DOI 10.18699/SBB-2022-673

Space-time correlation analysis in bio-macro-molecules, based on molecular dynamic trajectory data

Titov A.I.[1, 3]*, Shvetsov A.V.[1, 2, 3], Isaev-Ivanov V.V.[1]

[1] *Saint Petersburg Nuclear Physics Institute named by B.P. Konstantinov, NRC "Kurchatov Institute", Gatchina, Russia*

[2] *Peter the Great St. Petersburg Polytechnic University, St. Petersburg, Russia*

[3] *NRC "Kurchatov Institute", Moscow, Russia*

* *titov_ai@pnpi.nrcki.ru*

Key words: molecular dynamics, molecular modelling, GROMACS, algorithms, phytase, escherichia coli

Motivation and Aim: Since the end of the last century, biologists state a question: is it possible to model processes that happen in living beings at the molecular level? With rapid development of computing technologies and appearance of super-computer clusters, this question becomes even more actual. It is a pity that current computing levels are not enough for real-time modelling, even a cell-sized system is most likely not possible to recreate. On the other hand, calculation of inter-molecular interactions between biological molecules in real solvent for reasonable time durations is possible for the current-gen computing systems. However, modelling is not be-all and end-all, data analysis is even more important than getting it, but methods are lacking behind in development – such as dynamic properties analysis and inner-molecular interaction analysis.

Methods and Algorithms: There are a few programming packages for molecular modelling. Few should be mentioned among many: GROMACS[1] and AmberTools[2]. We have selected the first package, as it is open-source and allows to use its code and libraries for creation of personal programs to analyze computed trajectories of molecular dynamics (MD). We have created a few algorithms, in particular "Time and space correlation analysis in bio-macro-molecules from MD trajectory data"[3]. This work is about its application.

Results: We have selected *phytase* protein from *Escherichia* for inner-molecular analysis, since it is widely analysed in literature. All possible mutations were computed for 12 potential positions. During correlation analysis, we figured out that in more thermally stable variants the connectivity net is more spread, with relatively less connection nodes.

Conclusion: The obtained results allow us to consider that the created method of bio-macro-molecule analysis helps finding a list of potential thermo-stabilising variants. Based on the current data we plan to run a few *in vitro* experiments to prove this hypothesis.