Hierarchical multiscale modelling scheme from first principles to mesoscale

Alexander Lyubartsev and Aatto Laaksonen

Division of Physical Chemistry,
Stockholm University SE-10691 Stockholm, Sweden
sasha@physc.su.se, aatto@physc.su.se,

Abstract. We present a true hierarchical multiple scale modelling scheme which enables us to start from first principles *ab initio* computer simulations with electronic degrees of freedom and successively coarse-grain the model by leaving out uninteresting degrees of freedom. From *ab initio* simulations it gives us a set atomistic pair potentials to be used in subsequent classical all-atom simulations from which a coarse-grained model is constructed to carry out meso-scale simulations beyond the atomic resolution. We will show several examples on how this is done based on effective interaction potentials to tie together the various scales of physical description of molecular matter.

1 Realities & simulations

The three-dimensional operational real world for scientists performing computer modelling molecules is based on three painfully orthogonal axes with labels, **system size**, **motional time scale** and **accuracy of the model**. In moving inside this space we are repeatedly forced to choose the optimal lengths for the axes to make the computations of molecules feasible. This was so three decades ago it will be the same in times when the computers are three orders of magnitude more powerful than today. Compromising will be the key word. From the early days of computer simulations the methods did not evolve very much for some considerable period time. Possibly there was no need for it. Computers were just becoming faster following the Moores law pushing the development in many areas, including simulations, further. Only during the last two decades many new simulation methods have been published to stretch the time and length scales further. Also schemes were introduce to increase the accuracy by taking molecular dynamics to the domains of quantum mechanics (QM) and thereby to the first principles of physics. This allows MD simulations without any empirical parameters (such as potential functions or molecular mechanical (MM) force fields) as input. Hybrid methods mixing QM and MM based schemes are common tools today. The most interesting development may still be the schemes beyond atomistic resolution to model meso and nano size systems and soft matter. In other words there are now reliable simulation methods available to treat matter at three levels of physical description (QM, MM and soft matter), where the accuracy is successively decreasing while allowing the system length and time scales

to be increased. Examples of these are Car-Parrinello molecular dynamics (MD), classical atomistic MD and dissipative particle dynamics (DPD). In common terminology models beyond the atomistic resolution are result of coarse-graining (CG). There is no unique way to do coarse-graining within off-lattice framework. For heterogeneous systems like biological molecules anything from *ad hoc* to parameterize CG potentials have been used while for homogeneous systems, like in materials design, finite element and grid-based models are common. In the case of biological systems the water surrounding the biomolecules is a great challenge to coarse-grain. Such simplifications may include implicit description of the solvent with the help of solvent-mediated potentials and coarse-grained representation of molecules of interest. The problem is however how we can specify interaction potentials for such coarse-grained models.

2   Hierarchical multiscale modelling approach

The present contribution describes a general approach how to use results of a more fundamental, more accurate theory, to construct simplified, coarse-grained models to be used in large-scale simulations. The idea of the approach is the following: First, we carry out detailed simulations on a more fundamental, *ab-initio* level. Such simulations provide full information about the system, although on a limited length scale for reasons pointed out above. From these detailed simulations we can now determine a set of radial distribution functions (RDFs) between sites representing interesting degrees of freedom. Then we ask: which interaction potential within a coarse-grained model would reproduce the very same set of RDFs as we obtained within the full detailed model? To obtain the answer to this question, we need to solve an *inverse problem*: reconstruct interaction potential from the previously obtained RDFs. If the inverse problem can be solved, we have in hand an effective interaction potential, which for the coarse-grained model reproduces the same structural properties of the system as full, *ab initio* model. This allows us to increase the length scale of the problem and use the computed effective potentials for simulations on considerably larger scale.

It is clear that the essential point of this approach is solution of the inverse problem. This can be done with the help of the Inverse Monte Carlo method originally suggested by us in [1]. Some other approaches have appeared also (see review [2] for details). We will present several examples on application of the Inverse Monte Carlo (IMC) technique to derive effective potentials.

3   Site-site atomistic interaction potentials from *ab initio* simulations

Pair atom-atom potentials for classical molecular dynamics simulations may be derived from RDF computed in *ab initio*, for example Car-Parrinello molecular dynamics simulations. Now using high-performance computing (HPC) facilities

like those offered by the HPC2N computer centre here in Umeå it is possible to run Car-Parrinello simulations for a few hundred of atoms on a several tens of picoseconds. We have carried out Car-Parrinello simulations of water (32 and 64 molecules) in which oxygen-oxygen, oxygen-hydrogen and hydrogen-hydrogen RDFs were computed. Then these RDFs were sent as input to our IMC procedure and effective potentials for water have been obtained. They came out rather similar to those of the conventional SPC water model, with some differences at short distances between atoms. Thus it was confirmed that such *ab initio* parameterization of interaction potentials is feasible. In another example, effective potential between Li^+ ion and water has been computed from *ab initio* RDFs. It turned out that non-electrostatic part of this potential has a simple exponential shape, without a specific attractive interaction term.

4 Effective solvent-mediated ion-ion and ion-DNA potentials

On the next level of coarse-graining, the explicit description of water molecules may be removed and substituted by effective solvent-mediated ion-ion potentials. Such effective potentials were computed for Na^+ and Cl^- ions in water solutions. It turned out that at large distances (more that 10 Å) the effective $Na - Cl$ potential is very close to the Coulombic potential in a media with dielectric permittivity of water ($\epsilon = 80$), but on smaller distances makes a few oscillations, representing molecular structure of water. In further development, effective solvent mediated potentials between different alkali ions and DNA were computed. These ion-DNA effective potentials were used in coarse-grained simulations of ion atmosphere around DNA. In these simulations, relative binding affinities of different alkali ions to DNA were determined and found in good agreement with available experimental data.

5 CG phospholipid model for bilayers and vesicles

In a similar manner, effective solvent-mediated potentials for 10-sites CG model of DMPC lipid molecule have been constructed. Then several simulations of the coarse-grained model on longer length- and time-scale were performed, both within MC and MD simulations: a periodical sample of lipid molecules ordered in bilayer, a free sheet of such bilayer without periodic boundary conditions, formation of vesicle from a plain membrane, process of self-assembly of lipids randomly dispersed in volume. The number of lipid molecules in these coarse-grained simulations reached 3500, corresponding to several millions of atoms in the corresponding all-atom system. It was shown that the CG model, developed exclusively from all-atomic simulation data, reproduces well all the basic features of lipids in water solution.

4

6 Conclusions

A hierarchical true multiscale modelling approach presented here links together three levels of molecular modelling: *ab initio* molecular dynamics, classical molecular dynamics and meso-scale simulations. The method providing the link between these levels is the Inverse Monte Carlo approach. One of the problems hindering a wider use of this approach is that available software packages, implementing each of the steps of the multiscale modelling, are poorly compatible with each other. As long as the idea of bridging the length and time scales in molecular modelling is recognized to be more and more important, it should be reflected in a certain standardization of description of molecular models and simulation protocols which would provide better compatibility of already available software for simulations of molecular systems on different levels.

References

1. Lyubartsev, A.P., Laaksonen A.: Calculation of effective interaction potentials from radial distribution functions Phys. Rev. E **52** (1995) 3730–3737
2. Lyubartsev, A.P., Laaksonen A.: On the reduction of molecular degrees of freedom in computer simulations In: *Novel Methods in Soft Matter Simulations*, (Eds. M. Karttunen, I. Vattulainen, A. Lukkarinen), Springer, Berlin, 2004, pp.219–244.