Project PN-II-RU-TE-2014-4-1176

Scientific report for the period 01.10.2015 – 01.12.2016

A. During the first stage of the project (01.10.2015-31.12.2015), the following activities have been carried out:

1. Recruitment of a postdoctoral researcher

The postdoctoral researcher selected for the project is Dr Zoltan Antal, with experience in shape-based analysis and various molecular modelling methods. Dr Antal obtained his PhD degree at the University of Newfoundland, Canada, under the supervision of Professor Paul Mezey. Dr Antal started his activities in my group in January 2016.

2. Acquiring the first batch of compute servers, configuring and testing the platform

The first batch of compute servers has the following technical characteristics: 16 CPU cores (Haswell 8C E5-2630V3 2.4 GHz), 128 GB RAM, 3 TB hard disk, one Nvidia Tesla K40M GPU. We deployed the Ubuntu 14.04 LTS operating system, configured LDAP access and the NFS filesystem, and installed benchmarking software. We also installed and tested various molecular modelling packages (Rosetta, AmberTools, GMIN, OPTIM, PATHSAMPLE).

3. Global optimization of model hierarchically self-assembling systems

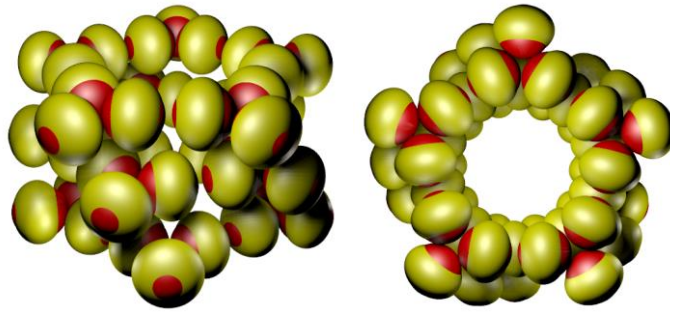

Figure 1. Global minimum for 45 particles. The structure is a chiral tube, consisting of pentamers of trimers.

This activity involved global optimization of a cluster of particles, which is capable of three distinct levels of self-assembly: (1) formation of trimers, (2) aggregation of trimers into rings, (3) stacking of rings into chiral tubes. The interactions necessary for this behaviour are encoded in the shape of the particle. One patch on the particle allows for two strong contacts, while the other one for one weak contact.

B. During the second stage of the project (01.01.2016-31.12.2016), the following activities have been carried out:

1. Recruitment of remaining team members

The two part-time PhD students selected for this project are Janos Szoverfi (student at the Polytehnic University of Bucharest) and Melinda Kakes (student at the Babes-Bolyai University in Cluj-Napoca).

2. Development of hierarchical coarse-graining methods

We continued optimizing the rigid body framework implemented in GMIN and OPTIM to allow for a better exploration of configurational space for anisotropic bodies. Given the high specificity and directionality of interactions, the energy landscapes associated with these systems are usually very complex. The basin-hopping algorithm needs specially adapted moves for such ellipsoidal bodies, ensuring that there is no overlap in any starting configuration, and also making the translational and rotational moves large enough to not get stuck in a local region of the configurational space. Also, ensuring no overlaps during DNEB interpolations (relevant to transition state search) is crucial for efficient path sampling. We implemented these enhanced moves into GMIN and OPTIM.

For the coarse-graining of protein surfaces, we chose a method of fitting a given number of ellipsoids to an arbitrary surface, which in our case will be the solvent-accessible surface of the CCMV capsid protein.

3. Studying the aggregation of CCMV capsid proteins

During this activity, we developed a model that has the potential to predict protein-protein binding configurations with a greater accuracy than standalone docking methods. We found that by using this method, we were able to correctly predict the structure of the ss-CCMV protein dimer, which is thought to be initiating the hierarchical self-assembly process. This result is particularly remarkable, because virus capsid proteins evolved to have more than one specific binding site, and our method could correctly point out the most favoured docking pose among these.

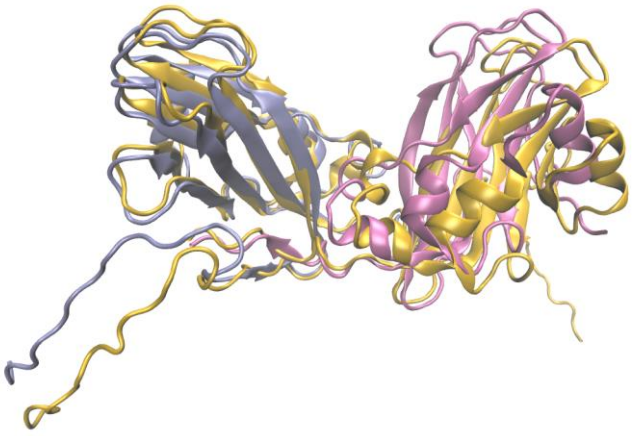

Figure 2. Predicted best dimer structure, overlaid with a CC type dimer taken from the crystal structure.

We also evaluated the relative strength of the two likely binding interfaces relevant for capsid assembly, using the docking protocol implemented in Rosetta. We found that the stability of the type II interface is much smaller than that of the type I interface (which is the predicted dimer interface). Also, docking two dimers along their common interface resulted in no improvement of the calculated interface score. This leads us to the conclusion that for CCMV, the second step of the hierarchical self-assembly is governed exclusively by the type II interface.

We prepared and submitted a publication from these results.

Now that we know the two most important interfaces of CCMV, we can coarse-grain the system by setting their relative binding affinities to match those calculated with Rosetta, and perhaps study the aggregation of these proteins in larger numbers as well. In order to do that, we implemented a rigid body molecular dynamics code into GMIN, which was originally developed for polyaromatic hydrocarbons. We adapted this code to correctly propagate translation and especially rotation for rigid bodies containing ellipsoids. A specific challenge was working out the correct implementation of ellipsoidal torques within the rigid body framework.

The Rosetta software suite has very useful algorithms for protein-specific moves, therefore we started working on the implementation of the Prepack algorithm into GMIN and OPTIM. The rotamer moves and other dihedral move sets can be incorporated to work with any protein, during basin-hopping global optimization or initial interpolation for discrete path sampling. We estimate to have a working version of this implementation by the end of December 2016.

4. New design principles for hierarchical self-assembly: macroions

Our implementation of the macroion model was devised with the aim of explaining an interesting aggregation behaviour encountered in dilute solution of highly charged polyoxometallates. Several systems are known for which hundreds or thousands of polyoxometallate macroions form closed shells. Such shells are stable, and the system is monodisperse. No current (discrete or continuous) electrostatic model can explain this behaviour. Our initial investigations revealed an intriguing design principle for obtaining anisotropic self-assembled structures from isotropic interactions. We allocated significant computing resources to this avenue of research, given that the potential energy landscapes of such systems are extremely complex.

Global optimization of clusters in which the particles have a large size mismatch is inefficient with the currently implemented Monte-Carlo moves (e.g. random perturbation of coordinates). A large MC move can cause ‘cold fusion’, if a counterion overlaps with a macroion, for example. We therefore implemented an algorithm for such systems, in which for each starting structure we determine dynamically which counterions are in the shell around a macroion, and move the shell cooperatively with the macroion (random Cartesian moves, rotational moves), with the possibility of random perturbation of the individual ions as well. We use this algorithm in conjunction with the PERCOLATE keyword, which diagnoses dissociation of particles and rejects such minima. This step taking algorithm largely improved convergence towards finding the global minimum for the clusters we investigated.

Conferences

My group attended three conferences this year:

a. Energy Landscapes conference, Porquerolles, France – oral presentation by myself, poster by Dr Zoltan Antal
b. International Conference on Chemistry, organized by the Hungarian Technical Scientific Society of Transylvania – oral presentations by Janos Szoverfi and myself, poster by Dr Zoltan Antal
c. MolMod2016, organized by UBB Cluj-Napoca – oral presentations by Janos Szoverfi and myself, poster by Dr Zoltan Antal

Summer school

I presented the latest results of my research and the methods we use for this project at the international Summer School for computational chemistry, organized at the University of Miskolc, Hungary.

Dr Szilard Fejer