ACS OMEGA

Open Access

http://pubs.acs.org/journal/acsodf

Article

Molecular Dynamics of Outer Membrane-Embedded Polysaccharide Secretion Porins Reveals Closed Resting-State Surface Gates Targetable by Virtual Fragment Screening for Drug Hotspot Identification

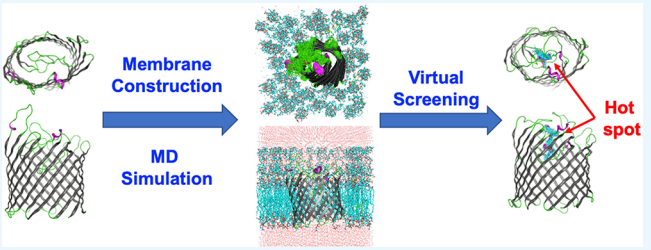
Tanos C. C. França, Fares Saïdi, Alain Ajamian, Salim T. Islam,* and Steven R. LaPlante*

Cite This: *ACS Omega* 2024, 9, 13217−13226

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Recent advances in iterative neural network analyses (e.g., AlphaFold2 and RoseTTA fold) have been revolutionary for protein 3D structure prediction, especially for difficult-to-manipulate α-helical/β-barrel integral membrane proteins. These model structures are calculated based on the coevolution of amino acids within the protein of interest and similarities to existing protein structures; the local effects of the membrane on folding and stability of the calculated model structures are not considered. We recently reported the discovery, 3D modeling, and characterization of 18-β-stranded outer-membrane (OM) WzpX, WzpS, and WzpB β-barrel secretion porins for the exopolysaccharide (EPS), major spore coat polysaccharide (MASC), and biosurfactant polysaccharide (BPS) pathways (respectively) in the Gram-negative social predatory bacterium *Myxococcus xanthus* DZ2. However, information was not obtained regarding the dynamic behavior of surface-gating WzpX/S/B loop domains or on potential treatments to inactivate these porins. Herein, we developed a molecular dynamics (MD) protocol to study the core stability and loop dynamism of neural network-based integral membrane protein structure models embedded in an asymmetric OM bilayer, using the *M. xanthus* WzpX, WzpS, and WzpB proteins as test candidates. This was accomplished through integration of the CHARMM-graphical user interface (GUI) and Molecular Operating Environment (MOE) workflows to allow for a rapid simulation system setup and facilitate data analysis. In addition to serving as a method of model structure validation, our molecular dynamics simulations revealed a minimal movement of extracellular WzpX/S/B loops in the absence of an external stimulus as well as druggable cavities between the loops. Virtual screening of a commercial fragment library against these cavities revealed putative fragment-binding hotspots on the cell-surface face of each β-barrel, along with key interacting residues, and identified promising hits for the design of potential binders capable of plugging the β-barrels and inhibiting polysaccharide secretion.

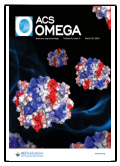


INTRODUCTION

Iterative neural network-based advances in protein folding such as AlphaFold2[1] and RoseTTA fold[2] have revolutionized the field of protein structure prediction. Nowhere has this been more apparent than with the 3D modeling of proteins that traverse membrane bilayers via α-helix or β-barrel architecture; the former are typically found in the inner (cytoplasmic) membrane of bacteria, eukaryotes, mitochondria, and chloroplasts,[3] while the latter are largely found in the outer membrane (OM) of Gram-negative bacteria, mitochondria, and chloroplasts.[4] Such integral membrane proteins are often recalcitrant to overexpression and purification, let alone in vitro manipulation, during various experiments. As such, neural network-based structure prediction has been a boon for the study of proteins localized to these subcellular compartments.

Protein model structures from applications such as AlphaFold2 and RoseTTA fold are largely based on the analysis of coevolving amino acids within the polypeptides of interest as well as similarities to existing protein structures.[1,2] However, for integral membrane proteins, the local effect of the membrane environment on the relative positioning and stability of the membrane-spanning α-helical or β-stranded tracts, as well as the extramembrane loops, is not accounted for prior to establishing the final model structure. Instead, molecular dynamics (MD) simulations of integral membrane proteins inserted in a membrane bilayer environment have proven to be a powerful complementary tool for examining

Received: December 13, 2023
Revised: February 6, 2024
Accepted: February 14, 2024
Published: March 4, 2024



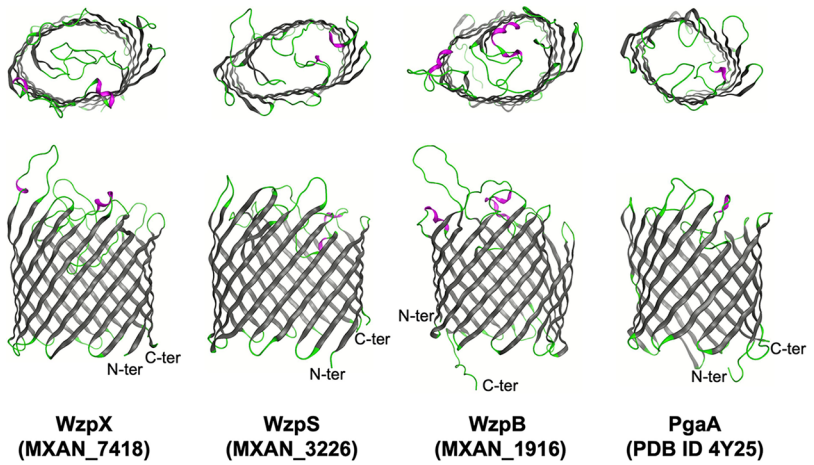


Figure 1. Three-dimensional model structures of EPS-pathway WzpX, MASC-pathway WzpS, and BPS-pathway WzpB porins,[6] as well as the X-ray crystal structure of the PNAG secretion porin PgaA (PDB 4Y25).[23] The cell-surface (*top panels*) and side (*bottom panels*) views of the proteins are presented. Legend: β-sheets, *dark gray*; α-helices, *magenta*; unstructured loops, *green*.

potential structural rearrangements/modifications due to nonsolvated hydrophobic surroundings.[5]

We recently reported the discovery, 3D modeling (via AlphaFold2), and characterization of the WzpX, WzpS, and WzpB outer-membrane (OM) β-barrel secretion porins[6] from the respective pathways for exopolysaccharide (EPS), major spore coat polysaccharide (MASC), and biosurfactant polysaccharide (BPS)[7] in Gram-negative[8] *Myxococcus xanthus* DZ2, a social and predatory soil bacterium.[9] This deltaproteobacterium is used as a model organism in which to study single-cell[10,11] and group motility,[7] developmental progression (via formation of spore-filled fruiting bodies),[7] drug tolerance,[12] and biofilm formation,[13] with the abovementioned secreted polysaccharides having outsized impacts on all of these physiological outcomes.

The WzpX/WzpS/WzpB OM porins[6,14] were respectively found to be part of the EPS/MASC/BPS Wzx/Wzy-dependent pathways[7,15,16] in which undecaprenyl pyrophosphate-linked sugar repeats are flipped across the inner membrane to its periplasmic leaflet (by Wzx),[17] where they are then polymerized[18] (by Wzy) and secreted across the cell envelope to the cell exterior. However, WzpX/WzpS/WzpB were identified as expanded structural homologues of the porin PgaA required for the secretion of poly-*N*-acetyl-D-glucosamine (PNAG) in *Escherichia coli*. Intriguingly, PNAG is made by a synthase-dependent pathway in which sugar addition in the cytoplasm is coupled to export of the growing chain across the OM by an equivalent amount. Prior to our investigation, synthase-dependent pathway secretion proteins had never before been identified in Wzx/Wzy-dependent pathways.[19]

WzpX/WzpS/WzpB were also found to be genetically and functionally coupled with their respective so-called outer-membrane polysaccharide export (OPX) proteins WzaX/WzaS/WzaB (required for EPS/MASC/BPS secretion); these OPX proteins are designated as "Class 3" as they do not have membrane-spanning domains and resemble the majority of OPX proteins found across all Gram-negative and Gram-positive bacteria.[6] Given the lack of OM-spanning capacity for WzpX/WzpS/WzpB, our characterization of the integral OM β-barrels WzpX/WzpS/WzpB finally provided an explanation for how the physiologically important EPS, MASC, and BPS polymers could be secreted across the OM in *M. xanthus* DZ2.[6] However, since the AlphaFold2-generated WzpX/WzpS/WzpB model structures (Figure 1) were of a specific conformation, no information was available regarding the stability of the core model structures or the dynamic behavior of surface-gating loop domains.

A popular approach to obtaining information on macromolecular protein dynamics is the use of the Molecular Operating Environment (MOE) software platform;[20] this provides a streamlined and unified environment for running MD simulations in addition to offering several other molecular modeling functionalities including (i) docking, (ii) virtual screening (VS), (iii) quantitative structure−activity relationship (QSAR), (iv) cheminformatic, and (v) computer-aided drug design (CADD) analyses. Through its versatile and user-friendly graphical user interface (GUI), molecular complexes can be constructed or imported to the MOE environment and automatically prepared for MD simulations using a set of applications for correcting molecular topologies and protonation states, generating and soaking complexes with defined periodic boundary conditions, as well as assigning partial charges and applying atomic restraints based on specified force field parameters. With prepared complexes in hand, MD simulation steps can be efficiently expressed using an easy MD protocol language to define temperature, pressure, and energy ramps. The MOE MD framework automatically generates topology, parameter, and script files for running MD simulations on a parallel cluster or GPU with full preservation of force field definitions from MOE, with no need of using complex command lines. Depending on the size of the system under study, this can save days of work. MD trajectories can then be imported and analyzed within the MOE environment using a set of molecular dynamics analysis methods that extract and summarize any information from the resulting trajectory files in spreadsheets from which the data can be easily plotted, making such analyses much easier to obtain by a larger segment of the scientific community. However, the MOE software platform currently lacks a tool to generate membrane-embedded protein models. Fortunately, this limitation can be overcome through prior use of the membrane-builder tool of the CHARMM (Chemistry at HARvard Macromolecular Mechanics) platform[21] and its associated GUI[22] which enables the construction of a variety of membrane-embedded models as well as the generation of input files for further MD simulations.

https://doi.org/10.1021/acsomega.3c09970
ACS Omega 2024, 9, 13217−13226

In this investigation, we integrated the CHARMM-GUI and MOE workflows and applied it toward the examination of WzpX/WzpS/WzpB model stability and dynamism, relative to that of the experimentally derived *E. coli* PgaA X-ray crystal structure, in a simulated asymmetric OM bilayer. These analyses reveal the minimal flexibility of extracellular loop domains in the absence of an external stimulus as well as the existence of apical cavities. Virtual screening (VS) studies against these extracellular cavities using a commercial fragment library enabled the determination of the cavities' fingerprint profiles, revealing the existence of promising druggable hotspots and ranked potential hits suitable for the inhibition of polysaccharide secretion via future drug design of targeted binders.

METHODOLOGY

Construction of the MD Model Systems. The MD model systems—consisting of PgaA or the related β-barrels WzpX (MXAN_7418), WzpS (MXAN_3226), and WzpB (MXAN_1916) embedded in an asymmetric *E. coli* OM bilayer—were constructed using the *Bilayer Builder* option of the *Membrane Builder* input generator tool of the CHARMM-GUI server (https://www.charmm-gui.org/).[22] First, the .pdb files of each model were uploaded to the server and the orientation option set as *"Run PPM 2.0"*, while water in the protein pore was generated with the *"Using the protein geometry"* option. Next, the *"Heterogeneous Lipid"* option was chosen since the *E. coli* OM is asymmetric. The box type chosen was *"Rectangular"*, and the *"Length of Z"* was based on *"Water thickness"* with a minimum water height of 22.5 Å on the top and bottom of the system. The *"Length of XY"* was based on the *"Ratios of lipid components"* and the *"Length of X and Y"* was set to 80 Å. The composition of the lower (periplasmic) leaflet of the OM was set to previously established values:[24] 89% of 1-palmitoyl(16:0)-2-palmitoleoyl-(16:1 cis-9)-phosphatidylethanolamine (PPPE), 1% of 1,10 -palmitoyl-2,20-vacenoyl cardiolipin with a net charge of _2e (PVCL2), and 11% of 1-palmitoyl(16:0)-2-vacenoyl(18:1 cis-11)-phosphatidylglycerol (PVPG), while the upper (surface) leaflet was 100% composed of lipopolysaccharide (LPS). The LPS chemical structure used was that from *E. coli* without *"O-units"* or *"chemical modification"*. As the behavior of the PgaA structure was to serve as the reference control, PgaA, as well as WzpX, WzpS, and WzpB, were all inserted in the same simulated OM. Though the LPS structure from *M. xanthus* is not identical to that of *E. coli*,[8,25] integral OM proteins from *M. xanthus* have been repeatedly shown to properly display and interact when heterologously expressed in the *E. coli* OM,[11,26] indicating a compatible environment.

The *"System Building Option"* chosen was the *"Replacement method"* with the *"Check lipid ring penetration"* option selected. The *"Component Building Options"* were as follows: *"Include Ions,"* the *"Ion replacing method"* was *"Distance"* with KCl at 0.15 M as the only *"Basic Ion Type"*. The .pdb file obtained was then uploaded into the MOE molecular modeling package (https://www.chemcomp.com/index.htm) for further preparation to run MD simulations.

MD Simulations. As mentioned above, MD simulations were performed through the MOE molecular modeling package. To begin, each system was opened in the MOE main window and visually inspected in order to fix the missing bonds and incorrect atom types and names. The systems were then optimized using the *"Protein−Structure Preparation"* tool to achieve the proper bond lengths, angles, and charges compatible with the physiological environment and neutralized through the replacement of water molecules by Na^+ or Cl^- ions. Once optimized, the files were prepared for the MD simulations using the MOE *"Compute/Simulations/Dynamics"* module. The files were then transferred via Linux terminal to an account at Calcul Quebec (https://www.calculquebec.ca/) where the MD simulations were run via command line according to Calcul Quebec protocols on GPU:v100:1 graphic cards with 500 MB of allocated memory. Following completion of the simulations, the output files were downloaded from Calcul Quebec to a local computer to be analyzed by using MOE tools.

As previously described,[27] simulations were performed in triplicate using the AMBER10:EHT[28] force field and NAMD software[29] with a cutoff of 10 kcal/mol for electrostatic interactions and a range between 8 and 10 kcal/mol for van der Waals interactions. The protocol before each production step included 5000 steps of energy minimization, followed by 100 ps of an isothermal−isobaric ensemble (NPT) simulation and 200 ps of an isothermal−isochoric ensemble (NVT). Considering that the models were quite complex owing to several different elements (protein, phospholipids, saccharides, ions, and solvent), packed close to each other, it was necessary to run three short 10 ns MD simulations before the principal production step, with a gradual release of the system. This was done to avoid crashing the system through many clashes caused by the sudden release of all elements at the same time. The first MD simulation was performed with a position restriction (PR) of the whole system except for the membrane. In the second MD simulation, only the protein was kept restrained. For the third MD simulation, only the protein backbones were restrained. Finally, for the fourth MD simulation, a production step of 1 μs was carried out with all components unrestrained.

The trajectories obtained were analyzed using the MOE "MD Analysis" extension and the Database Viewer (DBV). The root-mean-square deviation (RMSD), a measure of how much one atom deviates from its initial position, and root-mean-square fluctuation (RMSF), a measure of how much one amino acid deviates from its initial position, were calculated for the β-barrels in comparison to the first frames of each production step. Finally, plots of the MD results were created using GraphPad Prism, and figures were created with MOE and PyMol.[30]

Virtual Fragment Screening. Representative frames from the MD simulations corresponding to the averaged value of RMSD after stabilization were further submitted to VS of a fragment library designed and kindly provided by NMX Research and Solutions Inc. (https://www.fragmentresearch.com/) using the software Molegro Virtual Docker (MVD).[31] This library contains 1166 fragments previously curated for NMR binding experiments; it was designed to maximize conformational diversity and the number of fragments found in approved drugs, as well as excluding potential aggregating and pan assay interference (PAIN)[32] compounds. Fragment libraries are ideal for finding hotspots in molecular targets since fragments can serve as chemical biology probes to expose pockets and pocket features of potential interest.[33] This facilitates the identification of hits, as well as potential leads for initiating drug discovery projects. It was downloaded in the .sdf format, exported into a MOE database .mdb file where the fragments were "washed" to the appropriate bond lengths,

https://doi.org/10.1021/acsomega.3c09970
ACS Omega 2024, 9, 13217−13226

angles, charges, and protonation states compatible with the physiological environment, and then saved in .mol2 format. To run the VS, representative MD simulation frames of WzpX, WzpS, WzpB, and PgaA were opened in the MVD[31] main window together with the washed library. Each spherical search space was set to cover the whole external surface of the model including all loops, with the central coordinates and the corresponding radii summarized in Table S1. Thirty runs per model were performed with a maximum of 3000 interactions, and the percentage of returning poses was set to 20%. All poses obtained with "*moldockscore negative*" were selected for further analysis.

Fingerprint Analysis. The protein ligand interaction fingerprint (PLIF) application of the MOE software was used to map the most relevant interactions in the hotspots of each β-barrel. For this, we used the representative frames from the MD simulations mentioned above and databanks in the .mdb format containing only the poses observed in the hotspots. The fingerprints were prepared and generated using the PLIF setup panel with the weak and strong energy thresholds for the nine types of protein−ligand interaction, which normally comprise fingerprints, set as indicated in Table 1. The resulting fingerprint results were compiled in barcode plots.

Table 1. Energy Thresholds Used to Generate the Fingerprints of the Hotspots

type of interaction		weak (kcal mol^{-1})	strong (kcal mol^{-1})
side chain	H-donor	0.5	1.5
	H-acceptor	0.5	1.5
backbone	H-donor	0.5	1.5
	H-acceptor	0.5	1.5
solvent	H-donor	0.5	1.5
	H-acceptor	0.5	1.5
ionic attraction		0.5	3.5
metal ligation		0.5	3.5
arene attraction		0.5	1.0

RESULTS AND DISCUSSION

MD Simulation. Following the integration of the CHARMM-GUI and MOE workflows, MD simulations of WzpX, WzpS, WzpB, and PgaA inserted in an OM bilayer yielded RMSD plots that never exceeded 4.5 Å, showing very similar profiles for the experimental structure (PgaA) and the three models (WzpX/S/B; Figure 2). These results suggest an appropriate dynamic behavior of all systems and consistency of the 3D structures of the models when compared to PgaA, serving as an additional step of structure validation for the three β-barrel models (Figure 1). This also suggests a robust integration between each protein and the surrounding membrane components, corroborating the consistency of the protein−membrane models constructed herein.

Fluctuations in the position of each amino acid during the MD simulations were then probed via the analysis of the average RMSF plots (Figure 3). As expected, the highest fluctuations corresponded with the loop regions; however, the fluctuations did not surpass 5.0 Å, suggesting an overall low mobility of the loops. The only exceptions observed were for two external loops of the WzpX model structure (amino acids 101−137 and amino acids 252−268), which showed fluctuations up to 5.03 and 7.05 Å (respectively). The low fluctuations below 2.0 Å observed for the trans-membrane β-strands reflect their stability during the MD simulations and integration with the membrane. These results are illustrated by the overlapping of frames collected during the MD simulations (Figure 4), where for all systems the loops did not move enough on their own to open the extracellular face of the β-barrels. This is consistent with the hypothesis that the opening of WzpX/S/B to the extracellular milieu is not a spontaneous motion and instead depends on the binding of some triggering factor to the β-barrel lumen (e.g., the translocating polymer) and/or periplasmic face (e.g., WzaX/S/B Class-3 OPX protein)[6,14] capable of causing a major conformational change in the extracellular WzpX/S/B loops.

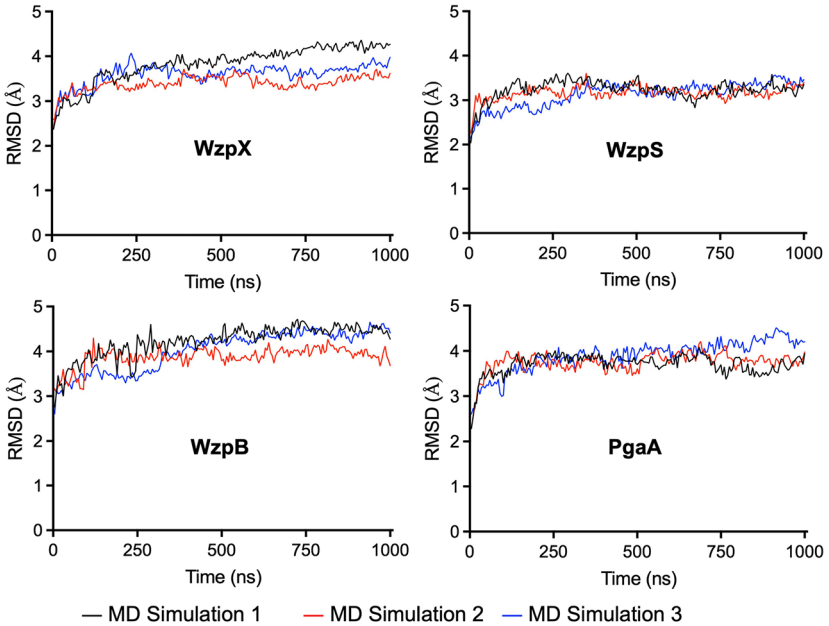


Figure 2. Plots of RMSD variation for WzpX, WzpS, WzpB, and PgaA embedded in an asymmetric OM bilayer after 1 μs of MD simulation in triplicate.

https://doi.org/10.1021/acsomega.3c09970
ACS Omega 2024, 9, 13217−13226

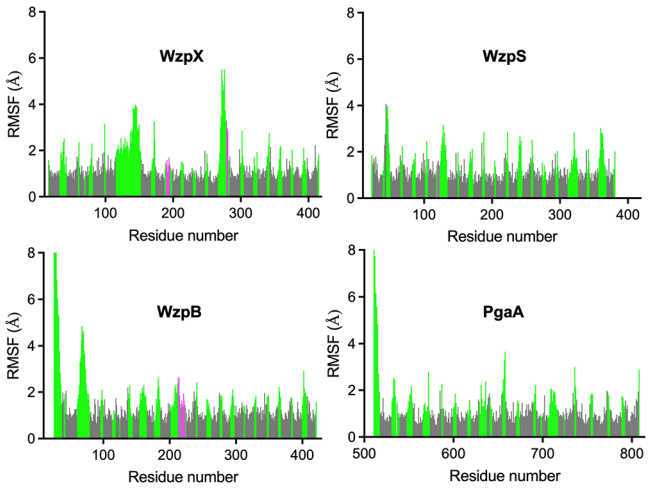


Figure 3. Plots of the average RMSF for WzpX, WzpS, WzpB, and PgaA after 1 μs of MD simulation in triplicate. Colors correspond to the secondary-structure elements from Figure 1, i.e., β-sheets (*dark gray*), α-helices (*magenta*), and unstructured loops (*green*).

Analysis of the extracellular views of representative frames of the MD simulations depicted as electrostatic surfaces (Figure 5) revealed the presence of several pockets among the loops that could be targeted in the future for the purpose of drug design. Moreover, each protein displayed a unique surface, showing distinct cavities and a different pattern of pockets. This suggests that it is possible to design and/or identify binding compounds capable of selectively blocking one specific β-barrel by exploring the different shapes and potential interactions with the pockets in question.

Fragment Library Virtual Screening. Having uncovered druggable pockets on the extracellular faces of WzpX, WzpS, WzpB, and PgaA, we applied this information toward the identification of compound fragments via VS with the potential to bind these various pockets and provide a basis for downstream drug development. The VS returned 1082 fragments for WzpX, 1001 for WzpS, 1143 for WzpB, and 599 for PgaA. The MolDockScores obtained were also significantly different among the β-barrels, as shown in the plot of MolDockScore versus fragment number (Figure 6). WzpX and WzpS were the proteins showing the best-ranked

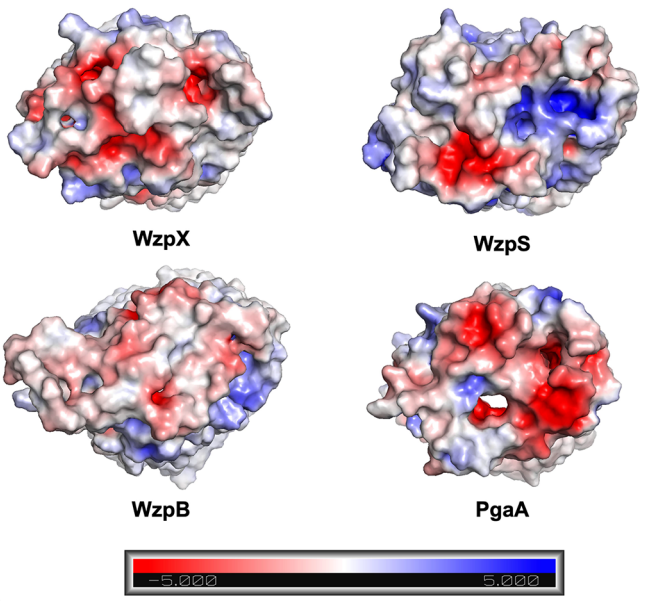


Figure 5. Extracellular view of the electrostatic surfaces of representative frames collected during the MD simulations (±5.000 kT/e). The regions with the most negative electrostatic potential are shown in *red*, while the regions with the most positive electrostatic potential are shown in *blue*.

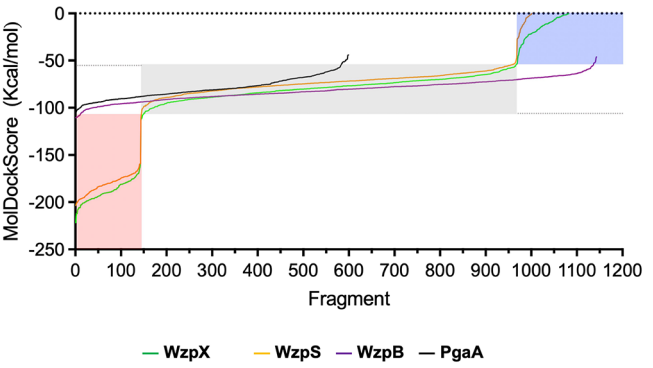


Figure 6. Plots of MolDockSocre (kcal/mol) vs fragment number obtained from VS. The colored areas correspond to the energy ranges of the negatively charged (*red*), neutral (*gray*), and positively charged (*blue*) fragments observed for the proteins.

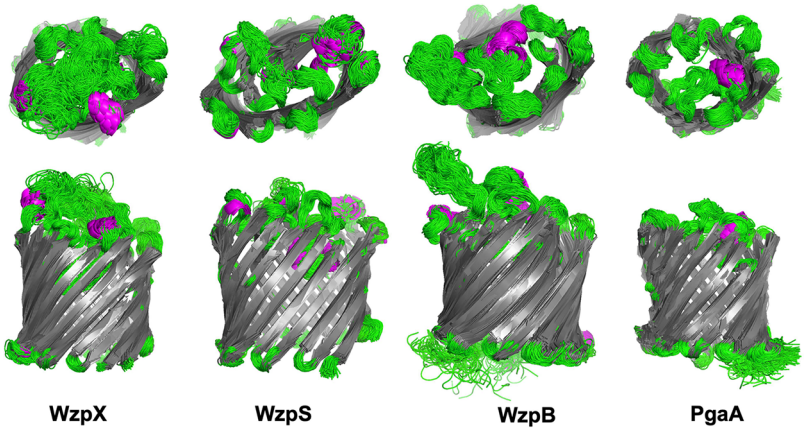


Figure 4. Superposition of 200 frames collected at 5 ns intervals during the MD simulations. Both the extracellular (*top row*) and side-facing (*bottom row*) views have been provided. Colors correspond to secondary-structure elements from Figures 1 and 3, i.e., β-sheets (*dark gray*), α-helices (*magenta*), and unstructured loops (*green*).

https://doi.org/10.1021/acsomega.3c09970
ACS Omega 2024, 9, 13217−13226

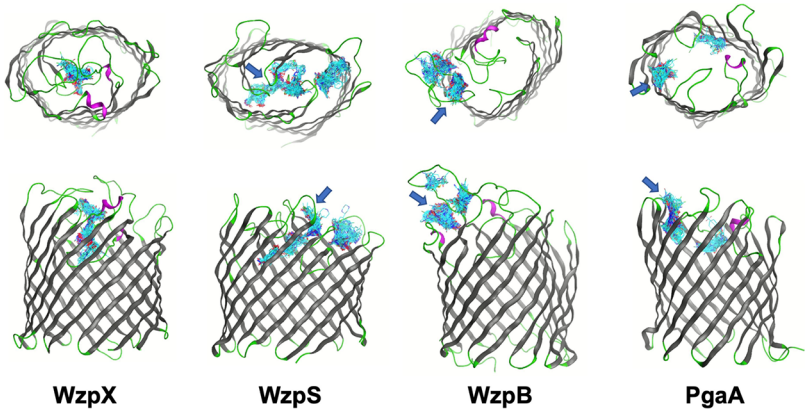


Figure 7. Concentration of fragment poses (*shown in cyan*) showcasing the most populated hotspots of the β-barrels. The arrows point to the hotspots concentrating the most poses.

Table 2. Hotspot Residues of Each β-Barrel

β-barrel	hotspot residues
WzpX	Leu101, Thr104, Leu106, Gly107, Ser108, Leu109, Asp112, Thr113, Pro114, Gly115, Thr116, Pro117, Asp118, Gly119, Pro120, Leu121, Arg122, Phe141, Ser143, Ser170, Ala180, Val181, Glu257, Thr303, Asp304, Arg305, Leu306
WzpS (hotspot 1)	Arg103, Ala104, Gly105, Leu106, Phe115, Arg137, Phe138, Glu139, Ala140, Val141, Gly151, Tyr152, Ala153, His154, Thr155, Gln177, Phe179, Val250, Gly251, Ala252, Asn253
WzpS (hotspot 2)	Arg15, Asp17, Asp19, Arg21, Met31, Arg60, Thr96, Asp97, Pro98, Arg108, Ser109, Thr110, Arg347
WzpB	Gly33, Val34, Tyr36, Phe37, Ser38, Pro39, Thr40, Asn95, Ser96, Ser97, Ala98, Ala99, Leu177, Ser178, Ser179, Val180
PgaA	Glu704, Tyr714, Asn715, Pro716, Ile717, Lys718, Thr719, Asp721, Ser751, Trp752, Gln753, Lys754, His755, Tyr756, Val761, Arg787, Pro788, Tyr789, Asp790

fragments with the lowest energy values of −221.77 and −203.53 kcal/mol, respectively. Conversely, for WzpB and PgaA, no fragment with energy below −110 kcal/mol was observed, and only 38 and 11 fragments, respectively, scored below −100 kcal/mol. These results suggest a higher affinity of the fragments for WzpX and WzpS, and that by extension, these two β-barrels might be more druggable compared to WzpB and PgaA.

The VS results also evidenced a unique charge-related energy distribution for the specific cavities of WzpX and WzpS; in both cases, negatively charged fragments ranked the best and positively charged fragments ranked the worst, while neutral fragments were distributed in between (Figure 6). This was not observed for WzpB and PgaA, which showed charged and neutral fragments distributed over the whole energy curve, with a slight prevalence of neutral and positively charged residues at the lowest energy regions. Interestingly, the range of energy where most of the fragment poses fell for these two proteins corresponded to the neutral range (Figure 6), which is the same energy range observed for the neutral fragments over WzpX and WzpS.

The distribution of fragment poses over the β-barrels' surfaces enabled the detection of hotspots of preferred binding among the external loops. For WzpX, the most populated binding zone was located in the central cavity between the main loops (Figure 7). This cavity concentrated 40.76% of the fragments, and also all fragments ranked below −200 kcal/mol. Similarly, WzpS also concentrated most of the fragments (39.56%) in a central hotspot. However, this porin also exhibited a second highly populated hotspot, which, together with the first one, concentrated the 112 best fragments, all ranking below −172 kcal/mol. For WzpB, most fragments (75.94%) concentrated at the three hotspots around the main loop at the extracellular face of the porin (Figure 7). These hotspots also hosted 33 of the 38 best-ranked fragments. Similar to WzpS, PgaA concentrated most fragments (85.47%) in the extracellular loops (see Figure 7), including the 11 fragments ranked below −100 kcal/mol.

To map the hotspot residues of interest for future drug design, we analyzed the MolDockScores and the corresponding amino acids involved in the potential H-bonds for the 10 best-ranked fragment poses from VS (Table S2 and Figures S1−S40) for each β-barrel. As mentioned above, all of the 10 best-ranked fragments from VS against WzpX and WzpS were negatively charged, while those for WzpB and PgaA were either neutral or positively charged (except for 3K-528S against PgaA). These data point to the differences in compound fragment specificity among the β-barrels that can be further explored in the design of selective inhibitors. Importantly (i) only one fragment was found to be a common putative binder between WzpX/WzpS and WzpB/PgaA, (ii) three fragments were common between WzpX and WzpS, (iii) one fragment was equivalent between WzpB and PgaA, and (iv) no common fragments were detected between WzpX/WzpS and WzpB (Table S2). These findings point to the differences in the hotspots to be explored for the downstream design of selective inhibitors.

For WzpX, up to 27 different amino acids were involved in some form of interaction with the screened fragments (Figures S1−S10 and Table 2). In particular, Arg305 stands out since this residue was observed to form H-bonds with all 10 of the best-ranked fragments. Residues Thr104, Ser108, Leu109, Asp112, Gly115, Thr116, Asp118, and Asp304 were identified as other potential sources of H-bonds, while residues Ser108, Thr116, and Asp118 showed potential arene−H interactions.

Five of the top 10 fragments identified for WzpS docked in the most populated hotspot (herein termed "hotspot 1"), while four docked in the second most populated zone ("hotspot 2")

https://doi.org/10.1021/acsomega.3c09970
ACS Omega 2024, 9, 13217−13226

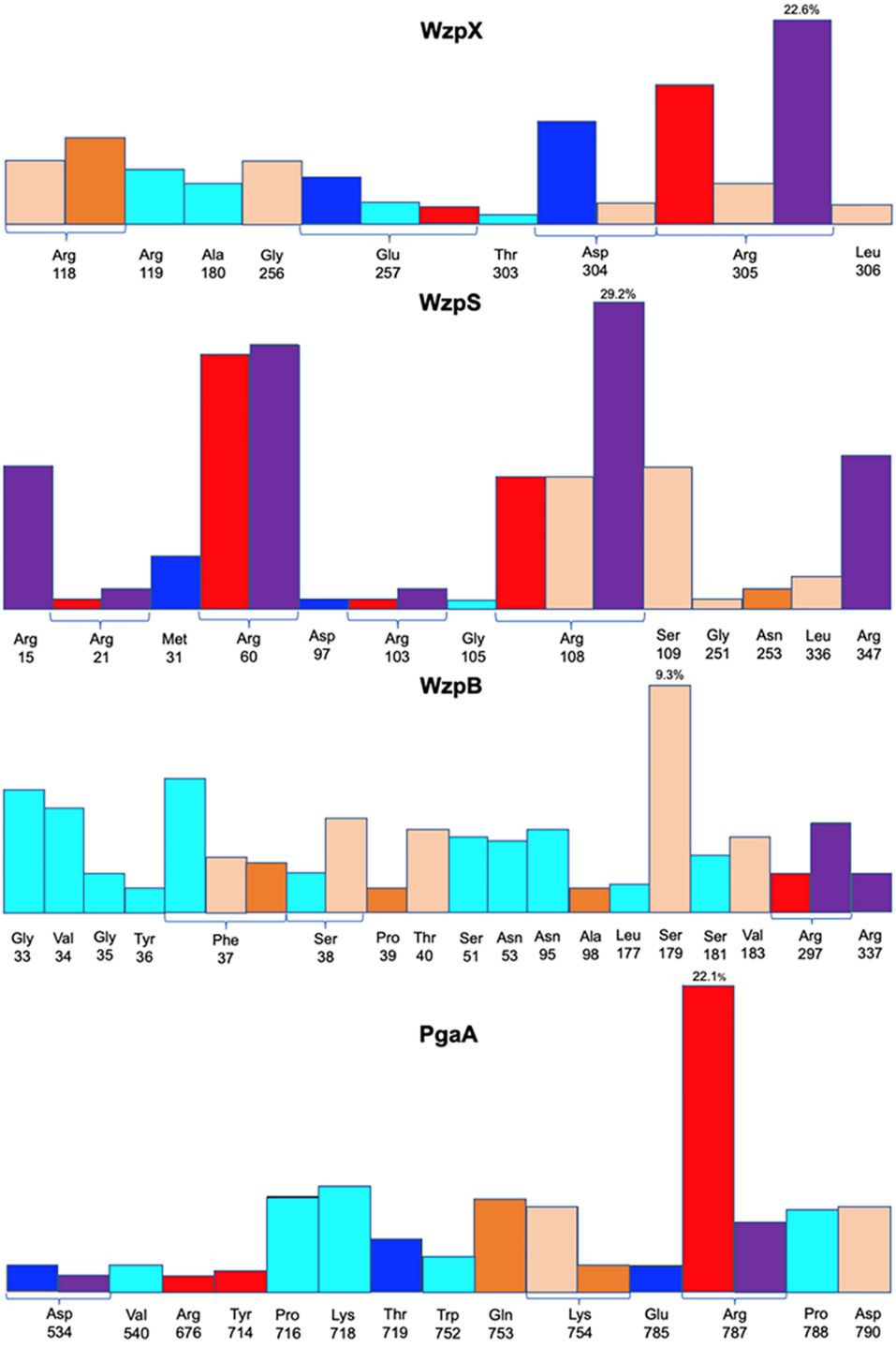


Figure 8. Fingerprints of the strong interactions on the hot spots of WzpX, WzpS, WzpB, and PgaA. Color barcode: red = side-chain hydrogen-bond acceptor; blue = side-chain hydrogen-bond donor; purple = ionic attraction; orange = arene attraction; salmon = backbone hydrogen-bond acceptor; cyan = backbone hydrogen-bond donor.

and one docked in the interface between the two hotspots. Twenty-one residues were observed in VS interactions with the fragments in hotspot 1, while 13 residues were identified in hotspot 2 (Figures S11−S20, Table 2). Arg103 and Asn253 formed potential H-bonds and arene−cation or arene−H interactions in hotspot 1, while for hotspot 2, the key H-bonding residues were Arg60 and Thr110, with the latter also showing potential for arene−H interactions. A fragment docking between the two hotspots (3K-528S) showed potential H-bonds and arene−H interactions with Gly105 (from hotspot 1) and Arg108 (from hotspot 2).

The 10 best fragments docked in WzpB concentrated in the most populated hotspot and showed potential interactions with 16 residues. Potential H-bonds were observed for residues Gly33, Val34, Tyr36, Thr40, Asn95, Ser96, Leu177, and Ser179, while arene−H interactions were observed for residues Pro39, Ala98, and Ser179. Arg40 and Ser179 were involved in

https://doi.org/10.1021/acsomega.3c09970
ACS Omega 2024, 9, 13217−13226

H-bonding for the majority of poses (Figures S21−S30 and Table 2).

Nineteen residues were observed to form potential interactions with the fragments in the main hotspot of PgaA; Pro716, Lys718, Gln753, Lys754, Arg787, Pro788, Tyr789, and Asp790 were implicated in H-bonding, while Lys754 and Tyr789 were observed in potential arene−H interactions. Lys754, Arg787, Pro788, and Tyr789 were those residues most frequently implicated in fragment binding (Figures S31−S40 and Table 2).

Fingerprint Analysis of the Hotspots. The barcode plots shown in Figure 8 reinforce the relevance of the arginines in the hotspots of WzpX (Arg118 and Arg305) and WzpS (Arg60 and Arg108). These residues performed mostly ionic attraction (purple bars), side chain hydrogen-bond acceptor (red bars), arene attraction (orange bars), and backbone hydrogen-bond acceptor (salmon bars) with the fragments. This corroborates the higher affinities for negatively charged residues observed in the VS studies for these two proteins. For WzpB and PgaA, on the other hand, ionic attractions (purple bars) did not show much relevance, and the most prevalent interactions were side-chain hydrogen-bond acceptor (red bars) and backbone hydrogen-bond acceptor (cyan bars). This is aligned with the trend observed in the VS studies, suggesting that these two proteins would be less selective to the charged compounds.

CONCLUSIONS

We have herein reported a new and straightforward protocol to run MD simulations of membrane-embedded systems merging resources from the CHARMM-GUI[22] server and the MOE molecular modeling package,[20] in order to take advantage of the more intuitive and user-friendly environment provided by the latter package to run MD simulations. This workflow will complement other tools already reported in the literature.[34]

Results of the MD simulations performed using our protocol on the experimental structure of PgaA as well as the neural network-based model structures of WzpX, WzpS, and WzpB corroborated the quality and consistency of the models and also revealed the low mobility of the extracellular loops of these four β-barrels. These findings support the hypothesis that an internal triggering mechanism is needed for the opening of these β-barrels; this would be consistent with nascent PNAG/EPS/MASC/BPS polymers (arriving from the periplasm) being translocated through the interior of OM-spanning PgaA/WzpX/WzpS/WzpB,[6] with this translocation step contributing to the opening of the extracellular loops of the respective porins. For WzpX/WzpS/WzpB, this opening may happen in conjunction with the proposed interactions[6,14] between the β-barrels and the periplasmic Class-3 OPX proteins WzaX/WzaS/WzaB for the respective EPS/MASC/BPS pathways.

Subsequent VS studies against the four β-barrels using a commercial fragment library resulted in the identification of compound-interaction hotspots among the extracellular loops; going forward, these sites can be targeted by inhibitors chosen to selectively bind each barrel. Potential drugs to be developed from these analyses could then incorporate one or more of the identified binding-fragment motifs. In this manner, the secretion of EPS, MASC, BPS, and PNAG could be selectively blocked in their respective organism backgrounds.

The results also suggested that the identified hotspots of WzpX and WzpS might have more affinity for negatively charged compounds, while those of WzpB and PgaA might interact better with neutral or positively charged compounds. Also, analysis of the binding modes of the top 10 fragment poses associated with the fingerprint analysis of all poses found in the hotspots revealed the key residues of each hotspot worth being explored for future drug design.

Importantly, our investigation has demonstrated the utility of using membrane-embedded integrated MD simulations to evaluate the quality of neural network-based membrane protein model structures, which are becoming widely reported in the scientific literature. In addition, our protocol will allow for the rigorous testing of more complex questions such as the proposed interaction dynamics between the EPS/MASC/BPS secretion porins (WzpX/S/B) and other proteins such as modeled octamers[14] of the WzaX/S/B Class-3 OPX proteins[6] required for the secretion of each polymer.[7,15]

ASSOCIATED CONTENT

Data Availability Statement

The following data are available as Supporting Information: The .pdb files of 4Y25 and the models MXAN_7418, MXAN_3226, and MXAN_1916 plus their corresponding output files obtained from CHARMM-GUI and the MOE files used to run the MD simulations.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.3c09970.

Grid data for the VS; 10 best-ranked fragments from the VS; and 2D interaction maps for the 10 best-ranked fragments (PDF)

AUTHOR INFORMATION

Corresponding Authors

Salim T. Islam − *Institut National de la Recherche Scientifique (INRS), Centre Armand-Frappier Santé Biotechnologie, Université du Québec, Institut Pasteur International Network, Laval, QC H7V 1B7, Canada; PROTEO, the Quebec Network for Research on Protein Function, Engineering, and Applications, Université Laval, Quebec, QC G1V 0A6, Canada;* orcid.org/0000-0001-6853-8446; Email: salim.islam@inrs.ca

Steven R. LaPlante − *Institut National de la Recherche Scientifique (INRS), Centre Armand-Frappier Santé Biotechnologie, Université du Québec, Institut Pasteur International Network, Laval, QC H7V 1B7, Canada; PROTEO, the Quebec Network for Research on Protein Function, Engineering, and Applications, Université Laval, Quebec, QC G1V 0A6, Canada;* orcid.org/0000-0003-2835-5789; Email: steven.laplante@inrs.ca

Authors

Tanos C. C. França − *Institut National de la Recherche Scientifique (INRS), Centre Armand-Frappier Santé Biotechnologie, Université du Québec, Institut Pasteur International Network, Laval, QC H7V 1B7, Canada; PROTEO, the Quebec Network for Research on Protein Function, Engineering, and Applications, Université Laval, Quebec, QC G1V 0A6, Canada; Laboratory of Molecular Modeling Applied to Chemical and Biological Defense, Military Institute of Engineering, Rio de Janeiro 22290-270, Brazil; Department of Chemistry, Faculty of Science,*

https://doi.org/10.1021/acsomega.3c09970
ACS Omega 2024, 9, 13217−13226

University of Hradec Kralove, 50003 Hradec Kralove, Czech Republic; orcid.org/0000-0002-6048-8103

Fares Saïdi − *Institut National de la Recherche Scientifique (INRS), Centre Armand-Frappier Santé Biotechnologie, Université du Québec, Institut Pasteur International Network, Laval, QC H7V 1B7, Canada; PROTEO, the Quebec Network for Research on Protein Function, Engineering, and Applications, Université Laval, Quebec, QC G1V 0A6, Canada*

Alain Ajamian − *Chemical Computing Group, Montreal, Quebec H3A 2R7, Canada*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.3c09970

Author Contributions

T.C.C.F.: Conceptualization, writing—original draft, formal analysis, writing—review and editing, and methodology F.S.: Writing—original draft A.A.: Software and formal analysis S.T.I.: Conceptualization, funding acquisition, supervision, writing—original draft, and writing—review and editing S.R.L.: Conceptualization, funding acquisition, supervision, and writing—review and editing.

Funding

T.C.C.F and S.R.L were funded by Quebec Consortium for Drug Discovery (CQDM), Ministère de l'Économie et de l'Innovation (MEI), NMX Research and Solutions Inc., and Mitacs. Discovery grants (RGPIN-2016-06637 and RGPIN-2023-05576) were provided to S.T.I. from the Natural Sciences and Engineering Research Council of Canada. F.S. was supported by these grants, as well as studentship funding from PROTEO (The Quebec Network for Research on Protein Function, Engineering, and Applications).

Notes

The authors declare the following competing financial interest(s): A.A. is an employee of Chemical Computing Group ULC, creators of MOE software used to perform most of the modeling work described herein.

■ ACKNOWLEDGMENTS

The authors wish to thank the Institut National de la Recherche Scientifique (INRS) and Prof Teodorico Ramalho from the Federal University of Lavras—Brazil, for software availability, and Calcul Québec for the availability of the supercomputer. This work was also supported by the Excellence project PrF UHK 2216/2023−2024.

■ REFERENCES

(1) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583−589.

(2) Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G. R.; Wang, J.; Cong, Q.; Kinch, L. N.; Schaeffer, R. D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373* (6557), 871−876.

(3) Islam, S. T.; Lam, J. S. Topological mapping methods for α-helical bacterial membrane proteins − an update and a guide. *MicrobiologyOpen* **2013**, *2* (2), 350−364. Gilmore, R.; Mandon, E. C. Understanding integration of α-helical membrane proteins: the next steps. *Trends Biochem. Sci.* **2012**, *37* (8), 303−308.

(4) Schleiff, E.; Soll, J. Membrane protein insertion: mixing eukaryotic and prokaryotic concepts. *EMBO Rep.* **2005**, *6* (11), 1023−1027.

(5) Chavent, M.; Duncan, A. L.; Sansom, M. S. P. Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. *Curr. Opin. Struct. Biol.* **2016**, *40*, 8−16.

(6) Saïdi, F.; Mahanta, U.; Panda, A.; Kezzo, A. A.; Jolivet, N. Y.; Bitazar, R.; John, G.; Martinez, M.; Mellouk, A.; Calmettes, C.; et al. Bacterial outer membrane polysaccharide export (OPX) proteins occupy three structural classes with selective β-barrel porin requirements for polymer secretion. *Microbiol. Spectr.* **2022**, *10* (5), e01290-22.

(7) Islam, S. T.; Vergara Alvarez, I.; Saïdi, F.; Guiseppi, A.; Vinogradov, E.; Sharma, G.; Espinosa, L.; Morrone, C.; Brasseur, G.; Guillemot, J.-F.; et al. Modulation of bacterial multicellularity via spatio-specific polysaccharide secretion. *PLOS Biol.* **2020**, *18*, e3000728.

(8) Saïdi, F.; Gamboa Marin, O. J.; Veytia-Bucheli, J. I.; Vinogradov, E.; Ravicoularamin, G.; Jolivet, N. Y.; Kezzo, A. A.; Ramirez Esquivel, E.; Panda, A.; Sharma, G.; et al. Evaluation of azido 3-deoxy-D-*manno*-oct-2-ulosonic acid (Kdo) analogues for click chemistry-mediated metabolic labeling of *Myxococcus xanthus* DZ2 lipopolysaccharide. *ACS Omega* **2022**, *7* (39), 34997−35013.

(9) Muñoz-Dorado, J.; Marcos-Torres, F. J.; García-Bravo, E.; Moraleda-Muñoz, A.; Pérez, J. Myxobacteria: moving, killing, feeding, and surviving together. *Front. Microbiol.* **2016**, *7*, 781.

(10) Faure, L. M.; Fiche, J.-B.; Espinosa, L.; Ducret, A.; Anantharaman, V.; Luciano, J.; Lhospice, S.; Islam, S. T.; Tréguier, J.; Sotes, M.; et al. The mechanism of force transmission at bacterial focal adhesion complexes. *Nature* **2016**, *539* (7630), 530−535. Jolivet, N. Y.; Han, E.; Belgrave, A. M.; Saïdi, F.; Koushki, N.; Lemon, D. J.; Faure, L. M.; Fleuchot, B.; Mahanta, U.; Jiang, H. Integrin-like adhesin CglD confers traction and stabilizes bacterial focal adhesions involved in myxobacterial gliding motility. *bioRxiv* **2023**, 1−41.

(11) Islam, S. T.; Jolivet, N. Y.; Cuzin, C.; Belgrave, A. M.; My, L.; Fleuchot, B.; Faure, L. M.; Mahanta, U.; Kezzo, A. A.; Saïdi, F.; et al. Unmasking of the von Willebrand A-domain surface adhesin CglB at bacterial focal adhesions mediates myxobacterial gliding motility. *Sci. Adv.* **2023**, *9*, No. eabq0619.

(12) Saïdi, F.; Bitazar, R.; Bradette, N.; Islam, S. T. Bacterial glycocalyx integrity impacts tolerance of *Myxococcus xanthus* to antibiotics and oxidative-stress agents. *Biomolecules* **2022**, *12* (4), 571.

(13) Saïdi, F.; Jolivet, N. Y.; Lemon, D. J.; Nakamura, A.; Belgrave, A. M.; Garza, A. G.; Veyrier, F. J.; Islam, S. T. Bacterial glycocalyx integrity drives multicellular swarm biofilm dynamism. *Mol. Microbiol.* **2021**, *116* (4), 1151−1172.

(14) Schwabe, J.; Pérez-Burgos, M.; Herfurth, M.; Glatter, T.; Søgaard-Andersen, L. Evidence for a widespread third system for bacterial polysaccharide export across the outer membrane comprising a composite OPX/β-barrel translocon. *mBio* **2022**, *13*, 5.

(15) Pérez-Burgos, M.; García-Romero, I.; Jung, J.; Schander, E.; Valvano, M. A.; Søgaard-Andersen, L. Characterization of the exopolysaccharide biosynthesis pathway in *Myxococcus xanthus*. *J. Bacteriol.* **2020**, *202*, e00335-20.

(16) Islam, S. T.; Lam, J. S. Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. *Can. J. Microbiol.* **2014**, *60* (11), 697−716.

(17) Islam, S. T.; Taylor, V. L.; Qi, M.; Lam, J. S. Membrane topology mapping of the O-antigen flippase (Wzx), polymerase (Wzy), and ligase (WaaL) from *Pseudomonas aeruginosa* PAO1 reveals novel domain architectures. *mBio* **2010**, *1* (3), e00189-10. Islam, S. T.; Eckford, P. D. W.; Jones, M. L.; Nugent, T.; Bear, C. E.; Vogel, C.; Lam, J. S. Proton-dependent gating and proton uptake by Wzx support O-antigen-subunit antiport across the bacterial inner membrane. *mBio* **2013**, *4* (5), No. e00678-13. Islam, S. T.; Lam, J. S. Wzx flippase-mediated membrane translocation of sugar polymer precursors in bacteria. *Environ. Microbiol.* **2013**, *15* (4), 1001−1015.

(18) Islam, S. T.; Gold, A. C.; Taylor, V. L.; Anderson, E. M.; Ford, R. C.; Lam, J. S. Dual conserved periplasmic loops possess essential

charge characteristics that support a catch-and-release mechanism of O-antigen polymerization by Wzy in *Pseudomonas aeruginosa* PAO1. *J. Biol. Chem.* **2011**, *286* (23), 20600−20605. Islam, S. T.; Huszczynski, S. M.; Nugent, T.; Gold, A. C.; Lam, J. S. Conserved-residue mutations in Wzy affect O-antigen polymerization and Wzz-mediated chain-length regulation in *Pseudomonas aeruginosa* PAO1. *Sci. Rep* **2013**, *3*, 3441.

(19) Whitney, J. C.; Howell, P. L. Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. *Trends Microbiol.* **2013**, *21* (2), 63−72.

(20) *Molecular Operating Environment (MOE)*; Chemical Computing Group ULC: Montreal, QC, Canada, 2023. https://www.chemcomp.com/Products.htm

(21) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4* (2), 187−217.

(22) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29* (11), 1859−1865.

(23) Wang, Y.; Andole Pannuri, A.; Ni, D.; Zhou, H.; Cao, X.; Lu, X.; Romeo, T.; Huang, Y. Structural basis for translocation of a biofilm-supporting exopolysaccharide across the bacterial outer membrane*. *J. Biol. Chem.* **2016**, *291* (19), 10046−10057.

(24) 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. *Biophys. J.* **2014**, *106* (11), 2493−2502. Dowhan, W. Molecular basis for membrane phospholipid diversity: why are there so many lipids? *Annu. Rev. Biochem.* **1997**, *66* (1), 199−232.

(25) MacLean, L.; Perry, M. B.; Nossova, L.; Kaplan, H.; Vinogradov, E. The structure of the carbohydrate backbone of the LPS from *Myxococcus xanthus* strain DK1622. *Carbohydr. Res.* **2007**, *342* (16), 2474−2480.

(26) Holkenbrink, C.; Hoiczyk, E.; Kahnt, J.; Higgs, P. I. Synthesis and assembly of a novel glycan layer in *Myxococcus xanthus* spores. *J. Biol. Chem.* **2014**, *289*, 32364. Jakobczak, B.; Keilberg, D.; Wuichet, K.; Søgaard-Andersen, L. Contact- and protein transfer-dependent stimulation of assembly of the gliding motility machinery in *Myxococcus xanthus*. *PLOS Genet.* **2015**, *11* (7), No. e1005341.

(27) França, T. C. C.; Botelho, F. D.; Drummond, M. L.; LaPlante, S. R. Theoretical investigation of repurposed drugs potentially capable of binding to the catalytic site and the secondary binding pocket of subunit A of ricin. *ACS Omega* **2022**, *7* (36), 32805−32815. Ali, Z.; Cardoza, J. V.; Basak, S.; Narsaria, U.; Singh, V. P.; Isaac, S. P.; França, T. C. C.; LaPlante, S. R.; George, S. S. Computational design of candidate multi-epitope vaccine against SARS-CoV-2 targeting structural (S and N) and non-structural (NSP3 and NSP12) proteins. *J. Biomol. Struct. Dyn.* **2023**, 1−20.

(28) Case, D. A.; Darden, T. A.; Cheatham, III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W. et al. *AMBER 10*. University of California: San Francisco, 2008.

(29) Nelson, M. T.; Humphrey, W.; Gursoy, A.; Dalke, A.; Kalé, L. V.; Skeel, R. D.; Schulten, K. NAMD: a parallel, object-oriented molecular dynamics program. *Int. J. Supercomp. Appl. High Perform. Comput.* **1996**, *10* (4), 251−268.

(30) DeLano, W. L.; Bromberg, S. *PyMOL User's Guide*. DeLano Scientific LLC, **2004**; p. 629.

(31) Thomsen, R.; Christensen, M. H. MolDock: a new technique for high-accuracy molecular docking. *J. Med. Chem.* **2006**, *49* (11), 3315−3321.

(32) Baell, J. B. Feeling nature's PAINS: natural products, natural product drugs, and pan assay interference compounds (PAINS). *J. Nat. Prod.* **2016**, *79* (3), 616−628. Baell, J. B.; Nissink, J. W. M. Seven Year Itch: Pan-Assay Interference Compounds (PAINS) in 2017—Utility and Limitations. *ACS Chem. Biol.* **2018**, *13* (1), 36−44.

(33) Huang, N.; Jacobson, M. P. Binding-site assessment by virtual fragment screening. *PLOS One* **2010**, *5* (4), No. e10109.

(34) Gao, Y.; Widmalm, G.; Im, W. Modeling and simulation of bacterial outer membranes with lipopolysaccharides and capsular polysaccharides. *J. Chem. Inf. Model.* **2023**, *63*, 1592. Gao, Y.; Lee, J.; Widmalm, G.; Im, W. Modeling and simulation of bacterial outer membranes with lipopolysaccharides and enterobacterial common antigen. *J. Phys. Chem. B* **2020**, *124* (28), 5948−5956. Patel, D. S.; Qi, Y.; Im, W. Modeling and simulation of bacterial outer membranes and interactions with membrane proteins. *Curr. Opin. Struct. Biol.* **2017**, *43*, 131−140. Pavlova, A.; Hwang, H.; Lundquist, K.; Balusek, C.; Gumbart, J. C. Living on the edge: simulations of bacterial outer-membrane proteins. *Biochim. Biophys. Acta* **2016**, *1858* (7, Part B), 1753−1759.

https://doi.org/10.1021/acsomega.3c09970
ACS Omega 2024, 9, 13217−13226