Comparison of order parameters and internal correlation times determined from experimental and simulated data

Anke Eisenmann[a], Finn Bauer[a,b], Kristian Schweimer[a], Heinrich Sticht[b], and Paul Rösch[a]

[a]Lehrstuhl für Biopolymere, Universität Bayreuth, D - 95440 Bayreuth
[b]Institut für Biochemie, Abt. Bioinformatik, Emil-Fischer-Zentrum, Universität Erlangen-Nürnberg, 91054 Erlangen

UNIVERSITÄT BAYREUTH

ABSTRACT

The early view of proteins as relatively rigid structures has been replaced by a dynamic model in which internal motions play an essential role for protein function [2]. Unfortunately, data describing internal dynamics microscopically are difficult to obtain by experimental techniques, such as NMR or fluoresence measurements. In contrast MD simulations provide an atomistic picture of conformational changes as a function of time and are routinely employed to gain insight into dynamic behaviour of proteins. However, until now the validation of simulated data has proven to be difficult, mostly because of the requirement of long simulation times. In this work, NMR relaxation data of the Lck-SH3 domain were used to identify residues showing motions on a 10-50 ps time scale. Back-calulation showed that similar order parameters could be obtained by simulation and experiment, while the time scales estimated from simulated data were considerably shorter.

ANALYSIS OF NMR DATA

Motions in proteins can span a timescale from picoseconds up to seconds. The overall reorientation of the protein is described by a rotational correlation time of about 5-10 ns, while fast internal motions of the backbone range from pico- to nanoseconds. Slow internal motions involve a microsecond to second timescale.

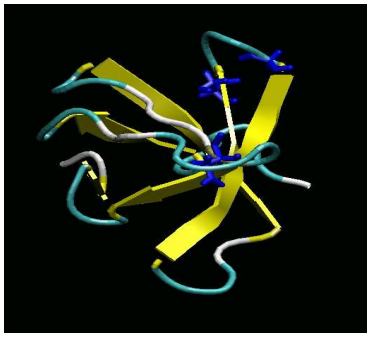

Fig 1. Solution structure of the Lck-SH3 domain. The domain consists of 63 residues forming a five stranded beta barrel (yellow). Residues selected for the back-calculation of motional parameters (G21, N44, T48) are shown as blue sticks. (Drawing with VMD)

A quantitative description of protein dynamics may be obtained by a formalism introduced by Lipari and Szabo [1] and extended by Clore et. al. [4]. Thereby the overall rotational correlation time τ_m is held fixed while internal motions are determined by optimizing the adjustable parameters of the following five motional models:

1 S^2
2 S^2, τ_e
3 S^2, k_{ex}
4 S^2, τ_e, k_{ex}
5 S^2_s, S^2_f, τ_e

In Models 1-4 the order parameter S^2 describes the amplitude of the internal motion and the internal correlation time τ_e its timescale. Model 3 and 4 incorporate an additional parameter k_{ex} accounting for exchange processes. A second timescale is introduced in Model 5 where S^2_s and $\tau_e = \tau_s$ characterize slow internal motions, S^2_f very fast internal motions and $S^2=S^2_s \cdot S^2_f$. For the evaluation of the global correlation time residues expected or found to take part in fast internal motions or exchange are omitted.

	S^2	$\tau_e \times 10^{-11}$ s
G21	0.79 ± 0.01	5.1 ± 1.3
N44	0.83 ± 0.01	4.7 ± 1.6
T48	0.80 ± 0.02	3.4 ± 1.4

Order parameters and internal correlation times of residues selected for back-calculation.

In this work, relaxation rates R1, R2, and the heteronuclear NOE acquired at 600 MHz were analysed using the software package TENSOR2. The internal motions were calculated assuming an isotropic diffusion model with a rotational correlation time of 4.6 ns. Of all residues that could be fitted to model 2 (corresponding to a rapid, low amplitude internal motion) three residues with an internal correlation time τ_e of 10 - 50 ps were selected for back-calculation. This time scale is slow compared to vibrational motions but fast compared to the overall rotational motion of the protein. Thus coupling of internal motions with the overall tumbling can be neglected and the SHAKE algorithm can be applied.

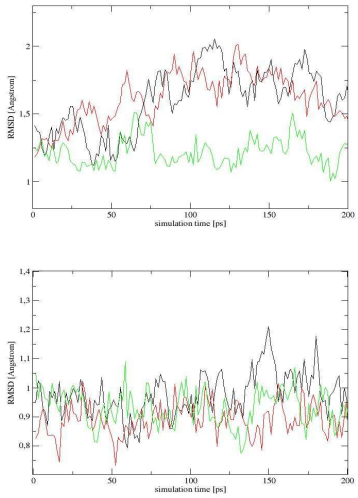

Fig 2. Backbone RMSD from the starting structure as a function of time in simulations performed with different initial configurations. (black: run #2; red: run #4; green: run #5). The structures were superimposed using the backbone atoms N, C', and C_α. RMSD values were computed for all residues in A), and for residues 7-61 in B), excluding the flexible N- and C-termius. The rather low RMSD of 0.8-1.2 Å in B) indicates a stable trajectory.

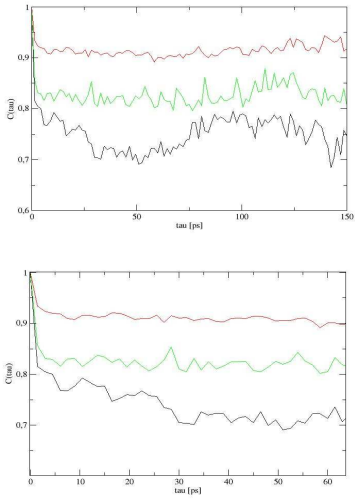

Fig 3. Time-correlation functions for three different residues of the Lck-SH3 domain calculated from run #4. A) over the full length of the simulation B) over a sector of the first 65 ps. The results for three backbone N-H vectors of G21 (black), N44 (red), and T48 (green) are shown.

MOLECULAR DYNAMICS SIMULATION

All molecular dynamics simulations were performed with the AMBER6.0 simulation package using the Cornell et. al. force field (parm94) [7] and the TIP3P water model. Six different starting structures were used, corresponding to the six lowest energy structures from the NMR ensemble (pdb accession code 1h92) [8]. The molecules were solvated in a ~60x60x60 Å waterbox and neutralized by adding counterions. After energy minimization the temperature of the system was raised gradually from 100 to 300 K, and the system was equilibrated at 300 K for 22.5 ps. For data collection a 202.5 ps MD simulation was performed with frames collected every 1.5 ps. A weak temperature and pressure coupling scheme was applied with coupling constants of 0.2 ps during equilibration and 0.5 ps during the production phase. The SHAKE procedure was used throughout the simulation to constrain hydrogen bonds and the integration step was 1.5 fs.

The MD snapshots were analysed using the ptraj program of the AMBER package. Prior to the calculation of the correlation coefficient, the overall rotational motion was eliminated by superimposing the frames on the principal axis system defined by the relatively rigid residues 9-16, 22-25, 28-34, 40-42, 51-60.

The correlation coefficient is given by

$$C(\tau) = \langle P_2 (\cos \chi_{t,t+\tau}) / r^3(t) r^3(t+\tau) \rangle$$

where P_2 is the second-order Legendre polynomial and $\chi_{t,t+\tau}$ is the angle between the interspin vector at the two timepoints t and t + τ [3]. The time correlation function C(τ) for a simple motion can be represented by a single decaying exponential with τ_e being the internal correlation time:

$$C(\tau) = \exp(-\tau/\tau_e)$$

For a spatially restricted internuclear vector, the time correlation function will not decay to zero, but reach a plateau value corresponding to the square of the order parameter S:

$$S^2 = \lim_{\tau \to \infty} C(\tau)$$

Figure 3 illustrates the time correlation functions estimated for the backbone N-H vectors of G21, N44, and T48 from a single molecular dynamics run (run #4). In the first 65 ps the correlation coefficients follow a steep mono-exponential decay with time constants in the range of 1-5 ps and plateau values of 0.91 for N44, 0.83 for T48, and 0.73 for N44 (Fig 3B). In the range of 65-150 ps the plateau values of T48 and N44 remain relatively constant, while the correlation coefficient for G21 increases slightly.

Figure 4 depicts the time correlation functions calculated for only one residue, T48, using three different starting structures (runs #2, #4 #5). The examination of the exponential decays observed in the first 65 ps yields squared order parameters from 0.72-0.83 and time constants from 1-5 ps. In the course of the simulation only correlation coefficients calculated from run #4 show a stable plateau value.

Though the magnitude of the order parameters agrees quite well with the data determined by the Lipari-Szabo analysis, the time scales estimates from simulated data are about ten times smaller than from experimental data. The further decrease of the plateau value observed in figure 4 for two of the correlation functions from simulations with different initial configurations is probably due to an insufficient convergence of the trajectory prior to data collection.

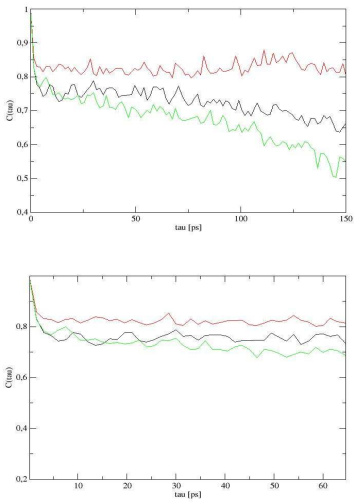

Fig 4. Time-correlation functions for T48 of the Lck-SH3 domain calculated from molecular dynamics runs using different starting structures. A) over the full length of the simulation B) over a sector of the first 65 ps. (black: run #2, red: run #4, green: run #5)

CONCLUSION

Back-calculation of NMR relaxation parameters has been reviewed recently (Case, 2002). Though the reproduction of order parameters has been successful in some cases, time scales observed in simulation are frequently shorter than time scales estimated from experimental data, as for example in the simulation of β-heptapeptides [3] and as observed in this work.

LITERATURE

[1] Lipari and Szabo, J Am Chem Soc, 1982, 104: 4546-4559
[2] Karplus et. al., Nat Struct Biol, 2002, 9(9): 646-652
[3] Peter et. al, J Biomol NMR, 2001, 20: 297-310
[4] Clore et. al., J Am Chem Soc, 1990, 112: 4989-4991
[5] Pawley et. al., J Biomol NMR, 2001, 20: 149-165
[6] Case, Acc Chem Res, 2002, 6: 325-331.
[7] Cornell et. al., J Am Chem Soc, 1995, 117:5179-5197
[8] Schweimer et. al., Biochemistry,2002, 41(16):5120-5130