE.

3.5. I/O requirements

We can flexibly adapt our computations to fit into large runs of several hours, which will be carried out with restart files, in progressive increments of reasonable lengths. No exceptional I/O is required for both simulation and analysis. Team members will spend most of their time on carrying out production runs, and developmental work needed to improve the code performance. The production jobs will be executed at a steady state throughout the year for simulations.

3.6. Optimizations for the requested resources

In practical REST2 and US/REST2 simulations, REST2 scaling factors and window parameters (window distribution with associated bias force constants) will be updated through a series of short

simulations. After several updates, the chosen window parameters and REST2 scaling factors will be kept throughout the production runs.

4. PARALLEL PERFORMANCE

US/REST2 simulation

The proposed US/REST2 simulations consist of ~100K or ~150K atoms per replica. One systematic scaling benchmark was generated with a representative system, consisting of membrane protein MlaDE, lipids, water and ions, with a total atom number of 157K. The benchmark time from NAMD on Theta for one replica is 0.5 node: 0.110 seconds per timestep (s/step); 1 node: 0.054 s/step; 2 nodes: 0.029 s/step; 4 nodes: 0.0017 s/step; 8 nodes: 0.0012 s/step; 16 nodes: 0.0009 s/step; 32 nodes: 0.0007 s/step (**Figure 5**). This means that the strong scaling persists up to 32 nodes. The other benchmark used a system of 100 K atoms, containing OccK5, DMPC lipids, water and ions. The scaling persists up to 16 nodes for this system. All node hours requested in **Table 3** is based on the benchmark time of 8 nodes per replica. All US/REST2 jobs will be run in generic parallel/parallel mode.

5. Developmental Work

NAMD is well optimized on Knights Landing architecture and US/REST2 is well supported. Currently, NAMD is one project of DOE Exascale Early Science Program (ESP) and will be a major MD application code on the forthcoming Exascale supercomputers.

6. Data storage

The total space needed is about 10 Terabytes (TB) per year (**Table 4**). We estimate that this should support data for all the windows that we propose to analyze with extensive free energy landscape analysis. The input files should occupy << 1 TB, which just account for a tiny amount of the required storage, while output data are mainly long MD trajectory data including coordinates and velocities as well as restart files. The output data dominates storage demands. We do not anticipate any challenges with relocating data for long-term storage.

Table 4. Data storage required for the proposed project.

FY	Primary Resources	Alternative Resources	Node Hours	Online Storage	Archive Storage
2023	Theta		2.484 M	13 TB	13 TB

References

1. Neu, H. C., Infections Due to Gram-Negative Bacteria: An Overview. *Reviews of Infectious Diseases* **1985**, 7 (Supplement_4), S778-S782.
2. Oliveira, J.; Reygaert, W. C., Gram Negative Bacteria. In *StatPearls*, Treasure Island (FL), 2021.
3. Nikaido, H., Multidrug resistance in bacteria. *Annu Rev Biochem* **2009**, 78, 119-46.
4. CDC, U. *Antibiotic Resistance Threats in the United States*; 2019.
5. Callaway, E., The coronavirus is mutating - does it matter? *Nature* **2020**, 585 (7824), 174-177.
6. Li, C.; Yang, Y.; Ren, L., Genetic evolution analysis of 2019 novel coronavirus and coronavirus from other species. *Infection, Genetics and Evolution* **2020**, 82, 104285.
7. Mijalovic, H.; Smith, S. G., Bacterial self-defence: how *Escherichia coli* evades serum killing. *FEMS Microbiology Letters* **2014**, 354 (1), 1-9.
8. Henderson, J. C.; Zimmerman, S. M.; Crofts, A. A.; Boll, J. M.; Kuhns, L. G.; Herrera, C. M.; Trent, M. S., The Power of Asymmetry: Architecture and Assembly of the Gram-Negative Outer Membrane Lipid Bilayer. *Annual Review of Microbiology* **2016**, 70 (1), 255-278.
9. May, K. L.; Grabowicz, M., The bacterial outer membrane is an evolving antibiotic barrier. *Proceedings of the National Academy of Sciences* **2018**, 115 (36), 8852-8854.
10. Bertani, B.; Ruiz, N.; Slauch, J. M., Function and Biogenesis of Lipopolysaccharides. *EcoSal Plus* **2018**, 8 (1).
11. Lee, J.; Pothula, K. R.; Kleinekathöfer, U.; Im, W., Simulation Study of Occk5 Functional Properties in *Pseudomonas aeruginosa* Outer Membranes. *The Journal of Physical Chemistry B* **2018**, 122 (34), 8185-8192.
12. Snyder, D. S.; McIntosh, T. J., The Lipopolysaccharide Barrier: Correlation of Antibiotic Susceptibility with Antibiotic Permeability and Fluorescent Probe Binding Kinetics. *Biochemistry* **2000**, 39 (38), 11777-11787.
13. Nikaido, H., Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiology and Molecular Biology Reviews* **2003**, 67 (4), 593-656.
14. Huang, H.; Hancock, R. E., Genetic definition of the substrate selectivity of outer membrane porin protein OprD of *Pseudomonas aeruginosa*. *Journal of Bacteriology* **1993**, 175 (24), 7793-7800.
15. Li, H.; Luo, Y.-F.; Williams, B. J.; Blackwell, T. S.; Xie, C.-M., Structure and function of OprD protein in *Pseudomonas aeruginosa*: From antibiotic resistance to novel therapies. *International Journal of Medical Microbiology* **2012**, 302 (2), 63-68.
16. Liu, J.; Eren, E.; Vijayaraghavan, J.; Cheneke, B. R.; Indic, M.; van den Berg, B.; Movileanu, L., OccK channels from *Pseudomonas aeruginosa* exhibit diverse single-channel electrical signatures but conserved anion selectivity. *Biochemistry* **2012**, 51 (11), 2319-30.
17. Soh, S. M.; Jang, H.; Mitchell, R. J., Loss of the lipopolysaccharide (LPS) inner core increases the electrocompetence of *Escherichia coli*. *Applied Microbiology and Biotechnology* **2020**, 104 (17), 7427-7435.
18. Vaara, M., Agents that increase the permeability of the outer membrane. *Microbiological Reviews* **1992**, 56 (3), 395-411.

19. Vaara, M., Outer membrane permeability barrier to azithromycin, clarithromycin, and roxithromycin in gram-negative enteric bacteria. *Antimicrobial Agents and Chemotherapy* **1993**, 37 (2), 354-356.
20. Lundstedt, E.; Kahne, D.; Ruiz, N., Assembly and Maintenance of Lipids at the Bacterial Outer Membrane. *Chemical Reviews* **2021**, 121 (9), 5098-5123.
21. Tang, X.; Chang, S.; Qiao, W.; Luo, Q.; Chen, Y.; Jia, Z.; Coleman, J.; Zhang, K.; Wang, T.; Zhang, Z.; Zhang, C.; Zhu, X.; Wei, X.; Dong, C.; Zhang, X.; Dong, H., Structural insights into outer membrane asymmetry maintenance in Gram-negative bacteria by MlaFEDB. *Nature Structural & Molecular Biology* **2021**, 28 (1), 81-91.
22. Zhou, C.; Shi, H.; Zhang, M.; Zhou, L.; Xiao, L.; Feng, S.; Im, W.; Zhou, M.; Zhang, X.; Huang, Y., Structural Insight into Phospholipid Transport by the MlaFEBD Complex from P. aeruginosa. *Journal of Molecular Biology* **2021**, 433 (13), 166986.
23. Kolich, L. R.; Chang, Y. T.; Coudray, N.; Giacometti, S. I.; MacRae, M. R.; Isom, G. L.; Teran, E. M.; Bhabha, G.; Ekiert, D. C., Structure of MlaFB uncovers novel mechanisms of ABC transporter regulation. *Elife* **2020**, 9.
24. Coudray, N.; Isom, G. L.; MacRae, M. R.; Saiduddin, M. N.; Bhabha, G.; Ekiert, D. C., Structure of bacterial phospholipid transporter MlaFEDB with substrate bound. *Elife* **2020**, 9.
25. Mann, D.; Fan, J.; Farrell, D. P.; Somboon, K.; Muenks, A.; Tzokov, S. B.; Khalid, S.; Dimaio, F.; Miller, S. I.; Bergeron, J. R. C., Structure and lipid dynamics in the A. baumannii maintenance of lipid asymmetry (MLA) inner membrane complex. *bioRxiv* **2020**, 2020.05.30.125013.
26. Chi, X.; Fan, Q.; Zhang, Y.; Liang, K.; Wan, L.; Zhou, Q.; Li, Y., Structural mechanism of phospholipids translocation by MlaFEDB complex. *Cell Research* **2020**, 30 (12), 1127-1135.
27. H.A., Z., ADMET Properties: Overview and Current Topics. In *Drug Design: Principles and Applications*, A., G., Ed. Springer, Singapore: 2017.
28. Prajapati, J. D.; Kleinekathöfer, U.; Winterhalter, M., How to Enter a Bacterium: Bacterial Porins and the Permeation of Antibiotics. *Chemical Reviews* **2021**, 121 (9), 5158-5192.
29. Fairman, J. W.; Noinaj, N.; Buchanan, S. K., The structural biology of β-barrel membrane proteins: a summary of recent reports. *Current Opinion in Structural Biology* **2011**, 21 (4), 523-531.
30. Sansom, M. S. P., Membrane proteins: A tale of barrels and corks. *Current Biology* **1999**, 9 (7), R254-R257.
31. Jo, S.; Jiang, W., A generic implementation of replica exchange with solute tempering (REST2) algorithm in NAMD for complex biophysical simulations. *Computer Physics Communications* **2015**, 197, 304-311.
32. Bajaj, H.; Scorciapino, M. A.; Moynié, L.; Page, M. G. P.; Naismith, J. H.; Ceccarelli, M.; Winterhalter, M., Molecular Basis of Filtering Carbapenems by Porins from β-Lactam-resistant Clinical Strains of Escherichia coli*. *Journal of Biological Chemistry* **2016**, 291 (6), 2837-2847.
33. Bartsch, A.; Ives, C. M.; Kattner, C.; Pein, F.; Diehn, M.; Tanabe, M.; Munk, A.; Zachariae, U.; Steinem, C.; Llabrés, S., An antibiotic-resistance conferring mutation in a neisserial porin: Structure, ion flux, and ampicillin binding. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2021**, 1863 (6), 183601.

34. Dai, Y.; Ma, H.; Wu, M.; Welsch, T. A.; Vora, S. R.; Ren, D.; Nangia, S., Development of the computational antibiotic screening platform (CLASP) to aid in the discovery of new antibiotics. *Soft Matter* **2021**, *17* (10), 2725-2736.
35. Samanta, S.; Scorciapino, M. A.; Ceccarelli, M., Molecular basis of substrate translocation through the outer membrane channel OprD of *Pseudomonas aeruginosa*. *Physical Chemistry Chemical Physics* **2015**, *17* (37), 23867-23876.
36. Pothula, K. R. Molecular Basis of Gating and Permeation in Bacterial Outer Membrane Channels. Jacobs Univ., 2017.
37. Golla, V. K.; Prajapati, J. D.; Joshi, M.; Kleinekathöfer, U., Exploration of Free Energy Surfaces Across a Membrane Channel Using Metadynamics and Umbrella Sampling. *Journal of Chemical Theory and Computation* **2020**, *16* (4), 2751-2765.
38. Golla, V. K.; Sans-Serramitjana, E.; Pothula, K. R.; Benier, L.; Bafna, J. A.; Winterhalter, M.; Kleinekathöfer, U., Fosfomycin Permeation through the Outer Membrane Porin OmpF. *Biophysical Journal* **2019**, *116* (2), 258-269.
39. Samsudin, F.; Khalid, S., Movement of Arginine through OprD: The Energetics of Permeation and the Role of Lipopolysaccharide in Directing Arginine to the Protein. *The Journal of Physical Chemistry B* **2019**, *123* (13), 2824-2832.
40. Cetuk, H.; Anishkin, A.; Scott, A. J.; Rempe, S. B.; Ernst, R. K.; Sukharev, S., Partitioning of Seven Different Classes of Antibiotics into LPS Monolayers Supports Three Different Permeation Mechanisms through the Outer Bacterial Membrane. *Langmuir* **2021**, *37* (4), 1372-1385.
41. Parkin, J.; Khalid, S., Atomistic Molecular-Dynamics Simulations Enable Prediction of the Arginine Permeation Pathway through OccD1/OprD from *Pseudomonas aeruginosa*. *Biophysical Journal* **2014**, *107* (8), 1853-1861.
42. Liu, J.; Eren, E.; Vijayaraghavan, J.; Cheneke, B. R.; Indic, M.; van den Berg, B.; Movileanu, L., OccK Channels from *Pseudomonas aeruginosa* Exhibit Diverse Single-Channel Electrical Signatures but Conserved Anion Selectivity. *Biochemistry* **2012**, *51* (11), 2319-2330.
43. Eren, E.; Vijayaraghavan, J.; Liu, J.; Cheneke, B. R.; Touw, D. S.; Lepore, B. W.; Indic, M.; Movileanu, L.; van den Berg, B., Substrate Specificity within a Family of Outer Membrane Carboxylate Channels. *PLOS Biology* **2012**, *10* (1), e1001242.
44. Biswas, S.; Mohammad, M. M.; Patel, D. R.; Movileanu, L.; van den Berg, B., Structural insight into OprD substrate specificity. *Nature Structural & Molecular Biology* **2007**, *14* (11), 1108-1109.
45. Luo, Q.; Yang, X.; Yu, S.; Shi, H.; Wang, K.; Xiao, L.; Zhu, G.; Sun, C.; Li, T.; Li, D.; Zhang, X.; Zhou, M.; Huang, Y., Structural basis for lipopolysaccharide extraction by ABC transporter LptB2FG. *Nature Structural & Molecular Biology* **2017**, *24* (5), 469-474.
46. Li, Y.; Orlando, B. J.; Liao, M., Structural basis of lipopolysaccharide extraction by the LptB2FGC complex. *Nature* **2019**, *567* (7749), 486-490.
47. Lee, J.; Patel, D. S.; Kucharska, I.; Tamm, L. K.; Im, W., Refinement of OprH-LPS Interactions by Molecular Simulations. *Biophysical Journal* **2017**, *112* (2), 346-355.
48. Patel, Dhilon S.; Re, S.; Wu, Emilia L.; Qi, Y.; Klebba, Phillip E.; Widmalm, G.; Yeom, Min S.; Sugita, Y.; Im, W., Dynamics and Interactions of OmpF and LPS: Influence on Pore Accessibility and Ion Permeability. *Biophysical Journal* **2016**, *110* (4), 930-938.

49. Wu, Emilia L.; Fleming, Patrick J.; Yeom, Min S.; Widmalm, G.; Klauda, Jeffery B.; Fleming, Karen G.; Im, W., E. coli Outer Membrane and Interactions with OmpLA. *Biophysical Journal* **2014**, *106* (11), 2493-2502.
50. Wu, Emilia L.; Engström, O.; Jo, S.; Stuhlsatz, D.; Yeom, Min S.; Klauda, Jeffery B.; Widmalm, G.; Im, W., Molecular Dynamics and NMR Spectroscopy Studies of E. coli Lipopolysaccharide Structure and Dynamics. *Biophysical Journal* **2013**, *105* (6), 1444-1455.
51. Patel, D. S.; Blasco, P.; Widmalm, G.; Im, W., Escherichia coli O176 LPS structure and dynamics: A NMR spectroscopy and MD simulation study. *Current Research in Structural Biology* **2020**, *2*, 79-88.
52. Kesireddy, A.; Pothula, K. R.; Lee, J.; Patel, D. S.; Pathania, M.; van den Berg, B.; Im, W.; Kleinekathöfer, U., Modeling of Specific Lipopolysaccharide Binding Sites on a Gram-Negative Porin. *The Journal of Physical Chemistry B* **2019**, *123* (27), 5700-5708.
53. Gao, Y.; Lee, J.; Widmalm, G.; Im, W., Modeling and Simulation of Bacterial Outer Membranes with Lipopolysaccharides and Enterobacterial Common Antigen. *The Journal of Physical Chemistry B* **2020**, *124* (28), 5948-5956.
54. Blasco, P.; Patel, D. S.; Engström, O.; Im, W.; Widmalm, G., Conformational Dynamics of the Lipopolysaccharide from Escherichia coli O91 Revealed by Nuclear Magnetic Resonance Spectroscopy and Molecular Simulations. *Biochemistry* **2017**, *56* (29), 3826-3839.
55. Luna, E.; Kim, S.; Gao, Y.; Widmalm, G.; Im, W., Influences of Vibrio cholerae Lipid A Types on LPS Bilayer Properties. *The Journal of Physical Chemistry B* **2021**, *125* (8), 2105-2112.
56. Hughes, A. V.; Patel, D. S.; Widmalm, G.; Klauda, J. B.; Clifton, L. A.; Im, W., Physical Properties of Bacterial Outer Membrane Models: Neutron Reflectometry & Molecular Simulation. *Biophysical Journal* **2019**, *116* (6), 1095-1104.
57. Kim, S.; Patel, Dhilon S.; Park, S.; Slusky, J.; Klauda, Jeffery B.; Widmalm, G.; Im, W., Bilayer Properties of Lipid A from Various Gram-Negative Bacteria. *Biophysical Journal* **2016**, *111* (8), 1750-1760.
58. Zhuang, X.; Dávila-Contreras, E. M.; Beaven, A. H.; Im, W.; Klauda, J. B., An extensive simulation study of lipid bilayer properties with different head groups, acyl chain lengths, and chain saturations. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2016**, *1858* (12), 3093-3104.
59. Boughter, C. T.; Monje-Galvan, V.; Im, W.; Klauda, J. B., Influence of Cholesterol on Phospholipid Bilayer Structure and Dynamics. *The Journal of Physical Chemistry B* **2016**, *120* (45), 11761-11772.
60. Feng, S.; Wang, R.; Pastor, R. W.; Klauda, J. B.; Im, W., Location and Conformational Ensemble of Menaquinone and Menaquinol, and Protein-Lipid Modulations in Archaeal Membranes. *The Journal of Physical Chemistry B* **2021**, *125* (18), 4714-4725.
61. Ren, Z.; Lee, J.; Moosa, M. M.; Nian, Y.; Hu, L.; Xu, Z.; McCoy, J. G.; Ferreon, A. C. M.; Im, W.; Zhou, M., Structure of an EIIC sugar transporter trapped in an inward-facing conformation. *Proceedings of the National Academy of Sciences* **2018**, *115* (23), 5962-5967.
62. McCoy, Jason G.; Ren, Z.; Stanevich, V.; Lee, J.; Mitra, S.; Levin, Elena J.; Poget, S.; Quick, M.; Im, W.; Zhou, M., The Structure of a Sugar Transporter of the Glucose EIIC Superfamily Provides Insight into the Elevator Mechanism of Membrane Transport. *Structure* **2016**, *24* (6), 956-964.

63. Lee, J.; Ren, Z.; Zhou, M.; Im, W., Molecular Simulation and Biochemical Studies Support an Elevator-type Transport Mechanism in EIIC. *Biophysical Journal* **2017**, *112* (11), 2249-2252.
64. Min, D.; Jefferson, R. E.; Qi, Y.; Wang, J. Y.; Arbing, M. A.; Im, W.; Bowie, J. U., Unfolding of a CIC chloride transporter retains memory of its evolutionary history. *Nature Chemical Biology* **2018**, *14* (5), 489-496.
65. Wang, L.; Friesner, R. A.; Berne, B. J., Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *The Journal of Physical Chemistry B* **2011**, *115* (30), 9431-9438.
66. Park, S.; Im, W., Theory of Adaptive Optimization for Umbrella Sampling. *Journal of Chemical Theory and Computation* **2014**, *10* (7), 2719-2728.
67. Acharya, A.; Prajapati, J. D.; Kleinekathöfer, U., Improved Sampling and Free Energy Estimates for Antibiotic Permeation through Bacterial Porins. *Journal of Chemical Theory and Computation* **2021**, *17* (7), 4564-4577.
68. Triantafilou, M.; Triantafilou, K.; Fernandez, N., Rough and smooth forms of fluorescein-labelled bacterial endotoxin exhibit CD14/LBP dependent and independent binding that is influenced by endotoxin concentration. *European Journal of Biochemistry* **2000**, *267* (8), 2218-2226.
69. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A., THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* **1992**, *13* (8), 1011-1021.
70. Shirts, M. R.; Chodera, J. D., Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of Chemical Physics* **2008**, *129* (12), 124105.
71. Stelzl, L. S.; Kells, A.; Rosta, E.; Hummer, G., Dynamic Histogram Analysis To Determine Free Energies and Rates from Biased Simulations. *Journal of Chemical Theory and Computation* **2017**, *13* (12), 6328-6342.
72. Suh, D.; Jo, S.; Jiang, W.; Chipot, C.; Roux, B., String Method for Protein–Protein Binding Free-Energy Calculations. *Journal of Chemical Theory and Computation* **2019**, *15* (11), 5829-5844.
73. Darve, E.; Pohorille, A., Calculating free energies using average force. *The Journal of Chemical Physics* **2001**, *115* (20), 9169-9183.
74. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K., Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **2005**, *26* (16), 1781-1802.
75. Maragakis, P.; van der Vaart, A.; Karplus, M., Gaussian-Mixture Umbrella Sampling. *The Journal of Physical Chemistry B* **2009**, *113* (14), 4664-4673.
76. Marsili, S.; Barducci, A.; Chelli, R.; Procacci, P.; Schettino, V., Self-healing Umbrella Sampling: A Non-equilibrium Approach for Quantitative Free Energy Calculations. *The Journal of Physical Chemistry B* **2006**, *110* (29), 14011-14013.
77. Wojtas-Niziurski, W.; Meng, Y.; Roux, B. t.; Bernèche, S., Self-Learning Adaptive Umbrella Sampling Method for the Determination of Free Energy Landscapes in Multiple Dimensions. *Journal of Chemical Theory and Computation* **2013**, *9* (4), 1885-1895.
78. Mezei, M., Adaptive umbrella sampling: Self-consistent determination of the non-Boltzmann bias. *Journal of Computational Physics* **1987**, *68* (1), 237-248.
79. Torrie, G. M.; Valleau, J. P., Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics* **1977**, *23* (2), 187-199.

80. Sugita, Y.; Okamoto, Y., Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters* **1999**, *314* (1), 141-151.
81. Swendsen, R. H.; Wang, J.-S., Replica Monte Carlo Simulation of Spin-Glasses. *Physical Review Letters* **1986**, *57* (21), 2607-2609.
82. Sugita, Y.; Kitao, A.; Okamoto, Y., Multidimensional replica-exchange method for free-energy calculations. *The Journal of Chemical Physics* **2000**, *113* (15), 6042-6051.
83. Pearlman, D. A.; Kollman, P. A., The lag between the Hamiltonian and the system configuration in free energy perturbation calculations. *The Journal of Chemical Physics* **1989**, *91* (12), 7831-7839.
84. Dickson, A.; Ahlstrom, L. S.; Brooks III, C. L., Coupled folding and binding with 2D Window-Exchange Umbrella Sampling. *Journal of Computational Chemistry* **2016**, *37* (6), 587-594.
85. Park, S.; Im, W., Two Dimensional Window Exchange Umbrella Sampling for Transmembrane Helix Assembly. *Journal of Chemical Theory and Computation* **2013**, *9* (1), 13-17.
86. Park, S.; Kim, T.; Im, W., Transmembrane Helix Assembly by Window Exchange Umbrella Sampling. *Physical Review Letters* **2012**, *108* (10), 108102.
87. Sabri Dashti, D.; Roitberg, A. E., Optimization of Umbrella Sampling Replica Exchange Molecular Dynamics by Replica Positioning. *Journal of Chemical Theory and Computation* **2013**, *9* (11), 4692-4699.
88. Huang, J.; MacKerell Jr, A. D., CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry* **2013**, *34* (25), 2135-2145.
89. Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W., Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *The Journal of Physical Chemistry B* **2010**, *114* (23), 7830-7843.
90. Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; MacKerell Jr, A. D., Additive empirical force field for hexopyranose monosaccharides. *Journal of Computational Chemistry* **2008**, *29* (15), 2543-2564.
91. Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D., CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *Journal of Chemical Theory and Computation* **2009**, *5* (9), 2353-2370.
92. Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W.; MacKerell, A. D., CHARMM Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate-Protein Modeling. *Journal of Chemical Theory and Computation* **2011**, *7* (10), 3162-3180.
93. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79* (2), 926-935.
94. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23* (3), 327-341.

Personal Justification and Management plan

A team of computational biophysicists led by Dr. Wonpil Im will perform this project. Dr. Im is Presidential Endowed Chair in Health, Science, and Engineering and a professor of Biological Sciences, Chemistry, Bioengineering, and Computer Science and Engineering at Lehigh University. His research has been funded by numerous NIH and NSF grants. The Im research group is specialized in developing community simulation software CHARMM-GUI, membrane protein and membrane properties research, bacterial outer membrane/LPS modeling, transporter and ion channel investigation. The PI is highly skilled, trained, and experienced in the field of biomolecular modeling and simulation and have a prolific record of publication.

Dr. Park is a research scientist in the Department of Biological Sciences at Lehigh University. He has expertise in umbrella sampling (US) and optimization of (replica exchange) US, where he obtained a simple analytic relation between the bias force constant and the window spacing, as well as a theory for adaptive optimization of US. Also, he has rich experience in the Hamiltonian replica exchange approach analogous to REST2. He is well acquainted with membranes, lipopolysaccharide (LPS) lipids, and membrane proteins.

Ms. Feng is a rising fifth-year graduate student in the Ph.D. program of Biology in the Department of Biological Sciences at Lehigh University. She has extensively studied lipid conformation ensemble and dynamics in the previous work of archaeal membranes, acquired in-depth knowledge of the Mla complex with microsecond-long simulations of MlaBDEF on Anton2 and GaMD, REST2, ABF simulations of MlaDE-lipid complex. She is versed in multiple main simulation programs, proficient in data analysis, and has received training on parallel computing and CUDA computing.

a. Personnel already in place

All personnel are in place.

Management plan

The project will be managed by Dr. Wonpil Im who will hold regular meetings in person or via zoom with Dr. Park and Ms. Feng. All final decisions about the research direction and progress will be Dr. Im's responsibility. The meetings will be bi-weekly between the three researchers, more often when necessary. The project progress will be reported and recorded on a shared document drive.

Dr. Park and Ms. Feng will work on system building, equilibration, running simulations, and the data analysis. The three researchers will be all responsible for preparation of manuscripts and contributing to discussion of progress, results, and research interpretation and direction. Dr. Im will be responsible for the timely progress of all Milestones described in the Milestones section of the project narrative document.

Year 1			Title
Milestone:	Details (as appropriate):	Dates:	
<p>1. Quantitative characterization of effects of LPS polysaccharide length on substrate transport</p> <p><u>1.a</u> Quantitative characterization of citrate translocation through OccK5 in Pa.G2 outer membrane;</p> <p><u>1.b</u> Quantitative characterization of citrate translocation through OccK5 in Pa.G2.O10 outer membrane;</p>	<p>Resource: Theta Node-hours: 1,624,400 Filesystem storage (TB and dates): 10 TB Archival storage (TB and dates): 10 TB Software Application: NAMD 3.0 Tasks: parallel/parallel computation on 2368 and 3648 nodes of Theta Dependencies: No dependencies, start immediately.</p>		01/01/2023- 12/31/2023
<p>2. Elucidating energetic landscape of lipid diffusion pathway in Mla</p> <p><u>2.a</u> Study the vertical diffusion pathway inside the MlaD crown</p> <p><u>2.b</u> Study the horizontal diffusion pathway along the MlaE vestibule</p>	<p>Resource: Theta Node-hours: 860,200 Filesystem storage (TB and dates): 3 TB Archival storage (TB and dates): 3 TB Software Application: NAMD 3.0 Tasks: parallel computation on 640 nodes of Theta Dependencies: No dependencies, start immediately.</p>		01/01/2023- 12/31/2023

WONPIL IM**EDUCATION & CAREER**

- 06/2002 — 08/2005 NSF CTBP (Center for Theoretical Biological Physics) Fellow
Postdoctoral Associate in the group of Professor Charles Brooks,
The Scripps Research Institute, La Jolla
- 02/1997 — 05/2002 Ph. D. in Biochemistry
Weill Medical College of Cornell University, New York
Thesis Supervisor: Professor Benoît Roux
- (02/1997 — 05/2000 Ph. D student in Chemistry, University of Montreal, Montreal)
- 03/1994 — 02/1996 M. Sc. in Chemistry, Hanyang University, Seoul
Thesis Supervisor: Professor Youngdo Won
- 03/1990 — 02/1994 B. Sc. in Chemistry, Hanyang University, Seoul

RESEARCH & PROFESSIONAL EXPERIENCE

- 08/2016 — Present Presidential Endowed Chair in Health - Science and Engineering
Professor of Biological Sciences, Chemistry, and Bioengineering
Lehigh University, Bethlehem
- 08/2015 — 07/2016 Professor
- 08/2011 — 07/2015 Associate Professor
- 08/2005 — 07/2011 Assistant Professor
Center for Computational Biology and
Department of Molecular Biosciences
The University of Kansas, Lawrence

PUBLICATIONS

1. C. Zhou, H. Shi, M. Zhang, L. Zhou, L. Xiao, S. Feng, W. Im, M. Zhou, X. Zhang, and Y. Huang, Structural Insight into Phospholipid Transport by the MlaFEBD Complex from *P. aeruginosa*. *J. Mol. Biol.* 433:166986 (2021).
2. S. Park, M.S. Yeom, O.S. Andersen, R.W. Pastor, and W. Im, Quantitative Characterization of Protein-Lipid Interactions by Free Energy Simulation Between Binary Bilayers. *J. Chem. Theory Comput.* 15:6491-6503 (2019).
3. J. Lee, D.S. Patel, J. Stähle, S-J. Park, N.R. Kern, S. Kim, J. Lee, X. Cheng, M.A. Valvano, O. Holst, Y. Knirel, Y. Qi, S. Jo, J.B. Klauda, G. Widmalm, and W. Im, CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *J. Chem. Theory Comput.* 15:775-786 (2019).
4. S. Park and W. Im, Theory of Adaptive Optimization for Umbrella Sampling. *J. Chem. Theory Comput.* 10:2719-2728 (2014).
5. S. Park, T. Kim, and W. Im, Transmembrane Helix Assembly with Window Exchange Umbrella Sampling. *Phys. Rev. Lett.* 108:108102 (2012).

RESEARCH INTEREST & EXPERTISE

1. Protein/peptide interactions with/in biological membranes
2. Transmembrane-induced signaling and regulation
3. NMR structure calculation & refinement
4. Modeling and simulation of glycoconjugates (<http://www.glycanstructure.org>)
5. Bacterial outer membranes and interactions with proteins

6. Protein-ligand and protein-protein interactions (<http://compbio.lehigh.edu/GLoSA>)
7. CHARMM-GUI development (<http://www.charmm-gui.org>)

SYNERGESTIC ACTIVITIES

1. CHARMM-GUI KIAS School, Seoul, Korea (2019)
2. CHARMM-GUI CECAM School, Lausanne, Switzerland (2018)
3. Membrane Protein Simulations and Free Energy Approaches at the ACS meeting, Boston, USA (2018)
4. Molecular Basis of Antibiotic Permeability in Gram-negative Bacteria, Braunschweig, Germany (2017)
5. CECAM Workshop: Tackling Complexity of the Nano/Bio Interface - Computational and Experimental Approaches, Bremen, Germany (2017)

COLLABORATORS

Alexey Aleksandrov	Ecole Polytechnique
Melissa J. Call	The Walter and Eliza Hall Institute of Medical Research
Matthew E. Call	The Walter and Eliza Hall Institute of Medical Research
Luke A. Clifton	Rutherford Appleton Laboratory
Tristan I. Croll	University of Cambridge
Richard D. Cummings	Harvard Medical School
Eric J. Deeds	University of California Los Angeles
Michael Feig	Michigan State University
Martin Frank	Biognos AB
Ya Gao	Shanghai University of Engineering Science
Howard C. Hang	Scripps Research
Hendrik Heinz	University of Colorado, Boulder
Yihua Huang	Chinese Academy of Sciences
Arwel V. Hughes	Rutherford Appleton Laboratory
Shahidul M. Islam	University of Illinois, Chicago
Wei Jiang	Argonne National Laboratory
Syma Khalid	University of Southampton
Seonghoon Kim	Korea Institute for Advanced Study
Jeffery B. Klauda	University of Maryland
Ulrich Kleinekathöfer	Jacobs University
Alexander D. MacKerell, Jr.	University of Maryland
Francesca M. Marassi	Sanford Burnham Prebys Medical Discovery Institute
Richard W. Pastor	National Institutes of Health
Marcos M. Pires	University of Virginia
Yifei Qi	East China Normal University
Benoit Roux	University of Chicago
Chaok Seok	Seoul National University
Yuji Sugita	RIKEN
Goran Widmalm	Stockholm University
Martin Zacharias	Technical University of Munich
Ming Zhou	Baylor College of Medicine

Section 6: Software Applications and Packages

Question #1

Please list any software packages used by the project, and indicate if they are on open source or export controlled.

Application Packages

Package Name

NAMD

Indicate whether Open Source or Export Controlled.

Open Source

Package Name

VMD

Indicate whether Open Source or Export Controlled.

Open Source

Section 7: Wrap-Up Questions

Question #1

National Security Decision Directive (NSDD) 189 defines Fundamental Research as "basic and applied research in science and engineering, the results of which ordinarily are published and shared broadly within the scientific community, as distinguished from proprietary research and from industrial development, design, production, and product utilization, the results of which ordinarily are restricted for proprietary or national security reasons." Publicly Available Information is defined as information obtainable free of charge (other than minor shipping or copying fees) and without restriction, which is available via the internet, journal publications, textbooks, articles, newspapers, magazines, etc.

The INCITE program distinguishes between the generation of proprietary information (deemed a proprietary project) and the use of proprietary information as input. In the latter, the project may be considered as Fundamental Research or nonproprietary under the terms of the nonproprietary user agreement. Proprietary information, including computer codes and data, brought into the LCF for use by the project - but not for generation of new intellectual property, etc., using the facility resources - may be protected under a nonproprietary user agreement.

Proprietary Information

Are the proposed project and its intended outcome considered Fundamental Research or Publicly Available Information?

Yes

Will the proposed project use proprietary information, intellectual property, or licensing?

No

Will the proposed project generate proprietary information, intellectual property, or licensing as the result of the work being proposed?

If the response is Yes, please contact the INCITE manager, INCITE@doeleadershipcomputing.org, prior to submittal to discuss the INCITE policy on proprietary work.

No

Question #2

The following questions are provided to determine whether research associated with an INCITE proposal may be export controlled. Responding to these questions can facilitate - but not substitute for - any export control review required for this proposal.

PIs are responsible for knowing whether their project uses or generates sensitive or restricted information. Department of Energy systems contain only data related to scientific research and do not contain personally identifiable information. Therefore, you should answer "Yes" if your project uses or generates data that fall under the Privacy Act of 1974 U.S.C. 552a. Use of high-performance computing resources to store, manipulate, or remotely access any national security information is prohibited. This includes, but is not limited to, classified information, unclassified controlled nuclear information (UCNI); naval nuclear propulsion information (NNPI); and the design or development of nuclear, biological, or chemical weapons or of any weapons of mass destruction. For more information contact the Office of Domestic and International Energy Policy, Department of Energy, Washington DC 20585, 202-586-9211.

Export Control

Does this project use or generate sensitive or restricted information?

No

Does the proposed project involve any of the following areas?

- i. Military, space craft, satellites, missiles, and associated hardware, software or technical data**
- ii. Nuclear reactors and components, nuclear material enrichment equipment, components (Trigger List) and associated hardware, software or technical data**
- iii. Encryption above 128 bit software (source and object code)**

- iv. Weapons of mass destruction or their precursors (nuclear, chemical and biological)**

No

Does the proposed project involve International Traffic in Arms Regulations (ITAR)?

No

Question #3

The following questions deal with health data. PIs are responsible for knowing if their project uses any health data and if that data is protected. Note that certain health data may fall both within these questions as well as be considered sensitive as per question #2. Questions regarding these answers to these questions should be directed to the centers or program manager prior to submission.

Health Data

Will this project use health data?

No

Will this project use human health data?

No

Will this project use Protected Health Information (PHI)?

No

Question #4

The PI and designated Project Manager agree to the following:

Monitor Agreement

I certify that the information provided herein contains no proprietary or export control material and is correct to the best of my knowledge.

Yes

I agree to provide periodic updates of research accomplishments and to acknowledge INCITE and the LCF in publications resulting from an INCITE award.

Yes

I agree to monitor the usage associated with an INCITE award to ensure that usage is only for the project being described herein and that all U. S. Export Controls are complied with.

Yes

I understand that the INCITE program reserves the right to periodically redistribute allocations from underutilized projects.

Yes

Section 8: Outreach and Suggested Reviewers

Question #1

By what sources (colleagues, web sites, email notices, other) have you heard about the INCITE program? This information will help refine our outreach efforts.

Outreach

Question #2

Suggested Reviewers

Section 9: Testbed Resources

Question #1

The ALCF and OLCF have test bed resources for new technologies, details below. If you would like access to these resources to support the work in this proposal, please provide the information below.

(1 Page Limit)

The OLCF Quantum Computing User Program is designed to enable research by providing a broad spectrum of user access to the best available quantum computing systems, evaluate technology by monitoring the breadth and performance of early quantum computing applications, and Engage the quantum computing community and support the growth of the quantum information science ecosystems. More information can be found here: <https://www.olcf.ornl.gov/olcf-resources/compute-systems/quantum-computing-user-program/quantum-computing-user-support-documentation>.

The ALCF AI Testbed provides access to next-generation of AI-accelerator machines to enable evaluation of both hardware and workflows. Current hardware available includes Cerebras C-2, Graphcore MK1, Groq, Habana Gaudi, and SambaNova Dataflow. New hardware is regularly acquired as it becomes available. Up to date information can be found here: <https://www.alcf.anl.gov/alcf-ai-testbed>.

Describe the experiments you would be interested in performing, resources required, and their relationship to the current proposal. Please note, these are smaller experimental resources and a large amount of resources are not available. Instead, these resources are to explore the possibilities for these technologies might innovate future work. This request does not contribute to the 15-page proposal limit.

No-Testbed.pdf

The attachment is on the following page.

Testbed Resources are not necessary for this project.