9 Concluding remarks

One of the main goals of this work was to understand long-term protein dynamics. In chapter 8 we have presented a simplified protein model that runs on the order of microseconds. The picture that emerges from this model is that proteins stay up to hundreds of nanoseconds in one conformation and then jump in a short time to a next conformation. This behavior can be observed on several time scales, for instance nanoseconds in MD simulations (see Figures 7.1 and 7.2), but also (sub-) picoseconds, as Troyer and Cohen showed [103]. Judging from the results of this model, a microsecond is not enough to estimate the conformational freedom of a relatively small protein, since the protein does not revisit previously sampled conformations. One could argue that this conclusion not is valid, since the force field for the simplified model is quite bad. Indeed one can certainly not expect qualitative results for a specific protein. But it is not unlikely that the jumping behavior, which is observed in MD simulations over a range of times scales of 10000, extends to microseconds or even a millisecond. Experimentally this can not be excluded since the only available experimental technique that can probe the dynamics at room temperature, NMR, can only distinguish conformations when they live longer than a millisecond. If the jumping behavior does extend another factor of 10000, then there is no hope that all the conformations in the native state of a protein can be probed by MD simulation. At least not in the near future.

In chapter 7 we have seen that proteins spend most of their time randomly diffusing on a locally quite smooth part of the energy landscape. In this aspect Brownian dynamics is an accurate description of protein dynamics. However, the change from Newtonian to Brownian dynamics does not increase the computational efficiency much, since the time step is still limited by the collisions of Lennard-Jones particles. One can average over particles, finally ending up with one particle per residue. It is not hard to make such a model that randomly diffuses within one conformation just like MD would do. But the problem is that it will not predict the correct jumps to the next conformation. The exact direction of the jump probably depends on very detailed interactions of atoms within the protein and also with the solvent.

When the residence times in each conformation are of the same order of magnitude, the exact spread within each conformation is not of interest, since the difference with a next conformation is much larger than the spread. The challenge is to invent a model that takes advantage of this, by skipping the random diffusion within each conforma-
Concluding remarks

One approach is filling up the minima in the energy landscape or conformational flooding [109]. By adding a potential which has approximately the same shape as the minimum of the current conformation, one can force a jump to a next conformation. Unfortunately it takes even longer to obtain a reasonable estimate of the shape of the minimum than it takes for the protein to spontaneously jump out of the minimum, as we have seen in chapter 7.

Another problem that the simplified model brings up is that there are no good experimental data available. Although X-ray crystallography and NMR can produce accurate structures of proteins in the native state, they give only a slight hint about the width of the distribution in the native state. The simplified model of HPr presented in chapter 8 surely samples many bad conformations, but the number of violations of distance bounds derived from NMR measurements is relatively low. This is due to the very non-linear $r^{-6}$ averaging in the experimental data. With just a few more "correct" structures in the ensemble, one can get rid of all violations. A more rigorous, but also much more elaborate (and almost never applied) validation is to check if the model predicts any short distances between atoms which are not observed in the NMR experiment. But there is still hope as a new technique in NMR is being developed. By putting long thin molecules or non-spherical aggregates into the NMR tube, which orient spontaneously in the magnetic field, hydrodynamic interaction can orient a non-spherical protein slightly [110]. This re-introduces the so called dipolar couplings, which are normally averaged out completely due to the isotropic tumbling of the protein. The dipolar couplings are directly related to the angles of bond vectors with the magnetic field. The big advantage of the dipolar couplings is that they average linearly and thus give much stricter bounds to the ensemble.

One question that remains to be answered is how accurate the description of a protein should be to be able to fold it, given that we have enough computer power to do so. The simplified model presented in chapter 8 is too simple, even if one would use an advanced implicit solvent model. My personal feeling is that an implicit solvent model is not accurate enough, since water has very specific interaction with the surface of the protein. Probably an all-atom description is required, including explicit solvent molecules. Maybe polarization needs to be incorporated into the model to correctly describe the interactions of chemical groups in different environments, such as the inside of the protein and the solvent. We have seen in chapter 4 that even a subtle quantum effect, such as going from rigid constraints to flexible constraints, can speed up the dynamics of water by 25%.