*Revue Roumaine de Chimie,* **2009**, *54*(10), 799–805

# CHARMM FORCE FIELD PARAMETERIZATION OF BACTERIAL LIPOPOLYSACCHARIDES

Alexandra MORARU, [a]* Istvan SVAB [a,b] and Dan Florin MIHĂILESCU [a]

[a] University of Bucharest, Faculty of Biology, Department of Anatomy, Animal Physiology and Biophysics, Splaiul Independentei 91-95, 050095, Bucharest, Roumania, Phone no. 0040213181569
[b]Institute of Cellular Biology and Pathology "N. Simionescu", Str B.P.Hasdeu 8, Bucharest, Roumania, Phone no. 0040213194518

*Received February 26, 2009*

Bacterial lipopolysaccharides (LPS) are major components of the Gram-negative bacteria outer membrane and are known to trigger a variety of inflammatory reactions in macrophages and other cells presenting a CD14 receptor. LPS consist of a hydrophobic domain known as lipid A (or endotoxin), a nonrepeating "core" oligosaccharide and a distal polysaccharide (or O-antigen). The purpose of this study was to obtain the parameters of the LPS molecule consistent with force fields that use an atomic detailed description of structure and dynamics like CHARMM and NAMD. The parameters for the lipid A region were determined *de novo*. All the quantum chemical calculations for this molecule were completed using the Gaussian03 software at the Restricted Hartree Fock (RHF) level of theory using the Pople 6-31G* basis set. The nonbonded terms, namely the van der Waals term and the electrostatic term, were available in the CHARMM force field and therefore were assigned by analogy. The "core" oligosaccharide contains unusual sugars with seven or eight carbon atoms. This structure was built using a patching method for which a flexible topology file was determined using VMD extension Molefacture. In our study we determined a set of parameters for the LPS molecule to be further used in molecular modeling simulations with CHARMM force field. These results can be easily adapted by the user to model different LPS molecules.

## INTRODUCTION

Gram-negative bacteria, like *Escherichia coli, Pseudomonas aeruginosa* or *Salmonella typhi,* are surrounded with a double membrane, the cytoplasmic membrane and the outer membrane. The cytoplasmic membrane is composed of phospholipids, whereas the outer membrane is an asymmetric structure containing primarily phospholipids in its inner monolayer and LPS in its outer monolayer. Gram-negative bacteria are less susceptible to specific antibiotics and their surfaces are less permeable than those of Gram-positive bacteria, that lack LPS in the membrane. LPS is thought to act as a permeability barrier, making the outer membrane relatively impermeable to hydrophobic antibiotics, detergeants, and host proteins.[1] Therefore it is important to understand the structural details of the bacterial membrane and its components.

Also, LPS molecule represents the endotoxin held responsible for the toxicity of Gram-negative bacteria. LPS are known to trigger a variety of inflammatory reactions in macrophages and other cells presenting a CD14 receptor. Medically relevant Gram-negative bacteria include organisms which cause sexually transmitted disease (*Neisseria gonorrhoeae*), meningitis (*Neisseria meningitidis*), respiratory symptoms (*Moraxella catarrhalis, Hemophilus influenzae*, *Klebsiella pneumoniae*, *Legionella pneumophila*, *Pseudomonas aeruginosa*), primarily urinary problems (*Escherichia coli*, *Proteus mirabilis*, *Enterobacter cloacae*, *Serratia marcescens*), and primarily gastrointestinal problems (*Helicobacter pylori*, *Salmonella enteritidis*, *Salmonella typhi*). There is considerable interest in understanding LPS properties because their release induces septic shock to the infected patients. A detailed insight of the tridimensional structure is very important in the development of strategies for preventing septic shock.

The structure of LPS differs from other phospholipids in that its backbone is formed of sugar residues, can have up to seven fatty acids per molecule, and has a sizable saccharide chain

* Corresponding author: alexandra.moraru@bio.unibuc.ro

extending from the lipid backbone. In a number of species, the O-antigenic polysaccharide chains are completely missing. In Figure 1 is a representation of the structure of LPS from *E.coli,* which lacks the region of the O-antigenic polysaccharide.

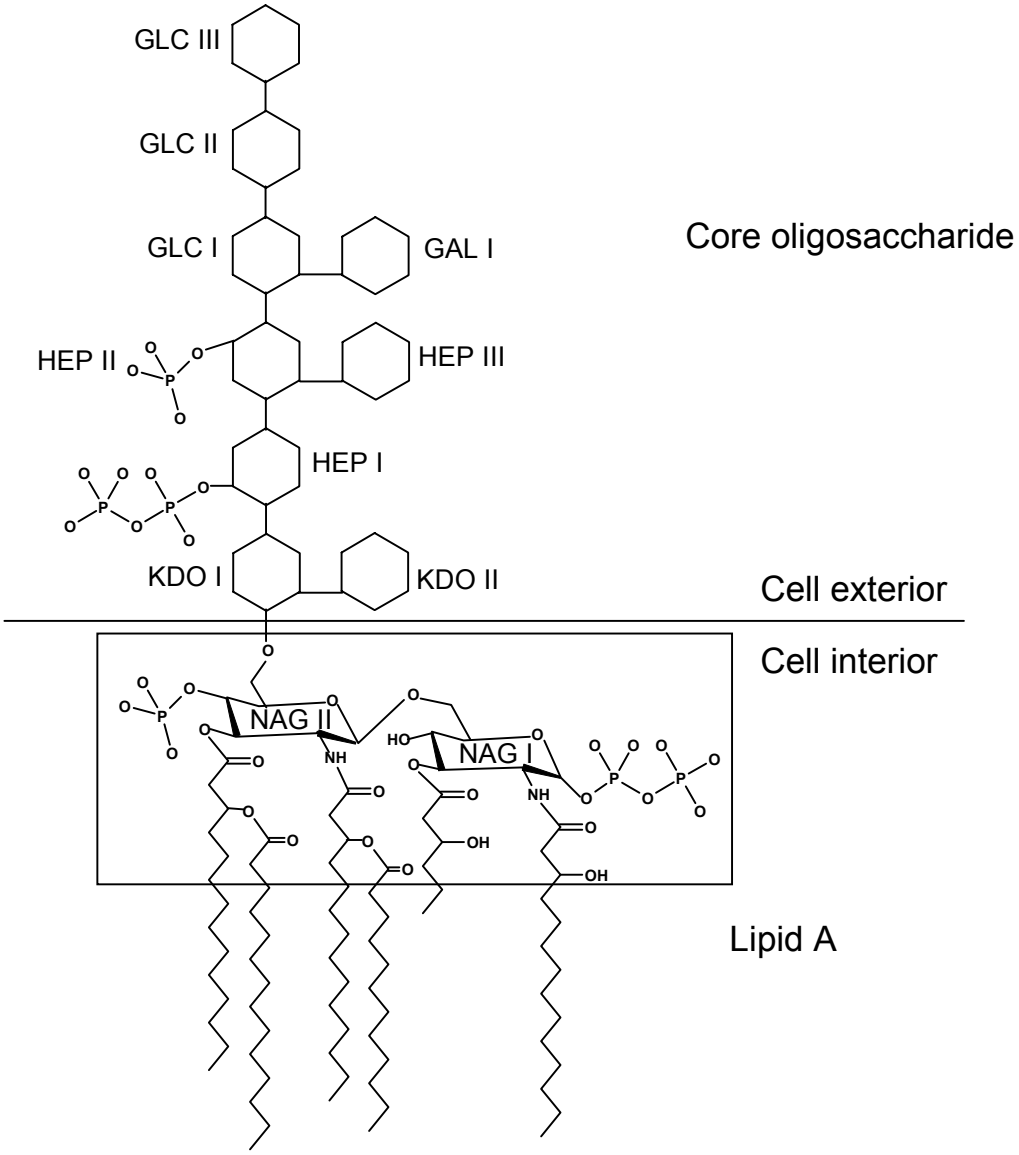


Fig. 1 – The structure of LPS from *E. coli* K-12: the lipid A part, the core oligosaccharide. LPSs are anchored in the outer membrane by their lipid A component, and their polysaccharide chain is presented at the cell surface.

The structural variability of the core oligosaccharide is higher than the structural variability of the lipid A.[2] This region contains unusual sugars with seven or eight carbon atoms like keto-deoxyoctulosonate (KDO) and heptose.

The minimal structure capable of expressing the endotoxic activities is considered to be a molecule that contains two D-glucosamine residues, two phosphoryl groups and six fatty acids in a defined arrangement.[2] The lipid A from *Escherichia coli* is the typical molecule with six fatty acid and endotoxic activity. Whereas the minimal LPS structure required to form a viable outer membrane consists of lipid A glycosylated with two KDO residues, often referred to as Re LPS for *E. coli.*[3] There is a large number of studies of Re LPS from *E.coli*, not only because it is the endotoxic molecule with the smallest molecular size and full endotoxic activities, but also because the genom of *E.coli* is easier to manipulate, and thus to determine the efects of mutations on the structure of LPS and the consequence on biological activity.[4]

Previous investigations into the supramolecular states of lipopolisaccharide and free lipid A have employed various techniques hydrolysis, mass

spectrometry, infrared spectroscopy, calorimetry, nuclear magnetic resonance spectroscopy and X-ray diffraction.[5,6,7] The reasons for the discrepancies in the proposed physical structures may be found in the difficulty in assigning the experimental results to definite three-dimensional structures.

The only 3D structure of an intact LPS molecule that is available was obtained by X-ray crystallography in an incidental complex with *E.coli* K-12 ferrichrome-iron receptor FhuA. There are two PDB codes for this structure 2FCP[8] and 1FI1.[9] The differences between the two crystals are due to the extraction methods, the one from 2FCP lost a phosphat group.

Lipid A and LPS from various bacteria have been previously modelled. The model of *Salmonella minnesota* S-form LPS by Labischinski et al[10] was based on a heptaacyl lipid A. Heptaacyl structure was originally assumed to be the typical component of *Salmonella* lipid A, but was later shown to have only weak biological activities.[11] In contrast, the hexaacyl lipid A determines the endotoxic activities of the LPS preparations from not only *E. coli* but also *Salmonella spp.*[12] Kastowsky et al[13] constructed the molecular models of *E. coli* Re LPS by the application of molecular modelling techniques. Although this seems to be the first conformational study of the complete structure of *E. coli* Re LPS, the models were not based on real analytical data, but exclusively on computer calculation. Therefore, these molecular models of *E. coli* Re LPS do not seem to be valid. Kastowski et al[14] also proposed the molecular model of *Salmonella* serogroup B S-form LPS that was constructed on the basis of their lowest energy *E.coli* lipid A model. Kato et al[2] provided the first molecular model of *E. coli* Re LPS based on analytical data from crystals of Re LPS. However the charges on atoms were ignored, so the not relevant enough considering the consequences of the highly charged inner core on biological activity of the LPS molecule in interaction with antibodies.

A molecular model for the rough LPS of *Pseudomonas aeruginosa* has been designed for a different force field than the one used in this study, based on experimentally determined structural information.[15] An electrostatic model was developed based on quantum chemical calculations of the complete LPS molecule to obtain partial atomic charges. However the molecular dynamics of a LPS membrane was run in a time scale that is too short to determine membrane properties like lateral diffusion of LPS or water permeation into the LPS membrane.[16]

The purpose of this study was to obtain parameters for LPS consistent with force fields that use an atomic detailed description of structure and dynamics like CHARMM,[17] NAMD[18] or LAMMPS.[19] The results can be introduced in these packages for larger sampling of conformational space and with less computational effort. They are compatible with parameters for proteins and lipids, so that more complicated systems are possible to simulate and for a longer period of time.

In the context of molecular mechanics, a force field refers to the functional form and parameter sets used to describe the potential energy of a system of particles. Force field functions and parameter sets are derived from both experimental work and high-level quantum mechanical calculations. In the atomistic approach, force fields provide parameters for every atom in a system, including hydrogen, while "united-atom" force fields treat groups of atoms as a single interaction center.

The basic functional form of a force field accounts for both bonded terms relating to atoms that are linked by covalent bonds, and nonbonded terms describing the electrostatic and van der Waals forces. A general form for the total energy in an additive force field can be written as $V=V_{bonded+}V_{nonbonded}$ where the components of the covalent and noncovalent contributions are given by the following summations: $V_{bonded} = V_{bond} + V_{angle} + V_{dihedral}$ and $V_{nonbonded} = V_{electrostatic} + V_{vanderWaals}$

In addition to the functional form of the potentials, a force field defines a set of parameters for each type of atom. The typical parameter set includes values for atomic mass, van der Waals radius, and partial charge for individual atoms, and equilibrium values of bond lengths, bond angles, and dihedral angles for pairs, triplets, and quandruplets of bonded atoms.

In our study we determined a set of parameters for the LPS molecule to be further used in molecular modeling simulations with CHARMM force field that uses the atomistic approach. These results can be easily adapted by the user to model different LPS molecules.

Topology and parameter files can be requested by email from the corresponding author.

## MATERIALS AND METHODS

The structure we used was the 3D structure of LPS obtained by X-ray crystallography in an incidental complex with *E.coli* K-12 ferrichrome-iron receptor FhuA(PDB code 1FI1).

Hydrogen atoms were added to this structure using the program software *The Extensible Computational Chemistry Environment* (ECCE).[20]

For the purpose of constructing models for fairly large molecules with various structures the whole molecule was divided into several fragments. This is allowed in additive force fields and it is preferred when there is a high variability in the structure of the molecule.

### Parameterization of lipid A region

In Figure 1, represented in a box, is the region from lipid A with missing parameters in CHARMM force field. The bonded terms of this set of parameters were determined by quantum chemistry calculations. The extensive computations were performed using quantum chemistry software program Gaussian 03.[21] The input for this software was obtained using the VMD[22] extension Paratool.

These calculations were set at the Restricted Hartree Fock (RHF) level of theory using the Pople 6-31G* basis set and they were completed on a quad dual-core Xeon computer equipped with 4GB RAM. The calculation of the bonded parameters required 30 days of CPU.

The connections between the atoms and the sets of internal coordinates from the topology file of lipid A were determined using the VMD extension Molefacture.

### Topology file for the "core" oligosaccharide

For the heptoses in the core oligosaccharide, that have the corresponding residues with six atoms in the saccharide topology file, [23] we have determined the topology for the atoms related to the seventh carbon atom using Molefacture. The same program was used to determine the complete topology for the KDO molecule, which is much more complex than the heptoses, having a carboxyl group bound in position C1. The parameters that correspond to these topologies were available in saccharide parameter file.[23]

### Construction of the LPS/DPPE bilayer

These parameters were tested by simulating a hydrated bilayer, with 16 molecules of LPS in one layer and 40 molecules of DPPE (1,2-dipalmitoylphosphoethanolamine) in the inner layer. In order to balance the charged functional groups on LPS counter-ions are required. 80 ions of $Ca^{2+}$ were added and the whole system was hydrated with 10986 TIP3 water molecules, the system reaching a total of 45878 atoms. NPT dynamics with periodic boundary conditions was run with a time step of 2 fs for a total of 5 ns. The target temperature of the Langevin thermostat was 310 K, with a damping coefficient of 1 $ps^{-1}$. The target pressure of 1 atmosphere was controlled by the Nosé-Hoover Langevin barostat.[24] Long-range electrostatic interactions contributions were evaluated using Ewald method.[25,26]

## RESULTS AND DISCUSSION

The nonbonded terms, namely the van der Waals term and the electrostatic term, were available in the CHARMM force field and therefore were assigned by analogy. The electric charge values for each atom in the two NAG residue and phosphate residue were assigned from the corresponding data in the saccharide topology file,[23] pirophosphate from the topology for NAD+, NADH and pirophosphat,[27] acyl chains from the topology for lipids.[28] These values were adjusted with less than 1% than the ones from similar atom types in CHARMM force field topology files,[23,27,28] so that the lipid chains were divided into neutral charge groups and the polar head into integral charge groups.

In order to make it easily to construct different versions of LPS molecules, the oligosaccharide region was built using the patching method. It is preferable to use parameters and topology of small fragments of the molecule and then covalently bind these fragments using the *patch* command. In this study we have separately obtained the parameters and topology files for the lipid A fragment and the topology for KDO and for new patches to build heptoses.

In Figure 2 is the charge distribution of the lipid A molecule. The numbers next to each charged atoms are the partial charges, in units of an electronic charge. The total charge of the lipid A molecule is -5, whereas the total charge of the LPS is -10. These charges correspond to the phosphate groups (-2) and to the pirophosphate groups (-3) and they are located in the inner core oligosaccharide as well as in the polar head of the lipid A, as shown in Figure 1. From the inspection of the model it can be observed that these phosphate groups are exposed to the surface of the molecule. This distribution of negative charge can explain the high affinity of the LPS for monoclonal antibodies unlike the low affinity of usual carbohydrates.

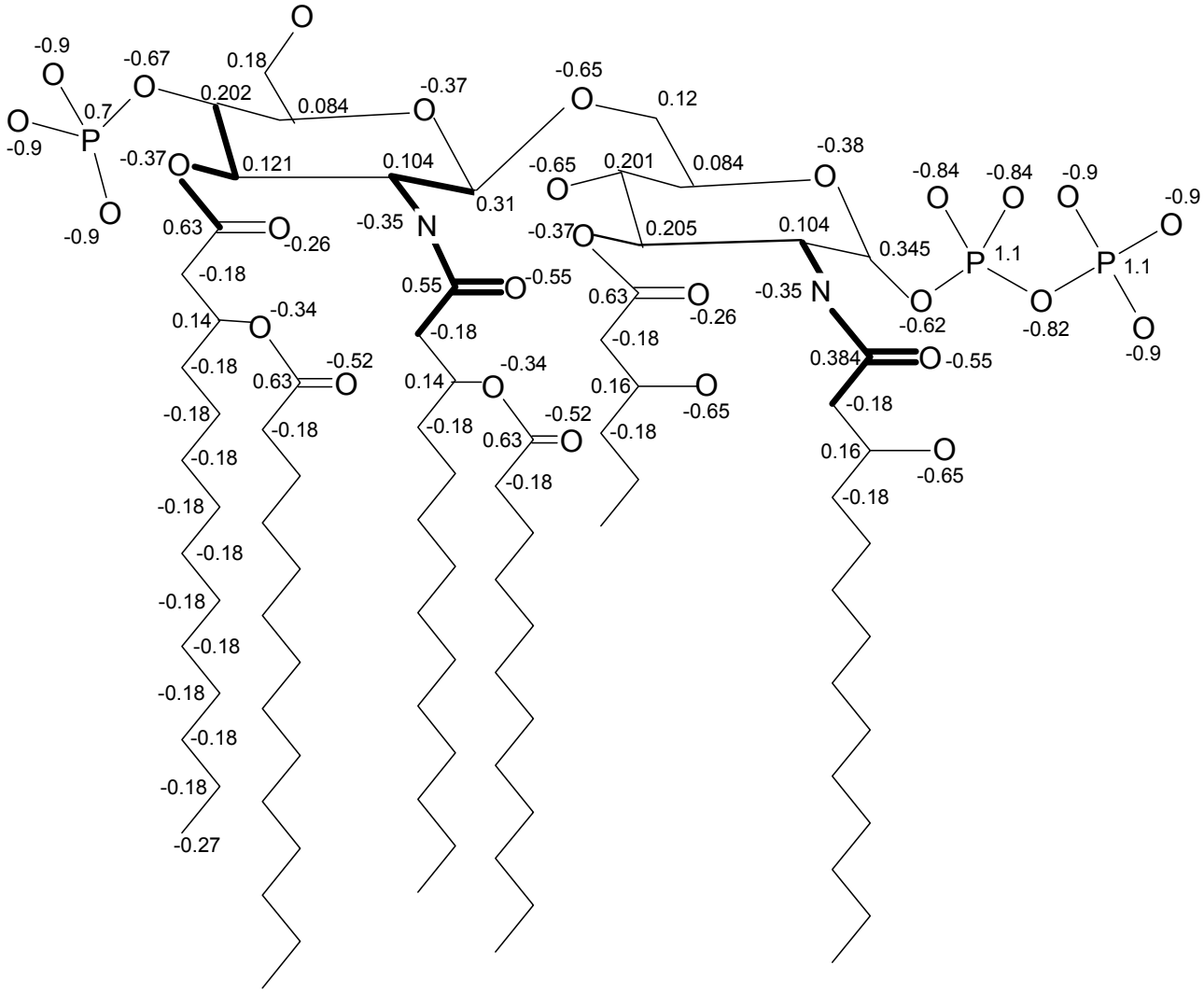


Fig. 2 – Charge distribution and torsion angles selected to be compared with experimental data, in thick line.

At the same time, due to the distribution of negative charge, the LPS molecule has a polyanionic structure can therefore bind divalent cations. The strong ionic interactions between the negatively charged polar head and the divalent counter ions is essential for the integrity of the membrane.[16]

The results of this study were used in a LPS/DPPE bilayer simulation with CHARMM and NAMD. The parameters were tested against experimental results. Here we present the comparison of 6 torsion angles (in bold lines in Figure 2) from the region near the polar groups, region recognized as the most dificult to parameterize, with the values determined experimentally through X-ray diffraction. For this purpose were extracted the values of the six torsion angles from the 20.000 conformations adopted by the 16 LPS molecules during 5 ns simulation. The histogram distributions of those dihedrals (in Figure 3) were obtained by splitting the -180 to 180 interval into $10^o$ bins. On the vertical is presented the frequency of the values for each bin. For the clarity of the graphs the results on vertical are arbitrarily scaled. Figure 3 shows different conformations of the chosen dihedral.

The crystal structure is a static structure and also restrained by the interaction with the protein and the rest of the crystal. Simulated dihedral distributions are consistent with the result from experiment. In more detail, Figure 3a shows the value from 2FCD, $121^o$, which is very similar with the value $127^o$ that corresponds to a local maximum and also the value from the crystal 1FI1, $109^o$, is similar with the local maximum $94^o$ from the distribution. The same for the value $98^o$ from 2FCD in Figure 3b with the local maximum $107^o$, $16^o$ in Figure 3c with $11^o$ , $168^o$ from 1FI1 in Figure 3e with $175^o$ and $-14^o$ from 2FCD with $-10^o$, $165^o$ from 2FCD in Figure 3f with $-172^o$.

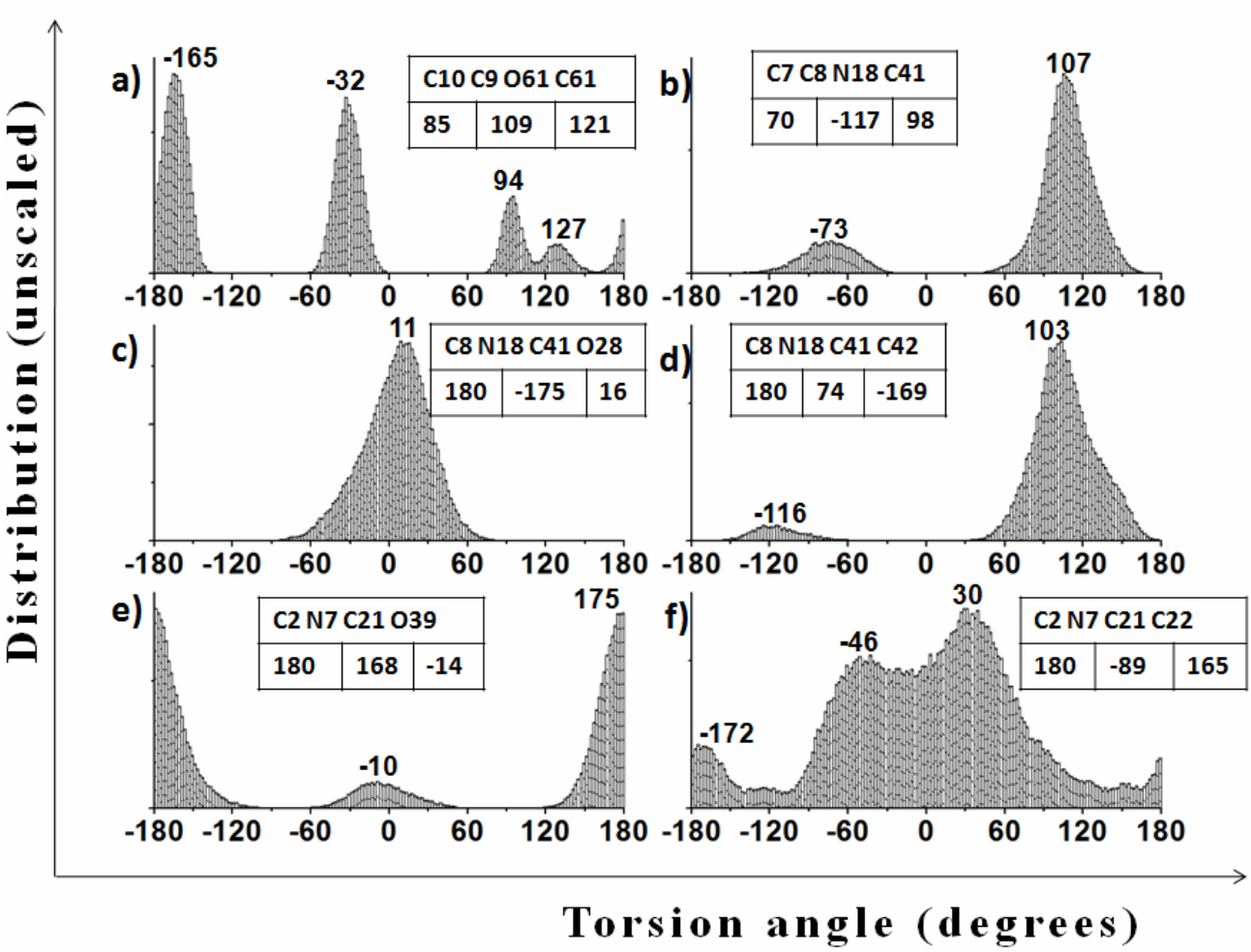


Fig. 3 – The distribution of the values for the selected torsion angles. First row of each table defines the torsion angle and the second row the equilibrium values determined in the parameter file (first column), the value from the crystal structure with PDB code 1FI1 (second column) and from 2FCD (third column).

There are some values in the distribution that do not correlate so well with the ones determined experimentally. A possible reason for that could be the short period of time used in the simulation. As support for this statement, it can be observed that dihedral in Figure 3c is defined by the same atom types as the dihedral from Figure 3e and thus should adopt similar conformations. Therefore the value -175 $^{o}$ from the crystal 1FI1 could be adopted by this dihedral, as shown by the local maximum 175 $^{o}$ in Figure 3e. There is a possibility this could happen in a simulation for a longer period of time.

The same situation can be observed in Figure 3d where the dihedral is defined by the same atom types as the one from Figure 3f and should adopt similar conformations. Therefore the value -169 $^{o}$ from the crystal 2FCD could be adopted by this torsion angle, as shown by the local maximum -172 $^{o}$ in Figure 3f.

In this study we have obtained a valid set of CHARMM parameters for the LPS that can be easily integrated into NAMD, CHARMM, LAMMPS packages to simulate complex systems with proteins and lipids. The topology is flexible for different molecules by using the newly determined patches.

Topology and parameter files can be obtained by email from the authors.

## REFERENCES

1. D. S. Snyder, and T. J. McIntosh, *Biochemistry*, **2000**, *39*, 11777-11787.
2. N. Kato, T. Sugiyama, S. Naito, Y. Arakawa, H. Ito, N. Kido, M. Ohta, and K. Sasaki, *Mol. Microbiol.*, **2000**, *36*, 796-805.
3. D. E. Heinrichs, J. A. Yethon, and C. Whitfield, *Mol. Microbiol.*, **1998**, *30*, 221-232.
4. J. A. Yethon, and C. Whitfield, *J. Biol. Chem.*, **2001**, *276*, 5498-5504.
5. M. Caroff, and D. Karibian, *Carbohydr. Res.*, **2003**, *338*, 2431-2447.
6. K. Brandenburg, and U. Seydel, *Eur. J. Biochem.*, **1990**, *191*, 229-236.

7. S.M. Strain, S.W. Fesik, and I.M. Armitage, *J. Biol. Chem.*, **1983**, *258*, 13466-13477.
8. A. D. Ferguson, E. Hofmann, J. W. Coulton, K. Diederichs, and W. Welte, *Science*, **1998**, *282*, 2215-2220.
9. A. D. Ferguson, J. Kodding, G. Walker, C. Bos, J. W. Coulton, K. Diederichs, V. Braun, and W. Welte, *Structure*, **2001**, *9*, 707-716.
10. H. Labischinski, G. Barnickel, H. Bradaczek, D. Naumann, E. T. Rietschel, and P. Giesbrecht, *J. Bacteriol.,* **1985**, *162*, 9-20.
11. S. Kotani, H. Takada, I. Takahashi, M. Tsujimoto, T. Ogawa, T. Ikeda, K. Harada, H. Okamura, T. Tamura, and S. Tanaka, S., *Infect. Immun.*, **1986**, *52*, 872-884.
12. S. Kotani, H., Takada, M. Tsujimoto, T. Ogawa, I. Takahashi, T. Ikeda, K. Otsuka, H. Shimauchi, N. Kasai, J. Mashimo, *Infect. Immun.*, **1985**, *49*, 225-237.
13. M. Kastowsky, A. Sabisch, T. Gutberlet, and H. Bradaczek*, Eur. J. Biochem*., **1991**, *197*, 707-716.
14. M. Kastowsky, T. Gutberlet, and H. Bradaczek, *J. Bacteriol.*, **1992**, *174*, 4798-4806.
15. R. D. Lins, and T. P. Straatsma, *Biophys. J.*, **2001**, *81*, 1037-1046.
16. T. A. Soares, and T. P. Straatsma, *Mol. Simulat.*, **2008**, *34*, 295–307.
17. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comp. Chem*., **1983**, *4*, 187-217.
18. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, *J. Comp. Chem.*, **2005**, *26*, 1781-1802.
19. S. J. Plimpton, *J. Comp. Phys.*, **1995**, *117*, 1-19.
20. G. Black, J. Chase, J. Chatterton, J. Daily, T. Elsethagen, D. Feller, D. Gracio, D. Jones, T. Keller, C. Lansing, S. Matsumoto, B. Palmer, M. Peterson, K. Schuchardt, E. Stephan, L. Sun, K. Swanson, H. Taylor, G. Thomas, E. Vorpagel, T. Windus, C. Winters, Pacific Northwest National Laboratory, Richland, Washington 99352-0999, USA.
21. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, Al M. A. Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople: Gaussian, Inc., Wallingford CT, 2004.
22. W.Humphrey, A. Dalke and K. Schulten, *J. Molec. Graphics*, **1996**, *14*, 33-38.
23. Molecular Simulations Inc. Parameter file and topology file for CHARMm version 22. Parameter file copyright 1993 and release February 1994. Topology file copyright 1992 and release September 1994).
24. S. E. Feller, Y. H. Zhang, R. W. Pastor, B. R. Brooks, *J. Chem. Phys.,* **1995,** *103*, 4613-4621.
25. T. Darden, D. York, L. Pedersen, *J. Chem. Phys.,* **1993**, *98*, 10089-10092.
26. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.,* **1995**, *103*, 8577-8593.
27. J. J. Pavelites, P. A. Bash, J. Gao, and A. D. MacKerell, *J. Comp. Chem.*, **1997**, *18,* 221-239.
28. S. Feller, and A. D. MacKerell, *J. Phys. Chem.*, **2000**, *104,* 7510-7515.