Improved description of complex plasticity and interactions
in protein-ligand simulations

Molecular simulations of biological systems allow to predict properties not accessible to experiment, to provide novel insights in experimental results, or to participate in improving the design of new experiments. Molecular Dynamics (MD) simulations are often used to study the interactions between a drug and a target protein on a detailed level: atomic interactions between the drug and the protein can be studied on a femtosecond time scale. In the last decades, the field of biomolecular simulation has seen a fast development, feeded by massively growing computer facilities and decreased costs to perform the simulations. The available computational power has been utilized to develop new methods, increase the detail of simulations and run simulations on longer time scales. However, although results from millisecond simulations of proteins have recently been reported, the field of biomolecular modeling is still faced with challenges restricting the reach of the current simulation efforts.

In this thesis two challenges have been addressed. First, possibilities to further improve the accuracy of describing (non-bonded) interactions in biomolecular simulations are investigated. Historically, the electrostatic interactions between molecules have been defined by assigning local, static, nuclei and partial charges to the atoms, which is an approximation of the molecular charge distribution due to its smeared-out electron density. As a consequence, molecular charge distributions in classical simulations are not able to respond to changes in their local electrostatic environment. This in turn can have a negative effect on the evaluation of non-bonded interactions for molecules that change environment, which is e.g. the case in the process of a drug binding to a protein. As a possible remedy, explicit inclusion of electronic polarization effects into classical force fields has been suggested before.

In Chapters 2 and 3 of this thesis, a major challenge in developing such a polarizable force field is addressed. Previously, it has been shown that proper assignment of atomic polarizabilities is far from trivial. In this work, we have proposed a QM/MM approach to obtain values for atomic polarizabilities that are directly suited to be implemented into a (bio)molecular force field. It was shown how these values for the atomic polarizability, together with a thorough reparametrization of other force field parameters, can be used to accurately reproduce relevant physico-chemical properties in simulation. In Chapter 4, an analytical method is proposed to systematically scan parts of parameter space, allowing to find multiple sets of parameters that can be used to reproduce experimental values for relevant pure-liquid properties. Using this approach, a minimal and transferable parameter set for primary alcohol molecules was proposed. In addition, Chapter 5 presents an attempt to use the QM/MM determined polarizabilities to analytically derive non-bonded parameters for the van der Waals force field term. Achieving to compute in a straightforward way the attractive van der Waals parameter would allow to use fully automated...
techniques to recalibrate a force field, or generate new force field parameters.

The second part of this thesis addresses the challenge of accurately and efficiently exploring the relevant parts of conformational space of a complex system such as a protein-ligand complex. In Chapters 6 and 7, a generic implementation of an iterative Linear Interaction Energy method is presented that allows to explore multiple, well-separated parts of conformational space of a ligand-protein complex in multiple parallel (short) simulations, and to integrate their results in a subsequent step based on Boltzmann weighting. For the flexible Cytochrome P450 2D6 enzyme, it is shown that improved accuracy in binding affinity prediction is obtained in this way. Moreover, the method is also shown to be efficient, making it suitable for (semi-)high throughput affinity prediction.

In addition to the attempt to improve the quality of predictions from biomolecular simulation, a special interest was placed in developing the presented and applied methods in such a way that they can be generalized and broadly distributed. For expert users, automatization carries the benefit of reducing overhead, allowing to spend more time on the analysis of the results and the design of new simulations rather than the setup of the simulations. More importantly, well-automated methods can also benefit non-expert users. The methods described in this thesis are proposed as new functionalities in existing online tools. The tools, such as eTOXsys and ATB, work with a (web)interface, and are designed to take over as much of the simulation setup and analysis efforts from the user, only requiring the user to give a basic description of the system as an input. In addition, through several methods to automatically check the validity of the simulations and analyses, the quality and applicability of the results is assessed.