Final report on:

Coarse-grained Biomolecular Modeling

Tutorial held at CECAM headquarters, October 17-21, 2011

(1) SUMMARY

The field of biomolecular coarse-grained modeling is currently receiving a lot of attention and enjoying increasing popularity. The main attraction to researchers is the expectation that within the next decade, models of sufficient chemical specificity to reliably relate to biologically relevant systems can be simulated on appropriate length- and timescales. More detailed models including all atoms (the classical molecular dynamics methods), possibly adding some electronic degrees of freedom treated quantum-chemically (QM-MD methods), will still be too limited in those respects to really address the questions that arise out of the latest experimental biophysical techniques such as single-molecule tracking.

In trying to address biomolecular problems, chemical specificity can be important, but complex behavior may also emerge from simple, more generic, interactions. It is here that coarse-graining techniques and approaches have tremendous potential in developing our understanding of the relationships between molecular interactions and the processes that define life.

With this tutorial we aim to introduce researchers to the field of biomolecular coarse-graining, through giving them basic lectures in simulation techniques and in the theories and ideas underlying coarse-graining and having them practice with some of the computational tools that are available to perform coarse-grained simulations in practice under the guidance of experts.

The field is already too large to be comprehensive, but the three approaches that form the current mainstreams in biomolecular modeling are represented: (1) hierarchical models deriving coarse-grained interactions from underlying atomistic simulations using generalized Yvon-Born-Green (YBG) theory; (2) the semi-quantitative, empirical, chemically specific models for time-scales up to a millisecond and length-scales up to 200 nm (MARTINI model); (3) the less specific mesoscopic models that are able to reproduce macroscopically measurable thermodynamic properties of e.g. bilayers and vesicles, modeled at length-scales in the micrometer regime (models due to Cooke et al and Shillcock et al). An advantage of the choices we made is also that the simulation software packages used in the tutorial are limited to GROMACS and Espresso.

The tutorial brought together mainly young researchers (PhD students, post-docs) and more experienced researchers starting in the field of coarse-graining for five days of intense learning and exchanging of ideas and discussions. The explicit opportunity to get feedback and practical advice from experts on the current projects of the participants on the last day was especially appreciated.

(2) SCIENTIFIC CONTENT

The hard work of dedicated expert scientists has provided the community with easy-to-use, freely available software packages for molecular simulations. The open-source philosophy has created a wealth of resources and knowledge available on the world-wide-web. The barriers to explore some of the computational tools are low, and many researchers are excited about the opportunity to gain molecular insights about their lab-work from the perspective of molecular models. However, proper use of the computational tools is not trivial and the user is not protected from making fundamental errors. Therefore, the need for training and guidance is acknowledged from the side of enthusiastic non-expert users and expert developers.

The tutorial Coarse-grained Biomolecular Modeling (CGBM), organized at the CECAM headquarters in Lausanne, Switzerland, aimed at providing basic training and guidance in some popular and interesting coarse-graining techniques with an emphasis on biomolecular systems.

Introductory lectures covered the general aspects of Molecular Dynamics (MD) [de Vries] and Dissipative Particle Dynamics (DPD) [Shillcock] simulations as means to generate representative ensembles of structures. Particular attention was given to the inevitable approximations involved in molecular modeling: in integrating the equations of motion, in working with relatively small systems and short timescales compared to most experimental techniques, and in the representation of the molecules themselves. Especially in coarse graining it is important to understand the limitations the model itself imposes on the questions that may be asked of the system under study. An emphasis was given to the meaning of time in coarse grained models and to the significance of distributions as a means to connect experimental data to appropriate model interactions (the force field) and as a means to compare and validate models at different spatial resolutions. Given the statistical nature of matter, comparisons is only really meaningful when distributions can be compared and distributions are closely connected to the underlying physical interactions.

The lectures brought to bear different approaches to use the information available in distributions. The semi-empirical approach takes experimental data reflecting statistical ensemble data, such as free energies of partitioning or free energies of solvation, and uses those to parameterize a family of chemically specific building blocks that can be strung together to form (bio)molecules. Since partitioning is such a prominent driving force in biomolecular phenomena, this approach promises to be predictive in the aggregation behavior of more complex molecules, once the interactions of the chemical building blocks have been parameterized based on their partitioning. The MARTINI coarse grained model fulfills this promise to a large extent for lipids in water, its first application [Marrink]. Extensions of the model are being developed and are proving useful in more and more areas of biomolecular and soft matter chemistry and physics, exemplified by peptides and proteins [Monticelli], carbon nano particles and polymers [Monticelli, Salonen]. The future directions and ideas concerning the MARTINI model were also discussed, including the polarizable water model and

carbohydrates [Marrink]. In all lectures insights were given into the scope and limitations of the model.

The second theme of the meeting concerned generic amphiphile models that emphasize the meso-scale behavior with less chemical specificity, although also in these models, the study of the relationship between the interactions between the building blocks and the mesoscopic properties (bending modulus, area compressibility, line tension, etc.) may be meaningfully connected to experimental data. The models presented were both solvent free [Deserno] and with explicit solvent [Shillcock] and the speakers urged the audience to think about what these models should and are able to achieve. In scaling up simulations to longer length-scales, it is essential to realize that equilibration requires much longer time-scales; thus, simulation of huge systems is meaningful only when allowing long enough time to properly equilibrate them. Examples of the effect of additives to bilayers were shown, including the induction of large deformations by protein mimics on the surface [Deserno] and solubilizing fullerene aggregates inside the bilayer [Shillcock].

Finally, the ideas of hierarchical coarse graining were introduced and illustrated [Noid]. In this approach, the aim is to use structural distributions and correlations directly to derive a computational model. It should be noted that this can be used to derive a coarse-grained model from a more detailed computational model, but that deriving a model from experimental data (such as those deposited in the Protein Data Bank) is amenable to this approach as well. The theory was carefully developed and the resulting quantities from a generalization of the Yvon-Born-Green equations discussed in detail.

All approaches were practised in tutorial sessions each afternoon. The available tutorials offered various level, from very basic for beginners, to more advanced for more experienced modelers. Practical guidance was given by assistants, PhD students that work with the programs on a daily basis.

Twelve participants brought a poster introducing their work to the meeting, which were discussed during and after the poster session.

(3) RESULTS AND FUTURE IMPACT

Given that the meeting was a tutorial, results in terms of scientific progress are few. The future impact of the meeting should be substantial, however. During the meeting, the participants were actively seeking guidance and often went beyond the standard tutorial material, trying to see if the approaches could be applied to their own research problem. Lively and in-depth discussions between participants, lecturers, and assistants during the meeting in some cases already led to collaborative projects and it may be expected that some of the techniques will find their way in existing or new research projects.
The feedback received from participants was altogether positive; the ample time reserved for the hands-on sessions and the level of the lectures were especially appreciated.

(4) FINAL PROGRAMME

Monday
09.00-09.20 Registration
09.20-09.30 Welcome by the organizers
09.30-10.30 Alex de Vries: Coarse Graining Basics
10.30-11.00 Coffee break
11.00-11.45 Alex de Vries: Coarse Graining Validation
11.45-12.30 Siewert Jan Marrink: The MARTINI model basics
12.30-13.30 Lunch
13.30-14.00 Introduction to PRACTICAL SESSIONS
14.00-18.30 PRACTICAL SESSION 1: MARTINI model basics

Tuesday
09.00-09.45 Luca Monticelli: The MARTINI extension to proteins
09.45-10.30 Luca Monticelli: MARTINI model for Carbon nano particles
10.30-11.00 Coffee Break
11.00-11.45 Emppu Salonen: Coarse-grained modelling of fullerenes, polymers, and their composites
11.45-12.30 Siewert Jan Marrink: New developments for MARTINI
12.30-13.30 Lunch
13.30-18.30 PRACTICAL SESSION 2: advanced MARTINI modeling
18.30-20.00 POSTER SESSION

Wednesday
09.00-10.30 Markus Deserno: Extreme coarse graining
10.30-11.00 Coffee Break
11.00-12.30 Julian Shillcock: Generic amphiphile models and Dissipative Particle Dynamics
12.30-13.30 Lunch
13.30-18.30 PRACTICAL SESSION 3: Extremely CG models using ESPRESSO

Thursday

09.00-10.30 Will Noid: Hierarchical coarse graining theory
10.30-11.00 Coffee Break
11.00-11.45 Will Noid: Hierarchical coarse graining examples
12.35-13.30 Lunch
13.30-18.30 PRACTICAL SESSION 4: Methods for hierarchical coarse graining
19.00- TUTORIAL DINNER

Friday
09.00-12.30 PRACTICAL SESSION 5: All tutorials and Participants' own systems of interest
12.30-13.30 Lunch and Departure
The CECAM facilities will be available until 17.00 hours