



HAL
open science

Structural insights into the membrane receptor ShuA in DDM micelles and in a model of gram-negative bacteria outer membrane as seen by SAXS and MD simulations

Stéphane Abel, Massimo Marchi, Justine Solier, Stéphanie Finet, Karl Brillet, Françoise Bonnete

► To cite this version:

Stéphane Abel, Massimo Marchi, Justine Solier, Stéphanie Finet, Karl Brillet, et al.. Structural insights into the membrane receptor ShuA in DDM micelles and in a model of gram-negative bacteria outer membrane as seen by SAXS and MD simulations. *Biochimica et Biophysica Acta: Biomembranes*, 2021, pp.183504. 10.1016/j.bbamem.2020.183504 . hal-02990445

HAL Id: hal-02990445

<https://hal.science/hal-02990445>

Submitted on 5 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Structural insights into the membrane receptor ShuA in DDM micelles and in a model of gram-negative bacteria outer membrane as seen by SAXS and MD simulations

Stéphane Abel^{a}, Massimo Marchi^a, Justine Solier^b, Stéphanie Finet^c, Karl Brillet^d and Françoise Bonneté^{e*}*

^a Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

^b Laboratoire d'Electrochimie et de Physico-chimie des Matériaux et des Interfaces, UMR 5279 CNRS Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, INP, F38000 Grenoble, France

^c Institut de Minéralogie, de Physique de Matériaux et de Cosmochimie, UMR 7590 CNRS-Sorbonne université, Bioinformatique et Biophysique, 4 Place Jussieu, F75005 Paris, France

^d Institut de Biologie Moléculaire et Cellulaire UPR 9002 CNRS, Architecture et Réactivité de l'ARN, 2 allée Konrad Roentgen, F67000 Strasbourg, France

^e Institut de Biologie Physico-Chimique (IBPC) UMR 7099 CNRS Université de Paris, Laboratoire de Biologie Physico-Chimique des Protéines Membranaires, 13 rue Pierre et Marie Curie, F75005 Paris, France

* Email: stephane.abel@cea.fr and bonnete@ibpc.fr

Keywords: Membrane transporter ShuA, DDM, Outer membrane model, SEC-MALLS, SEC-SAXS, Molecular dynamics simulations

Abstract

Successful crystallization of membrane proteins in detergent micelles depends on key factors such as conformational stability of the protein in micellar assemblies, the protein-detergent complex (PDC) monodispersity and favorable protein crystal contacts by suitable shielding of the protein hydrophobic surface by the detergent belt. With the aim of studying the influence of amphiphilic environment on membrane protein structure, stability and crystallizability, we combine molecular dynamics (MD) simulations with SEC-MALLS and SEC-SAXS (Size Exclusion Chromatography in line with Multi Angle Laser Light Scattering or Small Angle X-ray Scattering) experiments to describe the protein-detergent interactions that could help to rationalize PDC crystallization. In this context, we compare the protein-detergent interactions of ShuA from *Shigella dysenteriae* in n-Dodecyl- β -D-Maltopyranoside (DDM) with ShuA inserted in a realistic model of gram-negative bacteria outer membrane (OM) containing a mixture of bacterial lipopolysaccharide and phospholipids. To evaluate the quality of the PDC models, we compute the corresponding SAXS curves from the MD trajectories and compare with the experimental ones. We show that computed SAXS curves obtained from the MD trajectories reproduce better the SAXS obtained from the SEC-SAXS experiments for ShuA surrounded by 268 DDM molecules. The MD results show that the DDM molecules form around ShuA a closed belt whose the hydrophobic thickness appears slightly smaller (~ 22 Å) than the hydrophobic transmembrane domain of the protein (24.6 Å) suggested by Orientations of Proteins in Membranes (OPM) database. The simulations also show that ShuA transmembrane domain is remarkably stable in all the systems except for the extracellular and periplasmic loops that exhibit larger movements due to specific molecular interactions with lipopolysaccharides (LPS). We finally point out that this detergent behavior may lead to the occlusion of the periplasmic hydrophilic surface and poor crystal contacts leading to difficulties in crystallization of ShuA in DDM.

1. Introduction

Membrane proteins (MPs) play major roles in living organisms. They participate to exchanges and communications between cells and their immediate environment as receptors, transporters, ion channels, etc. MPs are also involved in a large number of pathologies (influenza, HIV, etc.) and genetic diseases (cystic fibrosis). Their physiological and biomedical relevance makes them major targets for at least 60 % of pharmaceutical molecules in development. Despite their importance, the three-dimensional structures at atomic level of MPs represent barely 2% of biological macromolecular structures in the Protein Data Bank (PDB). This deficit is due to difficulties in the preparation of suitable samples for structural studies by NMR, cryo-EM or crystallography. These structural techniques require the use of membrane-mimetic environments, such as detergents, amphipathic polymers or nanodiscs, to extract proteins from their native lipidic environments and to keep them stable, functional and monodisperse. In the case of crystallography, growing well-ordered diffracting crystals directly from detergent solutions (also named crystallization *in surfo*) remains the most largely used method, because of its ease of implementation similar to soluble proteins (vapor diffusion, batch, dialysis). However, this crystallization method is very dependent on the nature of the detergent used (tail length, size and charge of polar head). The detergent must both shield the hydrophobic surface of the MP to maintain it stable and avoid its inactivation and possible aggregation deleterious for crystallization, and favor polar contacts between MP-detergent complexes within the crystal [1][2], while keeping the MP-detergent complex monodisperse during the crystal growth. In that goal, information about the influence of detergents on the 3D structure of MPs, the homogeneity of the complex in detergent solution, and the size and shape of the detergent buoy around the MP is essential to select the suitable detergent that will best mimic the native MP structure, *i.e.* MP within the biological membrane, and that will help in the crystallization of the MP-detergent complex.

There are few experimental/computational studies describing both the MP-detergent complex and the detergent interactions with the protein. In the last ten years, some studies on structural properties of protein-detergent micelle complexes have been published describing either the protein conformation in amphiphilic environments or the detergent organization around the protein. Pioneer [3][4][5] and more recent works [6][7][8] using experimental and MD analysis focused on the dynamics or folding of MPs in detergent or lipid environments. Other studies specifically focused on the geometrical modeling of detergent organization around MPs by fitting of SAXS data [9][10][11]. For our part, we study the organization and the specific interactions of different amphiphilic environments around a MP as well as the conformational behavior of the protein in these environments by independent techniques: SEC-MALLS and SEC-SAXS on one hand, and MD simulations on the other hand. Indeed, molecular dynamics (MD) simulations have become a standard approach to study detergent/membrane protein (MP) interactions and help to better understand experimental data [12][13][14]. Besides the fact that the MD technique enables to study the temporal characteristics of a molecular system, it also gives access to structural and temporal characteristics of the detergent/MP interactions at atomic resolution and therefore provides valuable information that complete experiments. For instance, coupled to SAXS (or SANS), which gives access to the form factor of the complex and to SEC-MALLS, which permits to calculate mass fractions of protein and detergent in the complex and to deduce the homogeneity of the complex, MD could help to describe and understand the role played by detergent environments on MP structure and the ability of detergents to make MP crystallization possible.

Here, we report the structural and dynamics analyses of ShuA from *Shigella dysenteriae* in detergent micelles allowing us to better understand the mechanism of stabilization and crystallization. ShuA, which is known to cause severe forms of fever and gastrointestinal infections (such as acute diarrhea) in children in many developing countries [15][16][17], is a TonB-dependent heme outer membrane (OM) transporter (TBDT). The TBDT proteins share common structural features as antiparallel β -strands that form the rigid cylindrical barrel connected with extracellular (ECL) and periplasmic (PPL) loops, inserted in the

outer membrane. In addition, hydrophobic residues face the lipid bilayer exterior, and polar residues line the interior of the β -barrel. This transmembrane β -barrel protein enables passage of various polar ligands into the cell [18][19]. In particular, ShuA binds heme as an iron source from the methemoglobin. Iron is transported into the periplasm and then translocated into the cytoplasm through an ABC transporter [20]. ShuA is a monomeric MP of 70 kDa with 640 amino acids. Its 3D structure obtained at 2.6 Å, crystallized with n-octyl- β -D-glucopyranoside (OG) detergent (PDB ID: 3FHH) reveals 5 short helices and 33 β -strands with different sizes (including 22 β -strands forming the β -barrel transmembrane domain of the protein) filled by the N-terminal plug domain (residue 1-130) (Fig. 1 and Table 1) [21]. The β -strands forming the β -barrel are connected by 11 long extracellular and 11 periplasmic “short turn” flexible loops that coalesce at the mouth of the pore opening. The lengths, the flexibilities and the content of conserved amino-acids of the extracellular loops play an important role in the ligand recognition and transport as shown by several mutagenesis experiments [22]. In fact, the extracellular loops (ECL) contain mixture of hydrophobic, acidic and basic residues that vary in length from 6 to 32 residues [21]. Short regions of the extracellular loops of ShuA can interact with hemoglobin *via* accumulation of hydrogen bonds, electrostatic, and hydrophobic interactions [21]. Mutagenesis experiments showed that several histidine residues such as His86 in the plug domain or His428 located in one of the short β -stands in the extracellular loop 7 (ECL7) directly exposed to the solvent play an important role in the binding of the free heme [22][23]. In this paper, we study using atomistic MD simulations the interactions between ShuA and the n-Dodecyl- β -D-Maltopyranoside (DDM) detergent, the most widely used detergent for MP stabilization and crystallization [2]. To perform MD simulations with an appropriate number of detergent molecules around the MP, independent SEC-MALLS experiments were beforehand performed on ShuA solubilized in DDM. Then from comparison between SAXS profile from SEC-SAXS experiments and theoretical SAXS profiles from MD trajectories, we examine the impact of the amphiphilic environment on the structural stability of ShuA for its crystallization in detergent micelles (*in surfo*). The results are finally

compared with a realistic model of gram-negative bacteria outer membrane (OM) of *Shigella flexneri* (same bacterial family) containing a mixture of bacterial lipopolysaccharide and phospholipids.

2. Materials and methods

2.1. Detergent and protein solutions

All salt and buffer reagents were purchased from Sigma. n-Dodecyl- β -D-Maltopyranoside (DDM) was obtained from Anatrace (CAS number 69227-93-6; critical micelle concentration (CMC) 0.015 mM). The recombinant protein was prepared as described in Brillet et al. [24], purified in 1 % octylpolyoxyethylene from Bachem (Octyl-POE, CAS number 27252-75-1; CMC 0.23 %) with 50 mM Tris-HCl pH 8.0, 500 mM NaCl. For SEC-MALLS and SEC-SAXS experiments, the Octyl-POE was exchanged with the DDM by three successive dilution/concentration steps using Vivaspin 2 ultra-centrifugal filters (Sartorius) with a 30 kDa molecular weight cutoff. So, 50 μ L of ShuA in OPOE 1 % was diluted in 1 mL of DDM in 50 mM Tris pH 7.5 (named Tris buffer). Concentrations of DDM for SEC-MALLS and SEC-SAXS experiments were chosen above its CMC and high enough to cover the MP in an arbitrary ratio of 1 g of detergent/g of ShuA. The concentration of the new detergent in the sample was checked after each step of dilution/concentration by refractive index measurement. The filtrate was also checked by UV absorbance at 280 nm for evaluation of protein recovery. SEC-MALLS and SEC-SAXS were done using a silica gel Yarra SEC-3000 (300 X 4.6 mm) 3 μ m column (Phenomenex) with an elution buffer containing the detergent at concentration above CMC.

2.2. SEC-MALLS analysis of protein-free detergent micelles and protein-detergent complexes

SEC-MALLS experiments for absolute mass determination of protein-free detergent micelles (Mic) and protein (P) and detergent (D) mass fraction determination in protein-detergent complexes (PDC) were carried out on a Shimadzu HPLC coupled to an Optilab® T-rEX™ refractometer and a miniDawn™ TREOS Multi Angle Laser Light Scattering (MALLS) detector (Wyatt Technology). Molar mass of

molecules passing through a SEC column can be determined without calibration by measuring their static light scattering, knowing their concentration, by using the classical Rayleigh relationship:

$$R_{\theta,\text{sample}} - R_{\theta,\text{solvent}} = KcM \quad (1)$$

$$\text{with } K = \frac{4\pi^2 n_0^2}{N_A \lambda^4} \left(\frac{\partial n}{\partial c} \right)^2 \quad (2)$$

K is an optical constant, which depends on the refractive index (RI) of the buffer (n_0), the wavelength λ of the laser light, the refractive index increment $\partial n/\partial c$ of the scattering molecule and N_A the Avogadro number. θ is the angle between the incident and the scattered light. For the PDC, a two-component analysis must be carried out to determine both detergent and protein mass fractions. This methodology has been thoroughly described in recent articles [25][26]. Briefly, in this case, the accessible information from triple-detection SEC is the concentration (UV and RI) and the static light scattering (LS) of the PDC, which finally gives the molar mass of the complex through the equation:

$$M_{\text{PDC}} = \frac{R_{\theta,\text{PDC}} - R_{\theta,\text{solvent}}}{Kc_{\text{PDC}}} \quad (3)$$

PDC molar mass and concentration depend on the mass fraction (a) of the protein and (b) of the detergent or on the mass ratio ($\delta=b/a$) of bound detergent to protein through the expressions:

$$\begin{aligned} M_{\text{PDC}} &= (1 + \delta)M_{\text{P}} \\ c_{\text{PDC}} &= (1 + \delta)c_{\text{P}} = ac_{\text{P}} + bc_{\text{D}} \end{aligned} \quad (4)$$

Although c_{PDC} and a , b and δ are not directly known, the protein and detergent concentrations can be determined via UV absorbance and refractive index:

$$\begin{aligned} c_{\text{PDC}} &= \frac{\Delta Abs_{280\text{nm}}}{\varepsilon_{\text{PDC}}} = \frac{\Delta Abs_{280\text{nm}}}{a\varepsilon_{\text{P}} + (1-a)\varepsilon_{\text{D}}} \\ c_{\text{PDC}} &= \frac{\Delta n}{\left(\frac{\partial n}{\partial c}\right)_{\text{PDC}}} = \frac{\Delta n}{a\left(\frac{\partial n}{\partial c}\right)_{\text{P}} + (1-a)\left(\frac{\partial n}{\partial c}\right)_{\text{D}}} \end{aligned} \quad (5)$$

Knowing the respective extinction coefficients and refractive index increments of the protein and the detergent and thus the PDC concentration, the Rayleigh equation allows us to estimate the PDC molar

mass and the respective mass fraction of protein and detergent on each point of the SEC elution peak. The automatic procedure *Protein Conjugate* from Astra V permits the direct calculation of detergent and protein mass fraction using for ShuA $\epsilon^{0.1\%} = 1.664 \text{ L.g}^{-1}.\text{cm}^{-1}$ at 280 nm and $\partial n/\partial c = 0.185 \text{ mL.g}^{-1}$ and for DDM, $\epsilon^{0.1\%} = 0$ at 280 nm and $\partial n/\partial c = 0.1473 \text{ mL.g}^{-1}$. 20 μL of ShuA at 1.6 mg.mL^{-1} in 3 mM DDM was injected in triplicate through the size exclusion column (Yarra-SEC 3000, Phenomenex) and eluted at a flow rate of $400 \mu\text{L.min}^{-1}$.

2.3. SEC-SAXS analysis of protein-detergent complexes

Small Angle X-ray Scattering experiments were performed on the beamline SWING at the French synchrotron facility (SOLEIL, St Aubin). With a wavelength of 1.03 \AA and a sample-to-detector distance of 1,788 m, the achievable q -range was 0.0056 to 0.62 \AA^{-1} , with $q=4\pi\sin\theta/\lambda$ and 2θ the scattering angle. 5 and 50 μL of ShuA at 5.5 mg.mL^{-1} in 10 mM DDM were injected through the same size exclusion column (Yarra-SEC 3000, Phenomenex) on an Agilent HPLC system, a UV-Vis cell to measure protein concentration at 280 nm and the SAXS flow cell at a flow rate of $300 \mu\text{L.min}^{-1}$. 180 SAXS images were collected at the beginning of elution for buffer subtraction and 255 SAXS images for samples during the elution, each with a frame duration of 750 msec. The 2D-SAXS patterns were normalized to the transmitted intensity and azimuthally averaged by using the Foxtrot program [27]. The 180 frames corresponding to elution buffer signal were averaged to provide a buffer signal for subtraction of individual sample frames. Both forward intensity, $I(0)$, and radius of gyration, R_G , were determined from the Guinier approximation $I(q) = I(0)\exp(-(qR_G)^2/3)$ at very small angles, assuming that $qR_G < 1$ and then were plotted as a function of sample frames in order to determine ranges of frames where the radius of gyration presents no significant variation. The resulting scattering curves were analyzed, and a merged curve from the diluted ShuA_DDM at low q -range and the concentrated ShuA_DDM at high q -range was obtained and used for MD simulations.

2.4 Molecular dynamics simulations

2.4.1 Composition of the simulation box containing the protein in detergent or in outer membrane

In this work, as a starting structure for the protein, we used the unique crystal 3D-structure of ShuA obtained in n-Octyl- β -D-Glucopyranoside (OG) at 2.6 Å resolution (PDB entry: 3FHH) (Figure 1) [21]. The missing residues and atoms (i.e. N₂₈₀TGSSG₂₈₅, H₃₂₈HHHHHPGGA₃₃₇ and S₅₇₉SY₅₈₁) located in the extracellular loops of the protein, ECL4, ECL5 and ECL10, were rebuild with the ChimeraX [28] and SCWRL4.0 [29] programs. The protein, detergents and the bacterial phospholipids were modeled with the CHARMM36m [30] force fields for lipids and carbohydrates [31][32][33]. All the protein side chains were modeled with their typical protonation states at pH 7.

The initial configuration of ShuA in the two DDM detergent systems (ShuA_DDM) was constructed with a pre-assembled approach with the *Micelle Builder* module in CHARMM-GUI (www.charmm-gui.org) [34][35][36]. This tool places the desired number of detergent molecules (i.e. 268 or 294) obtained from SEC-MALLS analysis (see below) around the ShuA transmembrane domain delimited by 22 β -strand segments (Fig. 1). The PDC was then solvated in a cubic box containing 18 Na⁺ ions to neutralize the system and 90000 TIP3P water [37] to have enough room (\sim 20 Å) between the MP-complex and the box edge, leading to detergent concentrations between 0.17 and 0.19 M respectively for the two systems. We point out that to improve the sampling and examine the influence of the simulation starting conditions on the MD results, we performed two simulations (named hereafter, replica) with different random seeds of each ShuA_DDM system during 200 and 220 ns.

In case of ShuA in the OM system (ShuA_OM), we constructed a molecular model of the outer membrane formed with a mixture of lipopolysaccharide (LPS) of *Shigella flexneri* without O-antigen polysaccharide and chemical modifications (“user” core choice) and with bacterial phospholipids (PL) available in the CHARMM-GUI database. To do this, we used the *Membrane Builder* module in CHARMM-GUI and followed the steps described in Wu *et al.*[36]. Hence, the OM consisted of an asymmetric membrane where the outer layer is composed of a model of LPS and the inner layer with a mixture of PL. The LPS structure is quite complex and depends on the bacterial strains and the

environments [38]. Here, it contains a LipidA moiety formed with phosphorylated β -d-glucosaminyl-(1 \rightarrow 6)- α -N-glucosamine linked to a disaccharide backbone ester-linked and amide-linked to fatty acids with one C12 and five C14 chains (Fig. 2e). In case of the LPS inner and outer cores they contain a high proportion of rare saccharides such as 2-keto-3-deoxyoctulosonate (AKDO) and l-glycero-D-mannoheptose (AHEP) linked with 1 \rightarrow 2, 1 \rightarrow 5 to standard hexoses (such as α/β - glucose and α -galactose), respectively. In the case of the inner membrane leaflet, it is formed with a mixture of 1-palmitoyl(16:0)-2-palmitoleoyl(16:1cis-9)-phosphatidylethanolamine (PPPE, Fig. 2b), 1-palmitoyl(16:0)-2-vacenoyleoyl(18:1cis-11)-phosphatidylglycerol (PVPG, Fig. 2c) and 1,1'-palmitoyl-2,2'-vacenoyleoyl cardiolipin with a net charge of $-2e$ (PVCL2, Fig. 2d) at a ratio of 15:4:1, respectively [39]. To have enough lipids between the MP surface and the box edge, we constructed an outer membrane model with 48 LPS, 105 PPPE, 28 PVPG and 7 PVCL2, which gives a membrane with initial XY sizes around 106 X 106 Å. The ShuA_OM system was then solvated with 37970 TIP3P water. Since, the LipidA and the oligosaccharide core of the LPS are strongly negatively charged (the total charge of each LPS is $-9e$ in this study) and require divalent cations to form stable membrane [38][40], we have also added in the simulation box 216 Ca^{2+} , 163 K^{+} and 103 Cl^{-} ions to neutralize the system and have a KCl concentration around 150 mM. A summary of the characteristics and snapshots of all the systems extracted from the three simulations trajectories are given in Table 2 and in Fig. 3, respectively.

2.4.2 MD simulation parameters

The simulations discussed in this paper were carried out in the constant pressure and temperature (NPT) ensemble at 300 and 310 K (the latter temperature is consistent with the liquid-disordered structure of the OM, see ref. [36]) and 1.015 bar for the simulations carried out with the ShuA_DDM and ShuA_OM systems, respectively, with the GROMACS MD package (v2018.2) [41]. To prepare the system prior to the simulation productions, each system was minimized and equilibrated by using the CHARMM-GUI default simulation parameters. In brief, we minimized each initial molecular configuration with the

steepest descent algorithm and an energy tolerance lower than $1000 \text{ kJ.mol}^{-1}.\text{nm}^{-1}$. In the second stage, the minimized systems with restrained forces on the protein-detergent/membrane complex were equilibrated at constant volume and temperature (NVT) at 300 K (ShuA_DDM) and 310 K (ShuA_OM) for 400 ps using the Berendsen thermostat (with a coupling constant $\tau_T = 1.0 \text{ ps}$). This stage was followed by two restrained and unrestrained equilibration stages (30 ns in total) in the NPT ensemble with the temperature and the pressure controlled with the velocity-rescale (v-rescale) thermostat [42] (with $\tau_T = 1.0 \text{ ps}$) and the Parrinello–Rahman barostat [43,44] (with $\tau_p = 5.0 \text{ ps}$ and a compressibility value of $45 \times 10^{-6} \text{ bar}^{-1}$), respectively. During these stages, the restraint forces on the ShuA-detergent/lipids complex were gradually reduced from 1000 to $0 \text{ kJ.mol}^{-1}.\text{nm}^{-1}$. The MP-detergent/lipid complexes were then simulated with the Nose-Hoover thermostat [45,46] (with $\tau_T = 5 \text{ ps}$) and the Parrinello–Rahman barostat and with an isotropic (ShuA_DDM) and semi-isotropic (ShuA_OM) pressure coupling schemes. As previously [36][47], additional dihedral angle restraints were also applied to keep all the sugar rings in the LPS in the chair conformation and these restraints were kept in the production simulation. The electrostatic interactions were evaluated with the particle mesh Ewald (PME) algorithm [48] with a cutoff of 12 \AA and the van der Waals interactions were smoothly switched off at $10 - 12 \text{ \AA}$ by a force-switching function [49]. The P-LINCS algorithm [50] was applied to constrain bond lengths to their equilibration values. Periodic boundary conditions and a time step of 2 fs for integration of the equations of motion were used with the neighbor list updated every 100 fs . The atomic configurations of the five systems were collected every 5 ps and the structural properties analyzed and discussed in the main text or in the supporting information using different GROMACS tools and programs developed by us.

3. Results and discussion

3.1. SEC-MALLS and SEC-SAXS analysis

3.1.1. Quantification of detergent in ShuA_DDM complex for MD simulations

SEC-MALLS analysis has been performed on ShuA_DDM complex in order to characterize the homogeneity of ShuA_DDM complex, the oligomerization state of ShuA in the complex and determine the number of DDM molecules bound to the ShuA transmembrane domain, value that will be used in MD simulations. Fig. 4a shows both UV absorbance at 280 nm (blue line) and LS90° (red line) signals for a solution of ShuA at 1.6 mg/mL in 3 mM DDM eluted through the Yarra 3000 column with the Tris buffer containing 0.3 mM DDM (i.e. 2 CMC). The chromatograms display two main peaks corresponding to ShuA_DDM complex (peak 1) and DDM micelles (peak 2), the peak at $V_0=1.9$ mL corresponding to the void volume of the column. The light scattering (LS) signal from the MP complex is well separated from that of detergent micelles and aggregates meaning that this elution condition could be used for SEC-SAXS experiment. The molar mass (black symbols) of extra DDM micelles (~67 kg/mol for a corresponding number of 131 DDM per micelle) agrees well with literature (around 130) [51][52], and shows that the detergent exchange is achieved. The elution peak 1 presents an asymmetrical shape, suggesting the presence, at $V \approx 2.4 - 2.5$ mL, of inhomogeneous complexes of molar mass higher than 400.000 kg/mol (blue symbols). Molar masses and composition of ShuA_DDM complex are determined from excess Rayleigh ratio (equations 1 to 5) using the procedure Protein Conjugate from Astra V and are depicted on the Figure 4a. The mean value of the weight-average molar mass of ShuA (empty red dots) is $84.4 \pm 2\%$ kg/mol, ca. 19% above the expected value from its 6 his-tag sequence (i.e. 70.77 kg/mol), but suggesting mostly a monomeric form of ShuA. The discrepancy could be due to accuracy of the light scattering technique, which is about 5 %, but more likely to some oligomeric forms of ShuA in DDM, which could not be separated on the Yarra 3000 column and which overestimate the protein mass fraction. The detergent mass (full red symbols) bound to ShuA was evaluated to be between 137 and 150 kg/mol, corresponding to 268 and 294 DDM molecules and a mean mass ratio of bound DDM of 1.8 ± 0.2 g/g consistent with literature [9][26][53]. These minimum and maximum numbers of DDM molecules were then used in MD simulations.

3.1.2. Form factor of ShuA_DDM complex from SAXS

To validate further MD simulations, SEC-SAXS experiments have been performed by injecting two samples (5 and 50 μL) of ShuA_DDM at 5.5 mg/mL, exchanged in 10 mM DDM, through the same HPLC column as for SEC-MALLS experiment and eluting with a 0.1 % DDM Tris buffer. Figure 4b depicts the forward intensity $I(0)$ (left axis, empty red dots) and the radius of gyration R_G (right axis, empty blue square) for ShuA_DDM complex (5 μL) obtained from the Guinier analysis of each SAXS image recorded during elution. The SEC-SAXS elution profile shows as for SEC-MALLS two peaks, the first peak (frames 20 to 120) corresponding to the complex ShuA_DDM and the second peak (from frame 175) corresponding to the DDM micelles. As in SEC-MALLS experiment, the peak of ShuA_DDM complex presents an asymmetrical shape, suggesting presence of oligomers (ca. frames 25-50). We have analyzed SAXS data in the zone ($>$ frame 50) where R_G remains constant and corresponds to the SEC-MALLS analysis. For the two injection volumes, SAXS curves have been averaged in the stable R_G -zone (i.e. frames 110 to 120) and merged using the 5 μL sample in the low q -range ($q < 0.08\text{\AA}^{-1}$) and the 50 μL sample in the large q -range ($q > 0.08\text{\AA}^{-1}$) (Fig. 4c). The radius of gyration of the ShuA_DDM complex is $42 \pm 0.7\text{\AA}$ and the maximum dimension D_{max} from the pair distribution function using the program GNOM [54] is about 113 \AA . By using SEC-MALLS analysis and SEC-SAXS experiment that provide respectively the number of DDM molecules around ShuA and the form factor of the ShuA_DDM complex in solution, we conducted MD simulations.

3.2. MD simulations

The MD results will be used in the next sections to discuss the structural properties of the five systems in the DDM and OM environments. Before to analyze of the simulation results, we verified that each simulated system reached equilibrium. In case of the ShuA_DDM complexes, visualization of the MD trajectories shows that the integrity of the PDC complexes during the course of two production runs is maintained. We also noticed that all the DDM molecules formed a detergent corona with a typical

toroidal shape around the MP as previously observed for similar systems [5][9][55][56]. Moreover, all the detergent molecules remained attached to the corona during the course of the two systems and its replicas with their maltose headgroup facing the aqueous solution and their alkyl chain in contact with the ShuA TM domain (see discussion below). For instance, computation of the total PDC-water surface contacts between all the PDC non-hydrogen atoms and the water oxygen from Voronoi-Delaunay tessellation [57] with *trjVoronoi* [58][59][60], shows no drifts and fluctuates along the trajectories around an average value of 88200 and 92150 Å² indicating a stable structure for the PDC complex with 268 and 294 DDM molecules, respectively (Figure S2).

In case of the ShuA_OM system, we also checked the equilibration and the convergence of the membrane structure by plotting the lengths of X/Y box vectors (in Å). As shown in the Fig. S3, the membrane X/Y box vectors also shows no drifts during the course of the MD and fluctuates around values of 107 – 108.5 Å with an average value computed from the last 100 ns of 107.6 ± 2.6 Å, indicating that the membrane converged on the simulation timescale.

3.2.1 ShuA_DDM structure and interactions

3.2.1.1. SAXS analysis. To discuss the influence of the different environments on the structural properties of ShuA, we have firstly validated our models of ShuA_DDM complex against SEC-SAXS experiments. To this aim, we computed the theoretical SAXS intensity curves from collection of successive snapshots extracted from the four ShuA_DDM simulations using explicit-solvent calculations with *trjSaxs* [56][60]. As described in details in Ref. [56], the calculation of the theoretical SAXS intensity is based on 3D fast Fourier transforms and particle meshing similar to the one used in smooth PME for computing electrostatic energies and forces. The calculations of the solution $I_{\text{sol}}(q)$ is performed by computing the theoretical small angle scattering intensity as the difference: $I_{\text{sol}}(q) = I_{\text{PDC}}(q) - \alpha I_{\text{buf}}(q)$ [61][62]. $I_{\text{PDC}}(q)$, $I_{\text{buf}}(q)$ are the scattering intensities of the PDC and the buffer (i.e. the solution with only the water and ions), respectively, and α a rescaling factor very close to 1, used in the solution experiments

to take into account buffer over-subtraction [61]. For these calculations, the “solution” contained the ShuA_DDM complex and the “pure buffer” was composed by only 90000 TIP3P water and 18 Na⁺ ions and simulated in the same thermodynamic conditions of the protein complex for 30 ns. In both cases, the calculation of $I_{\text{sol}}(q)$ is carried out from conformational ensembles with 5000 individual frames as observed in last 150 – 200 ns with a grid of 288³ points, with a cubic B-spline interpolation of order 6 and a supercell $\sigma = 4$. The results of these calculations are shown in Fig. 4c in blue, red continue and violet and green dashed lines for the four ShuA_DDM complexes simulated with 268 and 294 DDM molecules, respectively, and with the experimental SAXS curve obtained from the SEC-SAXS experiment (empty black dots). The agreement with the experimental curves is rather good reproducing well the experimental SAXS profile characteristics for all the systems, i.e. minimum at $q = 0.1 \text{ \AA}^{-1}$ and the double-peak feature between $q \approx 0.12$ and 0.2 \AA^{-1} , which differs with a single broad peak observed for pure DDM micelle [63]. We point out that the use of different trajectory points at different times in the last 50 ns of each trajectory in these calculations give very similar results (not shown) indicating that averages are fully converged. Similar profiles were obtained, for instance, with the Aquaporin0 protein in DDM [14][56]. The averaged R_G of the ShuA complexed with 268 and 294 DDM, obtained from the computed SAXS curves for each replica with the Guinier approximation at small q 's, are similar and, on average, equal to 42.0 and 42.8 \AA , respectively. These values are in good agreement with the experimental value obtained from the Guinier analysis (42 \AA). However, at larger q values the experimental SAXS curve better fits with a monomeric ShuA complex surrounded with 268 DDM molecules.

3.2.1.2 Protein-Detergent complex characteristics. To obtain a more detailed picture of the spatial extents of the different components (i.e. the protein, protein detergent complex, the detergent, the DDM headgroup and tail) of the ShuA_DDM complex, we have computed the one-dimensional (1D) and two-dimensional (2D) density distributions along the three X , Y and Z axes from the last 150 ns of the simulations with 268 and 294 DDM molecules, using 0.5 \AA grid points (Figs. 5a-c and S4). The averaged

2D density distributions (Figs. 5a-b and S4) show some differences in the distribution of the DDM around the molecules between the two replicas indicating a slight dependence of the starting conditions in this structural property. However, in all the case, the detergent corona forms a closed belt that follows the protein shape leading to ellipsoidal objects for the two ShuA_DDM complexes. It is also evident from the Figures 5b and S4b and the MD snapshots extracted from the ShuA_DDM complex shown in Figures 3a-b that the increase in the number of DDM molecules leads to an accumulation of some detergent molecules around the MP. The density curves overlap indicating similar structure of the PDC along the Z axis for each replica and that the density variation of the DDM headgroup around -28 \AA compared to position $\sim 8 \text{ \AA}$ observed for one ShuA_294DDM simulation is caused by the difference of the distribution of the DDM around the protein. From the average 1D density distributions of the PDC oriented along the Z axis and centered at $Z = 0 \text{ \AA}$ (Fig. 5c-d) computed for the two replicas from the 150 ns of the runs, it is also possible to deduce the thickness of the hydrophobic (estimated as the distance where the red and green curves intersect, d_{thick} , violet arrow) domain of the DDM corona. For the four ShuA_DDM simulations d_{thick} , are similar and around 22 \AA . This value a bit smaller than the value estimated from the OPM to be $24.6 \pm 0.9 \text{ \AA}$ [64][65].

In Table 3, we report the corresponding average lengths of the semi-axes X , Y and Z for the protein (i.e. X_p , Y_p and Z_p), the overall complex (i.e. X_{PDC} , Y_{PDC} and Z_{PDC}) and the DDM corona (i.e. X_{DDM} , Y_{DDM} and Z_{DDM}) alone estimated from the 1D and 2D density distributions for the two ShuA systems and their replicas. The average dimensions of the MP alone are similar and do not change indicating no deformation of the protein with the increase of the number DDM molecules in the corona with X_p , Y_p and Z_p values around to 20.0 \AA , 20.0 \AA and 35 \AA , respectively. In case of the dimensions of the ShuA_DDM complexes, we obtained values that change a little between each replica and are on average of $42 - 46.0 \text{ \AA}$, $43.5 - 46.5 \text{ \AA}$ and 39.0 \AA that confirm that two complexes can be modeled as oblate ($X = Y > Z$) or tri-axials ellipsoids. In case of the average dimensions of the detergent belt (i.e. X_{DDM} , Y_{DDM} and Z_{DDM}), we obtained values around of $22.5 - 26.5 \text{ \AA}$, $23.5 - 26.5 \text{ \AA}$ and 16 \AA . The X_{DDM} , Y_{DDM} are not far from

the length of a DDM molecule (around 24 Å calculated with Pymol [66]). A more detailed analysis of the interactions between the detergents and the different protein domains will be described in the section *Structure of the DDM in the PDC corona*. There are few studies characterizing the properties of the detergents in the corona of membrane proteins and in DDM in particular. For instance, the works of Berthaud et al [9] showed that the DDM corona of the solubilized Aquaporin-0 tetramer could be fitted using parameters (i.e. overall thickness of the detergent layer and extent of hydrophobic/hydrophilic regions) that were in agreement with experimental parameters of protein-free detergent micelles [67]. But can we generalize this result to all MPs solubilized in DDM and compare MP-corona to protein-free micelles? To examine this aspect in case of the simulations with the ShuA_DDM complexes, we completed our analysis by investigating specifically the hydration and conformational properties of the DDM molecules forming the corona and compared with previous MD results obtained, for instance, with pure DDM micelles simulated with the same CHARMM36 force field (Tables S1-3) [31]. All these analyses were done for each ShuA_DDM and its replica and detailed in the sections S1 and S2 in supplementary information and a brief discussion of the overall results is given below.

Concerning the surface contacts properties between the water and DDM, they were investigated by computing the Voronoi surface with the *trjVoronoi* code [58][59][60]. The results reported in Table S1 show similar values independently of the replica and the number of DDM in the complex. The surface contacts shared between the different parts of the DDM molecule (namely the headgroup and alkyl chain) decrease a little with the increase of the number of DDM in the corona but remain close to the value found for pure DDM micelle (around 200.0 and 20.0 Å², respectively). Also in agreement with the micelle results [31], we found that the surface contact between the water and DDM headgroup in the two systems represents the major part (~92 %) of the total hydration of the DDM corona.

When one focuses specifically on the average number of hydration water of the DDM maltose headgroup and alkyl chain in each system, our calculations showed that the average number of water within 3.5 Å each DDM is around 10, which is similar to the values found for DDM micelle with MD [31] or estimated

from SAXS [51], fluorescence quenching (TRFQ) techniques [68] or DDM around a model of membrane protein (RC-LH1-PufX) [63] or OmpA [11] reported to be 8 and 12.5, respectively. Moreover, consistent with the calculated larger surface areas mentioned above, the results also show that the hydration of the DDM headgroup show some differences depending on the position of the glucose on the headgroup (see Tables S1 and S2 and discussion), in agreement with results obtained in DDM micelle [31].

In previous papers [31][69], we also showed that the slight changes of the detergent conformation can modify the detergent hydration or the interaction between adjacent headgroup at the micelle surface. To investigate this in more details, we focused on the conformational properties of the DDM headgroup and alkyl chain of the two ShuA_DDM systems. The detailed analysis are reported in the section S2 in supplementary information. Taking together, the results show that the conformations of the DDM in the corona are quite similar between the four systems and close to results we found previously in micelle [31]. In particular, we found that the alkyl chain of the surfactant are similarly folded (with 72.1 % of CCCC in trans conformation) lending a significant reduction of the DDM alkyl chain 12.0 Å (16.7 Å when stretched according to Tanford [70]), which is similar to the length of the DDM alkyl chain in micelle (12.3 Å) [31] or approximatively equal to the half of the d_{thick} value estimated above with the radial density functions. For the DDM headgroup, its thickness is also reduced (9.6 Å, to be compared to the length of the maltose in fully extended conformation (10.3 Å) computed with Pymol [66]) due to its partial folding and comparable to the thickness value found in micelle (9.8 Å) [31]. The slight differences in the two ShuA_DDM systems for the headgroup conformations (see Fig. S1) may explain the slight differences in the headgroup hydration mentioned above (see results in Table S3 and Fig. S1 and discussion).

3.2.2 ShuA structure in the micelle and the membrane environments

3.2.2.1. Stability and structure of ShuA with the DDM and in the outer membrane. In the following sections, we will now turn our attention to the influence of the DDM and OM environments on the

structure and stability of ShuA. To this aim, we firstly evaluated the stability of the protein by monitoring the root mean-square deviation (RMSD) of the overall protein C_α atoms (in black and grey colors) and those located in the 11 ECL (in red and magenta colors) and in 22 TM β-barrels (in green and dark green colors) from the minimized crystal structure (Figs 6a-c). Overall, one can first remark that the two replicas of each ShuA_DDM system give similar results and that the detergent corona and the OM imposed only minor structural changes on ShuA compared to crystal structure, since we see that the RMSD_{C_α} for the overall protein are stable during the course of the three MD runs with a plateau between 3.0 - 3.5 Å, which is indicative of the conformational similarity of our MD simulations to the protein minimized structure. The jumps of the RMSD for overall C_α atoms observed at ~180 and 160 ns in the simulations of the 268 and 294DDM systems, respectively, are mainly caused by larger structural fluctuations of protein EC loops (in red and magenta) and in less extent of the protein helices (not shown) as suggested by the corresponding RMSD_{C_α}. We also notice that the amplitude of the RMSD_{C_α} values of the C_α atoms in the 22 TM β-barrels slightly differ between the DDM and OM environments (RMSD_{C_α} = 1.0 vs. 2.0 Å) probably caused by the slight arrangement of the conformation of the ShuA TM domain induced by the membrane environment during the equilibration stages. But as reported elsewhere (for instance in refs [18][45][46][71]), the RMSD_{C_α} values of 22 TM β-barrels are found very stable indicating that the TM barrel domain of the protein is strongly stabilized by the detergent corona and the membrane lipids during the course of the simulations. These results contrast with RMSD_{C_α} of EC loops that shows significant differences between the two environments. This behavior can be explained by the fact that in DDM the ECLs are in direct contact with the solvent and can move freely whereas, in the OM environment, the loops can interact with the surrounding LPS reducing their structural changes. These behaviors are confirmed by the time evolution of the secondary structure of the ShuA computed with the *do_dssp* tool [72] of GROMACS and depicted in Figs S5-7. This stabilizing effect of the LPS sugar core on the EC loops stability was also observed in other MD simulations [36] and could be important in the function of ShuA and BtuB protein as noted previously [8][22].

The calculated root mean-square fluctuation (RMSF) of the protein C_{α} from the last 100 ns for the five systems and the crystal are plotted in Fig. 7f. In addition to these figures and to better visualize the different protein regions that have the larger flexibility in the different environments, we provided in Figs 7a-e, the 3D representations of ShuA with its different domains colored according to the residues $RMSF_{C_{\alpha}}$. We first point out that the amplitudes of the fluctuations of residues (e.g. around 50, 90, 135, etc.) located in some loop regions of the protein slightly vary depending on the replica and may be caused by the localization differences of the surfactant near these loop regions that affects their motilities. The computations of the correlation coefficients R for the protein α carbons from “Crystallographic vs. simulation RMSF” estimated for the former with the crystallographic Debye–Waller B-factor (with $RMSF_{C_{\alpha,i}} = \sqrt{(3B_i/8\pi^2)}$ where B_i is the B factor of the i^{th} protein C_{α} atom [73][74]) are shown in Fig. S8 for all the systems. The figures clearly identified the greater mobility of some residues located in the loops compared to the crystal and leading to a comparable weak correlation ($R = 0.335 - 0.526$) between the mobility of the ShuA residues in crystal and in the simulations. A more detailed analysis of the RMSF for all the simulations shows that the β -barrels are very rigid in both environments independently of the replica, with $RMSF_{C_{\alpha}}$ values smaller than 1.0 \AA , indicating, once again, a strong stabilizing effect on the detergent corona and the membrane lipids on this domain. In particular, the N-terminal plug domain (delimited with Thr1 to Gln130, not shown in the Figs), is also stable (low $RMSF_{C_{\alpha}}$) in the two environments. We also notice that some regions of the protein such as the 5th helix (e.g., H5 (residue 599-603)) or some extracellular loops (e.g. ECL2 (residues 176 – 189), ECL3 (residues 225 – 238)) or those missing in the crystal (e.g.; ECL4 (residues 279 – 286), ECL5 (residues 325 – 344) and ECL10 (residues 572 – 586)) or the long loop ECL7 (residues 416 – 447) and the periplasmic loops (e.g. PPL1 (residues 131 – 136), PPL5 (residues 304 – 308) PPL6 (residues 360 – 363) or PPL8 (residues 462 – 469)) can exhibit large mobilities depending on the replica due their different contacts with the DDM or membrane lipids (Fig. 7). Several of these loops contain important residues involved in the

protein function such as ECL7 that contains the NPNL (Asn434-Leu437) domain, which is highly conserved in all heme transporters.

3.2.3 ShuA detergent and membrane interactions

To highlight the different factors that may influence the conformational stability of ShuA in the detergent and outer membrane environments, it is important to consider how the detergent molecules and the membrane lipids interact with the different domains of the protein. To visualize the different interaction sites between the detergents and the protein, we performed a contact analysis and computed the average number of atomic contacts between the alkyl chain and headgroup of the detergent and each protein residue non-hydrogen atoms during the last 100 ns of each run (Fig. S9-10). For these calculations, to define a contact, we used a cutoff distance of 4.0 Å [31,36], i.e. a residue was considered a contact if any of its atoms was at a distance less than a cutoff from any of the detergent atoms. In addition, as for the RMSF_{C α} calculations, we provided a colored 3D representation showing the different contact sites of ShuA according to the average number of contacts between each residue and the *overall* DDM (Figs. 8 - 9 and S9). From these figures, one can first notice that the contact patterns for the two ShuA_DDM systems present similar features and the results are comparable between the replicas. In particular, as it would be expected from the ShuA structure where the interior of the β -barrel is polar and partially filled with water, most of the inward-facing β -barrel residues (colored in blue in the Figs.) as well as the N-terminal plug domain (Thr1- Gln130) including the TonB-box (Met4-TVTAT-Gly10) do not interact with any DDM molecules. Moreover, the residues in the β -strands mainly interact with the headgroup or the tail of the DDM. Another interesting feature is that, in agreement with the 3D representation (Figs. 8 - 9 and S9) and the 1D-density profiles (Fig. 5c-d), most of the ShuA periplasmic loops that are located near the limit of corona or that are very mobile (e.g., PPL5 (residues 304-308), PPL6 (residues 360-363), PPL8 (residues 462 – 469) or PPL11 (residues 604 – 611)) can interact with the detergent molecules.

This is in contrast with the ShuA extracellular loops where most of them interact only with water since the DDM corona is located a bit far from them.

The crystal structure of ShuA (PDB entry: 3FHH) crystallized with OG (Figure S12) suggests that specific residues, such as Trp566, Phe570, Gly588, Asp590 located in the outer-surface of the 19th and 20th β -strands, respectively, are in close contact (distance < 4.0 Å) and thus can interact with the OG. From the Figs. S6 and S7, one can see that these residues (highlighted with a star in the Figures) are also in close contact (recover) with the headgroup or the alkyl chain of the DDM in the four ShuA_DDM systems.

In the case of the ShuA_OM system, we built a model of an asymmetric membrane with an outer leaflet composed of LPS, and with a mixture of bacterial phospholipids (75% PPPE, 20 % PVPG and 5% cardiolipin) in the inner leaflet. Therefore, it would be expected that the ShuA TM domain traverses the two membrane leaflets and interacts with the LPS or the phospholipids. To investigate this in more details, we also computed the average 1D density profiles for the different components of the OM (i.e. LipidA, LPS sugar core, PPPE, PVPG, PVLC2, the K⁺/Cl⁻ and Ca²⁺ ions and water, Figure S13) and performed the same type of contact analysis as for ShuA_DDM systems for the lipids (Figure S14). From the 1D density profile and similarly to previous studies [36][75][76], one can first notice that the LipidA and the sugar cores of the LPS recover a large part of the protein extracellular domain (delimited region of 10 Å < Z < 30 Å) whereas the bacterial lipids are mainly located in the periplasmic region (Z < -10 Å) of ShuA. Some portions of the extracellular loops (e.g. ECL4 and 5) and β -strands (e.g. 19th and 21th) can interact with both the LipidA headgroup and the inner core or the water; the β -strands mainly interact with LPS or phospholipid acyl tails and finally the turns on the periplasmic side mainly interact with PPPE and PVPG phospholipids and water. Interestingly, a closer examination of the possible contact between the protein and the cardiolipin molecules show no significant interaction sites on the protein during the course of the simulation (Figs. 9 and S14). This may be explained by the fact that these molecules are too far from ShuA (not shown) and their lateral diffusion is too slow to see their

rapprochement towards the protein within the simulation time scale (~ 200 ns). In case of the K^+/Cl^- ions, they are widely distributed into the water phase or for few of them near the membrane (Fig. S13 left). The Figure S13 also shows the distribution of the Ca^{2+} ions and indicates that they are mainly localized in the LPS inner core domain (delimited region of $5 \text{ \AA} < Z < 30 \text{ \AA}$) with two well-defined peaks (a strong peak at $\sim 10 \text{ \AA}$ and smaller peak at $\sim 22 \text{ \AA}$). Few Ca^{2+} ions were also found in the periplasmic domain of the protein. The first peak corresponds to the binding sites of the LPS phosphate groups (maroon in Fig. S13 right) in agreement with refs. [38][40][77] and the second peak to possible Ca^{2+} - protein binding sites such as Asp380, 384 and 386 that are located at the protein surface in the ECL6 and at the beginning of the 21th β strand. These residues are also found to bind the Pb^{2+} ions in the crystal (Fig. S15)[21]. The water density profile also shows that water molecules penetrate deeply inside the inner and outer core sugar of the LPS and solvate the lipid headgroup regions (not shown) with a slight penetration into the hydrophobic hydrocarbon tail region. Water also fills the ShuA β -barrel and, as we found for ShuA in the DDM environment, the N-terminal plug domain interacts only with water and the ions.

In previous simulations of OMPs in various models of outer membranes [78], it was shown that the lengths of the lipidA and the phospholipids have a strong influence on the membrane properties such as the membrane hydrophobic thickness, T_{memb} . Here, we simulated a membrane formed with LipidA with four C14 and one C12 chains and with phospholipids with C16 and C18 chains. It is therefore expected that our membrane model will have a larger T_{memb} than other OM membranes such as for *P. aeruginosa* or *E. Coli*, which are made of LipidA with shorter alkyl chain (i.e.; C12) [36][76] of 20 – 22 \AA in length. To investigate this aspect, we computed the hydrophobic thickness of the membrane, T_{memb} , from the average distributions along the Z axis of C2 and C4 atoms (i.e., the carbon bonded to carbonyl group) of lipid A and the acyl chain C2 carbon atoms (i.e., the carbon bonded to carbonyl group) of PPPE, PVPG and PVCL2 phospholipids [76] and compare with the density profile of the ShuA TM β -barrel domain only (blue and red peaks and black curve, respectively in Figure S16a). We estimated a value for T_{memb} around 24.0 \AA that closely matches with the thickness of the TM hydrophobic domain of ShuA estimated

to be $24.6 \pm 0.9 \text{ \AA}$ [64][65] but also with length of the hydrophobic domain of the DDM corona computed above.

Finally, ShuA is rich in positively charged residues such as Lys and Arg located at its outer surface and in particular at the beginning of some β -strands (i.e., region $Z > 10 \text{ \AA}$) or some extracellular loops (Fig. S17 and Table 1). These preferential localizations may enhance their bindings with the surrounding negatively charged PO_3^{2-} and COO^- groups of the LipidA headgroup or the AHEP and AKDO sugars of the LPS inner core located nearly at the same level in the membrane. In the inner part, negatively charged residues such as Asp and Glu located in some periplasmic loops, which participate to polar contacts in the ShuA crystal lattice (Figure S18), are masked by DDM sugar heads, in contrast to OG where the headgroup and the alkyl chain of this surfactant are significantly smaller and gives to a surfactant corona with a smaller thickness (data not shown).

4. Conclusion

To shed light the impact of the amphiphilic environment on the structural stability of membrane proteins in view of their crystallization in detergent micelles, we investigated the large heme/hemoglobin outer membrane transporter, ShuA from *Shigella dysenteriae* by experimental (SEC-MALLS & SEC-SAXS) and explicit MD simulations. ShuA contains 640 residues and shares common structural features with other TBDTs such as a rigid cylindrical β -barrel formed with 22 antiparallel β -strands connected by long ECL loops and short PPL turns. The protein was first studied in presence of DDM micelles, which is one of the most commonly used detergent. Thus, by SEC-MALLS and SEC-SAXS experiments we could first determine, the number of the detergent molecules bound to the protein TM-domain (about 268 - 294 DDM / ShuA) and the shape of the ShuA_DDM complex, respectively. Using these experimental results, we could meaningfully model the protein detergent complex by atomistic MD simulations and compare its structural behavior with that of the protein inserted in a realistic model of gram-negative bacterial outer membrane formed by a mixture of lipopolysaccharide and long chain bacterial phospholipids. Our

model and simulations of the PDC were also validated by the calculation of the theoretical SAXS curves from the multiple MD trajectories, which were found in good agreement with the SAXS profile obtained from SEC-SAXS experiments.

Concerning the structural properties of ShuA in detergent micelles, we find that the DDM corona forms a large closed belt (with a thickness of ~ 22 Å) located near the periplasmic region of the protein and covers a large part of the ShuA TM domain. The study of the DDM properties in the corona indicates that the detergent molecules share similar structural features as already described for pure DDM micelles, such as, similar accessibility to the solvent, hydration degree and conformation of the detergent alkyl chain and headgroup. Concerning the structure of ShuA in the DDM and OM environments, our MD simulations reveal that the detergent and the membrane impose only minor structural changes to the inner core of the protein structure compared to the crystal structure. Indeed, in agreement with previous studies, we find that the TM domain delimited by the 22 β -barrels covered by the DDM molecules or the membrane lipids are very stable ($\text{RMSD}_{\text{C}\alpha} < 2.0$ Å) in contrast with the extracellular and periplasmic loops where we observe larger flexibilities in DDM environments than in OM. In both cases, a value of less than 4.0 Å for the $\text{RMSD}_{\text{C}\alpha}$ of the overall protein $\text{C}\alpha$ was found. The analysis of the interactions between the protein and the different lipids of the outer membrane shows that the LipidA and the sugar core of the LPS recover protein extracellular domains and that the bacterial lipids (PPPE, PVPG) interact with the periplasmic loops. The thickness of the membrane formed with the C14 and one C12 chains of the LPS and the C16 and C18 chains of the PPPE and PVPG is close to the thickness of the TM hydrophobic domain of ShuA (around 24.6 Å), thus favoring its stability. We also observe that the preferential localizations of the positively and negatively charged residues at the outer surface of the β -barrel or the ECL may enhance their bindings to the surrounding positive ions (such as Ca^{2+}) and negatively charged PO_3^{2-} and the COO^- groups of the LipidA headgroup or the AHEP and AKDO sugars in the LPS inner core, thus reducing significantly their movements. On the other hand, it may also favor polar contacts with negatively charged residues of the inner surface leading to possible crystallization.

To conclude, this combined study of experiments and MD simulations have been able to discriminate between effects observed on the structure and the interactions of ShuA due to detergent and native environments. Thus, our approach should be beneficial in future studies for choosing adequate detergents for crystallization.

Uncorrected/Edited Proof

List of Figures

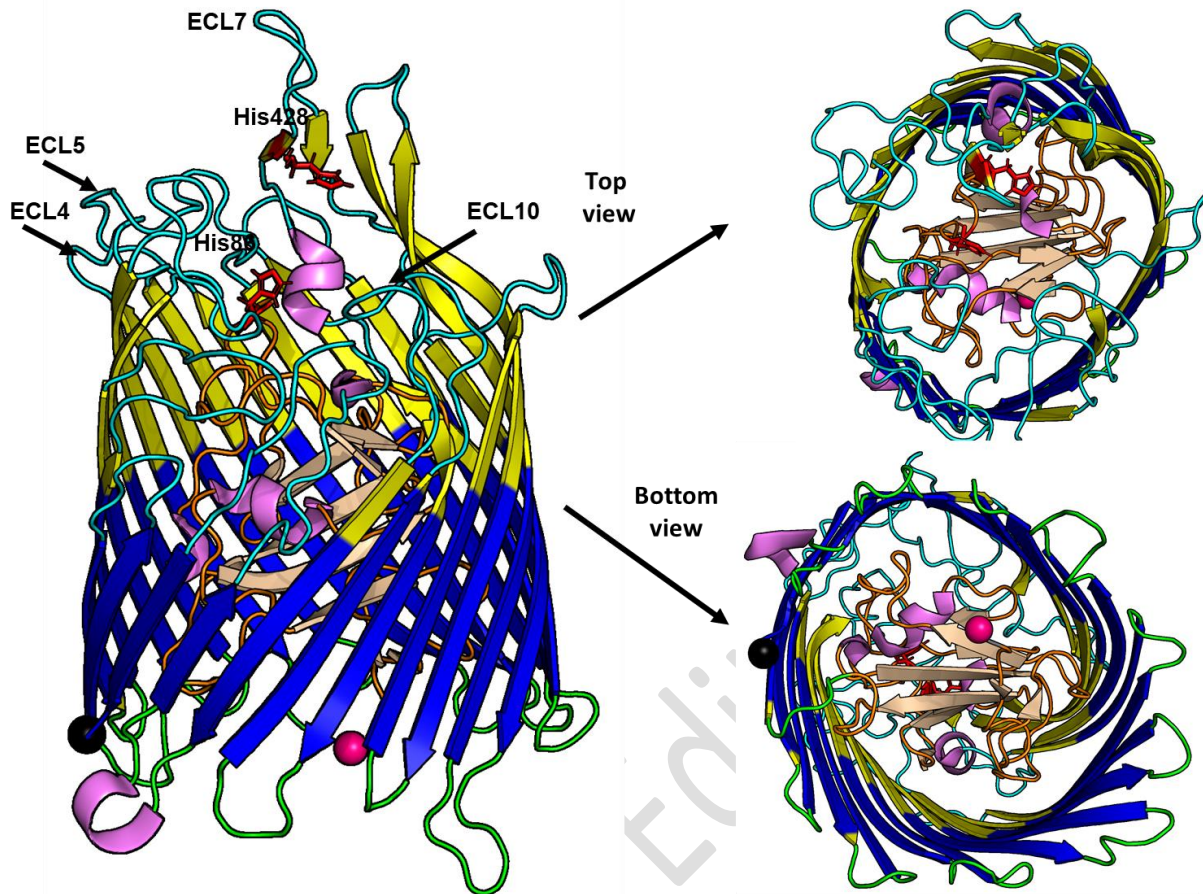


Figure 1: Representative views of the complete structure of ShuA from *Shigella dysenteriae* (PDB entry: 3FHH) with all the extracellular loops and missing residues (ECL4, ECL5 and ECL10 in the Figure) rebuild with *ChimeraX* [28] and *SCWRL4.0* [29]. The N and C term residues are highlighted in hot pink and black spheres, respectively. The 11 extra and periplasmic loops are in cyan and green colors respectively whereas the five helices are in pink. The 10 β -strands in the TonB plug domain as shown in cartoon and in the wheat color. The transmembrane domain region is highlighted in blue color and taken from Orientations of Proteins in Membranes (OPM) database (<http://opm.phar.umich.edu/>) [64]. The conserved histidine residues H86 and H428 in the ShuA rebuild structure are shown in red stick. The protein has 640 residues and contains 54.5 %, 3.1 % and 41.7 % of β -strands according to the DSSP program [79], (see also Table 1). Figures were produced with the PyMOL program [66].

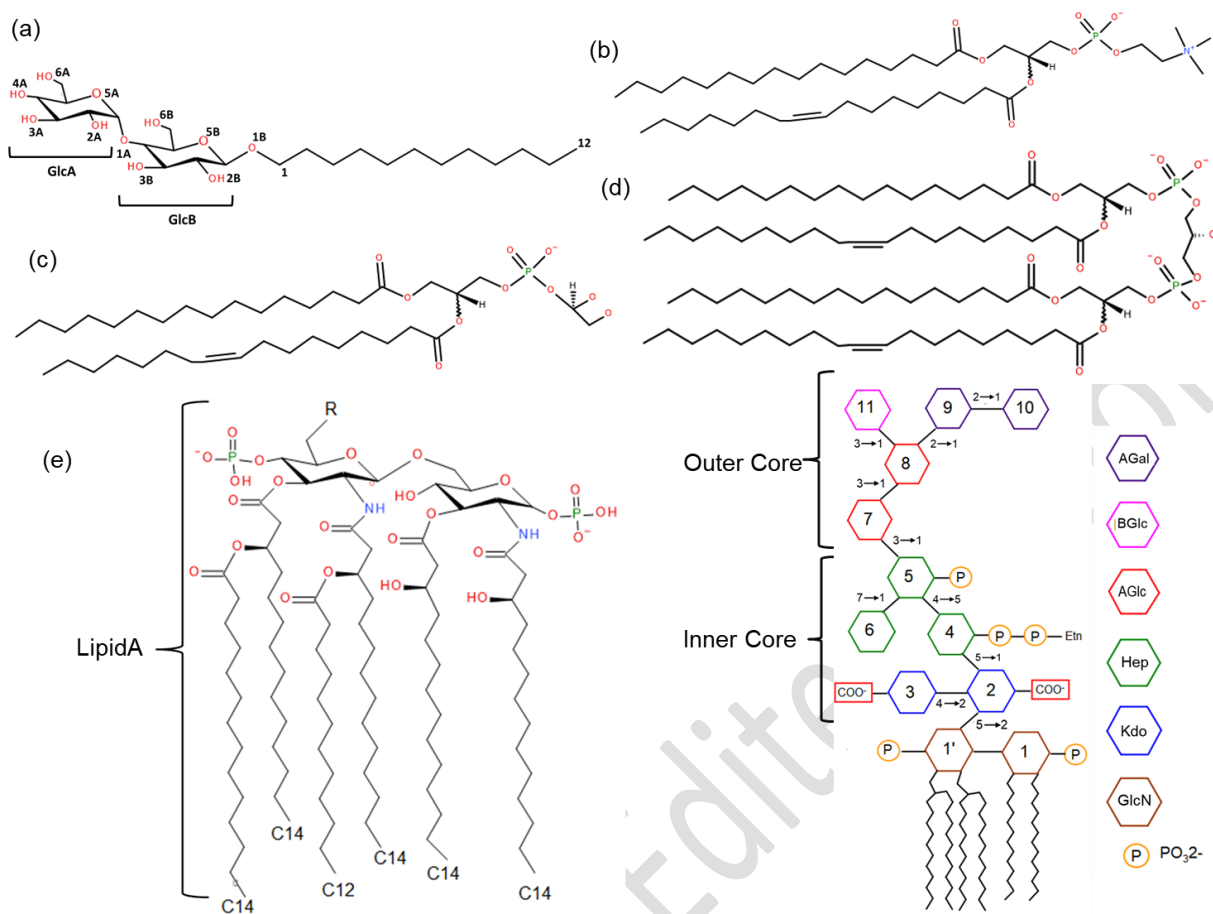


Figure 2: Chemical structures of the detergent, the bacterial lipids and the lipopolysaccharide (LPS) of *Shigella flexneri* strain used in this study. (a) n-Dodecyl-β-D-Maltopyranoside (DDM), (b) 1-palmitoyl(16:0)-2-palmitoleoyl(16:1cis-9)-phosphatidylethanolamine (PPPE), (c) 1-palmitoyl(16:0)-2-vacenoyleoyl(18:1cis-11)-phosphatidylglycerol (PVPG) (d) 1,1'-palmitoyl-2,2'-vacenoyleoyl cardiolipin (PVCL2) and (e) lipid A and the inner and core region denoted by colored hexagons. EtN, ethanolamine; Glc, glucose; Hep, heptose; KDO, 2-keto-3-deoxyoctulosonic acid; P, phosphate. The lipid A of *Shigella flexneri* consists of two d-glucosamine residues joined by a β-(1→6)-linkage, two monophosphoester groups at O1 and O4', and six amide/ester-linked fatty acids with different chain lengths. The inner R1 core has two Kdo residues and two Hep residues, one of which have a diester-phospho-ester group at their O4 position and bonded to an ethanolamine (EtN). The outer core consists of five hexopyranoses, three D-glucoses, and two D-galactoses, all of which are α-linked, except for the β-linkage between the two glucose 7 and 8 residues. Figure adapted from [47,80].

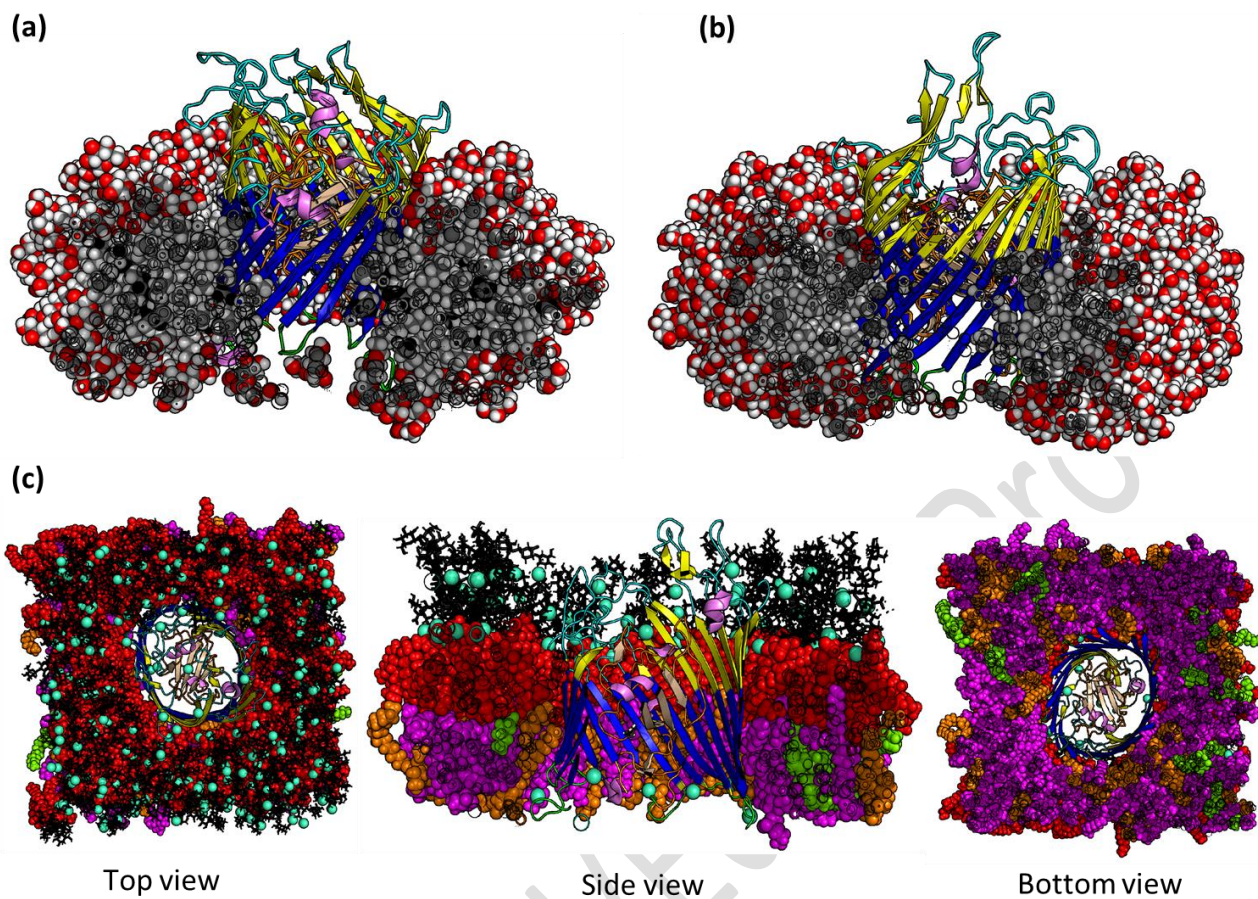


Figure 3: Representative snapshots of ShuA in the 268DDM (a), 294DDM (b) and embedded the outer membrane (c) with Lipid A, sugar core, PPPE, PVPG, and PVCL2 in red, black (sticks), magenta, orange, and lime green spheres, respectively. The different regions of the proteins were colored as in Figure 1. Calcium ions are green cyan sphere. Water molecules and KCl ions are omitted for clarity (see main text for details). Figures were produced with the PyMOL program [66].

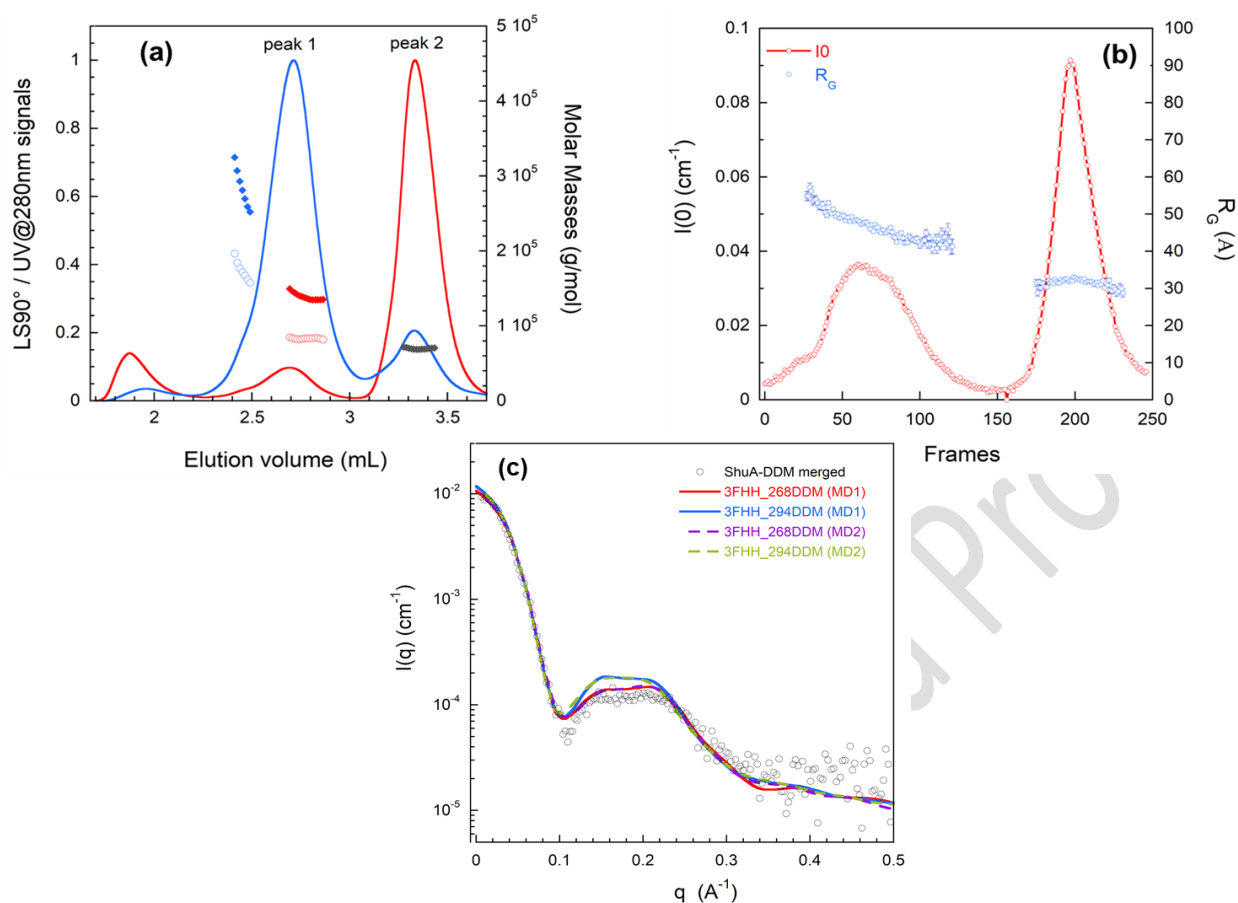


Figure 4: (a) SEC-MALLS elution profile of ShuA_DDM, 20 μL injected at 1.6 $\text{mg}\cdot\text{mL}^{-1}$ in 3mM DDM, eluted on a silica gel Yarra SEC-3000 at flowrate of 400 $\mu\text{L}\cdot\text{min}^{-1}$, with 50mM Tris pH 7.5, 0.3mM DDM: UV absorbance at 280 nm in blue (solid line), Rayleigh ratio in red (solid line), symbols (\circ and \blacklozenge) represent the calculated molar masses for the protein and the detergent respectively. (b) SEC-SAXS profile of ShuA_DDM, 5 μL injected at 5.5 $\text{mg}\cdot\text{mL}^{-1}$ in 10 mM DDM, eluted on a silica gel Yarra SEC-3000 at flowrate of 300 $\mu\text{L}\cdot\text{min}^{-1}$, with 50mM Tris pH 7.5, 0.1% DDM: forward intensity $I(0)$ (empty red dots), radius of gyration (empty blue dots). (c) The merged SAXS curve of ShuA_DDM complex is obtained from the 5 μL -sample at 5.5 mg/mL in the low q -range and the 50 μL -sample at 5.5 mg/mL in the large q -range. Comparison between the experimental SAXS (empty black dots) and the calculated SAXS curves for the four ShuA_DDM systems.

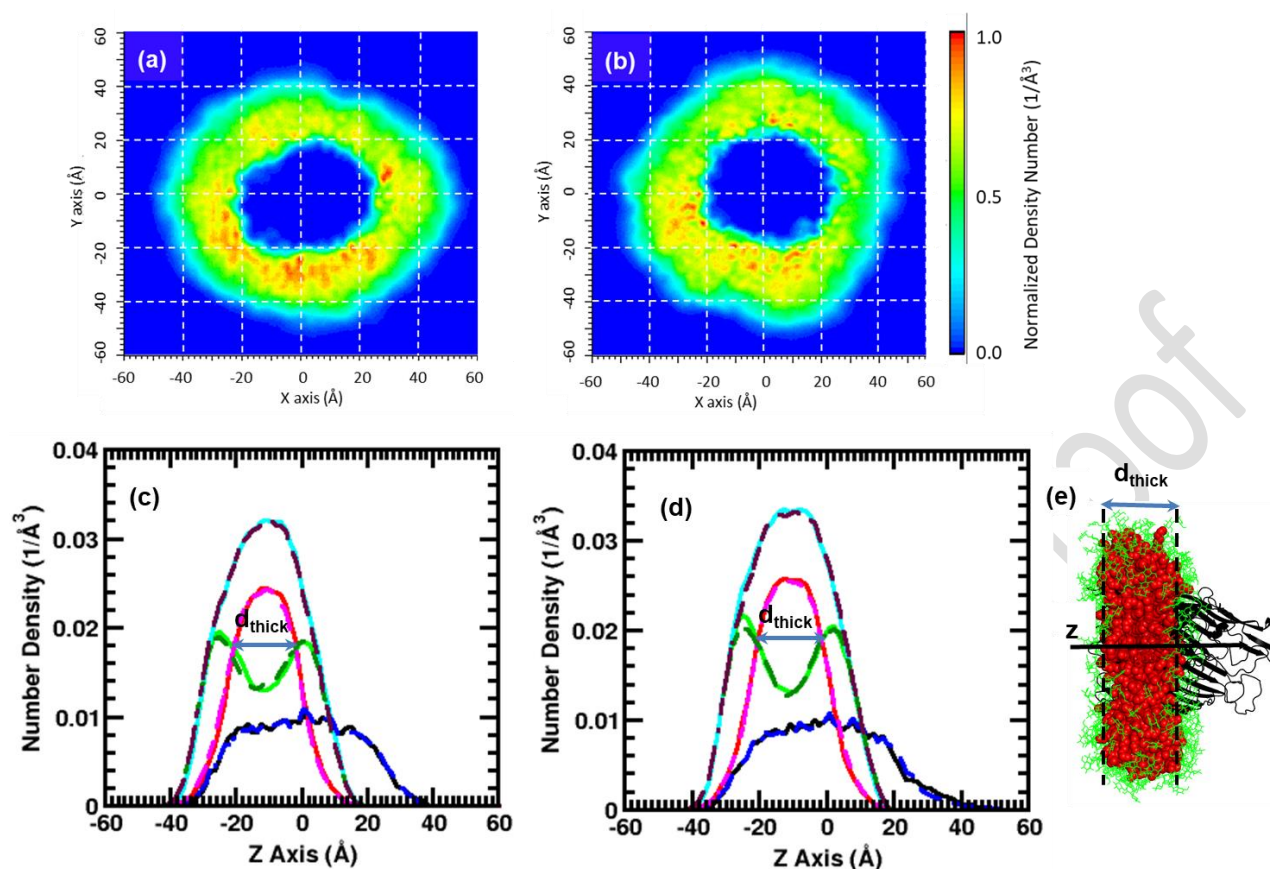


Figure 5: Averaged 1D density profiles (bottom) and 2D density maps (top) of the DDM detergent atoms. Here the ShuA_DDM complex was centered at $Z = 0 \text{ \AA}$. The Z-axis is the principal axis of the protein. A grid spacing of 0.5 \AA was used. In the 1D profiles, the protein, the alkyl chain, the headgroup and the whole DDM are shown in black, red, green, violet and blue, magenta, dark green, cyan lines for the first and second replicas, respectively. (a and c) 268 DDM, (b and d) 294 DDM. The 2D density maps of the second replicas are shown in Fig. S4 in the supporting information.

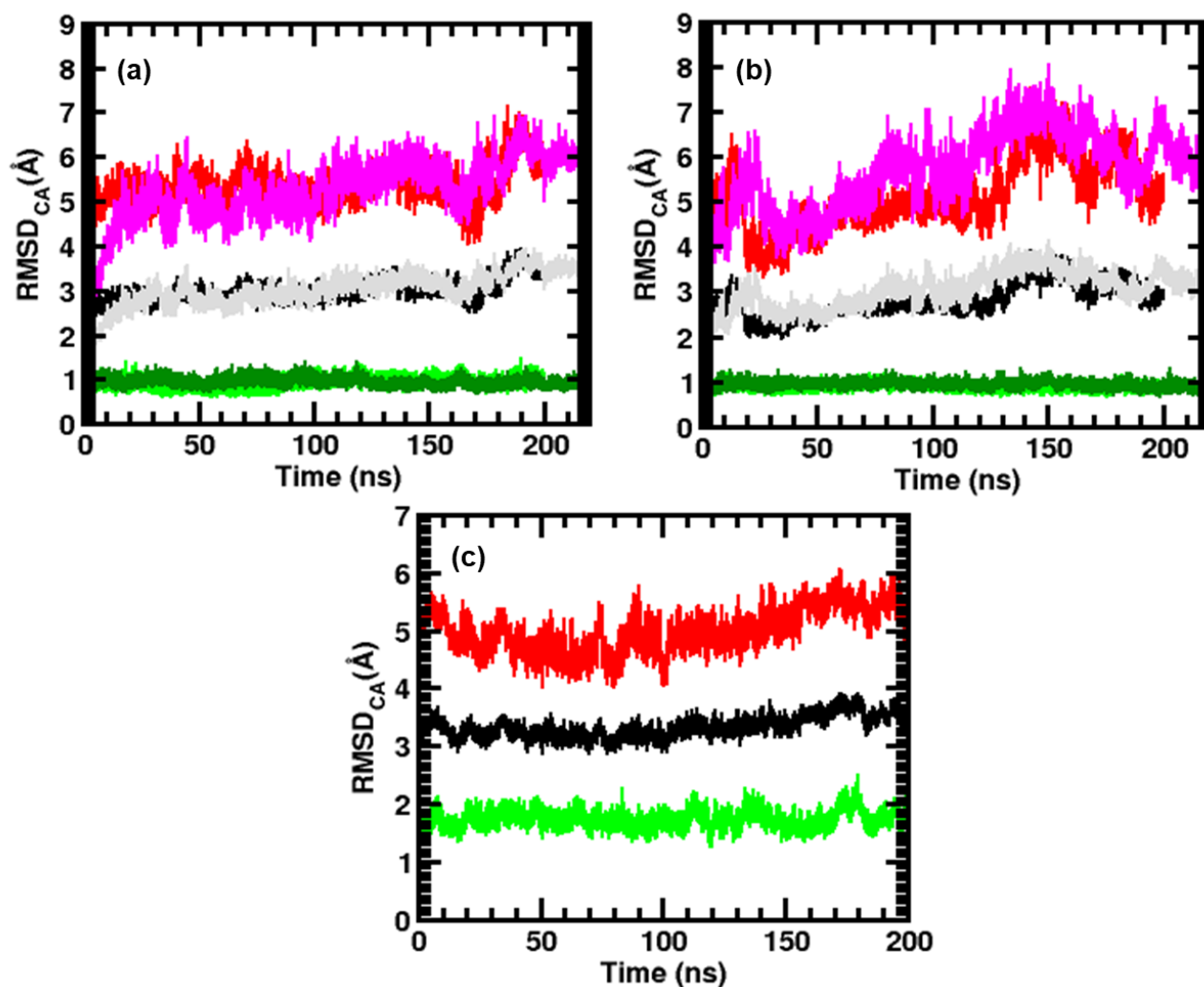


Figure 6: Root mean square deviations (RMSD) computed from the minimized X rays structure of ShuA for all C α (black and grey colors), extracellular loops (red and magenta colors) and TM barrels (green and dark green colors) in the micelle with 268 (a) and 294 (b) DDM molecules and their replicas and in the outer membrane (c).

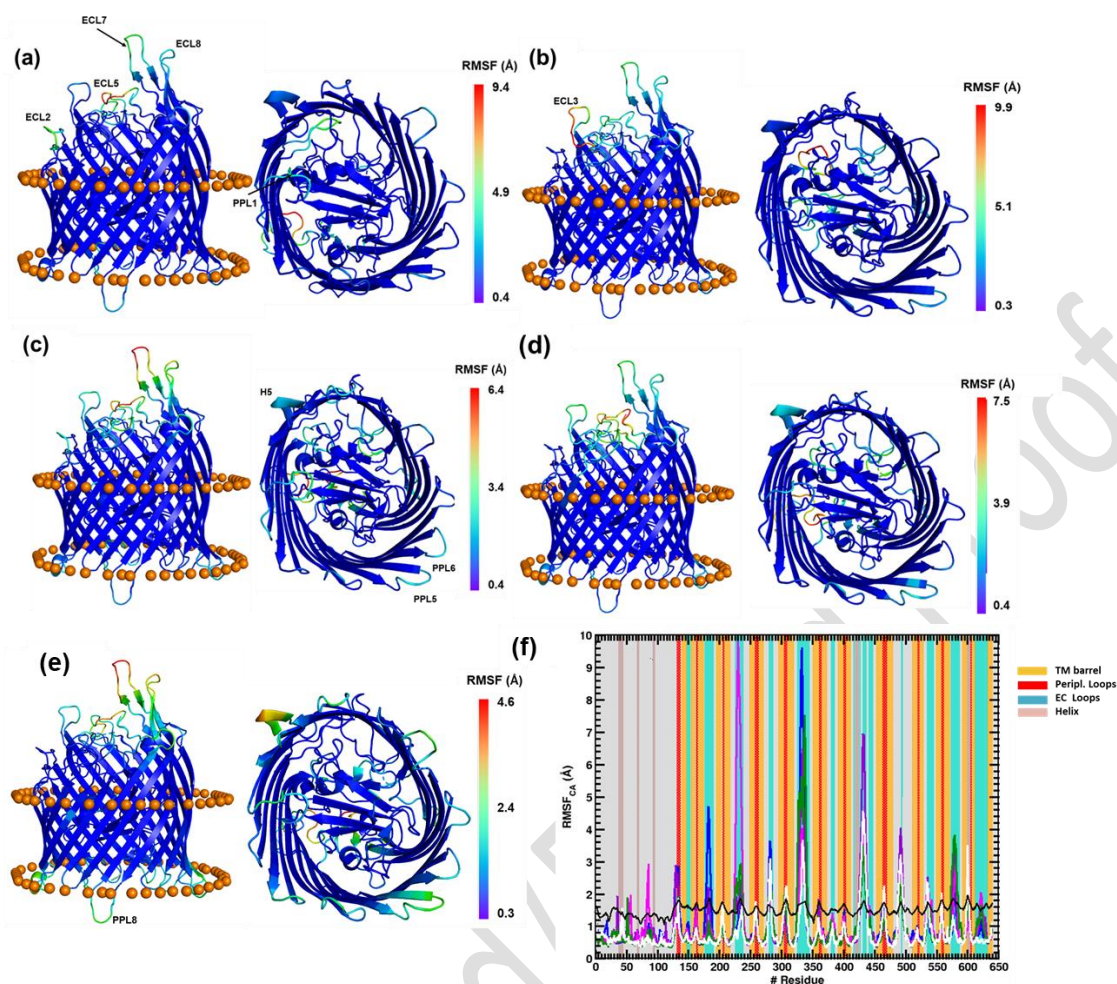


Figure 7: Colored 3D representation of ShuA (a and c) with 268, (b and d) 294 DDM and (e) in the outer membrane environments according to the $RMSF_{C\alpha}$. In the panel (d) $RMSF_{C\alpha}$ calculated from the MD simulations of ShuA with 268DDM (black), 294DDM (red) and in the bacterial membrane (blue) and from the experimental B-factor (green). For each picture, the minimum and maximum $RMSF_{C\alpha}$ values are indicated in the right of each color bars. The ECLX and PPLX labels indicate the extracellular and periplasmic loops with the highest $RMSF_{C\alpha}$. The orange spheres delimit the transmembrane domain according to OPM [64]. See main text for details. The $RMSF_{C\alpha}$ fluctuations of the i^{th} protein atom were estimated from the crystallographic Debye-Waller B-factors, (with $RMSF_{C\alpha} = \sqrt{(3B_i/8\pi^2)}$ where i^{th} protein atom) [73][74]. The orange, red, cyan and maroon vertical bars in RMSF figures delimit the TM barrels, periplasmic, helix, and extracellular loop regions according to the data listed in the Table 1. Figures were produced with PyMOL [66].

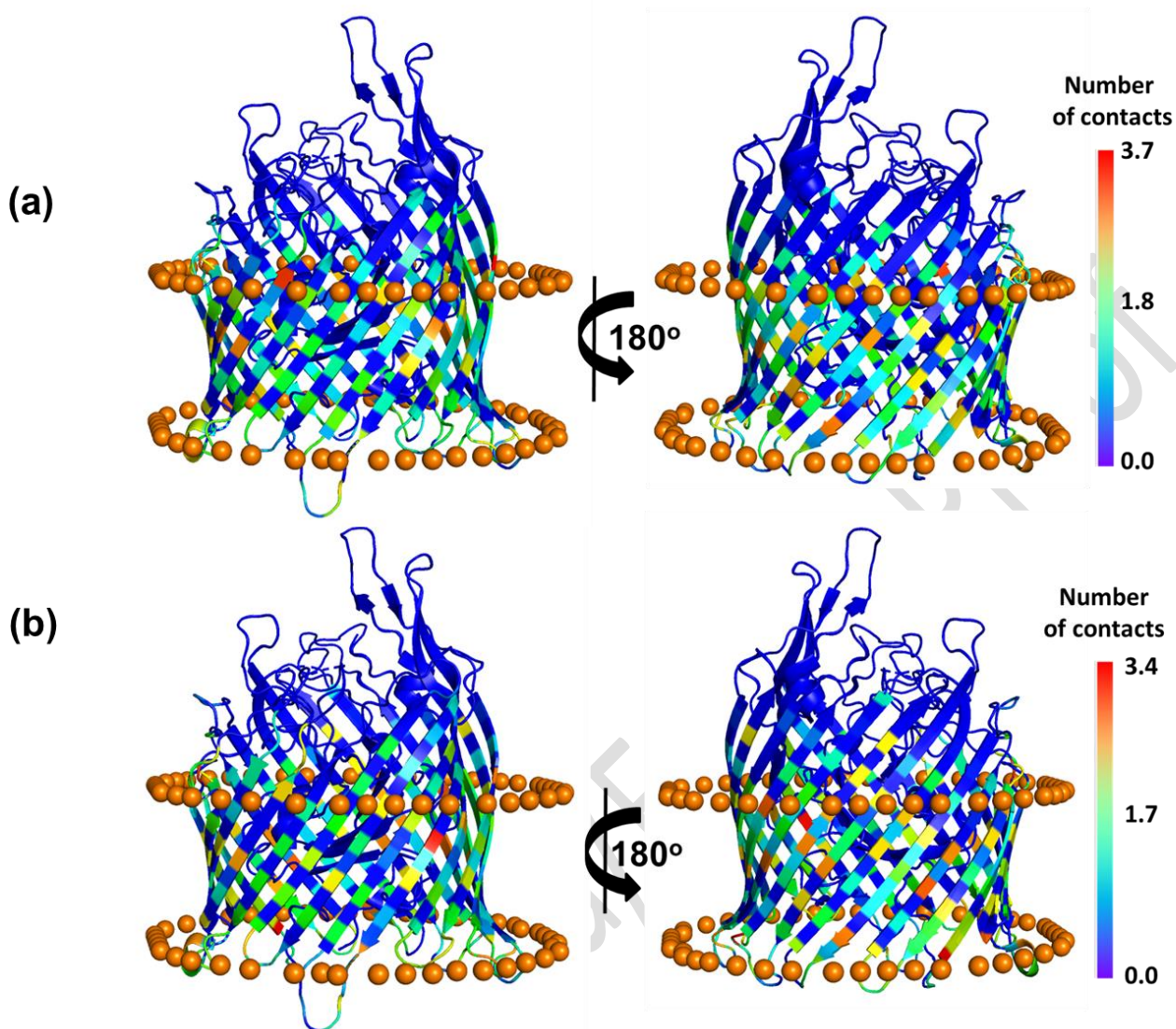


Figure 8: Colored 3D representation of ShuA in the DDM environment (first replica) with (a) 268 and (b) 294 molecules according to the number of contacts between the protein heavy atoms and the DDM heavy atoms. The contact analysis was carried out using a cutoff of 4.0 Å and each residue colored based on its respective number of contacts from low (blue) to high (red). The orange spheres delimit the transmembrane domain according to OPM webserver [64]. See main text for details. The corresponding figures for the second replica are shown in Fig. S9. Figures were produced with PyMOL [66].

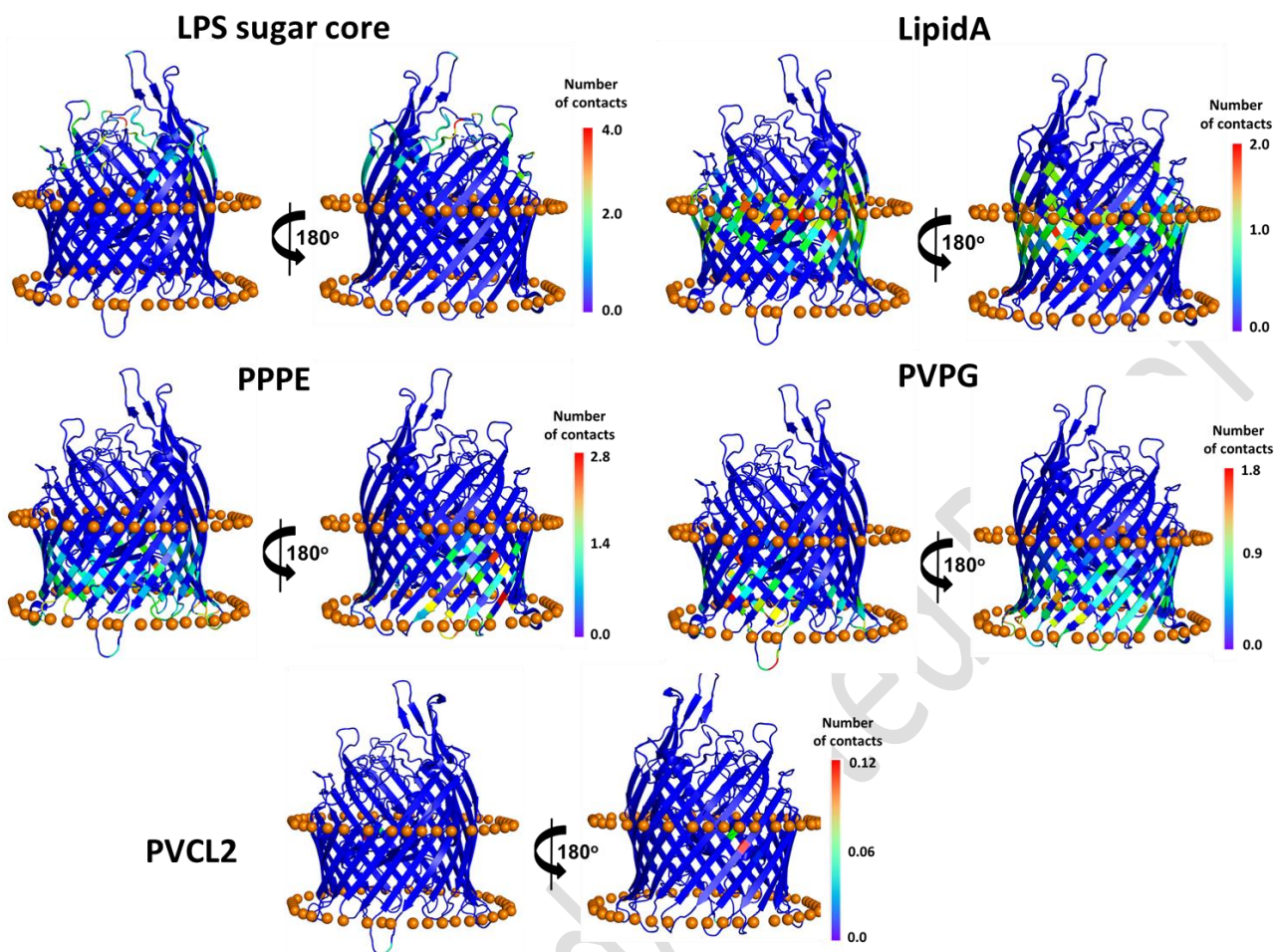


Figure 9: Colored 3D representation of ShuA in the outer membrane according to the number of contacts between the protein and the lipid heavy atoms of the bacterial membrane. The contact analysis was carried out using a cutoff of 4.0 Å and each residue colored based on its respective number of contacts from low (blue) to high (red). The orange spheres delimit the transmembrane domain according to OPM webserver [64]. See main text for details. Figures were produced with PyMOL [66].

List of Tables

| Secondary Structures | Number | Residue Positions |
|-----------------------------|-------------|--|
| Helix | 5 | (37 – 43), (67 – 69), (93 – 95) , (418 – 424) ⁺ , (599 – 603) |
| All β – Strands | 9+24 | (4 – 5), (13 – 14), (22 – 26), (48 – 50), (61 – 63), (71 – 75), (78,79), (97 – 104), (118 – 124) , (137 – 146), (151 – 161), (165 – 176), (190 – 203), (208 – 224), (239 – 255), (265 – 278), (287 – 303), (309 – 325), (344 – 359), (364 – 379), (385 – 398), (404 – 415), (427 – 429) ⁺ , (437 – 439) ⁺ , (448 – 461), (470 – 491), (496 – 517), (522 – 534), (545 – 555), (562 – 571), (587 – 597), (605 – 611), (632 – 640) |
| Transmembrane Domain Region | 22 | (138 – 146), (151 – 160), (165 – 174), (194 – 202), (208 – 216), (248 – 256), (263 – 271), (294 – 303), (309 – 319), (350 – 359), (364 – 372), (390 – 399), (404 – 411), (452 – 461), (470 – 478), (510 – 518), (521 – 528), (549 – 556), (562 – 570), (590 – 598), (605 – 612), (632 – 639) |
| Internal Loop (BtuB domain) | 10 | (2 – 3), (6 – 12), (15 – 21), (27 – 36), (44 – 47), (51 – 60), (64 – 66), (80 – 92), (105 – 117), (125 – 130) |
| Extracellular Loops (ECL) | 11 | (147 – 152), (176 – 189), (225 – 238), (279 – 286), (325 – 344), (379 – 384), (416 – 417)*, (425 – 426)*, (430 – 436)*, (440 – 447)*, (492 – 495), (535 – 545), (572 – 586), (612 – 631) |
| Periplasmic loops (PPL) | 11 | (131 – 136), (162 – 164), (204 – 207), (257 – 262), (304 – 308), (360 – 363), (400 – 403), (462 – 469), (519 – 520), (557 – 561), (604 – 605) |

Table 1: List of the residues involved in the secondary structure domains of ShuA (PDB ID: 3FHH). The 3 helix and 9 β -strands in bold are considered to belong to the TonB plug domain. The transmembrane domain region lists 22 transmembrane β -Strands in contact with the membrane according to OPM webserver [64] and that belongs to the 33 β -Strands of the protein. The loops in italic contains the missing residues and atoms missing the protein crystal due to their flexibilities and those highlighted with an asterisk belong to 7th extracellular loop and are separated with 2 β -strands and one helix highlighted with an upper cross.

| System | Det/Lipids | N _{det} /N _{lipids} | N _{Not} | N _{water} | N _{ions} | L _{box} | t _{sim} |
|-------------|------------------------------|---------------------------------------|------------------|--------------------|--|-----------------------|------------------|
| ShuA_268DDM | DDM | 268 | 301398 | 90000 | 18 Na ⁺ | 143.0 ³ | 200/220 |
| ShuA_294DDM | | 294 | 303504 | | | | 200/220 |
| ShuA_OM | LPSA/PPPE/ PVPG/PVCL 2 | 48/ 105/ 28/7 | 168147 | 37970 | 216Ca ²⁺ , 163K ⁺ , 103Cl ⁻ | 106 ² x160 | 200 |

Table 2: A summary of the simulated system. N_{tot} , N_{ions} , L_{box} and t_{sim} are the total number of atoms in the system, numbers of ions box lengths (Å) and the simulation time (ns), respectively.

| System | ShuA_268DDM | ShuA_294DDM |
|------------------|-------------|-------------|
| X_{PDC} | 47/45* | 45/40* |
| Y_{PDC} | 43/42* | 46/47* |
| Z_{PDC} | 39/39* | 39/39* |
| X_{P} | 20/20* | 20/20* |
| Y_{P} | 20/20* | 20/20* |
| Z_{P} | 35/35* | 35/35* |
| X_{DDM} | 27/25* | 25/20* |
| Y_{DDM} | 23/22* | 26/27* |
| Z_{DDM} | 16/16* | 16/16* |

Table 3: Average length (in Å) of the ShuA_DDM complex systems (PDC), the protein alone (P) and the surfactant corona (DDM) for the first and the second replicas (*) estimated from the 1D and 2D density distributions. See main text for details.

Corresponding Authors

§ Stéphane Abel (stephane.abel@cea.fr) & Françoise Bonneté (bonnete@ibpc.fr)

Address where the experimental work was done

†Institut des Biomolécules Max Mousseron (IBMM) UMR 5247 CNRS-UM-ENSCM, Chimie Bio-organique et Systèmes Amphiphiles, Université d'Avignon, 301 rue Baruch de Spinoza, F84000 Avignon, France

Author Contributions

K.B. purified and supplied the protein for experiments; J.S., S.F and F.B performed the experiments, analyzed the data; S.A and M.M performed simulations; S.A, S.F, M.M and F.B conceived, designed the project and wrote the manuscript.

Funding Sources

This project was supported by the ANR programs « Investissements d'Avenir ANR-10-LABX-05-01 » and “Laboratoires communs organismes de recherche publics – PME/ETI” ANR-14-LAB7-0002

CHEM2STAB. This work was also granted for access to the HPC resources of [CINES/IDRIS/TGCC] under the allocation 2018-A0040710413 made by GENCI. We also acknowledge the CCRT for accessing to the supercomputer Cobalt where some analyses were performed.

Supporting Information

Additional results, tables and figures

Acknowledgements: We acknowledge the French Synchrotron Radiation Facility (SOLEIL) for the provision of synchrotron radiation facilities and Javier Perez on SWING (SOLEIL) for assistance in using the beamline. We also acknowledge Jean-Jacques Lacapère (LBC Paris) and Sophie Combet-Jeancel (LLB Saclay, France) for fruitful discussions.

Uncorrected/Edited Proof

References

- [1] C.G. Tate, Practical considerations of membrane protein instability during purification and crystallisation, in: *Methods Mol. Biol.*, 2010: pp. 187–203. doi:10.1007/978-1-60761-344-2_12.
- [2] A. Stetsenko, A. Guskov, An overview of the top ten detergents used for membrane protein crystallization, *Crystals*. 7 (2017) 197. doi:10.3390/cryst7070197.
- [3] R.K. Le, B.J. Harris, I.J. Iwuchukwu, B.D. Bruce, X. Cheng, S. Qian, W.T. Heller, H. O'Neill, P.D. Frymier, H. O'Neill, P.D. Frymier, Analysis of the solution structure of *Thermosynechococcus elongatus* photosystem I in n-dodecyl- β -d-maltoside using small-angle neutron scattering and molecular dynamics simulation, *Arch. Biochem. Biophys.* 550–551 (2014) 50–57. doi:10.1016/j.abb.2014.04.005.
- [4] A. Calcutta, C.M. Jessen, M.A. Behrens, C.L.P. Oliveira, M.L. Renart, J.M. González-Ros, D.E. Otzen, J.S. Pedersen, A. Malmendal, N.C. Nielsen, Mapping of unfolding states of integral helical membrane proteins by GPS-NMR and scattering techniques: TFE-induced unfolding of KcsA in DDM surfactant, *Biochim. Biophys. Acta - Biomembr.* 1818 (2012) 2290–2301. doi:10.1016/j.bbamem.2012.04.005.
- [5] G. Khelashvili, M. V. LeVine, L. Shi, M. Quick, J.A. Javitch, H. Weinstein, The membrane protein LeuT in micellar systems: aggregation dynamics and detergent binding to the S2 site, *J. Am. Chem. Soc.* 135 (2013) 14266–14275. doi:10.1021/ja405984v.
- [6] S.L. Rouse, M.S.P. Sansom, Interactions of lipids and detergents with a viral ion channel protein: molecular dynamics simulation studies, *J. Phys. Chem. B.* 119 (2015) 764–772. doi:10.1021/jp505127y.
- [7] I. Josts, J. Nitsche, S. Maric, H.D. Mertens, M. Moulin, M. Haertlein, S. Prevost, D.I. Svergun, S. Busch, V.T. Forsyth, H. Tidow, Conformational States of ABC Transporter MsbA in a Lipid Environment Investigated by Small-Angle Scattering Using Stealth Carrier Nanodiscs, *Structure*. 26 (2018) 1072-1079.e4. doi:10.1016/j.str.2018.05.007.
- [8] T. Pieńko, J. Trylska, Extracellular loops of BtuB facilitate transport of vitamin B12 through the outer membrane of *E. coli*, *PLOS Comput. Biol.* 16 (2020) e1008024. doi:10.1371/journal.pcbi.1008024.
- [9] A. Berthaud, J. Manzi, J. Pérez, S. Mangenot, Modeling Detergent Organization around Aquaporin-0 Using Small-Angle X-ray Scattering, *J. Am. Chem. Soc.* 134 (2012) 10080–10088. doi:10.1021/ja301667n.
- [10] A. Koutsioubas, A. Berthaud, S. Mangenot, J. Pérez, Ab Initio and All-Atom Modeling of

- Detergent Organization around Aquaporin-0 Based on SAXS Data, *J. Phys. Chem. B.* 117 (2013) 13588–13594. doi:10.1021/jp407688x.
- [11] J. Døvling Kaspersen, C. Moestrup Jessen, B. Stougaard Vad, E. Skipper Sørensen, K. Kleiner Andersen, M. Glasius, C.L. Pinto Oliveira, D.E. Otzen, J.S. Pedersen, Low-Resolution Structures of OmpA·DDM Protein-Detergent Complexes, *ChemBioChem.* 15 (2014) 2113–2124. doi:10.1002/cbic.201402162.
- [12] J.S. Hub, Interpreting solution X-ray scattering data using molecular simulations, *Curr. Opin. Struct. Biol.* 49 (2018) 18–26. doi:10.1016/j.sbi.2017.11.002.
- [13] L. Boldon, F. Laliberte, L. Liu, Review of the fundamental theories behind small angle X-ray scattering, molecular dynamics simulations, and relevant integrated application, *Nano Rev.* 6 (2015) 25661–25681. doi:10.3402/nano.v6.25661.
- [14] P.C. Chen, J.S. Hub, Structural Properties of Protein-Detergent Complexes from SAXS and MD Simulations, *J. Phys. Chem. Lett.* 6 (2015) 5116–5121. doi:10.1021/acs.jpcclett.5b02399.
- [15] S.K. Niyogi, Shigellosis., *J. Microbiol.* 43 (2005) 133–143.
<http://www.ncbi.nlm.nih.gov/pubmed/15880088>.
- [16] M. Drancourt, Acute Diarrhea, in: *Infect. Dis. (Auckl).*, Elsevier, 2017: pp. 335-340.e2. doi:10.1016/B978-0-7020-6285-8.00038-1.
- [17] S.K. Bhattacharya, D. Sur, Shigellosis, in: *Int. Encycl. Public Heal.*, Elsevier, 2017: pp. 499–501. doi:10.1016/B978-0-12-803678-5.00406-9.
- [18] A. Mills, H.-T. Le, F. Duong, TonB-dependent ligand trapping in the BtuB transporter, *Biochim. Biophys. Acta - Biomembr.* 1858 (2016) 3105–3112. doi:10.1016/j.bbamem.2016.09.019.
- [19] J.M. LaLonde, D.A. Bernlohr, L.J. Banaszak, The up-and-down beta-barrel proteins., *FASEB J.* 8 (1994) 1240–1247. doi:10.1096/fasebj.8.15.8001736.
- [20] V. Braun, K. Hantke, Recent insights into iron import by bacteria, *Curr. Opin. Chem. Biol.* 15 (2011) 328–334. doi:10.1016/j.cbpa.2011.01.005.
- [21] D. Cobessi, A. Meksem, K. Brillet, Structure of the heme/hemoglobin outer membrane receptor ShuA from *Shigella dysenteriae* : Heme binding by an induced fit mechanism, *Proteins Struct. Funct. Bioinforma.* 78 (2010) 286–294. doi:10.1002/prot.22539.
- [22] L.M. Meneghini, S. Tripathi, M.A. Woodworth, S. Majumdar, T.L. Poulos, G.A. Weiss, Dissecting binding of a β -barrel membrane protein by phage display, *Mol. Biosyst.* 13 (2017) 1438–1447. doi:10.1039/C7MB00163K.
- [23] K.A. Burkhard, A. Wilks, Characterization of the Outer Membrane Receptor ShuA from the Heme Uptake System of *Shigella dysenteriae*: Substrate specificity and identification of the

- heme protein ligands, *J. Biol. Chem.* 282 (2007) 15126–15136. doi:10.1074/jbc.M611121200.
- [24] K. Brillet, A. Meksem, A. Thompson, D. Cobessi, Expression, purification, crystallization and preliminary X-ray diffraction analysis of the TonB-dependent haem outer membrane transporter ShuA from *Shigella dysenteriae*, *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65 (2009) 402–405. doi:10.1107/S1744309109008148.
- [25] P. Strop, A.T. Brunger, Refractive index-based determination of detergent concentration and its application to the study of membrane proteins, *Protein Sci.* 14 (2005) 2207–2211. doi:10.1110/ps.051543805.
- [26] K. Gimpl, J. Klement, S. Keller, Characterising protein/detergent complexes by triple-detection size-exclusion chromatography, *Biol. Proced. Online.* 18 (2016) 4. doi:10.1186/s12575-015-0031-9.
- [27] R. Girardot, FOXTROT: A JAVA-Based Application to Reduce and Analyse SAXS and WAXS Piles of 2D Data at Synchrotron Soleil, (2015).
http://www.cansas.org/wgwiki/images/f/f3/Abstract_JavierPerez.pdf.
- [28] T.D. Goddard, C.C. Huang, E.C. Meng, E.F. Pettersen, G.S. Couch, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis, *Protein Sci.* 27 (2018) 14–25. doi:10.1002/pro.3235.
- [29] G.G. Krivov, M. V. Shapovalov, R.L. Dunbrack, Improved prediction of protein side-chain conformations with SCWRL4, *Proteins Struct. Funct. Bioinforma.* 77 (2009) 778–795. doi:10.1002/prot.22488.
- [30] J.B. Klauda, R.M. Venable, J.A. Freites, J.W. O'Connor, D.J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A.D. MacKerell, R.W. Pastor, Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types, *J. Phys. Chem. B.* 114 (2010) 7830–7843. doi:10.1021/jp101759q.
- [31] S. Abel, F.-Y.Y. Dupradeau, E.P. Raman, A.D. MacKerell, M. Marchi, Molecular simulations of dodecyl- β -maltoside micelles in water: influence of the headgroup conformation and force field parameters, *J. Phys. Chem. B.* 115 (2011) 487–499. doi:10.1021/jp109545v.
- [32] O. Guvench, S.N. Greenr, G. Kamath, J.W. Brady, R.M. Venable, R.W. Pastor, A.D. Mackerell, S.N. Greene, G. Kamath, J.W. Brady, R.M. Venable, R.W. Pastor, A.D. Mackerell, Additive empirical force field for hexopyranose monosaccharides, *J. Comput. Chem.* 29 (2008) 2543–2564. doi:10.1002/jcc.21004.
- [33] E. Hatcher, O. Guvench, A.D. MacKerell, CHARMM Additive All-Atom Force Field for Aldopentofuranoses, Methyl-aldopentofuranosides, and Fructofuranose, *J. Phys. Chem. B.* 113

- (2009) 12466–12476. doi:10.1021/jp905496e.
- [34] X. Cheng, S. Jo, H.S. Lee, J.B. Klauda, W. Im, CHARMM-GUI Micelle Builder for Pure/Mixed Micelle and Protein/Micelle Complex Systems, *J. Chem. Inf. Model.* 53 (2013) 2171–2180. doi:10.1021/ci4002684.
- [35] S. Jo, X. Cheng, J. Lee, S. Kim, S.J. Park, D.S. Patel, A.H. Beaven, K. Il Lee, H. Rui, S. Park, H.S. Lee, B. Roux, A.D. MacKerell, J.B. Klauda, Y. Qi, W. Im, CHARMM-GUI 10 years for biomolecular modeling and simulation, *J. Comput. Chem.* 38 (2017) 1114–1124. doi:10.1002/jcc.24660.
- [36] E.L. Wu, P.J. Fleming, M.S. Yeom, G. Widmalm, J.B. Klauda, K.G. Fleming, W. Im, E. coli Outer Membrane and Interactions with OmpLA, *Biophys. J.* 106 (2014) 2493–2502. doi:10.1016/j.bpj.2014.04.024.
- [37] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* 79 (1983) 926–935. doi:10.1063/1.445869.
- [38] F.J.S. Pontes, V.H. Rusu, T.A. Soares, R.D. Lins, The Effect of Temperature, Cations, and Number of Acyl Chains on the Lamellar to Non-Lamellar Transition in Lipid-A Membranes: A Microscopic View, *J. Chem. Theory Comput.* 8 (2012) 3830–3838. doi:10.1021/ct300084v.
- [39] J.E. Vance, D. Vance, *Biochemistry of Lipids, Lipoproteins and Membranes*, Elsevier Science B.V., 2002.
- [40] A. Nascimento, F.J.S. Pontes, R.D. Lins, T.A. Soares, Hydration, ionic valence and cross-linking propensities of cations determine the stability of lipopolysaccharide (LPS) membranes, *Chem. Commun.* 50 (2014) 231–233. doi:10.1039/C3CC46918B.
- [41] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, E. Lindahl, E. Lindahl, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX.* 1–2 (2015) 19–25. doi:10.1016/j.softx.2015.06.001.
- [42] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.* 126 (2007) 014101. doi:10.1063/1.2408420.
- [43] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.* 52 (1981) 7182–7190. doi:10.1063/1.328693.
- [44] A. Rahman, F.H. Stillinger, A. Rahman, Molecular Dynamics Study of Liquid Water, *J. Chem. Phys.* 55 (1971) 3336–3359. doi:10.1063/1.1676585.
- [45] W.G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, *Phys. Rev. A.* 31

- (1985) 1695–1697. doi:10.1103/PhysRevA.31.1695.
- [46] S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, *Mol. Phys.* 52 (1984) 255–268. doi:10.1080/00268978400101201.
- [47] E.L. Wu, O. Engström, S. Jo, D. Stuhlsatz, M.S. Yeom, J.B. Klauda, G. Widmalm, W. Im, Molecular Dynamics and NMR Spectroscopy Studies of *E. coli* Lipopolysaccharide Structure and Dynamics, *Biophys. J.* 105 (2013) 1444–1455. doi:10.1016/j.bpj.2013.08.002.
- [48] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.* 103 (1995) 8577–8593. doi:10.1063/1.470117.
- [49] P.J. Steinbach, B.R. Brooks, New spherical-cutoff methods for long-range forces in macromolecular simulation, *J. Comput. Chem.* 15 (1994) 667–683. doi:10.1002/jcc.540150702.
- [50] B. Hess, P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation, *J. Chem. Theory Comput.* 4 (2008) 116–122. doi:10.1021/ct700200b.
- [51] C. Dupuy, X. Auvray, C. Petipas, I. Rico-Lattes, A. Lattes, Anomeric Effects on the Structure of Micelles of Alkyl Maltosides in Water, *Langmuir.* 13 (1997) 3965–3967. doi:10.1021/la9604285.
- [52] M. Aoudia, R. Zana, Aggregation Behavior of Sugar Surfactants in Aqueous Solutions: Effects of Temperature and the Addition of Nonionic Polymers, *J. Colloid Interface Sci.* 206 (1998) 158–167. doi:10.1006/jcis.1998.5627.
- [53] D.J. Slotboom, R.H. Duurkens, K. Olieman, G.B. Erkens, Static light scattering to characterize membrane proteins in detergent solution, *Methods.* 46 (2008) 73–82. doi:10.1016/j.ymeth.2008.06.012.
- [54] D.I. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, *J. Appl. Crystallogr.* 25 (1992) 495–503. doi:10.1107/S0021889892001663.
- [55] P.C. Chen, J.S. Hub, Interpretation of solution x-ray scattering by explicit-solvent molecular dynamics, *Biophys. J.* 108 (2015) 2573–2584. doi:10.1016/j.bpj.2015.03.062.
- [56] M. Marchi, A first principle particle mesh method for solution SAXS of large bio-molecular systems, *J. Chem. Phys.* 145 (2016) 045101. doi:10.1063/1.4959123.
- [57] G. Voronoi, Nouvelles applications des paramètres continus à théorie des formes quadratiques. Deuxième Mémoire. Recherches sur les paralléloèdres primitifs, *J. Fur Die Reine Und Angew. Math.* (1909) 198–287. doi:10.1515/crll.1909.136.67.
- [58] S. Abel, F.-Y. Dupradeau, M. Marchi, Molecular Dynamics Simulations of a Characteristic DPC Micelle in Water, *J. Chem. Theory Comput.* 8 (2012) 4610–4623. doi:10.1021/ct3003207.
- [59] C.H. Rycroft, VORO++: A three-dimensional Voronoi cell library in C++, *Chaos An*

- Interdiscip. J. Nonlinear Sci. 19 (2009) 041111. doi:10.1063/1.3215722.
- [60] M. Marchi, openTRJ, (2019) <https://github.com/octupole/openTRJ>.
<https://github.com/octupole/openTRJ>.
- [61] A. Grishaev, L. Guo, T. Irving, A. Bax, Improved Fitting of Solution X-ray Scattering Data to Macromolecular Structures and Structural Ensembles by Explicit Water Modeling, *J. Am. Chem. Soc.* 132 (2010) 15484–15486. doi:10.1021/ja106173n.
- [62] J. Köfinger, G. Hummer, Atomic-resolution structural information from scattering experiments on macromolecules in solution, *Phys. Rev. E.* 87 (2013) 052712.
doi:10.1103/PhysRevE.87.052712.
- [63] L.-A. Barret, C. Barrot-Ivolot, S. Raynal, C. Jungas, A. Polidori, F. Bonneté, Influence of Hydrophobic Micelle Structure on Crystallization of the Photosynthetic RC-LH1-PufX Complex from *Rhodobacter blasticus*, *J. Phys. Chem. B.* 117 (2013) 8770–8781. doi:10.1021/jp403483q.
- [64] M.A. Lomize, I.D. Pogozheva, H. Joo, H.I. Mosberg, A.L. Lomize, OPM database and PPM web server: resources for positioning of proteins in membranes, *Nucleic Acids Res.* 40 (2012) D370–D376. doi:10.1093/nar/gkr703.
- [65] A.L. Lomize, I.D. Pogozheva, H.I. Mosberg, Anisotropic Solvent Model of the Lipid Bilayer. 2. Energetics of Insertion of Small Molecules, Peptides, and Proteins in Membranes, *J. Chem. Inf. Model.* 51 (2011) 930–946. doi:10.1021/ci200020k.
- [66] W.L. DeLano, The PyMOL Molecular Graphics System, Version 2.3, (2020).
- [67] J. Lipfert, L. Columbus, V.B. Chu, S.A. Lesley, S. Doniach, Size and Shape of Detergent Micelles Determined by Small-Angle X-ray Scattering, *J. Phys. Chem. B.* 111 (2007) 12427–12438. doi:10.1021/jp073016l.
- [68] G.G. Warr, C.J. Drummond, F. Grieser, B.W. Ninham, D.F. Evans, Aqueous solution properties of nonionic n-dodecyl-beta-D-maltoside micelles, *J. Phys. Chem.* 90 (1986) 4581–4586.
doi:10.1021/j100410a022.
- [69] S. Manet, A.S. Cuvier, C. Valotteau, G.C. Fadda, J. Perez, E. Karakas, S. Abel, N. Baccile, structure of bolaamphiphile Sophorolipid micelles characterized with SAXS, SANS, and MD simulations, *J. Phys. Chem. B.* 119 (2015) 13113–13133. doi:10.1021/acs.jpcc.5b05374.
- [70] C. Tanford, Micelle shape and size, *J. Phys. Chem.* 76 (1972) 3020–3024.
doi:10.1021/j100665a018.
- [71] T.P. Straatsma, T.A. Soares, Characterization of the outer membrane protein OprF of *Pseudomonas aeruginosa* in a lipopolysaccharide membrane by computer simulation, *Proteins Struct. Funct. Bioinforma.* 74 (2009) 475–488. doi:10.1002/prot.22165.

- [72] D. Frishman, P. Argos, Knowledge-based protein secondary structure assignment., *Proteins*. 23 (1995) 566–579.
- [73] D.A. O'Connor, *Thermal Vibrations in Crystallography*, Cambridge, London, 1975.
doi:10.1088/0031-9112/26/11/033.
- [74] A. Kuzmanic, B. Zagrovic, Determination of Ensemble-Average Pairwise Root Mean-Square Deviation from Experimental B-Factors, *Biophys. J.* 98 (2010) 861–871.
doi:10.1016/j.bpj.2009.11.011.
- [75] P.J. Fleming, D.S. Patel, E.L. Wu, Y. Qi, M.S. Yeom, M.C. Sousa, K.G. Fleming, W. Im, BamA POTRA Domain Interacts with a Native Lipid Membrane Surface, *Biophys. J.* 110 (2016) 2698–2709. doi:10.1016/j.bpj.2016.05.010.
- [76] J. Lee, D.S. Patel, I. Kucharska, L.K. Tamm, W. Im, Refinement of OprH-LPS Interactions by Molecular Simulations, *Biophys. J.* 112 (2017) 346–355. doi:10.1016/j.bpj.2016.12.006.
- [77] J.M.R. Albano, N. Mussini, R. Toriano, J.C. Facelli, M.B. Ferraro, M. Pickholz, Calcium interactions with Cx26 hemmichannel: Spatial association between MD simulations biding sites and variant pathogenicity, *Comput. Biol. Chem.* 77 (2018) 331–342.
doi:10.1016/j.combiolchem.2018.11.004.
- [78] S. Kim, D.S. Patel, S. Park, J. Slusky, J.B. Klauda, G. Widmalm, W. Im, Bilayer Properties of Lipid A from Various Gram-Negative Bacteria, *Biophys. J.* 111 (2016) 1750–1760.
doi:10.1016/j.bpj.2016.09.001.
- [79] W. Kabsch, C. Sander, Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers*. 22 (1983) 2577–2637.
doi:10.1002/bip.360221211.
- [80] A. Rice, J. Wereszczynski, Atomistic Scale Effects of Lipopolysaccharide Modifications on Bacterial Outer Membrane Defenses, *Biophys. J.* 114 (2018) 1389–1399.
doi:10.1016/j.bpj.2018.02.006.

Supporting Information

Structural insights into the membrane receptor ShuA in DDM micelles and in a model of gram-negative bacteria outer membrane as seen by SAXS and MD simulations

Stéphane Abel^{a*}, *Massimo Marchi*^a, *Justine Solier*^b, *Stéphanie Finet*^c, *Karl Brillet*^d and *Françoise Bonneté*^{e*}

^a Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

^b Laboratoire d'Electrochimie et de Physico-chimie des Matériaux et des Interfaces, UMR 5279 CNRS Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, INP, F38000 Grenoble, France

^c Institut de Minéralogie, de Physique de Matériaux et de Cosmochimie, UMR 7590 CNRS-Sorbonne université, Bioinformatique et Biophysique, 4 Place Jussieu, F75005 Paris, France

^d Institut de Biologie Moléculaire et Cellulaire UPR 9002 CNRS, Architecture et Réactivité de l'ARN, 2 allée Konrad Roentgen, F67000 Strasbourg, France

^e Institut de Biologie Physico-Chimique (IBPC) UMR 7099 CNRS Université de Paris, Laboratoire de Biologie Physico-Chimique des Protéines Membranaires, 13 rue Pierre et Marie Curie, F75005 Paris, France

*Email: stephane.abel@cea.fr and bonnete@ibpc.fr

S1. Surface and hydration properties of the DDM in the corona of the two ShuA_DDM complexes

The surface area (SA) properties for the whole DDM (SA_{DDM}), the headgroup (SA_{HG}) the dodecane alkyl chain (SA_{DOD}) for the two ShuA_DDM systems were computed from the last 150 ns of the simulations with *trjVoronoi* code [1][2][3] by excluding, as in previous works [4,5], the hydrogen atoms. From these calculations, we estimated the average SA ratio, f_{tail} for the DDM surface alkyl chain of the corona with the relation: $f_{tail} = 1 - (1 - (SA_{HG}/SA_{DDM}))$. The results obtained for each ShuA_DDM system and replicas are similar and reported in Table S1.

| System | ShuA_268DDM | ShuA_294DDM | Micelle |
|------------|--------------|--------------|---------|
| SA_{DDM} | 220.4/220.5* | 214.7/214.5* | 222.9 |
| SA_{HG} | 202.2/202.3* | 197.6/197.2* | 202.8 |
| SA_{DOD} | 18.2/18.3* | 17.1/17.3* | 20.1 |
| f_{tail} | 8.3/8.3 % | 8.0/8.1* % | 9.0 % |

Table S1: Average surface areas between a detergent molecule and water for the entire DDM (SA_{DDM}), its headgroup (SA_{HG}^V) and its dodecane alkyl tail (SA_{DOD}). All values are in \AA^2 . f_{tail} is the average surface fraction shared between the water and the DDM alkyl chain heavy atoms. The statistical errors (maximum errors) are always lower than 3.0 and 0.2% for the surface areas and f_{tail} values. “Micelle” values are taken from ref. [4] for pure micelle with 132 DDM with the CHARMM36 parameters. For the first and the second replicas (*).

The low f_{tail} values for the alkyl chain (~ 8 % of the total DDM surface contact) indicate, in agreement with previous results [4], that in corona the DDM alkyl chain is protected from the water by DDM headgroup and/or are buried in the hydrophobic core of the corona. Concerning the surface contacts of the DDM headgroup, SA_{HG} , it represents a large part (~ 92 %) of the total SA to water of the corona and these values tend to slightly decrease (202.2 \AA^2 to 197.6 \AA^2) with the increase in DDM in the corona.

We also estimated the hydration level of each DDM in the corona using a simple distance criterion and consider that the water oxygen is in contact with any heavy atoms of the DDM if its distance is less than 3.5 \AA [4]. The calculations were carried out for the whole DDM molecule ($\langle n_{DDM}^W \rangle$), the maltose headgroup ($\langle n_{MALT}^W \rangle$) and for the dodecane alkyl chain ($\langle n_{C12}^W \rangle$). Again, each corresponding value for the 268 and 294 DDM systems was compared with those obtained for the pure DDM micelle [4]. The average numbers of water for the overall DDM, ($\langle n_{DDM}^W \rangle$), are found to be around 10. For the alkyl chain and the maltose headgroup, the values are found, on average, to be 9.5 and 1.0 waters, respectively, close to

the values obtained from MD DDM or sophorose (glucose β 1-2) micelles [4][5].

The hydration of the different oxygen atoms in the detergent headgroup is also given and were computed from the averaged radial pair density functions (RDFs) from the last 100 ns of each run of maltose-oxygen, or $\langle n_{O_x-O_w} \rangle$ and by integrating the RDF up to their first minima at $r \approx 3.5$ Å. The results for the water-maltose-oxygens (i.e. O_{2A-6A} to O_{1B-6B}, see Figure 2a for localization of these atoms on the molecule) pairs are listed in the Table S2, The results of these calculations show similar behaviors between the two ShuA_DDM complexes, with a total value around of 16 waters, which is not far from the value in DDM micelle (17.8 waters) [4].

| $\langle n_{O_x-O_w} \rangle$ | ShuA_268DDM | ShuA_294DDM | Micelle |
|---------------------------------|-------------------|-------------------|-------------|
| O _{2A} -O _w | 2.2/2.1* | 2.1/2.1* | 2.3 |
| O _{3A} -O _w | 2.8/2.7* | 2.8/2.7* | 3.2 |
| O _{4A} -O _w | 2.3/2.3* | 2.3/2.2* | 2.6 |
| O _{5A} -O _w | 0.9/0.9* | 0.9/0.9* | 1.0 |
| O _{6A} -O _w | 2.2/2.2* | 2.2/2.1* | 2.3 |
| Total GlcA | 10.4/10.4* | 10.3/10.0* | 11.4 |
| O _{1B} -O _w | 0.6/0.6* | 0.7/0.6* | 0.3 |
| O _{2B} -O _w | 1.5/1.5* | 1.5/1.4* | 1.6 |
| O _{3B} -O _w | 1.7/1.7* | 1.6/1.6* | 1.8 |
| O _{4B} -O _w | 0.3/0.3* | 0.3/0.3* | 0.5 |
| O _{5B} -O _w | 0.4/0.3* | 0.5/0.4* | 0.6 |
| O _{6B} -O _w | 1.4/1.3* | 1.4/1.3* | 1.6 |
| Total GlcB | 5.9/5.9* | 6.0/5.6* | 6.4 |
| Total Headgroup | 16.3/16.3* | 16.3/15.6* | 17.8 |

Table S2: Nearest neighbors for the maltose oxygens. The DDM oxygen atoms are labeled as follows: O_{4B} and O_{1B} are the acetalic oxygens, O_{5A} and O_{5B} are the maltose ring oxygens, and all the remaining are hydroxyl oxygens (see Figure 2a in the main text). For the first system and its replica (*). “Total GlcA” and “Total GlcB” values give the average nearest neighbors for all the oxygen atoms of the outermost and innermost glucose unit, respectively. “Micelle” values are taken from [4] for pure micelle with 132 DDM simulated with the CHARMM36 parameters.

More specifically, we also found that the highest $\langle n_{O_x-O_w} \rangle$ values are obtained for the hydroxyl oxygen atoms O_{2A}, O_{3A}, O_{4A} and O_{6A} (2.8 – 2.3) of the outermost glucose (GlcA)

whereas the equivalent oxygen atoms in the innermost glucose (GlcB) (which is linked to the C12 alkyl chain) is ~ 1.0 water smaller. The hydration of the ring oxygen atoms of the two glucose rings (i.e. O_{5A} and O_{5B}) is significantly lower than the hydroxyl oxygens (0.9 and 0.4, respectively) in agreement with previous findings [4][5]. Finally, for the oxygen atoms involved in the 1-4 glycosidic bond (O_{4A}) and the ether link between the headgroup and alkyl chain (O_{1B}), the corresponding hydration values are low values (< 0.3 and < 0.6) close to the values we found for glycolipids in micelles [4][5].

S2. Conformational properties of the DDM in the corona of the two the ShuA_DDM complexes

The conformations of the DDM headgroup and tail for the surfactant were also investigated by computing the average end-to-end distance for the dodecane alkyl chain (d_{C12}) (i.e. between the first (C₁) and last carbon (C₁₂) atoms) and the headgroup (d_{G2}) (i.e. between the O_{4A} and O_{1B} oxygens of the maltose headgroup, see Figure 2a). These calculations were repeated for each replica. The d_{C12} values for DDM in the two ShuA_DDM simulations are found similar and around 12.0 Å. This is not far from the value we found for DDM in micelle (12.3 Å) [4]. A smaller value for d_{C12} is known to be the consequence of the partial folding of the alkyl chain and existence of “*gauche defaults*” [6][7] and also found in micelles with a same C12 chain (see for instance refs. [1][4][8][9]). This assumption is confirmed with the computation of the percentage of *trans* conformation (p_{trans}) extracted from the averaged and normalized dihedral distribution curves (we used a bin width of 1°). The average p_{trans} values for the CCCC dihedral angles of the detergent alkyl chain (first 3 columns in Table S3) indicate that these dihedral angles are mostly in their *trans* state with relative populations of 72.1 %, in agreement with previous findings (e.g. refs [4][8][9]). Further analysis of the CCCC dihedral angles along the alkyl chain also show that the inner and outermost CCCC (i.e. C₁C₂C₃C₄ and C₉C₁₀C₁₁C₁₂) have a smaller *trans* population, $p_{trans} \approx 65$ % as we found in micelle (67.7 % and 68.5 %, respectively) [4]. For the dihedral angles involving the atoms in the ether link (i.e. C_{1B}O_{1B}C₁C₂ and O_{1B}C₁C₂C₃) these angles are also in *gauche* and *trans* conformations with a similar relative trans populations of 10.5 % and 80.2 % as we found in micelle (5.0 % and 76.4 %) [4].

| | ShuA_268DDM | ShuA_294DDM | Micelle |
|---|-------------|-------------|---------|
| Total CCCC | 72.1/72.1* | 72.1/72.8* | 72.1 |
| C ₁ C ₂ C ₃ C ₄ | 65.1/64.5* | 65.4/64.6* | 67.7 |
| C ₉ C ₁₀ C ₁₁ C ₁₂ | 66.6/66.1* | 69.5/66.8* | 68.5 |
| O _{1B} C ₁ C ₂ C ₃ | 10.5/10.6* | 10.6/9.8* | 4.9 |
| C _{1B} O _{1B} C ₁ C ₂ | 80.1/81.1* | 80.5/79.9* | 76.4 |
| O _{6A} C _{6A} C _{5A} O _{5A} (ω 1) | 35.5/36.0* | 36.0/38.0* | 49.4 |
| O _{6B} C _{6B} C _{5B} O _{5B} (ω 2) | 55.5/55.9* | 58.0/57.9* | 41.2 |

Table S3: Trans Populations, p_{trans} for selected dihedrals of the DDM alkyl chain and headgroup. Dihedral angles between -120° and $+120^\circ$ are considered as in the *trans* conformation. For the first system and its replica (*). See Figure 2a in the main text) for localization of the atoms in DDM molecule. “Micelle” values are taken from [4] for pure micelle with 132 DDM simulated with the CHARMM36 parameters.

Concerning the maltose head end-to-end distance and conformation, the d_{G2} value (9.6 Å) are found similar to the value found for DDM in micelle (9.8 Å) and 4.8 % lower than the value found for maltose in fully extended conformation (10.3 Å with Φ_{H} : H_{1A}C_{1A}O_{4B}C_{4B} = 0.0° and Ψ_{H} : C_{1A}O_{4B}C_{4B}H_{4B} = 0.0°). In Fig. S1, we have plotted the normalized distributions β 1-4 glycosidic dihedral angles $\Phi_{\text{H}}/\Psi_{\text{H}}$ ($P(\Phi_{\text{H}},\Psi_{\text{H}})$) for DDM in the corona and found that these distributions are similar for two ShuA_DDM simulations. We observe a maximum at around $(-20^\circ, -10^\circ)$ that differ slightly from the results obtained for DDM in the protein-free micelle (maximum near $(-40^\circ, -25^\circ) \pm 10^\circ$) [4].

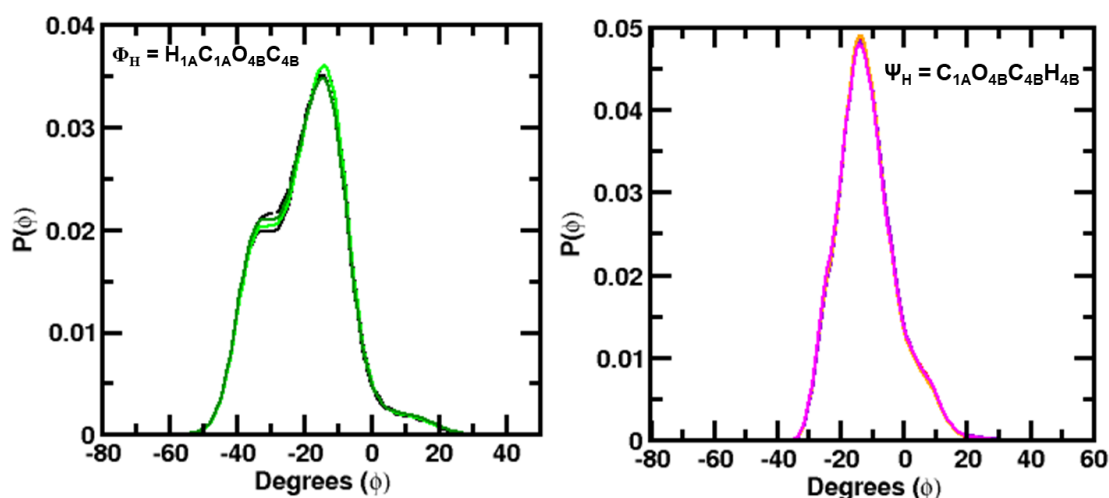


Figure S1: Normalized distributions of the ϕ_{H} (left panel) and ψ_{H} (right panel) of glycosidic dihedral angle pairs of the DDM headgroup. In black and green continue/dashed lines and red and magenta continue/dashed lines for the systems with 268 and 294 DDM, respectively and their replicas. Note that for ψ_{H} curves are superposed.

Finally, in the last two rows of Table S3, we have also listed the average relative populations of the two ω dihedral angles of the reducing (GlcA) and non-reducing (GlcB) glucose unit (i.e. $\omega_1 = O_{6A}C_{6A}C_{5A}O_{5A}$ and $\omega_2 = O_{6B}C_{6B}C_{5B}O_{5B}$, respectively). Their conformations are found similar to that of DDM micelle i.e. *trans*, but with different populations (~36.0 % vs. 49.4 %) and (55.0 – 58.0 % vs. 41.2 %). Taking together, all these differences may explain the differences in the headgroup hydration and conformation of the DDM in the corona with the DDM in pure micelle.

S3. Additional Figures

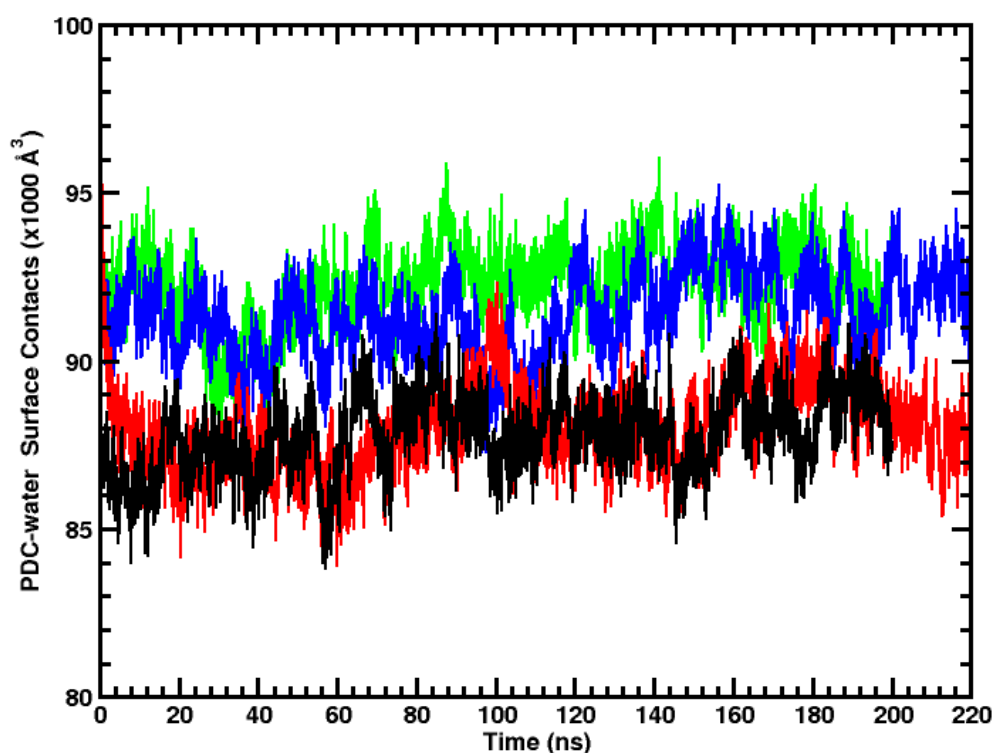


Figure S2: Time evolution of the protein detergent complex-water surface contacts of the two ShuA_268DDM (in black and red colors) and ShuA_294DDM (in green and blue colors) simulations computed with the *trjVoronoi* program [1][2][3].

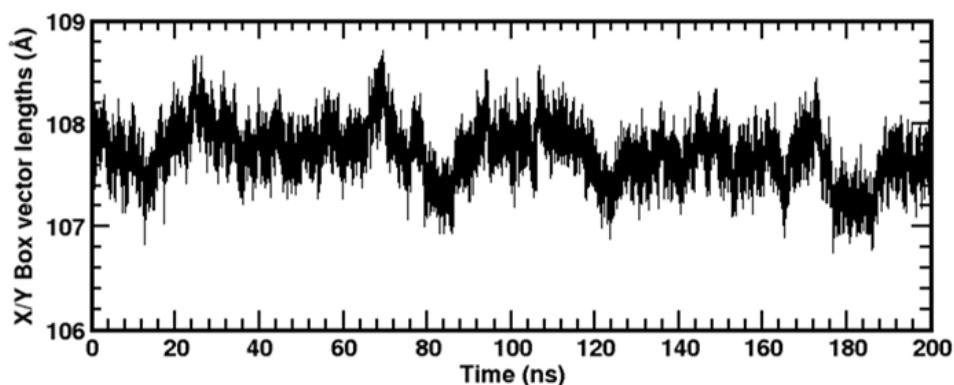


Figure S3: Evolution of the lengths of the X and Y box vectors during the course of the OM

simulation.

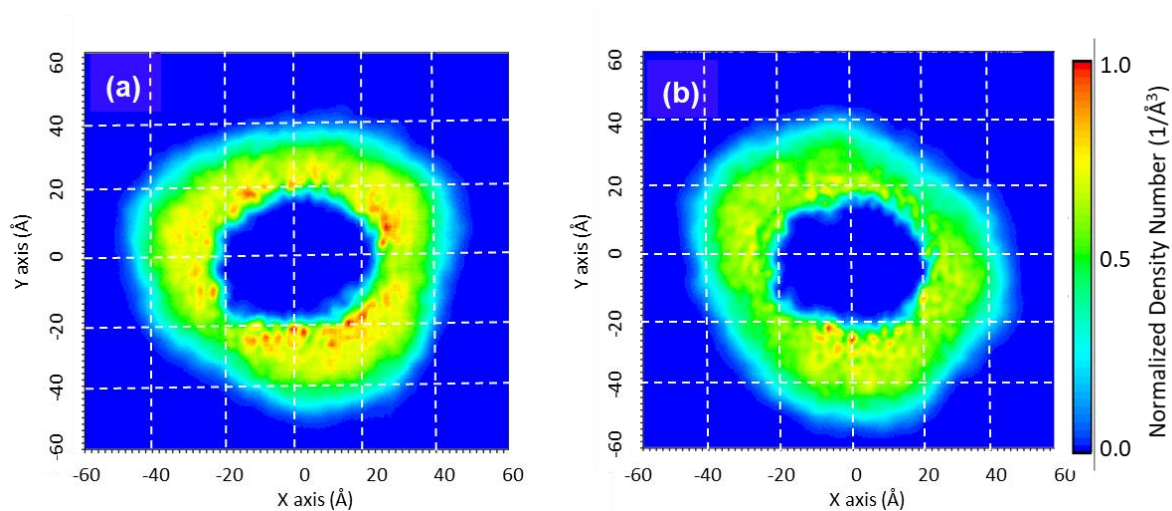


Figure S4: Averaged 2D density maps for the systems with (a) 268 and (b) 294 DDM molecules (replica). Here the ShuA_DDM complex was centered at $Z = 0$ Å.

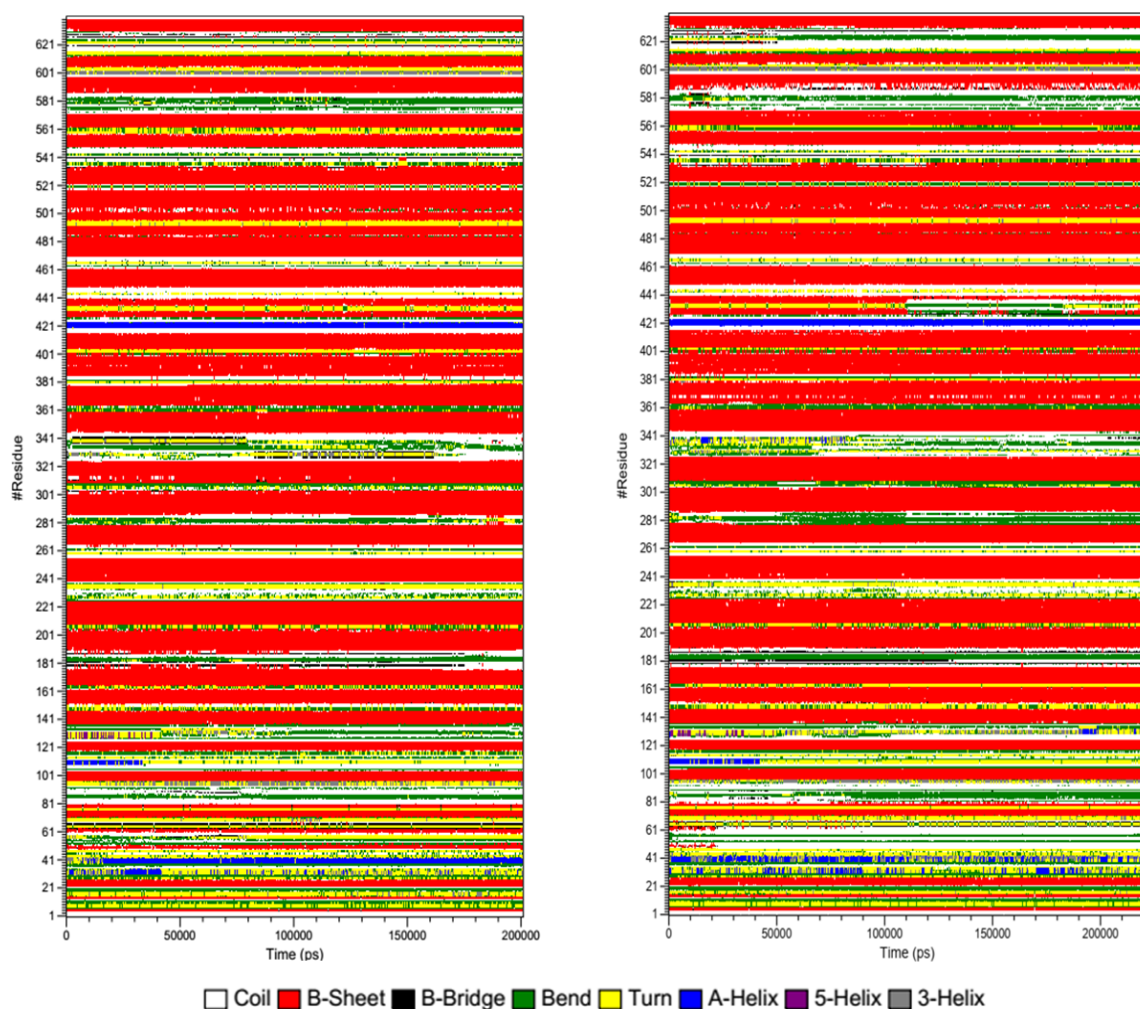


Figure S5: Time evolution of the secondary structures of ShuA in presence of 268 DDM for the first simulation and its replica (left and right panels, respectively).

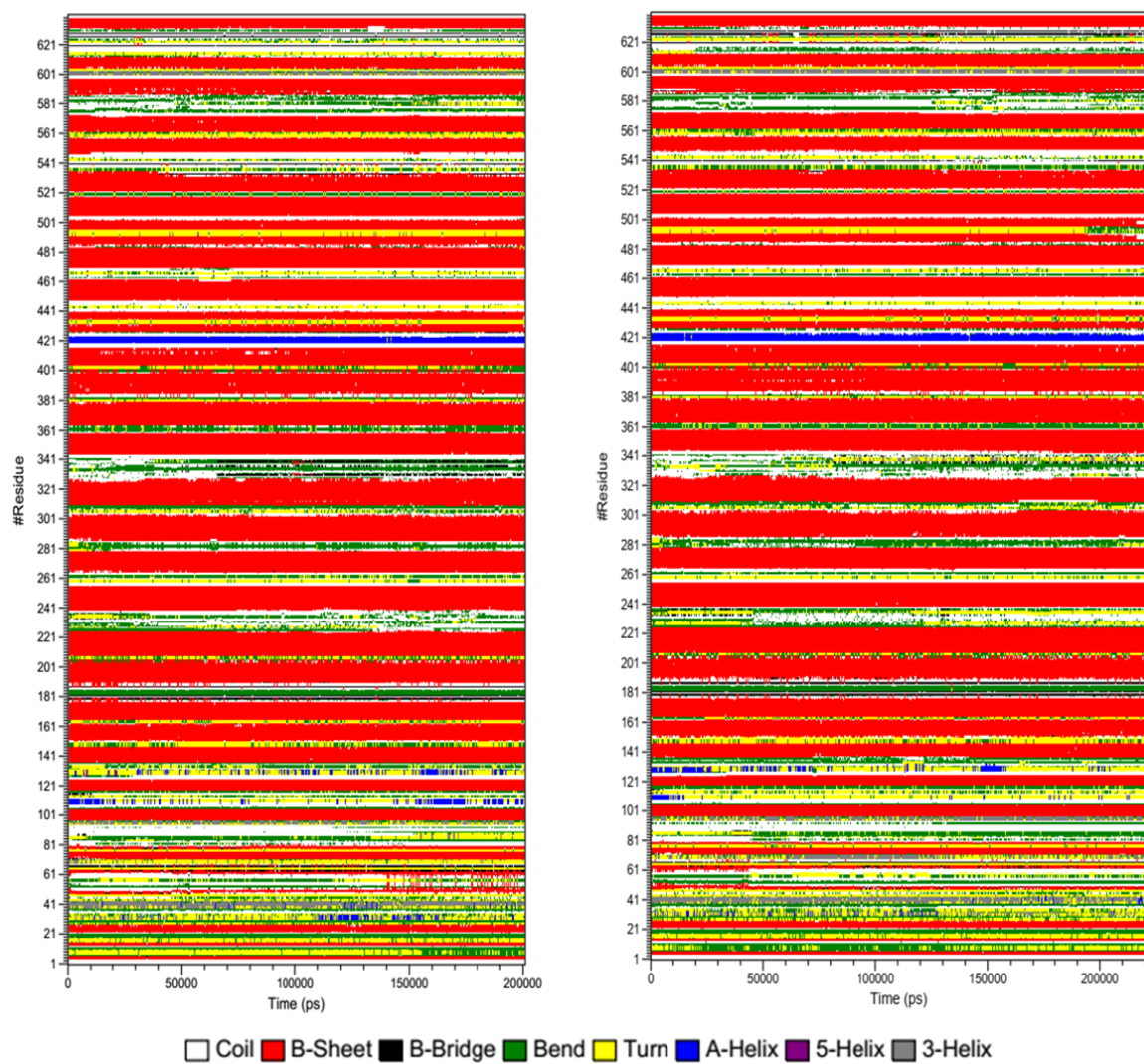
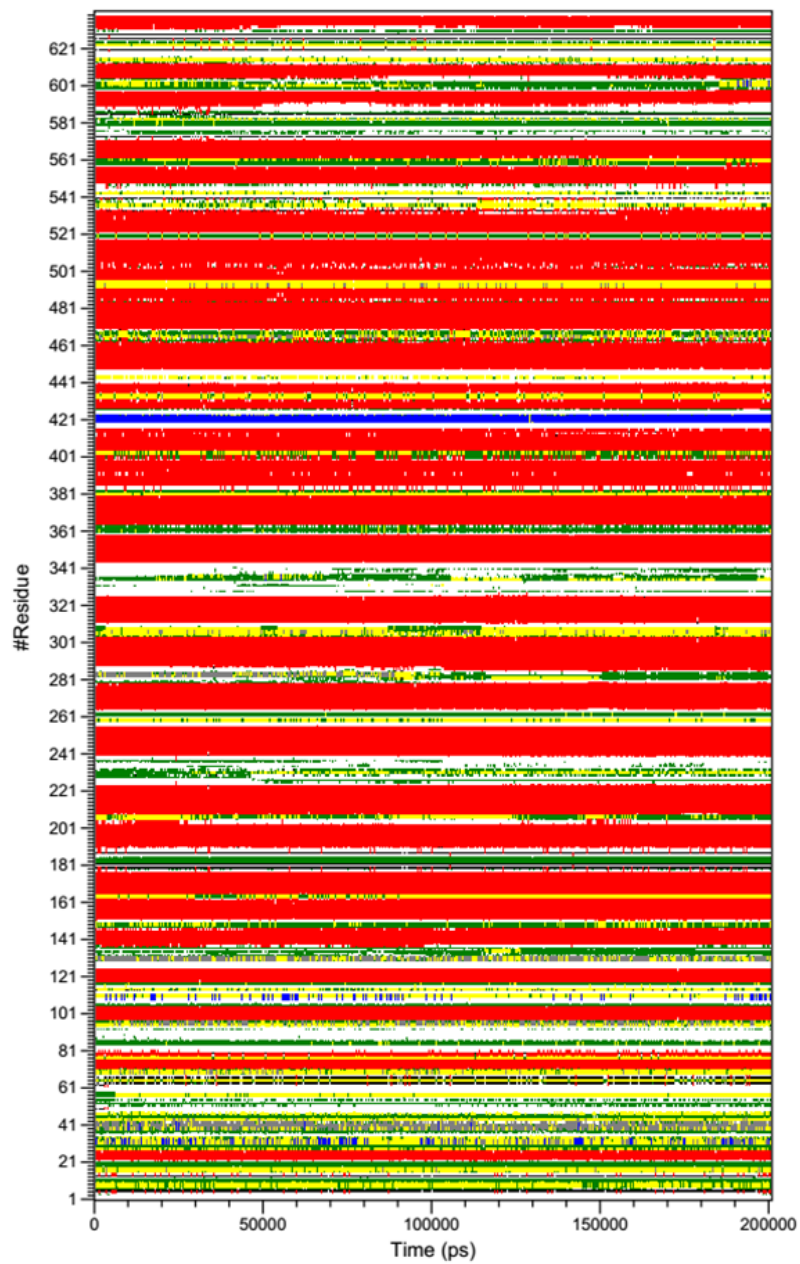


Figure S6: Time evolution of the secondary structures of ShuA in presence of 294 DDM for the first simulation and its replica (left and right panels, respectively).



Coil
 B-Sheet
 B-Bridge
 Bend
 Turn
 A-Helix
 5-Helix
 3-Helix

Figure S7: Time evolution of the secondary structures of ShuA in the outer bacterial membrane.

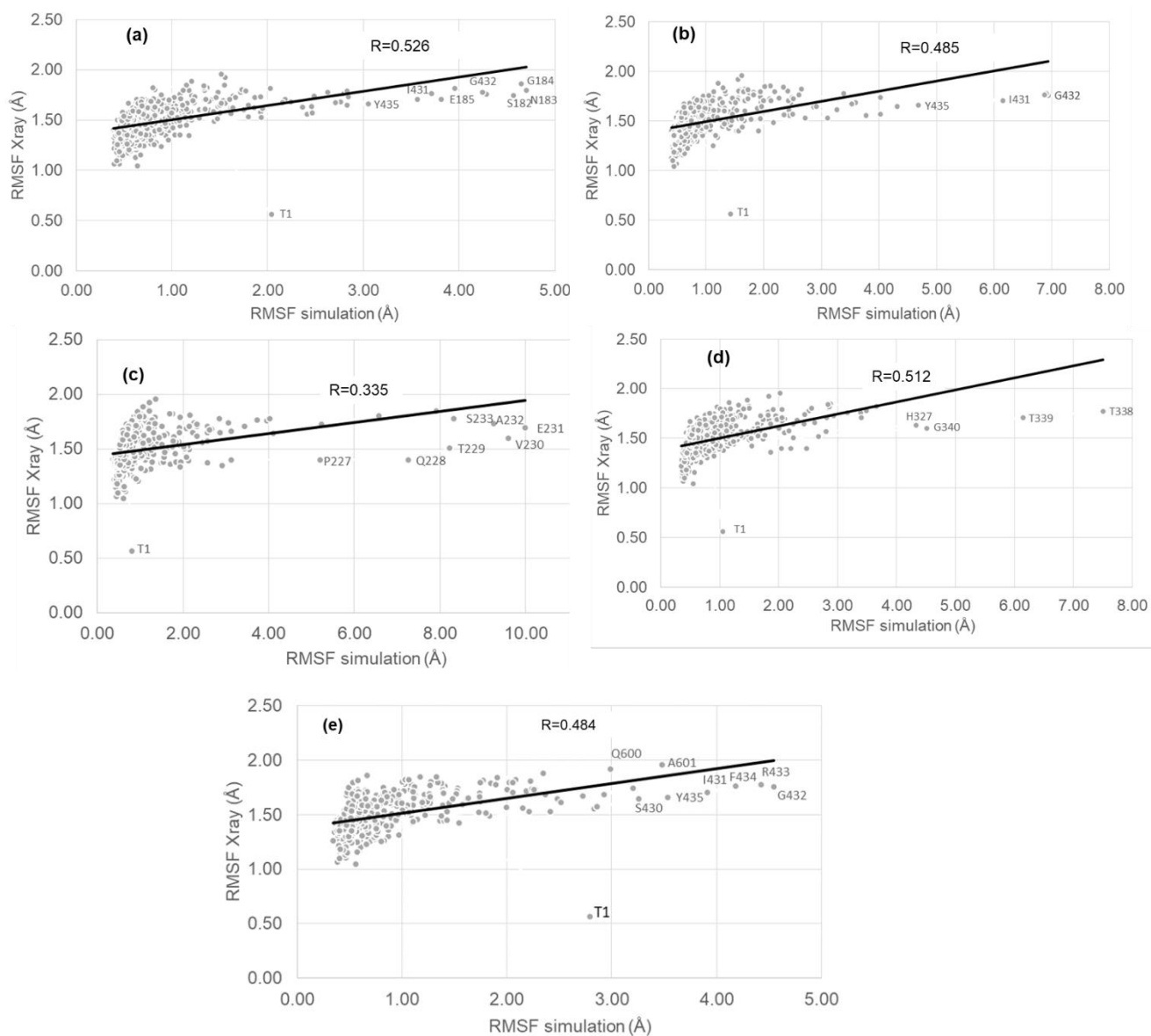


Figure S8: Crystallographic vs. simulation RMSF for all the protein α carbons. The correlation coefficients R was computed by excluding the missing residues in the crystal structure. Panel a-b: ShuA_268DDM and its replica, c-d: ShuA_294DDM and its replica and e: ShuA_OM. See also Figure 7f in the main text to see the localizations of higher mobile residues in the protein.

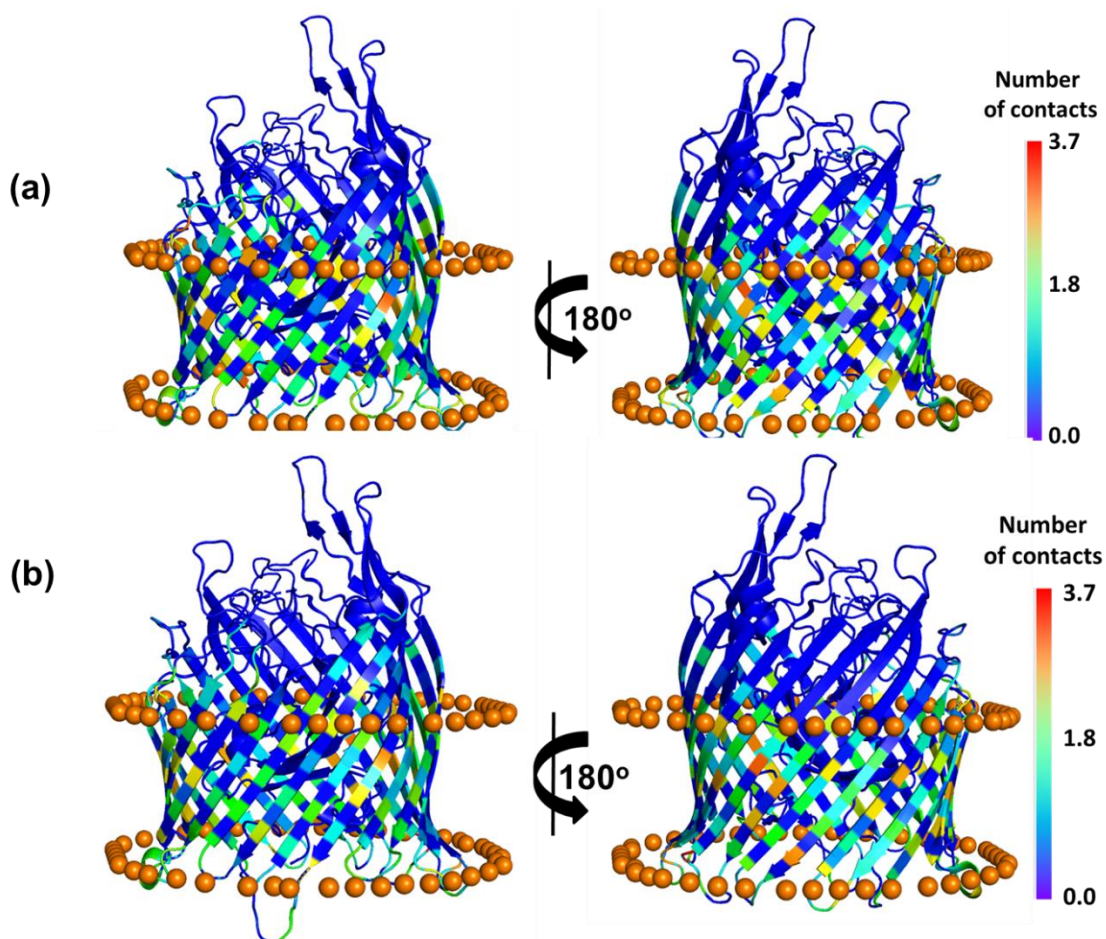


Figure S9: Colored 3D representation of ShuA in the DDM environment (replica) with (a) 268 and (b) 294 surfactants according to the number of contacts between the protein heavy atoms and the DDM heavy atoms. See Fig. 8 in the main text for details. Figure was produced with the PyMOL program [10].

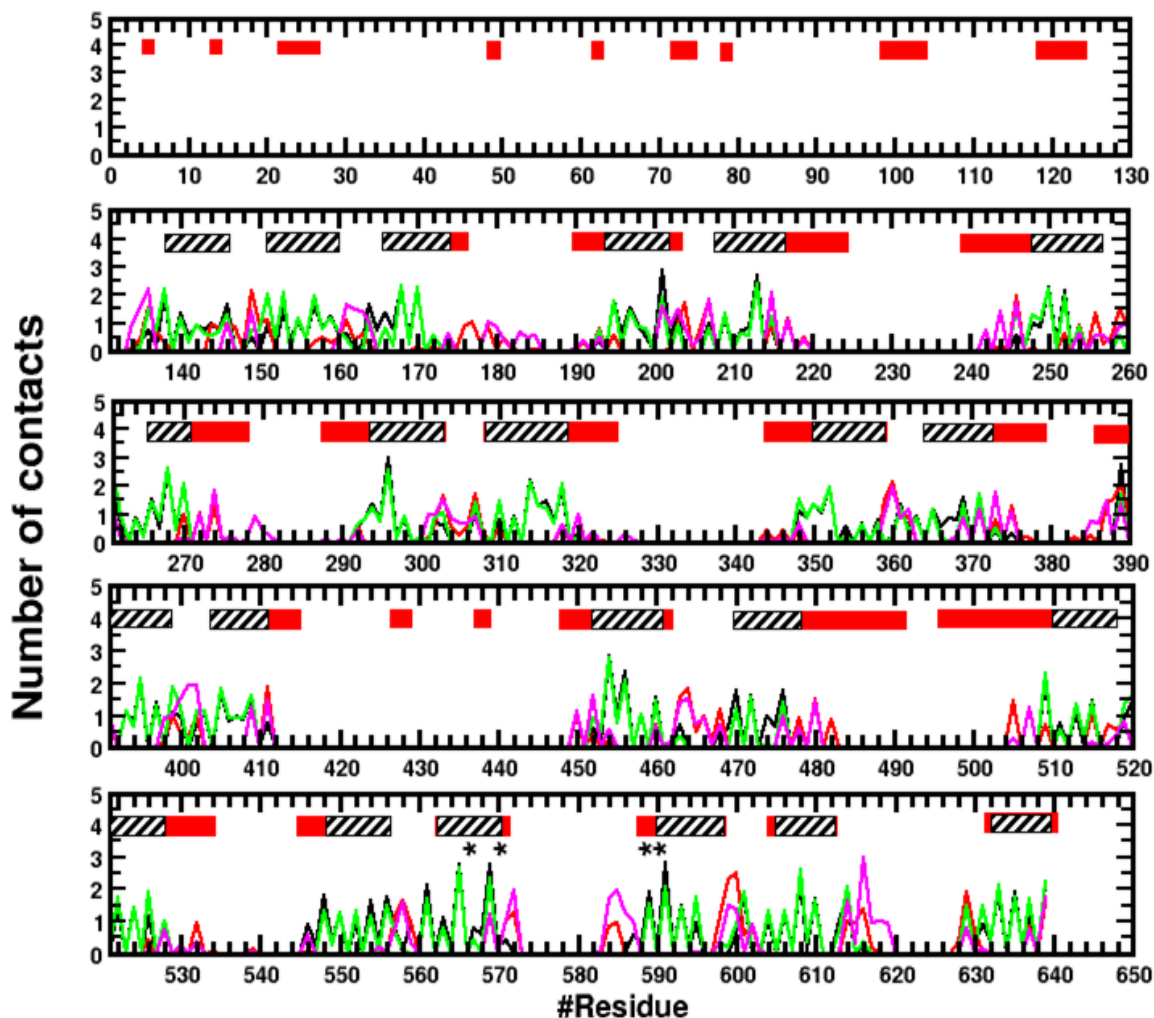


Figure S10: Patterns of interactions of protein residues with the DDM for ShuA_268DDM system and its replica. The graph shows, for each residue, the average number of contacts with the DDM alkyl chain (in black and green colors) and the headgroup (red and magenta colors). A contact is first counted when the distance of between the heavy atoms of a ShuA residue and the DDM is ≤ 4 Å. Multiple contacts between each DDM and protein residues heavy atoms are counted as one. The red and dashed bars above show the limits of the 33 β -strands including the 22 TM β -strands, respectively. The four asterisks highlight the residues interacting with the N-octyl- β -D-glucopyranoside (OG) in the crystal (PDB entry: 3FHH) [11][10].

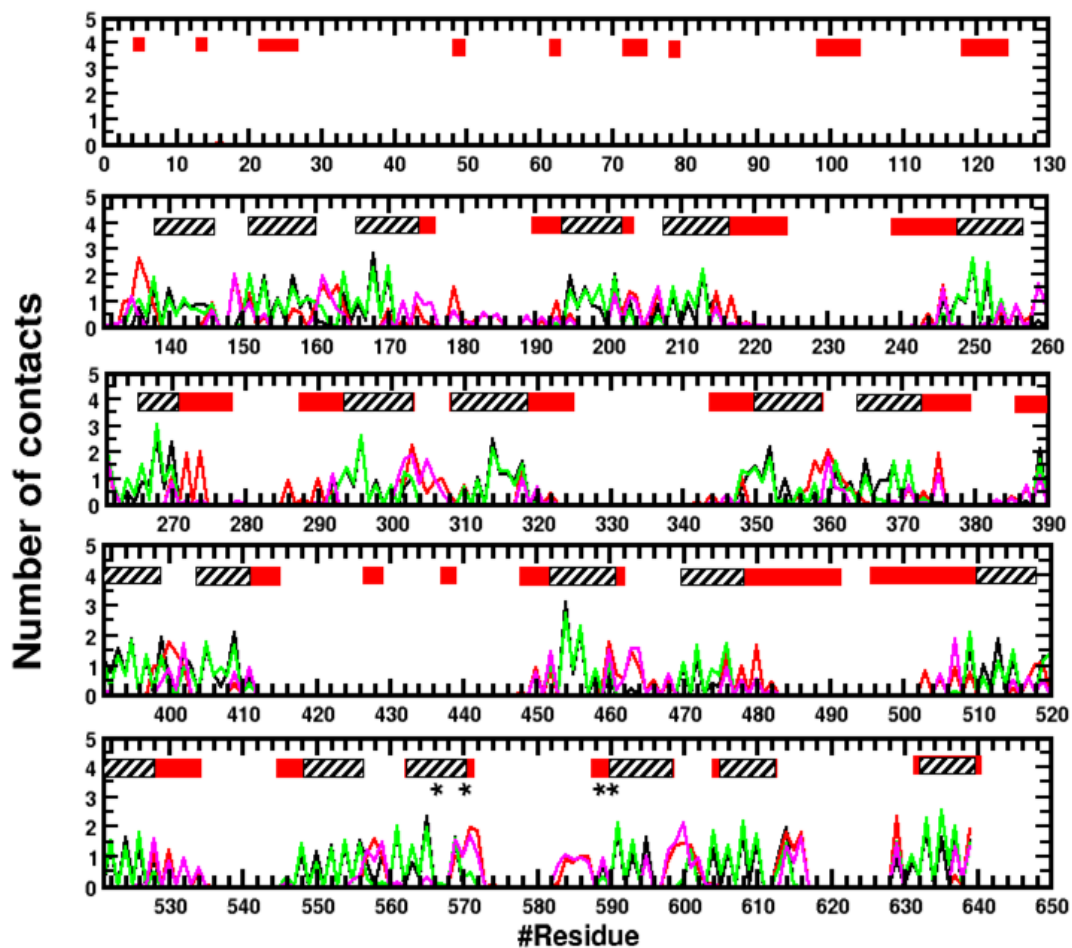


Figure S11: Same legend as Figure S10 for the ShuA_294DDM system.

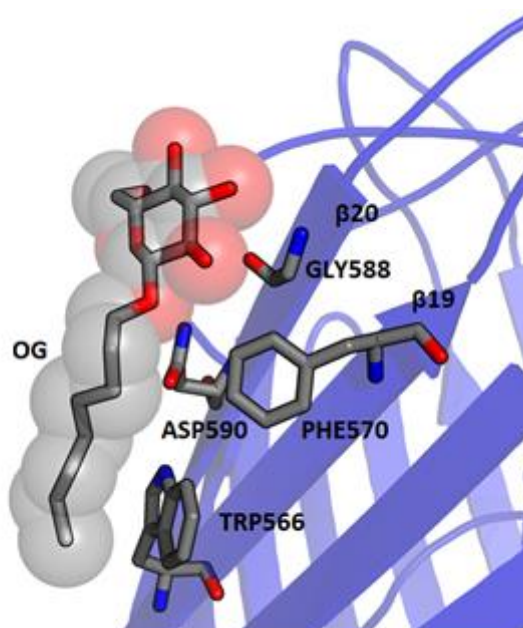


Figure S12: Snapshot showing selected outer residues of ShuA in close contact (distance < 4.0 Å) with n-octyl-β-D-glucopyranoside (βOG) heavy atoms presents in the crystal (PDB entry: 3FHH)[12]. Figure was produced with the PyMOL program [10].

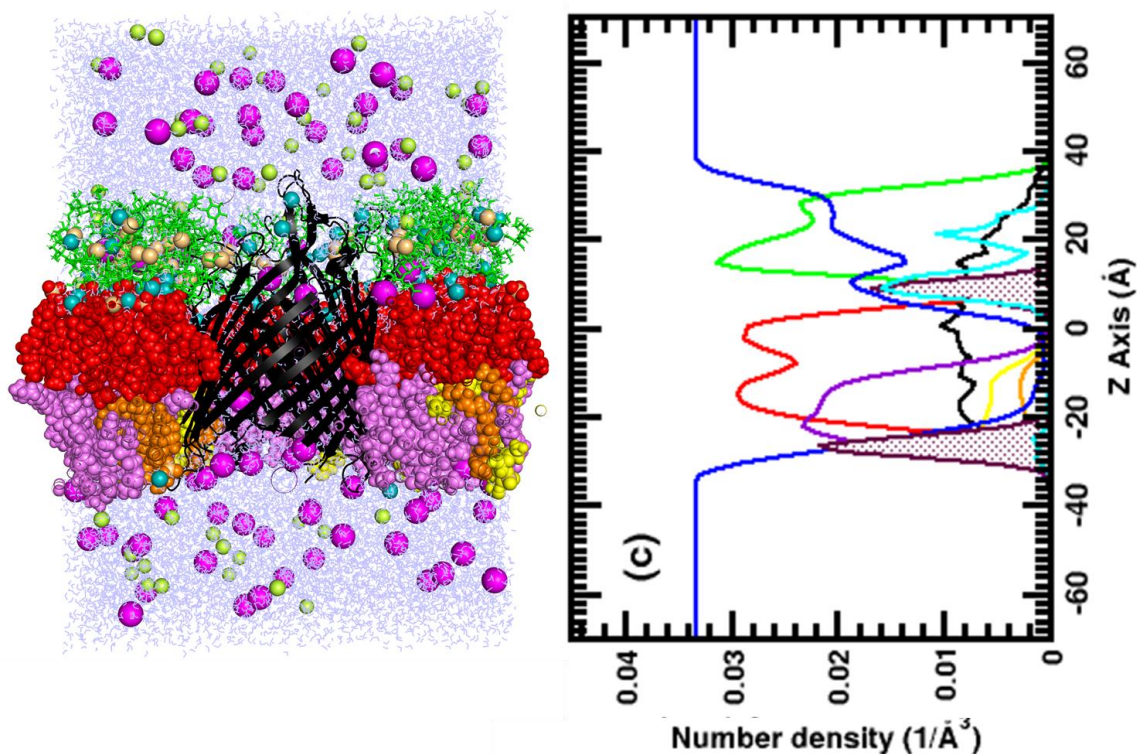


Figure S13: (Left) Representative snapshot of the ShuA_OM system associated with the number density profiles (right) for the protein (black), the LPS alkyl chain (red), the inner and outer sugar cores (green), PPPE (violet), PVPG (yellow), PVCL2 (orange), phospholipid phosphate group (maroon), calcium ion (cyan) and the water (blue). The protein is centered at $Z = 0 \text{ \AA}$ and the OM normal is the Z axis. Note for visual clarity the phosphate shown in left Figure and K^+/Cl^- ions in the 1D density profiles. The density profiles are averaged over the last 100 ns of the MD simulations. Figure were produced with the PyMOL program [10].

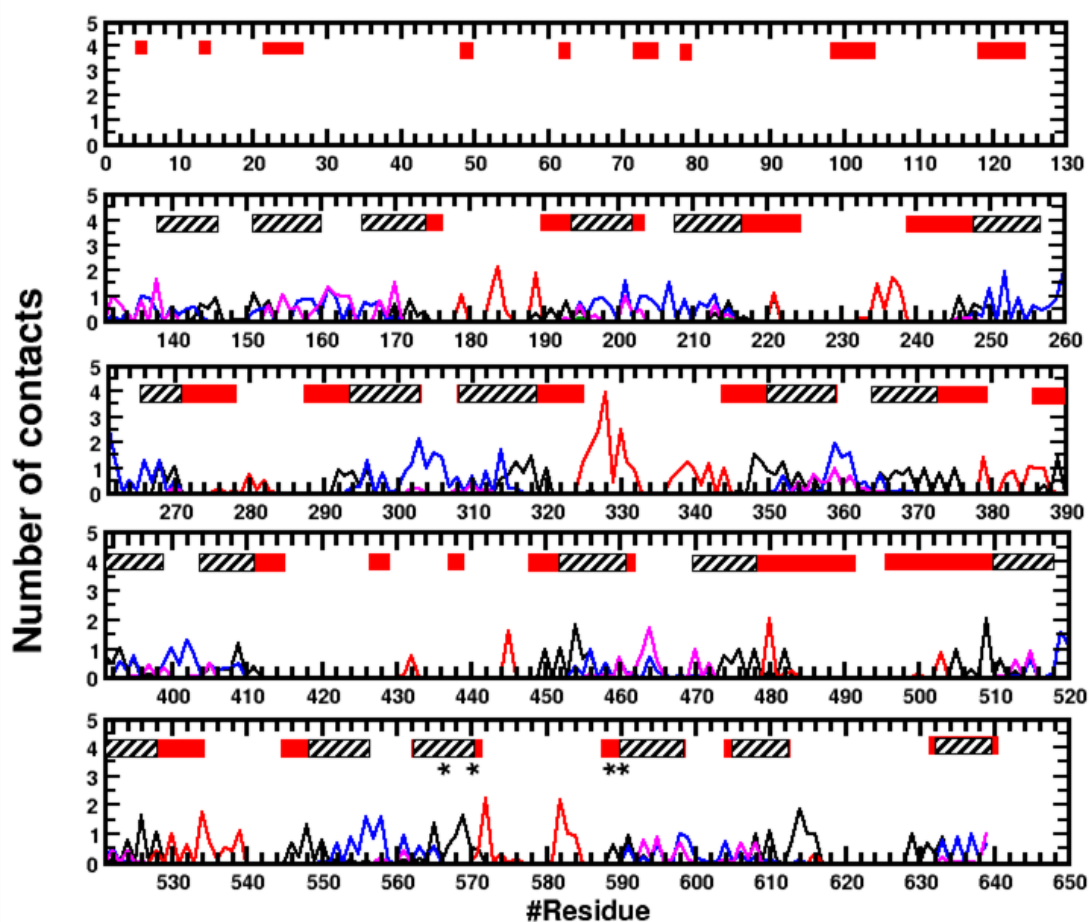


Figure S14: Patterns of interactions of protein residues with the different lipid constituents in the outer membrane. The graph shows, for each residue, the average number of contacts with the LipidA alkyl chain (in black), sugar core (in red), PPPE (blue), PVPG (in pink) and PVCL2 (in green) the headgroup (red). A contact is first counted when the distance of between the heavy atoms of a ShuA residue and its partner is $< 4 \text{ \AA}$. Multiple contacts between the each lipids and the protein residues heavy atoms are counted as one. The red and dashed bars above show the limits of the 33 β -strands including the 22 TM β -strands, respectively. The four asterisks highlight the residues interacting with the OG molecule presents in the crystal (PDB entry: 3FHH) [11].

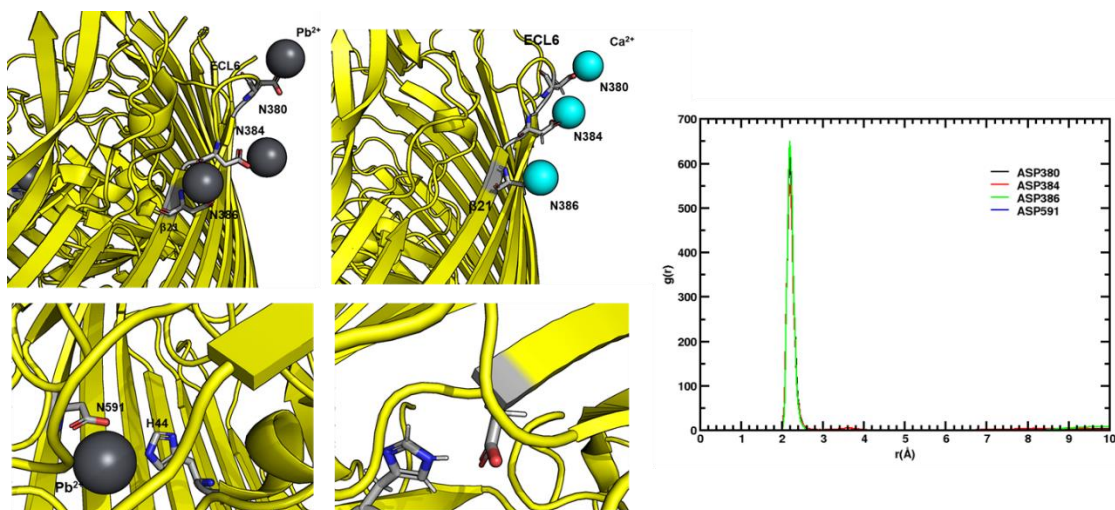


Figure S15: Representative snapshots showing the binding sites of the Pb²⁺ and Ca²⁺ ions and some ASP residues (i.e. Asparagine residue 380, 384, 386, 591) in the crystal (left panel) and in the membrane simulation (middle panel). In the simulations, only Asp 380, 384, 386 interact strongly with Ca²⁺ ion and not Asp 591. This interactions are very stable during the course of the membrane simulation and occurs at an equilibrium distance of $\sim 2.3\text{\AA}$ (computed from the $g(r)$, left panel) and in agreement with Albano et al. [12].

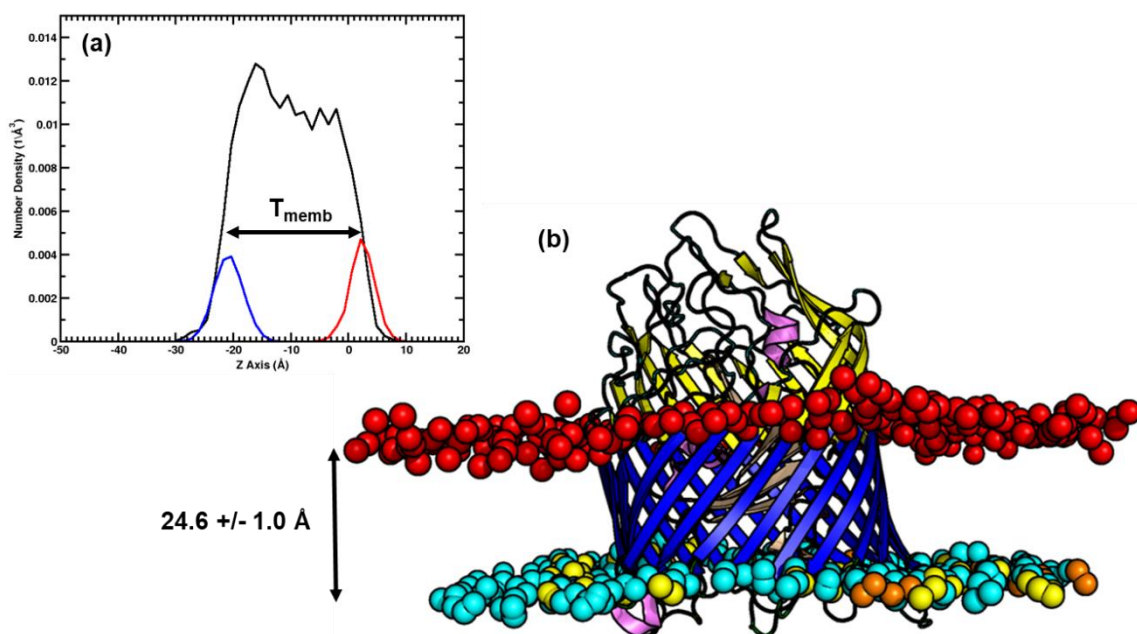


Figure S16: (a) Distributions of the Z position of the ShuA TM β -barrels (black) with the C2 and C4 atoms (i.e., the carbon bonded to carbonyl group) of lipid A (red) and C2 atoms (i.e., the carbon-bonded-to-carbonyl group) of PPPE, PVPG and PVCL2 phospholipids (blue). (b) Snapshot showing the limit of the hydrophobic thickness of the outer membrane. The beads for the C2-C4 and C2 atoms of the lipid A and phospholipids are in red, cyan, yellow and oranges colors, respectively. The colors scheme for the protein is the same as in Figure 1. The distance between the two peaks “ T_{memb} ” highlighted with a horizontal arrow in (a) were used to calculate the hydrophobic thickness of the membrane. The OM is centered at $Z = 0\text{\AA}$ and the OM normal is the Z axis. Figures were produced with the PyMOL program [10].

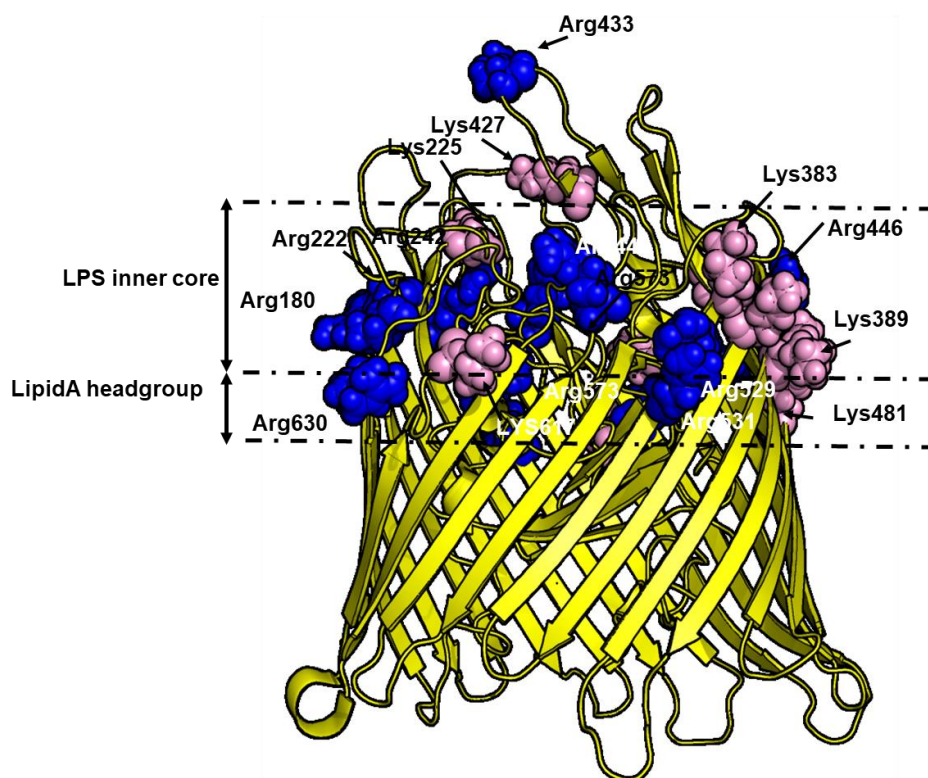


Figure S17: Selected Lys (pink) and Arg (blue) outer residues located in the $Z > 10 \text{ \AA}$ region of ShuA that can interact with the negatively charged lipidA headgroup and the LPS inner sugar core (see also Figure 2e) in the main text. Figure was produced with the PyMOL program [10].

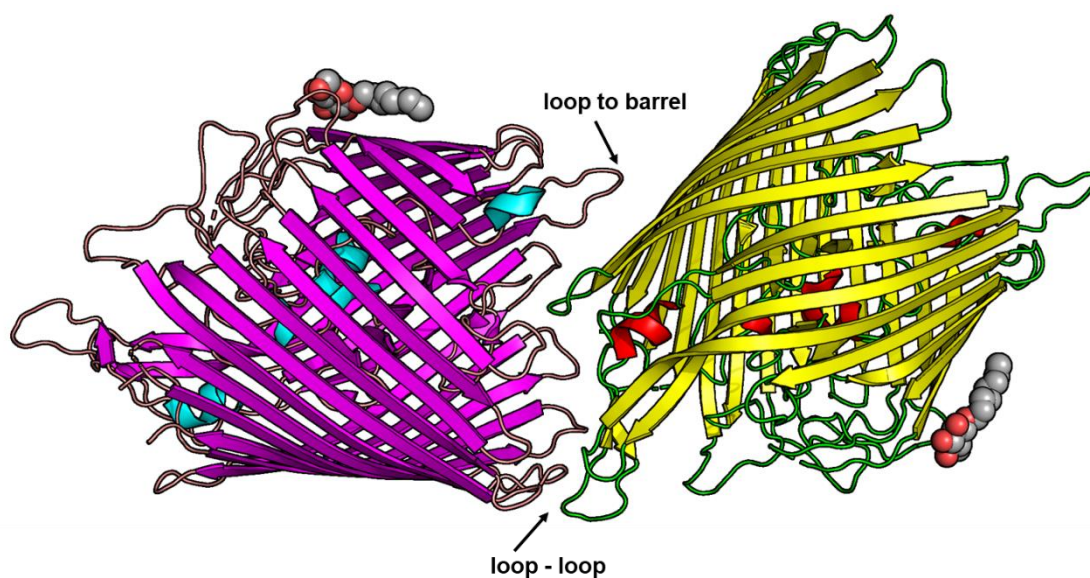


Figure S18: Contacts in apo-ShuA crystal (PDB entry 3FHH) [12] with the OG molecules represented as a sphere. Figure was produced with the PyMOL program [10].

S4. References

- [1] S. Abel, F.-Y. Dupradeau, M. Marchi, Molecular Dynamics Simulations of a Characteristic DPC Micelle in Water, *J. Chem. Theory Comput.* 8 (2012) 4610–4623. doi:10.1021/ct3003207.
- [2] C.H. Rycroft, VORO++: A three-dimensional Voronoi cell library in C++, *Chaos An Interdiscip. J. Nonlinear Sci.* 19 (2009) 041111. doi:10.1063/1.3215722.
- [3] M. Marchi, openTRJ, (2019). <https://github.com/octupole/openTRJ>.
- [4] S. Abel, F.-Y.Y. Dupradeau, E.P. Raman, A.D. MacKerell, M. Marchi, Molecular Simulations of Dodecyl- β -maltoside Micelles in Water: Influence of the Headgroup Conformation and Force Field Parameters, *J. Phys. Chem. B.* 115 (2011) 487–499. doi:10.1021/jp109545v.
- [5] S. Abel, A. Lorieau, B. de Foresta, F.-Y. Dupradeau, M. Marchi, Bindings of hMRP1 transmembrane peptides with dodecylphosphocholine and dodecyl- β -d-maltoside micelles: A molecular dynamics simulation study, *Biochim. Biophys. Acta - Biomembr.* 1838 (2014) 493–509. doi:10.1016/j.bbamem.2013.10.012.
- [6] S. Manet, A.S. Cuvier, C. Valotteau, G.C. Fadda, J. Perez, E. Karakas, S. Abel, N. Baccile, Structure of Bolaamphiphile Sophorolipid Micelles Characterized with SAXS, SANS, and MD Simulations, *J. Phys. Chem. B.* 119 (2015) 13113–13133. doi:10.1021/acs.jpcc.5b05374.
- [7] F. Holler, J.B. Callis, Conformation of the hydrocarbon chains of sodium dodecyl sulfate molecules in micelles: an FTIR study, *J. Phys. Chem.* 93 (1989) 2053–2058. doi:10.1021/j100342a068.
- [8] K.A. Dill, Configurations of the amphiphilic molecules in micelles, *J. Phys. Chem.* 86 (1982) 1498–1500. doi:10.1021/j100206a006.
- [9] X. Liang, M. Marchi, C. Guo, Z. Dang, S. Abel, Atomistic Simulation of Solubilization of Polycyclic Aromatic Hydrocarbons in a Sodium Dodecyl Sulfate Micelle, *Langmuir.* 32 (2016) 3645–3654. doi:10.1021/acs.langmuir.6b00182.
- [10] E. Karakas, C. Taveneau, S. Bressanelli, M. Marchi, B. Robert, S. Abel, Derivation of original RESP atomic partial charges for MD simulations of the LDAO surfactant with AMBER: applications to a model of micelle and a fragment of the lipid kinase PI4KA, *J. Biomol. Struct. Dyn.* 35 (2017) 159–181. doi:10.1080/07391102.2015.1135822.
- [11] The PyMOL Molecular Graphics System, Version 1.7.4, (n.d.).
- [12] D. Cobessi, A. Meksem, K. Brillet, Structure of the heme/hemoglobin outer membrane

receptor ShuA from *Shigella dysenteriae* : Heme binding by an induced fit mechanism, *Proteins Struct. Funct. Bioinforma.* 78 (2010) 286–294. doi:10.1002/prot.22539.

- [13] J.M.R. Albano, N. Mussini, R. Toriano, J.C. Facelli, M.B. Ferraro, M. Pickholz, Calcium interactions with Cx26 hemmichannel: Spatial association between MD simulations biding sites and variant pathogenicity, *Comput. Biol. Chem.* 77 (2018) 331–342. doi:10.1016/j.compbiolchem.2018.11.004.