8 A model for microsecond protein dynamics

We present a coarse grained model for simulating proteins. A Brownian Dynamics algorithm in combination with a novel method for constraining bond lengths allows for large time steps. The efficiency is further increased by decreasing the number of degrees of freedom. The model was tested on the 85-residue Histidine containing Phosphocarrier protein. We have generated and analyzed an ensemble of trajectories of several microseconds in total. These trajectories reveal that features of the energy landscape found in nanosecond Molecular Dynamics simulations reappear on larger time and length scales.
8.1 Introduction

Biologically relevant processes in proteins span a large range of time scales. From picoseconds for electron transfer, through nanoseconds for local motions to micro or milliseconds for large scale motions and protein folding. In this chapter we probe microsecond fluctuations in a coarse grained manner by way of simulation. In principle, a number of coarse graining methods can be envisioned. For example Monte Carlo [79], Dissipative Particle Dynamics [80], Landau models [81, 82] and for proteins, folding [83, 84, 85, 86] and lattice models [87, 86]. Here we address the problem of mode coupling between fast local motions on an atomic scale and large ’mesoscale’ fluctuations on the scale of the size of the protein. To this end we have developed a Brownian Dynamics or Position Langevin Dynamics method [35] for simulating proteins with simplified force fields. In the model inertial forces are neglected. The bonds lengths are constrained with a novel projection algorithm [10] (see chapter 3). We have tried to specify the problem by following the behavior of the Histidine containing Phosphocarrier protein HPr from Escherichia coli. HPr is a small 85-residue protein with an $\alpha/\beta$ fold. The structure of HPr has been solved by both NMR [88] and X-ray crystallography [76].

Our main conclusion is that the protein exhibits a hierarchical scaling behavior, in which phenomena on various length and time scales actually are much alike. The results further demonstrate the relevance of the energy landscape picture [89], in which proteins ’hop’ from one state to the other. Small scale hops occur more frequently than large scale hops. This behavior does not depend on the precise choice of the parameters and the forcefield. However, for mimicking the behavior of a specific protein the parameters need more critical evaluation.

We have found that the simulated protein follows completely different time trajectories, given similar initial configurations. The trajectories span over several microseconds and therefore already probe a significant amount of the thermodynamically stable states.

The chapter is organized as follows. In section 8.2 we briefly describe the Brownian Dynamics (BD) algorithm and the force field. In section 8.3 we analyze the trajectories and the hopping behavior and compare the results with NMR data. Section 8.4 contains the discussion and conclusions.

8.2 The Model

Protein folding models have shown that two or three atoms per side chain are sufficient to correctly reproduce side-chain packing [84]. Several coarse grained models for protein folding have been proposed with only one interaction site per residue and a backbone consisting of C$_\alpha$’s connected by virtual bonds [90, 91]. The forcefield for the complex side chain interactions can be derived from a database of folded protein structures [92].
8.2 The Model

Figure 8.1: Simplified side chains, similar to the side chains in the LINUS model [84]. The side chains that are not shown here are all-atom.

While such models are able to distinguish correct folds from incorrect folds, they are not suitable for prediction the energy difference between folded and partially unfolded conformations, since there is no information on unfolded conformations in the database.

The model presented in this chapter has a large reduction of the number of degrees of freedom compared with full atomic-detail Molecular Dynamics (MD), but it contains enough detail to simulate the dynamic behavior of specific proteins. All the larger side chains have been simplified (Figure 8.1). This reduces the number of particles as well as the number density of the particles. The backbone is all-atom to ensure the correct representation of secondary structure elements.

A similar Brownian Dynamics method as we propose here has recently been applied to poly-alanine [93]. In general, the friction matrix has non-zero off-diagonal elements, since the protein atoms influence each other through the solvent. We will neglect these effects and assume that the matrix is diagonal and constant and that the friction is the same for each particle. The position Langevin equation with isotropic homogeneous friction (eq. (2.44)) for $N$ particles is:

$$d\mathbf{r}(t) = \frac{1}{\gamma} \mathbf{f}(\mathbf{r}(t)) \, dt + d\mathbf{W}$$

where $\mathbf{r}$ is the position $3N$-vector, $\mathbf{f}$ is the force, $\gamma$ is the friction-coefficient and $\mathbf{W}(t)$ is a Wiener process. The noise term $d\mathbf{W}$ has a Gaussian distribution according to the...
fluctuation-dissipation theorem \[6, 94\]. The equation can be discretized using an Euler scheme (eq. (2.60)):

\[
r^{n+1} = r^n + \frac{\Delta t}{\gamma} \mathbf{f}(r^n) + \eta^n, \quad \left\langle \eta^m_i \eta^m_j \right\rangle = \frac{2k_B T \Delta t}{\gamma} \delta_{nm} \delta_{ij}
\]

(8.2)

where \(\Delta t\) is the time step, \(i\) and \(j\) indicate vector components, \(k_B\) is the Boltzmann constant, \(T\) is the temperature and \(\delta\) is the Kronecker delta function. Time scales linearly with \(\gamma\), because \(\Delta t\) and \(\gamma\) only occur together. This means that one simulation is valid for any value of \(\gamma\) when the time is scaled accordingly.

Constraints are implemented using the LINCS algorithm \[10\] (see chapter 3). This is a non-iterative matrix method, which projects out the forces that alter the bond lengths. The projection algorithm allows for larger time steps than the SHAKE algorithm \[9\]. We have implemented the BD and LINCS algorithms in the GROMACS package \[3\].

The forcefield should have a mean-field correction for the water interactions and a soft-core potential for the interactions between the unified atoms in the side chains. At present a very rough approximation is used. The electrostatics are screened with a distance-dependent dielectric constant, which goes from 1 at short distance to 78 at infinity, following a sigmoidal curve. The Lennard-Jones (LJ) interactions are scaled down compared with the GROMOS-87 forcefield \[95, 96, 97\] with explicit solvent, the details are explained in Table 8.1. The main term which is missing in this forcefield is a solvation term, which adds a favorable energy for charges that interact with the solvent. A first approximation of this term could be based on solvent accessible surface.

The friction coefficient for particles at the surface of the protein can be approximated by applying Stokes law to a sphere with size of a carbon atom in water: \(5 \cdot 10^{-12}\) kg/s. For simplicity we also used this value for particles inside the protein. When comparing MD simulations with solvent and the BD method without explicit solvent this seems a reasonable value. We have compared several MD simulations of the central helix of Calmodulin \[98\] with BD simulations and the bending of the helix occur on the same time scale of several nanoseconds (results not shown).

### 8.3 Results

#### Structural changes

To test the model we used HPr, the Histidine-containing Phosphocarrier protein from *Escherichia coli*. We choose this protein because it is relatively small (85 residues) and contains 3 \(\alpha\)-helices as well as a 4 stranded \(\beta\)-sheet. Also a lot of experimental and simulation data are available. A cluster of 30 possible solution structures was determined by NMR and restrained MD simulations \[88\].

After simplifying the side chains, 12 BD simulations were started from structures from the NMR cluster (simulations 1-12 in Table 8.1). Every run is 1 \(\mu\)s, \(8 \cdot 10^7\) time
### 8.3 Results

<table>
<thead>
<tr>
<th>simulation nr.</th>
<th>1-6</th>
<th>7-12</th>
<th>13-15</th>
</tr>
</thead>
<tbody>
<tr>
<td>LJ-factor PP/PH</td>
<td>0.2</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>LJ-factor HH</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>charges</td>
<td>full</td>
<td>full</td>
<td>halved</td>
</tr>
<tr>
<td>temperature (K)</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>starting conf.</td>
<td>(a)</td>
<td>(a)</td>
<td>(b)</td>
</tr>
<tr>
<td>$\sim \Delta t$ (fs)</td>
<td>12.5</td>
<td>12.5</td>
<td>12.5</td>
</tr>
<tr>
<td>$\sim$ length (ns)</td>
<td>1000</td>
<td>1000</td>
<td>500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>simulation nr.</th>
<th>16-21</th>
<th>22-27</th>
<th>28-33</th>
<th>34-39</th>
<th>40-45</th>
<th>46-51</th>
</tr>
</thead>
<tbody>
<tr>
<td>LJ-factor PP/PH</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>LJ-factor HH</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>charges</td>
<td>full</td>
<td>full</td>
<td>full</td>
<td>full</td>
<td>full</td>
<td>full</td>
</tr>
<tr>
<td>temperature (K)</td>
<td>150</td>
<td>200</td>
<td>250</td>
<td>350</td>
<td>400</td>
<td>450</td>
</tr>
<tr>
<td>starting conf.</td>
<td>(a)</td>
<td>(a)</td>
<td>(a)</td>
<td>(a)</td>
<td>(a)</td>
<td>(a)</td>
</tr>
<tr>
<td>number of steps ($10^6$)</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>$\sim \Delta t$ (fs)</td>
<td>12.5</td>
<td>12.5</td>
<td>12.5</td>
<td>12.5</td>
<td>10.0</td>
<td>8.0</td>
</tr>
<tr>
<td>$\sim$ length (ns)</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 8.1: Details of the 51 Brownian Dynamics simulations of HPr. The force field used in the simulations is based on the GROMOS-87 forcefield. All the parameters are given in appendix 8.A. Lennard-Jones parameters for the simplified side chains were derived by fitting one LJ function to the total LJ field of several atoms. All LJ parameters were scaled, based on the particle types. LJ-factor PP/PH is the scaling factor for the interactions between polar and between polar and hydrophobic particles. LJ-factor HH is the scaling factor for the interactions between hydrophobic particles. The charges row indicates if full charges or halved charges were used on charged side chains. All simulations were performed with a twin-range cut-off of 0.8/1.4 nm. The Coulomb forces between 0.8 and 1.4 nm were updated every 25 steps. The distance-dependent dielectric constant $\epsilon$ is defined as $\epsilon = D - (D - 1)(0.5rs^2 + rs + 1)e^{-rs}$, where $D = 78$ and $s = 1.6$ nm$^{-1}$. The temperature is the temperature which was used to determine the noise level. The starting configurations were taken from 1hdn.pdb [88], (a) is structures 1, 2, 12, 17, 19 and 26, (b) is structures 1, 2 and 12. Each simulation was started with a different random seed.
steps of 12.5 fs, note that this time is not accurate since it scales linearly with the friction coefficient. Each run took one week CPU time on one SGI R10000 processor.

Two time scales can be distinguished in which the structures adapt to the forcefield (Figure 8.2). In a few nanoseconds the backbone root mean square (RMS) deviation from the starting structure rises rapidly and half of the secondary structure is lost according to the DSSP program [99]. After 50 ns half of the runs reach their maximum RMS deviation, the other half reach their maximum between 50 and 500 ns.

Closer inspection of the structures reveals that the strength of the Coulomb interactions of charged side chains is overestimated. These side chains are mostly located on the outside of the protein, exposed to the solvent. The charges are screened by the solvent, resulting in lower effective Coulomb interactions with the rest of the protein.
the model oppositely charged side chains attract each other too strongly and disrupt the structure of the protein. The screening can be approximated by halving all the charges in the charged side chains, while leaving the dipole charges unchanged. Three simulations of 500 ns were performed with this new forcefield (simulations 13-15 in Table 8.1). After 500 ns less than a quarter of the secondary structure is lost, which is a significant improvement over the forcefield with full charges. The backbone RMS deviation reaches plateaus of 0.35 and 0.42 for runs 13 and 15 respectively. Initially run 14 has the lowest RMS deviation of all the runs at 300 K, but the protein starts to unfolds after 200 ns, when the C-terminal helix looses contact with the $\beta$-sheet.

It is not yet clear to what extent the simulation structures correspond to real solution structures. This agreement depends of course critically on the choice of parameters. Recent developments in NMR spectroscopy [100] and the analysis of X-ray data [101] indicate that proteins might exhibit larger conformational changes than previously expected from nanosecond MD simulations.

**Comparison with NMR data**

Distances between hydrogens can be compared with upper bounds derived from Nuclear Overhauser Effects (NOE’s) [88]. The strength of a NOE depends on the distance between one or more pairs of hydrogen atoms. When the distance between hydrogens fluctuates on time scales larger than the rotation correlation time of the protein, the NOE is proportional to $\langle d^{-6} \rangle$, where the averaging is over time and over all the proteins in the ensemble. Measured NOE’s can be converted into upper bounds for the average distance and thus provide a means to verify the simulation results [102].

We have transformed upper bounds with non-polar hydrogens to upper bounds on carbon atoms. From the simulation structures, we calculated the average distances as $\langle d^{-6} \rangle^{-1/6}$, where the averaging is over time and trajectories (Table 8.2). A convenient measure for the disparity of experimental and theoretical values is the so-called 'sum of violations'. A violation occurs when the difference between the measured distance and NMR derived upper-bound is larger than zero, when this is the case, the value of the violation is equal to this difference. The sum of violations over single trajectories is between 18 and 50 nm, but averaging over the ensemble of trajectories 1-12 reduces this to 3.23 nm. The large difference between single trajectories and the whole ensemble is caused by the $\langle d^{-6} \rangle$ averaging. A satisfied upper-bound in one or two trajectories can compensate for violations of the same upper-bound in all the other trajectories.

To get an idea of the significance of these violations, we compared results of a 'narrow' and a 'broad' ensemble (Table 8.2). The 'narrow' ensemble consists of two MD trajectories of 5 ns of HPr in explicit solvent. The 'broad' ensemble consists of 69 MD trajectories, which were performed in vacuo starting from a completely extended $\beta$-strand. Because of the attractive Lennard-Jones forces the structures collapse to a 'frozen' random state in 40 ps. The ensemble of structures, each of which is almost as
A model for microsecond protein dynamics

Table 8.2: Comparison of $\langle d^{-6} \rangle^{-1/6}$ averaged distances in 4 HPr ensembles with upper-bounds derived from NMR measurements. Only 744 of the 1108 restraints were taken into account, because of the simplified side chains in the BD model. The ensembles are: BD: Brownian Dynamics trajectories 1-12 with a simplified protein model and without explicit solvent; MD: 2 Molecular Dynamics trajectories of HPr in a box of 5315 water molecules, with a twin-range cut-off of 1.0 and 1.8 and a time step of 2 fs; random: 69 MD trajectories in vacuo starting from a completely extended structure; 1hdn.pdb: 30 structures modeled from NMR data [88].

<table>
<thead>
<tr>
<th></th>
<th>BD</th>
<th>MD</th>
<th>random</th>
<th>1hdn.pdb</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of runs</td>
<td>12</td>
<td>2</td>
<td>69</td>
<td>-</td>
</tr>
<tr>
<td>start time for averaging (ns)</td>
<td>100</td>
<td>1</td>
<td>0.042</td>
<td>-</td>
</tr>
<tr>
<td>end time for averaging (ns)</td>
<td>1000</td>
<td>5</td>
<td>0.050</td>
<td>-</td>
</tr>
<tr>
<td>number of structures</td>
<td>10812</td>
<td>8002</td>
<td>345</td>
<td>30</td>
</tr>
<tr>
<td>ensemble averages:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>radius of gyration (nm)</td>
<td>1.08</td>
<td>1.09</td>
<td>1.17</td>
<td>1.12</td>
</tr>
<tr>
<td>sum of violations (nm)</td>
<td>3.23</td>
<td>0.68</td>
<td>127.41</td>
<td>0.18</td>
</tr>
<tr>
<td>number of violations</td>
<td>41</td>
<td>11</td>
<td>376</td>
<td>4</td>
</tr>
<tr>
<td>largest violation (nm)</td>
<td>0.33</td>
<td>0.14</td>
<td>1.25</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Temperature effects

The temperature in a BD model scales the noise with respect to strength of the interactions. Since the number of degrees of freedom has been reduced, the potential is actually a potential of mean force, which is the result of averaging over an equilibrium distribution of the degrees of freedom that have been removed. Therefore a potential of mean force is temperature dependent and might not be valid for all temperatures. Scaling the temperature is a simple method to verify if the strength of the interactions is correct and also to check if the potential of mean force is still valid at higher temperatures where an unfolding transition should take place.

To determine the temperature dependence of the model runs of 200 ns were per-
formed at 6 different temperatures (simulations 16-51 in Table 8.1). Results are shown in Figure 8.2. The amount of secondary structure decreases with increasing temperature, the largest difference is between 300 and 350 K. If we consider structures with an RMS deviation of 0.55 nm from the starting structure folded, then the average RMS deviation of the folded structures does not change with temperature, but the number of unfolded structures does increase with temperature. After 200 ns unfolding occurs in runs 29, 37, 49 and 50, with an RMS deviation of about 0.8 nm and very little secondary structure. Runs 30, 43, 44, 46, 48 and 51 show partial unfolding, with high RMS deviations of 0.58 up to 0.7 nm, although most of the structures still have an average amount of secondary structure. In all ten unfolding runs the two larger $\alpha$-helices loose contact with the $\beta$-sheet and the small $\alpha$-helix. The model does not exhibit a sharp, cooperative unfolding transition. It might not be possible to reproduce this with a potential of mean force which is strictly only valid at one temperature. However, the general trend is correct, as a lot of structure is lost at and above 350 K where also most unfolding events occur. Furthermore, the simulations clearly demonstrate the relevance of combining the results from an ensemble of simulations. The individual trajectories may behave quite differently: some unfold, others do not.

The energy landscape

To give a visual impression of the ensemble, the trajectories can be projected on the two degrees of freedom with the largest fluctuations. To obtain these degrees of freedom a covariance matrix of the ensemble was built and diagonalized [60]. The ensemble was projected on the two eigenvectors with the largest eigenvalues. Figure 8.3 shows the projections of the simulations 1-12. The runs were started from six structures in the middle of the picture. This two dimensional plot shows that the conformations do not change gradually, but rather jump from one conformation or energy minimum to another. This behavior can be seen more clearly in an RMSD matrix, in which the RMS deviation between all pairs of frames in a trajectory is plotted. Figure 8.4 shows RMSD matrices for trajectories 8 and 6. Matrix (a) is a typical example of an BD trajectory. The light blocks on the diagonal indicate that the protein stays in one conformation or energy minimum for a certain length of time. The light blocks are embedded in a hierarchy of darker blocks revealing a hierarchy of energy minima. Matrix (b) has large off-diagonal blocks with low RMS deviation. This means that the protein revisits a conformation it had already sampled earlier in the run. These light off-diagonal blocks are very scarce in the twelve runs, thus the sampling is far from complete. A detailed view of a transition between two energy minima in run 14 is shown in Figure 8.5. The major part of the difference between the two conformations is caused by the rearrangement of the C-terminal end of the smallest helix and the loop connecting this helix to a $\beta$-strand. The path between the two conformations goes through a state (from 56 to 58 ns) which is very similar to both conformations. Within both minima several smaller minima can be distinguished.
12 trajectories of 1 μs each, projected onto the two coordinates (eigenvectors) with the largest variation. A third of the displacement in the runs is in these two coordinates. Two runs start from each of the six starting structures in the middle. The gray lines are trajectories 1-6, the black lines trajectories 7-12 (see Table 8.1). Each line is a step of 1 ns. Areas of high density are connected by a few lines, indicating that the protein quickly jumps from one conformation to another.
Figure 8.4: RMSD Matrices for trajectories 8 (top) and 6 (bottom), showing the backbone RMS deviation between every pair of frames in the trajectory. The light off-diagonal blocks in the bottom matrix indicate that the protein revisits a conformation it has already sampled before.
Figure 8.5: RMS deviation matrix for part of trajectory 14. A transition between two energy minima (the two light blocks) is visible, which goes through a state (from 56 to 58 ns) which is similar to the conformations in both minima. Both light blocks contain a gray-scaled block pattern, the small light squares correspond to small energy minima.
8.4 Discussion and conclusions

In principle the free energy landscape can be reconstructed from RMSD matrices, since the relative sizes of the blocks are determined by the free energy differences between the conformations. However, the results show that the sampling is incomplete. The block sizes range over all the time scales which are available from the trajectories, i.e., from a nanosecond to a microsecond. The distances between the minima vary from 0.1 to 0.4 nm. Troyer and Cohen analyzed RMSD matrices for nanosecond MD simulations of bpti [103]. They found minima from the sub-picosecond up to the nanosecond time scale, separated by 0.01 up to 0.07 nm. In the picosecond range the minima are invisible due to the thermal noise. The RMSD matrices of MD and BD trajectories show a scaling behavior over 6 orders of magnitude in time, which will probably persist for another few orders of magnitude, until full sampling has been reached.

8.4 Discussion and conclusions

The ultimate challenge in the field of theoretical protein research is to predict the fold of a protein given its amino-acid sequence. It is important to realize that a correct prediction does not necessarily imply further insight into the process of folding itself, or related to this, insight in large-scale motions of proteins in solution. Several computer algorithms have been developed to predict the structure of proteins. These algorithms search for the minimum of an energy function, which can use a force field with a physical basis [85] or parameters derived from a database of known protein structures [86], assuming Boltzmann statistics [104]. Although search algorithms might be the most efficient way to find the fold of a protein, they do not provide insight into the process of protein folding and the nature of folding intermediates. A dynamical approach does provide this insight.

With the present model a significant part of the energy landscape close to the native state of a protein can be sampled by a dynamical approach. The results re-confirm the picture of a flat energy landscape with many shallow minima as described by Bryngelson et al. [105] and Chan [89]. In addition, we have shown that important physical properties such as landscape hopping and hierarchical dynamical scaling are present in the model. The model can also provide a better characterization of intermediate states through which a protein passes from one conformational sub-state to the other. The RMSD matrices show that the time spent in intermediate states is on the order of nanoseconds, which is close to the time estimated by Mössbauer experiments on myoglobin [106].

We believe that the model is not only relevant for the physics of complex systems, but also for gaining insight into the dynamics of specific proteins. The general results are not very sensitive to the details of the forcefield. For answering questions about a specific protein the parameters and forcefield need to be improved. That this is possible, at least to some extent, is indicated by the results with the forcefield with halved charges.

Proteins are fragile structures, with a stability of 10-20 \( k_B T \) compared to the denatured state. Our model reproduces this marginal stability, with a transition, although not
at sharp as in real proteins, between 300 and 350 K. A few runs show unfolding. Atomic distances in the generated ensemble are in reasonable agreement with upper bounds derived from NMR data. Although the method reaches time scales which are three orders of magnitude larger than those of regular MD simulations, the sampling is far from complete. The large number of runs presented in this chapter all sample different parts of the conformational space. The way in which the space is sampled is quite similar for all the different runs. The protein samples a local minimum for a long time and jumps to the next minimum in a short time.

Doing only one run of a few microseconds [107, 108] is not sufficient for the calculation of ensemble properties. As we have shown here, a number of microsecond trajectories need to be combined.
8.A Appendix: parameters

[ atomtypes ]

; The Lennard-Jones parameters.
; The parameters for two
; different particle-types
; a and b are calculated as:
; sqrt(c_a * c_b)
; LJ and Coulomb interactions
; are not present for particles
; which are separated by one or
; two bonds.
;
; name    unscaled c6    unscaled c12
; (kJ nm^{-6})  (kJ nm^{-12})
; Gromos-87 atomtypes
O  0.22617E-02  0.74158E-06
OM 0.22617E-02  0.74158E-06
OA 0.22617E-02  0.15062E-05
N  0.24362E-02  0.16924E-05
NT 0.24362E-02  0.16924E-05
NL 0.24362E-02  0.16924E-05
C  0.23402E-02  0.33740E-05

CH1 0.12496E-01  0.71747E-04
CH2 0.90975E-02  0.35333E-04
CH3 0.88765E-02  0.33740E-05
CB 0.23402E-02  0.33740E-05
H  0.00000E+00  0.00000E+00
S  0.99844E-02  0.13078E-04
NZ 0.24362E-02  0.16924E-05
NE 0.24362E-02  0.16924E-05
NA 0.72059E-04  0.21014E-07
CA 0.10561E-01  0.21499E-04
SD 0.10561E-01  0.21499E-04
CD 0.90507E-02  0.21758E-04

; phenyl ring
BEN1 0.66430E-01  0.55162E-03

; histidine side-chain
W5-R 0.66430E-01  0.55162E-03

; two carbons (Arg,Lys,Ile)
C2H4 0.19761E-01  0.10847E-03

; lysine NH3 group
NH3T 0.29160E-02  0.21258E-05

; end of arginine side-chain
N3CT 0.28150E-01  0.99052E-04

; acid group (Glu,Asp)
C-O2 0.28150E-01  0.99052E-04

; amide group (Gln,Asn)
C-ON 0.28150E-01  0.99052E-04

; hydroxyl group
OH-U 0.22617E-02  0.15062E-05

[ bondtypes ]

; Bond-lengths.
;
; i   j   bond-length
; (nm)

C O 0.123
C OM 0.125
C OA 0.136
C N 0.133
C NT 0.133
C NL 0.133
CH1 OA 0.143
CH1 N 0.147
CH1 NT 0.147
CH1 NL 0.147
CH1 C 0.153
CH1 CH1 0.153
CH2 OM 0.143
CH2 OA 0.143
CH2 N 0.147
CH2 NT 0.147
CH2 NL 0.147
CH2 C 0.153
CH2 CH1 0.153
CD2 CH2 0.153
CD3 C 0.153
CD3 CH1 0.153
CD4 CH3 0.153
CD5 CH3 0.153
CD6 CH3 0.153
A model for microsecond protein dynamics

CH3 CH2 0.153  
CB O 0.123
CB OA 0.136
CB NT 0.133
CB C 0.153
CB CH2 0.153
CB CH3 0.153
CB CB 0.139
H N 0.100
H NT 0.100
H NL 0.100
S O 0.150
S CH2 0.183
S CH3 0.178
S S 0.204
NZ C 0.134
NZ H 0.100
NE C 0.134
NE CH2 0.147
NE H 0.100
CH2 BEN1 0.292
CH1 CDUM 0.153
CDUM C4-L 0.153
CH2 W5-R 0.270
W5-R BEN1 0.217
BEN1 OA 0.275
BEN1 OH-U 0.275
CH1 C2H4 0.200
CH2 C2H4 0.200
N C2H4 0.194
C2H4 C2H4 0.250
C2H4 NH3T 0.187
C2H4 N3CT 0.187
CH2 C-02 0.196
CH2 C-ON 0.196
CH1 OH-U 0.143
CH2 OH-U 0.143

[ angletypes ]

; The potential is harmonic in
; the angle.
;
; i  j  k  angle force-c.
;
; (deg.) (kJ/mol)
;
N CH1 C2H4 109.5  20.00  C CH1 N 109.5  460.24
C CH1 C2H4 109.5  20.00  C CH1 NT 109.5  460.24
8.A Appendix: parameters

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>109.5</th>
<th>460.24</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>CH1</td>
<td>NL</td>
<td></td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>OA</td>
<td>109.5</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>N</td>
<td>109.5</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>NT</td>
<td>109.5</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>NL</td>
<td>109.5</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>C</td>
<td>109.5</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>CH1</td>
<td>111.0</td>
</tr>
<tr>
<td>CH1</td>
<td>OA</td>
<td>109.5</td>
<td></td>
</tr>
<tr>
<td>CH1</td>
<td>N</td>
<td>109.5</td>
<td></td>
</tr>
<tr>
<td>CH1</td>
<td>NT</td>
<td>109.5</td>
<td></td>
</tr>
<tr>
<td>CH1</td>
<td>NL</td>
<td>109.5</td>
<td></td>
</tr>
<tr>
<td>CH1</td>
<td>C</td>
<td>109.5</td>
<td></td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>OA</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>N</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>NT</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>NL</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>C</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>CH1</td>
<td>111.0</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>OM</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>OA</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>N</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>NT</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>C</td>
<td>111.0</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>CH1</td>
<td>111.0</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>CH2</td>
<td>111.0</td>
</tr>
<tr>
<td>CH2</td>
<td>CH3</td>
<td>OM</td>
<td>109.5</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>OA</td>
<td>111.0</td>
</tr>
<tr>
<td>CH2</td>
<td>CH2</td>
<td>CH2</td>
<td>111.0</td>
</tr>
<tr>
<td>CH3</td>
<td>CH2</td>
<td>CH1</td>
<td>111.0</td>
</tr>
<tr>
<td>CH3</td>
<td>CH2</td>
<td>CH2</td>
<td>111.0</td>
</tr>
<tr>
<td>CH4</td>
<td>CH1</td>
<td>OA</td>
<td>109.5</td>
</tr>
<tr>
<td>CH3</td>
<td>CH2</td>
<td>CH1</td>
<td>111.0</td>
</tr>
<tr>
<td>CB</td>
<td>CH2</td>
<td>OM</td>
<td>109.5</td>
</tr>
<tr>
<td>CB</td>
<td>CH2</td>
<td>CH1</td>
<td>111.0</td>
</tr>
<tr>
<td>CB</td>
<td>CH2</td>
<td>CH2</td>
<td>111.0</td>
</tr>
<tr>
<td>CB</td>
<td>CH2</td>
<td>CH1</td>
<td>113.0</td>
</tr>
<tr>
<td>CB</td>
<td>CH2</td>
<td>CH2</td>
<td>113.0</td>
</tr>
<tr>
<td>NE</td>
<td>CH2</td>
<td>CH2</td>
<td>109.5</td>
</tr>
<tr>
<td>CB</td>
<td>CB</td>
<td>O</td>
<td>120.0</td>
</tr>
<tr>
<td>CB</td>
<td>CB</td>
<td>NT</td>
<td>120.0</td>
</tr>
</tbody>
</table>

[ improper dihedral types ]

; The parameters for an improper dihedral on atoms i,j,k,l, the type is determined by the outer two atoms. The potential is harmonic in the angle. The angle is the angle between the normals of the planes i-j-k.
A model for microsecond protein dynamics

; and j-k-l.
; i l angle force-c.
; (degrees) (kJ/mol
; rad^-2)

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CH1</td>
<td>C2H4</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CH3</td>
<td>C2H4</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>C</td>
<td>OM</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>C</td>
<td>NT</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CH1</td>
<td>OA</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CH2</td>
<td>N</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CH2</td>
<td>NT</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CH2</td>
<td>C</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CH3</td>
<td>CH1</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CB</td>
<td>O</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CB</td>
<td>OA</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CB</td>
<td>NT</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CB</td>
<td>C</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CB</td>
<td>CH2</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CB</td>
<td>CH3</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CB</td>
<td>CB</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>H</td>
<td>N</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>H</td>
<td>NT</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>H</td>
<td>NL</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>NZ</td>
<td>C</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>NE</td>
<td>C</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>NE</td>
<td>H</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>C</td>
<td>CH1</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>0.000</td>
<td>167.36</td>
</tr>
<tr>
<td>CH1</td>
<td>OA</td>
<td>35.264</td>
<td>334.72</td>
</tr>
</tbody>
</table>

[ proper dihedral types ]

; The parameters for a proper
dihedral on atoms i,j,k,l, the
type is determined by the
; inner two atoms. The potential
; is a cosine function. The
; angle is the angle between the
; normals to the planes i-j-k
; and j-k-l.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CH1</td>
<td>OA</td>
<td>0.000</td>
<td>1.255</td>
</tr>
<tr>
<td>CH1</td>
<td>N</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH1</td>
<td>NT</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH1</td>
<td>NL</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH1</td>
<td>C</td>
<td>0.000</td>
<td>0.418</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>0.000</td>
<td>5.858</td>
</tr>
<tr>
<td>CH2</td>
<td>OA</td>
<td>0.000</td>
<td>1.255</td>
</tr>
<tr>
<td>CH2</td>
<td>N</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH2</td>
<td>NT</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH2</td>
<td>NL</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH2</td>
<td>C</td>
<td>0.000</td>
<td>0.418</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>0.000</td>
<td>5.858</td>
</tr>
<tr>
<td>CB</td>
<td>OA</td>
<td>0.000</td>
<td>7.113</td>
</tr>
<tr>
<td>CB</td>
<td>NT</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>CB</td>
<td>C</td>
<td>0.000</td>
<td>2.929</td>
</tr>
<tr>
<td>CB</td>
<td>CH2</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>CB</td>
<td>CH3</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>CB</td>
<td>CB</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>H</td>
<td>N</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>H</td>
<td>NT</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>H</td>
<td>NL</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>NZ</td>
<td>C</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>NE</td>
<td>C</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>NE</td>
<td>H</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>C</td>
<td>CH1</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>CH1</td>
<td>OA</td>
<td>35.264</td>
<td>16.736</td>
</tr>
<tr>
<td>NE</td>
<td>H</td>
<td>180.0</td>
<td>33.472</td>
</tr>
<tr>
<td>C</td>
<td>CH1</td>
<td>180.0</td>
<td>33.472</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>0.000</td>
<td>0.418</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>0.000</td>
<td>2.929</td>
</tr>
<tr>
<td>CH1</td>
<td>OA</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CH2</td>
<td>N</td>
<td>180.0</td>
<td>33.472</td>
</tr>
<tr>
<td>CH2</td>
<td>NT</td>
<td>180.0</td>
<td>33.472</td>
</tr>
<tr>
<td>CH2</td>
<td>C</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CH3</td>
<td>CH1</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>CH1</td>
<td>N</td>
<td>180.0</td>
<td>0.418</td>
</tr>
<tr>
<td>CH1</td>
<td>NT</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH1</td>
<td>NL</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH1</td>
<td>C</td>
<td>0.000</td>
<td>0.418</td>
</tr>
<tr>
<td>CH1</td>
<td>CH1</td>
<td>0.000</td>
<td>5.858</td>
</tr>
<tr>
<td>CH2</td>
<td>OA</td>
<td>0.000</td>
<td>1.255</td>
</tr>
<tr>
<td>CH2</td>
<td>N</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH2</td>
<td>NT</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH2</td>
<td>NL</td>
<td>0.000</td>
<td>3.766</td>
</tr>
<tr>
<td>CH2</td>
<td>C</td>
<td>0.000</td>
<td>0.418</td>
</tr>
<tr>
<td>CH2</td>
<td>CH1</td>
<td>0.000</td>
<td>5.858</td>
</tr>
<tr>
<td>CB</td>
<td>OA</td>
<td>0.000</td>
<td>7.113</td>
</tr>
<tr>
<td>CB</td>
<td>NT</td>
<td>180.0</td>
<td>33.472</td>
</tr>
<tr>
<td>CB</td>
<td>C</td>
<td>180.0</td>
<td>5.858</td>
</tr>
<tr>
<td>CB</td>
<td>CH2</td>
<td>0.000</td>
<td>0.418</td>
</tr>
<tr>
<td>CB</td>
<td>CH3</td>
<td>0.000</td>
<td>2.929</td>
</tr>
<tr>
<td>CB</td>
<td>CB</td>
<td>0.000</td>
<td>16.736</td>
</tr>
<tr>
<td>NE</td>
<td>C</td>
<td>180.0</td>
<td>33.472</td>
</tr>
<tr>
<td>NE</td>
<td>H</td>
<td>180.0</td>
<td>33.472</td>
</tr>
<tr>
<td>C</td>
<td>CH1</td>
<td>180.0</td>
<td>5.858</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>0.000</td>
<td>0.418</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>0.000</td>
<td>2.929</td>
</tr>
<tr>
<td>CH1</td>
<td>OA</td>
<td>35.264</td>
<td>334.72</td>
</tr>
<tr>
<td>NE</td>
<td>CH2</td>
<td>180.0</td>
<td>0.418</td>
</tr>
<tr>
<td>CDUM</td>
<td>CH1</td>
<td>0.000</td>
<td>5.858</td>
</tr>
</tbody>
</table>