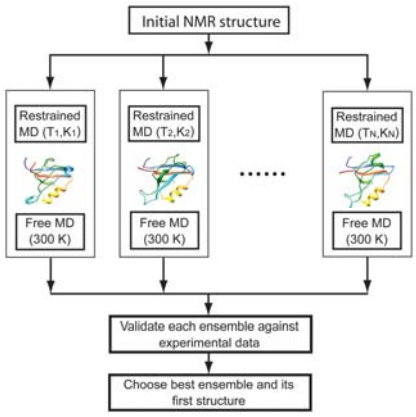
Robust Protocol to Make Protein NMR Structures Amenable to Long MD Simulations

D.-W. Li (NHMFL); R. Brüschweiler (FSU, Chemistry & Biochemistry, NHMFL)

Introduction

Molecular dynamics (MD) computer simulations of proteins provide important new insights into protein behavior and function. In recent years, all-atom molecular dynamics (MD) simulations of proteins in explicit solvent have made major strides in their ability to quantitatively reproduce a wide range of experimental structural dynamic parameters of proteins. These advances were made possible by the development of improved all-atom molecular mechanics force fields and the availability of ever more powerful computer hardware, which have turned microsecond timescale simulations in explicit solvent practically feasible. A requirement for the success of MD, besides accurate force fields and powerful computers, is the availability of a suitable 3D protein structure as a starting point. The vast majority of protein starting structures for MD have been determined by x-ray crystallography with a resolution preferably ~2 Å or better. Although over 8500 protein structures determined by NMR spectroscopy have been deposited in the Protein Data Bank (PDB), because the quality of NMR structures is on average lower and more heterogeneous, NMR models are often not the first choice for protein structure data mining and MD simulations. Their sporadic use for MD is due to their tendency to lead to unstable MD trajectories that drift rapidly away from the initial structure. This situation is unfortunate since for many of these proteins no crystal structure is available.



Flowchart of NMR structure refinement protocol using extended restraint-free MD simulations.

Experimental

We have used 4 different proteins, which fulfill the following two criteria: (1) a NMR structure is available along with the experimental NOE constraint list, chemical shifts, and residual dipolar couplings; (2) an x-ray crystal structure is available at a resolution of 2 Å or better. The flow chart of our refinement protocol is depicted in Figure. To refine the NMR structures by restrained MD, we first ran MD simulations for 6 different combinations of simulation temperatures (300 K, 330 K, and 360 K) and 2 different NOE restraint force constants for 100 or 200 ns length for each protein, which are termed "restrained MD". Next, the final conformation of each of the 6 restrained MD trajectories corresponds to a refined NMR structure candidate, which is taken as a starting structure for a 100 ns or 200 ns long free MD simulation (termed "free MD") in the absence of any experimental restraints. This protocol was repeated multiple times per protein with different randomly assigned initial atomic velocities. Chemical shifts, NOE constraints, and RDCs were back-calculated from the free MD trajectories and compared with experiment to determine the highest quality MD ensemble. The refined NMR structure candidate that served as starting point of the highest quality free MD ensemble is then designated as the "refined NMR structure".

Results and Discussion

Without exception, the quality of the refined NMR structures improved substantially compared with the original NMR model. The quality of the refined structures is comparable or only slightly worse than the corresponding x-ray structures. The quality of MD simulations is now reaching a level, which permits protein structure refinement at nearly experimental accuracy. Previously, we found that due to the presence of sampling-related bottlenecks this type of structure refinement only works for a subset of cases. Here, we demonstrate how the inclusion of a large number of NOE constraints into the MD simulations can address this challenge. With the proper choice of simulation temperature and restraint strength, the MD simulations make the structures native-like state within less than a microsecond simulation time thereby considerably improving the quality of the structure.

Acknowledgements

This work was supported by the National Science Foundation (grant MCB0918362).