CAUSALITY ANALYSIS OF MOLECULAR DYNAMICS EVENTS. STRATEGY TO UNDERSTANDING MECHANISMS OF BIOMOLECULAR SYSTEMS AND THE LOGIC OF THEIR FUNCTIONING

B. LESYNG

CoE BioExploratorium and Faculty of Physics, Warsaw University, Poland

Functions of complex biomolecular systems are determined by atomic and molecular motions, which can be simulated using molecular dynamics (MD) simulation methods. An MD simulation generates from a few thousands to a few hundred thousands atomic trajectories localized in the configuration space of a biomolecular system. In the case of a quantum-classical system (e.g. an enzyme macromolecule) a set of classical trajectories is supplemented with quantum degrees of freedom, for example with time-dependent parameters of an enzyme active-site wave function – see e.g. results of quantum-classical molecular dynamics (QCMD) studies of an enzymatic system (Bala P, Grochowski P., Nowinski K., Lesyng B. & McCammon J. A., 2000, Biophys. J, 79, 1253–1262).

Often we are capable to predict time-evolution of a studied system, however, it does not automatically mean that we understand how the system operates. The situation is a little similar to numerical weather forecasts – namely, based on physical principles we can predict evolution of parameterized fields determining the state of the atmosphere, it doesn't however mean that we understand *why* the weather changes as it is predicted by the simulation techniques. Here we assume that weather forecasts are sufficiently precise (from the point of view of this lecture it is a technical problem). Similar situation appears in studies of complex economic systems.

From the mathematical point of view simulation results are time series (signals). Understanding the functioning of a biomolecular system requires finding out causal relations between characteristic events. Events are elements of evolution of the system in its phase space. This refers for example to causal relations between conformational transitions. In 2003 C. W. J. Granger got the Nobel prize for his contribution to the causality analysis of complex economic systems, see e.g. Granger C. W. J., 1969, Econometrica, 37, 424 –438. Similar models are applied also by biomedical physicists or neurophysiologists in studies of EEG signals, see e.g. Blinowska K. J., Kus R. & Kaminski M. J., 2004, Phys. Rev. E, 70, 050902.

We applied similar methodologies in studies of MD simulation data for a few model molecular systems with the proton transfer processes, as well as for HIV-1 protease. An important problem in these studies is reduction of dimensionality of the MD trajectory sets. The proposed approach is presented in Gorecki A., Trylska J. & Lesyng B., 2006, Europhys. Lett., 75, 503–509 and Gorecki A., Trylska J. & Lesyng B., 2007, *Causality and Correlation Analyses of Molecular Dynamics Simulation Data*, in: From Computational Biophysics to Systems Biology (CBSB07), Publication Series of the John von Neumann Institute for Computing, vol. 36, pp. 25–30. One applies either a transformation to selected internal degrees freedom, or a principal component analysis (PCA) followed by projection of the MD trajectories onto the principal directions, and next analysis of the amplitudes of these projections treated as the time series. The proposed strategy is quite general and allows to understand mechanisms of complex biomolecular systems and the logic of their functioning. There are problems which require further studies, in particular, nonlinear couplings between PCA modes with a time-delay.

Acknowledgements: These studies are supported by CoE BioExploratorium, University of Warsaw.