COGRIMEN 2.0 - a toolkit for coarse-grained and all-atom modeling of biomolecules in implicit environments

Aim of the project

The aim of the project is to build a toolkit for coarse-grained (CG) and all-atom (AA) molecular modeling of all types of biomolecules (DNA/RNA/proteins) and their complexes, in aqueous and membrane implicit environments. Then, using the built and tested modeling and analysis tools, to perform long 100 μs molecular dynamics (MD) simulations of biologically important complexes and to investigate their dynamic properties.

Description of research

On a selected set of proteins, DNA and RNA, the solvation parameters will be optimized for both CG and AA force fields using a multi-criteria optimization algorithm. Solvation parameters will also be determined for selected post-translational modifications in proteins (phosphorylation, palmitoylation, prenylation), because these modifications strongly affect the dynamics of the systems in which they are bound. GaMD (Gaussian accelerated MD) and REMD (replica exchange MD) methods will also be used for more efficient sampling of the conformational space in both CG and AA modeling. Using these tools we will study the kinetics of large systems, e.g. selected GPCRs (G protein-coupled receptors) with their effector proteins (G protein and arrestin) with necessary post-translational modifications, the membrane γ-secretase complex with selected substrates, including β-amyloid peptides, and the ribosome protein-nucleic acid complex, which will be the largest system studied.

Reasons for undertaking this research topic

Long MD simulations are necessary to understand the dynamic properties of the studied biological systems, including the processes of complex assembly and disassembly. The use of continuous environments in both CG and AA methods will significantly extend the time scale of calculations due to the removal of solvent molecules (membrane/water) but also accelerate the occurrence of conformational changes due to the smoothing of the potential energy hypersurface.

The most important expected effects

The use of CG and implicit solvent calculations will allow for a speed increase of about three orders of magnitude compared to all-atom calculations in the full environment. In turn, implicit AA calculations, adapted by us to the most up-to-date CHARMM36m force field for biomolecules, will also allow for a speedup of about one order of magnitude compared to all-atom calculations in the full environment.

Thanks to modern methods of microcrystallography and cryo-EM (low-temperature electron microscopy), a large number of structures of complexes important for the functioning of our organism are already known, while the study of their dynamics, including binding and unbinding of their complexes is far behind, mainly due to the necessary, very long of the order of 100 μs time scale simulations. Using the COGRIMEN 2.0 toolkit, the simulations of biologically important systems will be performed: (1) GPCRs are responsible for the most of signal transduction processes across the cell membrane; (2) γ-secretase is an important membrane protease that cuts over 150 peptide substrates; (3) and the ribosome which is a huge complex of RNA and proteins where new copies of proteins are produced. Based on the obtained results, it will be possible to determine the dynamics of individual parts of these complexes and how it may affect their functions.

The COGRIMEN 2.0, after parameterization and testing, will be made available free of charge to all research groups. A server will also be built based on this toolkit, where it will be possible to perform online simulations and analyses using the tools created.