Post-Processing Challenges in Simulations of Lipopolysaccharide Membranes and Aggregates

Denys E. S. Santos, Victor H. Rusu, Frederico J. S. Pontes, Roberta P. Dias, Keila C. Cunha, Laércio Pol-Fachin, Thereza A. Soares and Roberto D. Lins

Fundamental Chemistry Department, Federal University of Pernambuco
Recife, PE, 50.740-560, Brazil

denys.ewerton@ufpe.br, victorusu@gmail.com, roberto.lins@ufpe.br

Abstract. *Computer simulation of Gram-negative bacteria membrane is a computational intense task that requires massive parallel computing. In addition, these simulations require post-processing tools that are in general not parallelized. Here we comment on the types of parallelizable analysis tools and propose two new types of analyses that can be applied on such systems.*

1. Introduction

One of the fundamental goals of the study of matter by computer simulations is the description of measurable macroscopic properties in terms of microscopic ones. The successful application of such methods strongly relies on post-processing tools to fully understand and interpret these simulations. The main goal of post-processing tools is to gather quantitative and qualitative information in order to boost the understanding and interpretation of the simulation data. In the field of molecular dynamics (MD), the simulation engines have been extended to take advantage of improved sampling methods that, combined to, scalable algorithms allow for faster and longer simulations. Since membrane fluctuations and motions span a wide range of time scales, large trajectory files are generated. This problem is aggravated when simulating Gram-negative bacterial membranes. These membranes are formed by molecules of lipopolysaccharides (LPS) anchored to a phospholipid leaflet. LPS comprises of a highly-charged carbohydrate moiety covalently bound to a variable number (4 to 7) of acyl chains. Its shortest head group is formed by a couple of phosphorylated N-acetyl-glucosamine molecules distributed along the lamellar plane. The dynamics of the LPS membrane tends to be slower compared to cell membranes due to its heavily charged carbohydrate moiety. In some cases, microsecond simulations are needed to produce meaningful data (Kirschner, 2012). To assess these phenomena *in silico,* new types of analysis and faster post-processing tools are needed.

One of our research goals is to develop analysis tools to extend the understanding of Gram-negative bacterial membrane structure and polymorphism from MD simulations. Bacteria biological activity depends on the intrinsic conformation of its cell wall molecules (Figure 1), allowing for drug resistance and/or pathogenic and immunological behaviors (Pontes, 2012). The quantitative structural classification of bacterial membranes is a difficult task to be performed experimentally. Despite its potential toxicity and therefore handling restrictions, these quantities have large intrinsic errors associated that can reach up to 240% (Seydel, 2000). To overcome this limitation, we have been developing novel analysis tools capable of evaluating the influence of

physical chemical factors (ions, temperature, pH) as well as chemical variations in LPS on the conformational polymorphism of Gram-negative bacterial membranes.

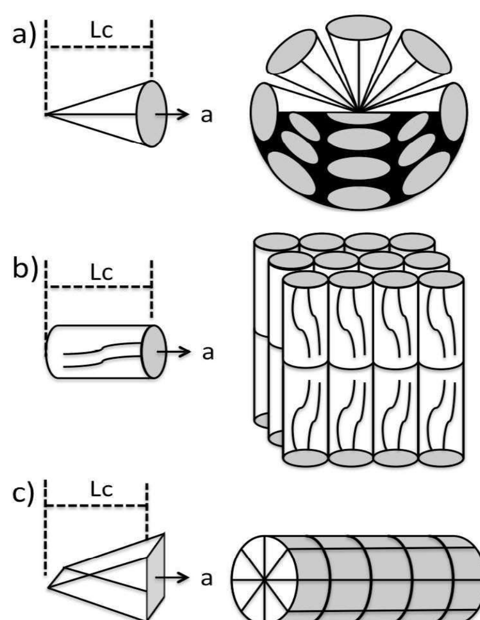

Figure 1. Typical membrane conformational states: a) lamellar; b) micelle c) planar aggregate.

As it can be seen in Figure 1, lamellar to non-lamellar transition involves either negative or positive curvature of the LPS membrane plane. To assess these transitions in molecular dynamics trajectories, we have developed a sequential post-processing tool written in C language and based on GROMACS library for full compatibility with GROMACS input and output files (Hess, 2008). The program measures the distribution of the sugar-sugar-membrane plane and head-tail-membrane normal vector angles. It has been applied to trajectories of hundreds of nanoseconds for six different Gram-negative membrane phenotypes. Our results showed shifts towards higher angle values as a function of increasing in temperature for both angles. The expected result allowed us to quantify membrane disorder as a function of LPS phenotypical variation (Pontes, 2012). To date, the code has been used on this single occasion and therefore systematic performance benchmarks have not been carried out.

2. Analysis tools performance

Sequential post-processing tools do not scale nor have a performance compatible with very long MD trajectories (Peterson, 2008; Kumar 2012a & Kumar 2012b). Nowadays, MD engines are able to produce terabytes of simulation trajectories and analyzing these data imposes a stress on CPU, memory and I/O. To illustrate, radial distribution function (RDF) and linear mass density are typical analyses performed on membrane simulations. The RDF algorithms require computation of three nested loops: i. the outer-loop over the entire trajectory; ii. the middle-loop over the reference set of atoms; and, iii. the inner-loop over the complementary set of atoms. Both tools were run in a single node since they are not parallelized. The ratio between the times spent on the RDF calculation between solute and solvent molecules divided by the total time spent by the program is shown in Figure 2.

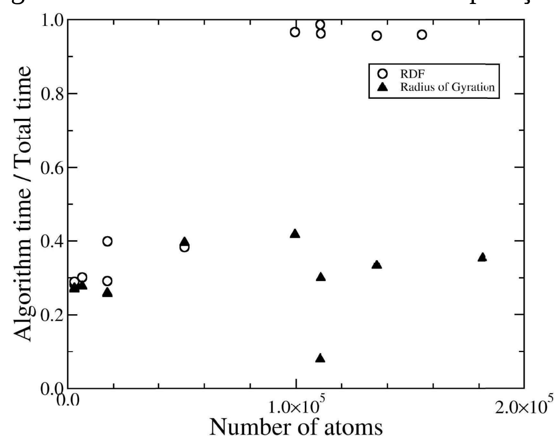

Figure 2. Algorithm time divided by the total program time versus the number of atoms for the RDF and Radius of gyration analyses.

For small systems the parallelization of the RDF code would not improve the performance significantly, however with the increase in the number of particles the RDF calculation spends 98% of the computing time.

3. Membrane curvature analysis

Depending on the ionic strength and temperature Gram-negative bacterial membranes undergo phase transitions that promote membrane shape changes (Brandenburg, 1988; Brandenburg, 1993; Brandenburg, 1997 & Seydel, 2000). The simple analysis of the foretold angles is not enough to classify the membrane conformational state. Ripple-like shapes (large negative followed by large positive curvatures) cannot be quantitatively expressed in terms of the average tilt angles. To address this issue, we propose two alternatives to measure membrane curvatures: using the Voronoi tessellation method and by measuring the deviation of the membrane leaflets packed to a cylinder.

In the Voronoi tessellation method the seeds are the membrane head-group atoms and the curvature can be measured by calculating the angle formed between two consecutives seeds and the membrane normal vector (transmembrane plane) along all the axis parallel to the membrane (Franz, 1991). Other properties measurements can also take advantage of this approach, e.g., one can calculate the curvature dependent density profile of the system along the membrane normal vector. The ratio between the effective volume of the molecular subunit that constitute a membrane and a theoretical cylinder-like shape dictates how far the membrane can be in the lamellar state (ratio < 0.3 and < 0.5) (Figure 1a), the micelle state (ratio < 0.3) (Figure 1b) and planar aggregate (ratio > 0.5 and < 1) (Figure 1c). However, once the computation of the effective volume of the molecular subunit is carried out for each phospholipid composing the membrane and over several million frames, this task becomes highly demanding. Therefore, a parallel implementation of this algorithm is a well sought-after feature. Different approaches can be used like the Gauss-quadrature or the Monte Carlo method.

4. Conclusions

Simulation of LPS membranes and aggregates is a computational intense task that is performed on parallel clusters. Depending on the bacterial membrane phenotype, μs simulations are required to achieve an equilibrated system (Kirschner, 2012). Typically it is necessary between 30,000 to 60,000 core hours to simulate a single system. However, the tools used for post-processing are not parallel, which results in an inefficient time to solution for the user, as demonstrated for the calculation of the radial distribution function. One of the most used analyses that can take advantage of parallelization is the radial distribution function, Voronoi tessellation and the membrane volume analysis. The latter have never been applied for membrane systems. The development of these analyses tools is an ongoing effort in our group. Taken together, they will be capable to characterize membrane structural polymorphism in an unambiguously manner.

4. Acknowledgments

This work was developed with financial support from FACEPE, NanoBiotec-BR/CAPES, CNPq, INCT-INAMI and nBioNet. Computational resources were provided by Argonne Leadership Computing Facility, located at Argonne National Laboratory, the Swiss Supercomputing Centre (CSCS), the University of Umeå Supercomputing Centre (HPC2N) and CENAPAD-PE.

5. References

Aurenhammer, F. (1991). *ACM Computing Surveys*, 23, pages 345-405.

Brandenburg, K. and Seydel, U. (1988). *European Journal of Biochemistry* 16, pages 83-87.

Brandenburg, K., Mayer, H., Koch, M. H., Weckesser, J., Rietschel, E. T. and Seydel, U. (1993). European Journal of Biochemistry 218, pages 555-559.

Brandenburg, K., Kusumoto, S. and Seydel, U. (1997). *Biochimica et Biophysica Acta* 1329, pages 183-201.

Hess, B., Kutzner, C., van der Spoel, D. and Lindahl, E. (2008). *Journal of Chemical Theory Computation,* 4, pages 435-447.

Kirschner, K. N., Lins, R. D., Maass, A. and Soares, T. A. (2012). *Journal of Chemical Theory and Computation*, 11, pages 4719-4731.

Kumar, A., Grupcev, V., Yuan, Y., Tu, Y-C. and Shen, G. (2012a). *EDTB*, pages 26-30.

Tu, Y-C., Chen, S., Pandit, S., Kumar, A. and Grupcev, V. (2012b). *ACM Conference on Bioinformatics, Computational Biology and Biomedicine*, pages 527-529.

Peterson, E. S. ; William, L. ; Stephan, E. R. ; Rose, S. ; Corrighan, A. ; Lai, C. ; Lins, Roberto D., Critchlow, T., Straatsma, T. P., Soares, Thereza A., Scarberry, R. (2008). *BIOCOMP*, pages 558-562.

Pontes, F. J. S., Rusu, V. H., Soares, T. A. and Lins, R. D. (2012). *Journal of Chemical Theory and Computation*, 8, pages 3830-3838.

Seydel, U., Oikawa, M., Fukase, K., Kusumoto, S. and Brandenburg, K. (2000). *European Journal of Biochemistry*, 267, pages 3032-3039.