

Modelling Biological Polymeric Nanostructures in Physiological Fluids: Focus on Ribonucleic Acid Nanotubes

Shyam Badu¹, Roderick Melnik^{1,2}, Sanjay Prabhakar^{1,2}

¹MS2Discovery Interdisciplinary Research Institute, M²NeT Laboratory
Wilfrid Laurier University, 75 University Ave W, Waterloo, ON, Canada N2L 3C5
m2netlab@gmail.com; rmelnik@wlu.ca

²BCAM - Basque Center for Applied Mathematics
Mazarredo 14, E48009 Bilbao, Basque Country - Spain
rmelnik@bcamath.org

Abstract - In this paper, we study biological polymeric molecular nanostructures, focusing on ribonucleic acids (RNA) and their behaviour in fluids, in particular in physiological solutions. First, we give a brief overview of the most recent results in this field related to the properties of these structures that can be characterised by root mean square deviation, radius of gyration and radial distribution function. Among other things, we provide insight into typical distributions of ²³Na⁺ and ³⁵Cl⁻ ions around the RNA nanotubes as a function of time within a distance of 5 Å from their surface. Finally, we provide details of the recent achievements in the coarse grain modeling of the such biological nanotubes.

Keywords: Biological polymeric molecular nanostructures, Ribonucleic acids, Molecular dynamics simulation, Fluids and physiological solutions

1. Introduction

The self assembly of biological molecules such as DNA and RNA, and a range of different proteins can assist in creating the nanobelts in fluidic solutions which can be used for many therapeutic applications [1]. Due to the importance of therapeutic applications, several kinds of polygon-shaped self assemblies have been developed by using the deoxyribonucleic acid [2, 3].

By now it is clear that the self assembly of biomolecules [4, 5], can be used for a wide range of novel bionanodevices in nanobiotechnology. It has been demonstrated that the stability of the ribonucleic acid (RNA) assemblies is higher than that of the DNA self assembled nanoparticles in fluidic solutions [6, 7]. The differently shaped structures of the RNA molecules have been formed from the RNA building blocks as well as from their complexes with other biomolecules [8, 9, 10]. Given the increasing importance of such RNA structures for nanobiotechnology, and in particular in scaffolding and drug delivery applications, our focus in this paper is on RNA nanostructures. In particular, the building blocks of nanorings, studied in earlier works [11, 12], are engineered in such a way that the RNAI and RNAII ends are complementary to each other. The complexation and coordination of these two RNA fragments via sticky ends is an important feature for constructing nanoclusters like RNA nanorings and nanotubes. By using six helical building blocks of either one or two types (RNAI/RNAII) the nanoring is formed by self-assembling them via base pairing hydrogen bonds. Multiple nanorings with their links are used to form the nanotube which can operate in fluidic physiological solutions. The latter nanostructures will be in the main focus of the rest of this paper.

2. Boltzmann Inversion Method in Multiscale Modeling: Dynamic and Thermal Characteristics

The main methodology used here for the analysis of behaviour of RNA nanostructures is based on molecular dynamics simulation. The methodology provides a key to the dynamic behaviour of the biomolecules in fluidic physiological solutions. It also provides the important information on the thermal properties of the biomolecules under study and prediction of the thermally compatible conformation of the molecules [13]. The basis of the developed methodology molecular lies with statistical thermodynamics models. Given the multiscale nature of the problem, the calculation of the properties of the system at one level is done using the models from a different level. While there is a range of multiscale modeling methodologies, the most popular are:

1. The Boltzmann inversion method
2. The force matching method for developing the coarse-grained modeling
3. Multiscale coupling method for direct transfer of the information from mesoscopic and atomic scales during the simulation.

In the present work we have applied all three above methods at the different stages of the modeling process. The Boltzmann inversion method, in particular, has been used for the coarse-grained modeling of the RNA nanoclusters which is described in the next section. Here we will recall that in the force matching process [14], the objective function, depending upon the parameter α , is defined follows

$$Z(\alpha) = Z_F(\alpha) + Z_c(\alpha), \quad (1)$$

$$Z_F(\alpha) = \left(3 \sum_{k=1}^M N_k \right)^{-1} \sum_{k=1}^M \sum_{i=1}^{N_k} |F_{ki}(\alpha) - F_{ki}^0|^2, \quad (2)$$

$$Z_C(\alpha) = \sum_{r=1}^{N_C} W_r |A_r(\alpha) - A_r^0|^2. \quad (3)$$

The integer M in Z_F (the force objective function) is the number of configurations, N_k is the number of atoms in the k -th configuration and $F_{ki}(\alpha)$ is the force on the i th atom in the k th configuration which is obtained from the parametrization of α , and the F_{ki}^0 is the corresponding reference force obtained from the first principles calculations. In the constraint objective function Z_C the quantities $A_r(\alpha)$ are also physical parameters obtained from parametrization, A_r^0 are experimental values or the values calculated from the first principles methods and W_r is the weight factor. The force objective function defined in equation (1) is minimized for given α to calculate the classical force parameters by using the force and physical quantities obtained from *ab initio* calculations. Therefore, in order to calculate the force objective function, and constraint objective function it is necessary to do the *ab initio* calculations. The parameters α defined in the above equations (1.1) -(1.3) are calculated by matching the forces obtained by using the first-principles calculations of the several configurations of the molecular system and the classical potentials [15]. Clearly that at the mesoscopic level, one can take advantage of the existing effective field theories of continuum mechanics, but our main interest in this paper lies the smaller scales.

3. Computational Coarse-Grained Models, Results, and Discussions

In modeling biological systems such as ours, it is inevitable to resort to coarse-grained models which should be designed to explain information about the system at larger scales from the smaller scale that are modeled from the atomistic classical approach. The developed coarse-grained models should be easy enough to simulate accurately enough the physical characteristics of the system. In the coarse-grained modeling we represent the sum of atoms as a pseudo atom and then define an effective energy function U_{CG} that determines the thermodynamical properties, which should be identical to the system's properties once the proper energy function is predicted.

Recently, the modeling of the coarse-grained structure of RNA and RNA-Protein using the fluctuation matching method has been performed [16], in which the authors also followed the assumptions used in [17]. There are also several other investigations done on coarse grained modeling of RNA for the prediction of the tertiary structures [18]. In one of the earlier studies on the coarse-grained modeling of the RNA 3D structure, a single nucleobase has been approximated by five pseudo atoms [19]. In order to determine the forcefield parameters the 688 experimentally determined structures of RNA have been used.

3.1. Computational Model Implementation

The entire system is integrated in time. The potential of the system used during such integration uses the CHARMM force field. It is defined as follows:

$$\begin{aligned} V_{total} = & \sum_{bond} K_b(r-r_0)^2 + \sum_{angle} K_\theta(\theta-\theta_0)^2 + \sum_{dihedral} K_\phi(1+\cos(n\phi-\gamma)) \\ & + \sum_{Hbond} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + \sum_{impropers} K_w(w-w_0)^2 \\ & + \sum_{Urey-Bradley} K_u(u-u_0)^2 + \sum_{Vanderwaals} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{10}} \right) + \sum \frac{q_i q_j}{\epsilon r_{ij}} \end{aligned} \quad (4)$$

In (4), the first term corresponds to bonds, second corresponding to angle parameters, the third term corresponds to the potential energy and interactions arising from the dihedral angles in the molecular system, the fourth term defines the interaction coming from the hydrogen bonds which includes the base pairing as well as the hydrogen bonding between the RNA and the water molecules. The fifth term known as the improper term that arises due to out-of-plane bending of molecular system and the sixth term is the Urey-Bradley contribution. The improper term is included in the potential energy expression to maintain the planarity of the molecule. Finally, the last term in the potential expression represents the long distance interactions known as the van der Waals' interactions. We have performed all-atom molecular dynamics simulations of RNA nanotubes by using the CHARMM27 force field [20] implemented in the NAMD package [21] as it was done for the nanoring [11, 12].

In the studies [22, 23] the development of the CHARMM force field was carried out and its compatibility for the DNA and the RNA has been tested. It was found that the results are close to the experimental results.

The modeling of the nanotube, visualization and the analysis of the simulation outputs have been performed using the software visual molecular dynamics (VMD) [24]. The VMD is a molecular graphics software developed to display the biomolecular systems like biopolymers and proteins interactively. In this program the molecules can be viewed in several colors and several kind of representations. This can allow us to modify a particular protein structure as needed. In particular, one can do mutation, deletion or addition of bonds between the atoms using VMD tool. Furthermore, we can display several structures at the same time using VMD. The trajectory of the molecular dynamics simulation can be displayed as well as analysed to study the molecular properties. The VMD program is written in C++ and is provided with the complete documentation with instructions to use it. This program is compatible with several kinds of molecular dynamics simulation packages including NAMD, LAMMPP, GROMACS etc.

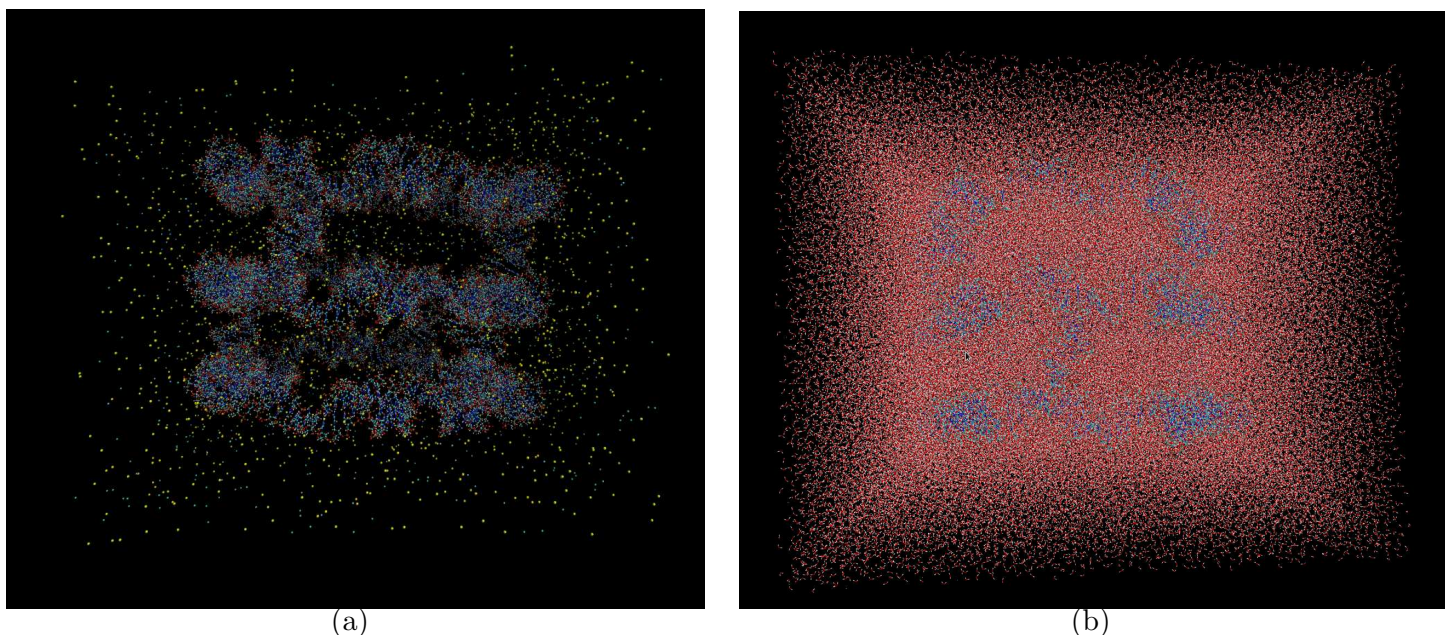


Fig. 1: (a) Modeling RNA nanotubes and ions without water and (b) modeling RNA nanotubes in fluidic physiological solutions.

In our typical runs, the RNA-nanotube has been solvated in a water box. The size of the box is taken in such a way that the distance from the surface of the nanocluster to the wall is slightly larger than the cut off radius used in the molecular dynamics simulation. In order to make the system neutral we have added 924, and 1254 $^{23}\text{Na}^+$ ions for three ring and four ring nanotubes, respectively. Furthermore, to make the solution equivalent to physiological solutions we have added extra 924 and 1254 $^{23}\text{Na}^+$ and $^{35}\text{Cl}^-$ ions to the three ring and four ring nanotube, respectively. The resulting system has been first simulated at constant temperature and pressure using the NAMD software package. The temperature in the system has been controlled by using Langevin's method with damping $\eta = 5 \text{ ps}^{-1}$. For adding chemical bonds between the segments in the nanoclusters we have used the topotools available in the VMD.

3.2. Representative Examples and Their Discussions

In what follows, we will describe a few typical results of modeling the RNA nanotubes with multiple nanorings. Some of the earlier results have been presented in our papers [25, 12]. For the modeling of RNA nanotubes the hexagonal nanorings were connected to each other by using the links between them as described in our earlier studies [11, 25, 12]. Typical sample structures of the three ring nanotube without water and with water are presented in Figure 1(a) and (b) respectively. The six helical segments are constructed from RNAI and RNAII building blocks. Also, the tails used to connect the RNA nanorings are the double strand RNAs with the length of 22 nucleotides.

As we discussed earlier, the RNAI and RNAII are the double strand RNAs. By using the VMD tools we were able to connect multiple rings via three links at junctions presented in Figure 1(b). The links used in connecting the multiple numbers of nanorings (to build the RNA nanotube) are composed of helical double strand RNAs with 22 nucleobases. Three links are used in between two consecutive rings to connect them to form the nanotubes as shown in Figure 1(b). The chemical bonds between the ring and the links are mediated through the phosphorous of the phosphate group in the ring and the oxygen in the sugar ring of the corresponding link or vice versa. Using NAMD, we optimized the chemical bonds added between different segments of the RNA nanoclusters.

Here at first, we present the results for the RNA nanotube of different sizes obtained from the molecular dynamics simulation. The results for the simulation of three the ring nanotube have been analyzed for the variation of the energy and temperature as a function of simulation time, and they can be summarized as follows. At the beginning of the simulation the energy of the system varies and then becomes stable once the system becomes stabilized.

Other characteristics that were analyzed included also the calculated properties such as the number of ions around the RNA nanotube within the distance of 5 \AA at different temperatures, the number of bonds per basepairs, the radius of gyration and the root mean square deviation (RMSD) at two temperatures, 310K and 510K. The results corresponding to variations of the parameters are similar to the results obtained for the other nanoclusters described in our earlier studies [11, 25, 12].

The radial distribution functions have been calculated for our RNA nanotubes for phosphorous-phosphorous, phosphorous-water, phosphorous-sodium and phosphorous-chlorine. For example, from the P-P RDF analysis, we have concluded that there are three well-pronounced peaks around the same positions at it was observed for other nanoclusters studied in our earlier paper [12]. These peaks actually show the first, second and third nearest neighbours of the phosphorous atom respectively. The intensity of the peaks is increased on going from 310K to 510K. The position of the first peak is at the same position, whereas the second and third peaks are shifted slightly to the lower distances at 510K in comparison to their positions at 310K.

Likewise, from the P-OH2 RDF analysis (carried out for two different temperatures of 310K and 510K), it has been concluded that for each RDF there is a peak around the distance of 4 \AA . This first peak indicates the first solvation shell around the phosphorous

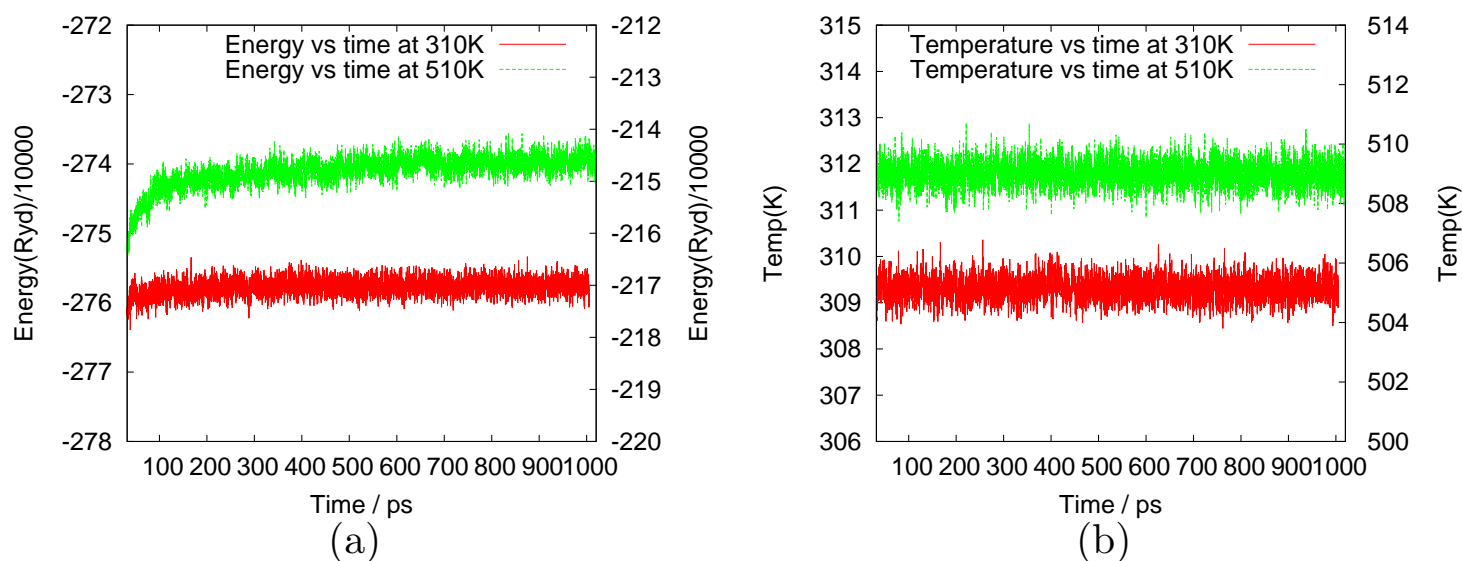


Fig. 2: The evolutions of energy (a) and temperature (b) as functions of simulation time for the all-atom molecular dynamics (the example shown is for a four-ring RNA nanotube in water and salt).

atom taken from the surface of the RNA nanotube. Similarly, the second small peak shows the second solvation peak for the phosphorous atom in the phosphate backbone of RNA strands that builds the RNA nanotubes. In the rest of the range, the nature of the P-OH2 RDF plots remained more or less stable showing that the water molecules are distributed uniformly after certain distance from the surface of the RNA nanotube. In spite of showing a similar trend at both temperatures, the height of the first peak is significantly dropped on going from 310K to 510K. This indicates that a significant amount of water molecules are expelled out from the surface of the RNA nanocluster at higher temperatures as demonstrated in our earlier papers [11, 12].

It is important to note that the results for larger RNA nanotubes (with larger numbers of nanorings) demonstrated the nature of solvation and the ionic distribution during the molecular dynamics simulation to be similar to those found in the case of the three-ring nanotube, as well as to the results described in our earlier work [11, 12]. A typical example of modeling for a four-ring nanotube is presented in Figure 2.

For all of these systems we see that the peaks for the P-P RDF remain almost the same, for P-OH2 the intensity at the peak is decreased on increasing the temperature, but in the P-Na and P-Cl RDF plots the intensity of the first peaks is significantly increased on going from 310K to 510K temperature. From these observations we can conclude that the $^{23}\text{Na}^+$ and $^{35}\text{Cl}^-$ ions are attracted toward the surface of the nanocluster and the water molecules are pushed away from the surface as the temperature is increased. In short, we observe that the ions are being precipitated around the surface of the RNA nanocluster as the temperature of the system is increased. This phenomenon of self stabilization was first discovered in [11] and has recently been observed for a larger class of RNA nanoclusters in our earlier work [12].

4. Conclusion

In this paper we discussed some of the key aspects of modeling the ribonucleic acid nanostructures in fluidic physiological solutions. Our motivation was based on the fact that the RNA is more versatile and compatible for the biomedical applications in the human body, compared to other alternatives, due to its flexibility in structure as well as the absence of toxicity.

Our typical examples included the optimized structures of nanotubes up to the size of 40nm. The individual RNA nanorings were connected via double-helical rings mediated by the bonds between the phosphate group and sugar rings. The newly added bond lengths have been optimized by using algorithms available in NAMD. Our analysis of RNA nanotube structures was based on the root mean square deviation, radius of gyration, number of hydrogen bonds per basepair, ion accumulation around the tube, and the radial distribution functions. For drug delivery and other biomedical applications further computational efforts will be dependent on the shape of the nanocluster, on the drug particle to be delivered, as well as the type of fluidic physiological solution to be used. All these factors will influence the complexity of the problems to be addressed in the future.

Acknowledgments

Authors are grateful to the NSERC and CRC Programs for their support. R. M. was also supported by the Bizkaia Talent Grant under the Basque Government through the BERC 2014-2017 program and by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2013-0323. The authors are also gratefully acknowledge the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca) for providing the computational facilities.

References

- [1] H. Cui, T. Muraoka, A. G. Cheetham, S. I. Stupp, *Nano Lett*, vol. 9, no. 3, pp. 945, 2009.
- [2] F. A. Aldaye and H. F. Sleiman, *J. Am. Chem. Soc.*, vol. 129, no. 44, pp. 13376, 2007.

- [3] S. H. Park, R. Barish, H. Li, J. H. Reif, G. Finkelstein, H. Yan, and T. H. LaBean, *Nano Lett.*, vol. 5, no. 4, pp. 693, 2005.
- [4] T. P. J. Knowles, T. W. Oppenheim, A. K. Buell, D. Y. Chirgadze, and M. E. Welland, *Nat Nano*, vol. 5, no. 3, pp. 204, 2010.
- [5] M. Endo, T. Sugita, Y. Katsuda, K. Hidaka, and H. Sugiyama, *Chem. Eur. J.*, vol. 16, no. 18, pp. 5362, 2010.
- [6] W. W. Grabow, L. Jaeger, *Acc. Chem. Res.*, vol. 47, no. 6, pp. 1871, 2014.
- [7] K. A. Afonin, W. Kasprzak, E. Bindewald, P. S. Puppala, A. R. Diehl, K. T. Hall, T. J. Kim, M. T. Zimmermann, R. L. Jernigan, L. Jaeger, and B. A. Shapiro, *Methods*, vol. 67, no. 2, pp. 256, 2014.
- [8] M. Anokhina, S. Bessonov, Z. Miao, E. Westhof, K. Hartmuth, and R. Lhrmann, *The EMBO Journal*, vol. 32, no. 21, pp. 2804, 2013.
- [9] E. Osada, Y. Suzuki, K. Hidaka, H. Ohno, H. Sugiyama, M. Endo, and H. Saito, *ACS Nano*, vol. 8, no. 8, pp. 8130, 2014.
- [10] N. B. Leontis and E. Westhof, *Science*, vol. 345, no. 6198, pp. 732, 2014.
- [11] M. Paliy, R. Melnik, and B. A. Shapiro, *Phys. Biol.*, vol. 6, no. 4, pp. 046003, 2009.
- [12] S. R. Badu, R. Melnik, M. Paliy, S. Prabhakar, A. Sebetci, and B. A. Shapiro, *Eur Biophys J.*, vol. 43, no. 10-11, pp. 555, 2014.
- [13] T. Hansson, C. Oostenbrink, and W. van Gunsteren, *Current Opinion in Structural Biology*, vol. 12, no. 2, pp. 190, 2002.
- [14] F. Ercolessi and J. B. Adams, *Euro. Phys. Lett.*, vol. 26, no. 8, pp. 583, 1994.
- [15] F. Ercolessi, E. Tosatti, and M. Parrinello, *Phys. Rev. Lett.* vol. 57, pp. 719, 1986.
- [16] N. Hori and S. Takada, *J. Chem. Theory Comput.*, vol. 8, no. 9, pp. 3384, 2012.
- [17] M. Paliy, R. Melnik, and B. A. Shapiro, *Phys. Biol.*, vol. 7, no. 3, pp. 036001, 2010.
- [18] J. Bida and R. Das, *Current Opinion in Structural Biology*, vol. 22, no. 4, pp. 457, 2012.
- [19] Z. Xia, D. P. Gardner, R. R. Gutell, P. Ren, *J. Phys. Chem. B.*, vol. 114, no. 42, pp. 13497, 2010.
- [20] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kal, and K. Schulten, *J. Comput. Chem.*, vol. 26, no. 16, pp. 1781, 2005.
- [21] MacKerell, D. Bashford, Bellott, Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wirkiewicz-Kuczera, D. Yin, and M. Karplus, *J. Phys. Chem. B.*, vol. 102, no. 18, pp. 3586, 1998.
- [22] N. Foloppe and A. D. MacKerell Jr., *Journal of Computational Chemistry*, vol. 21, no. 2, pp. 86, 2000.
- [23] A. D. MacKerell and N. K. Banavali, *Journal of Computational Chemistry*, vol. 21, no. 2, pp. 105, 2000.
- [24] W. Humphrey, A. Dalke, and K. Schulten, *Journal of Molecular Graphics*, vol. 14, no. 1, pp. 33, 1996.
- [25] S. R. Badu, R. Melnik, M. Paliy, S. Prabhakar, A. Sebetci, and B. A. Shapiro, in *Proceedings of IWBBIO-2014 (International Work-Conference on Bioinformatics and Biomedical Engineering)*, pp. 601–607, 2014.