



Calculation of NMR-relaxation parameters for flexible molecules from molecular dynamics simulations

Christine Peter, Xavier Daura & Wilfred F. van Gunsteren*

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology Zürich, ETH-Zentrum, CH-8092 Zürich, Switzerland

Received 22 December 2000; Accepted 10 May 2001

Key words: internal dynamics, molecular dynamics simulation, NMR relaxation, NOESY, peptides, relaxation matrix, ROESY, spectral density functions

Abstract

Comparatively small molecules such as peptides can show a high internal mobility with transitions between several conformational minima and sometimes coupling between rotational and internal degrees of freedom. In those cases the interpretation of NMR relaxation data is difficult and the use of standard methods for structure determination is questionable. On the other hand, in the case of those system sizes, the timescale of both rotational and internal motions is accessible by molecular dynamics (MD) simulations using explicit solvent. Thus a comparison of distance averages ($\langle r^{-6} \rangle^{-1/6}$ or $\langle r^{-3} \rangle^{1/3}$) over the MD trajectory with NOE (or ROE) derived distances is no longer necessary, the (back)calculation of the complete spectra becomes possible. In the present study we use two 200 ns trajectories of a heptapeptide of β -amino acids in methanol at two different temperatures to obtain theoretical ROESY spectra by calculating the exact spectral densities for the interproton vectors and the full relaxation matrix. Those data are then compared with the experimental ones. This analysis permits to test some of the assumptions and approximations that generally have to be made to interpret NMR spectra, and to make a more reliable prediction of the conformational equilibrium that leads to the experimental spectrum.

Introduction

In the standard methods of structure determination based on NMR relaxation data, interatomic distances in a molecular model are compared to NMR-derived distances. The model can be a single structure or a set of structures, e.g., from a molecular dynamics (MD) simulation. In this case the interatomic distances in the set of structures are usually averaged according to the method that describes best the relationship between fluctuating distances and NMR relaxation parameters, i.e. $\langle r^{-6} \rangle$ - or $\langle r^{-3} \rangle$ -averaging (Tropp, 1980).

This procedure is for two reasons problematic in the case of small molecules that show a high internal mobility with transitions between several conformational minima. First, the averaging method is not very

sensitive to distance fluctuations. A small fraction of structures with short distances between two atoms can make a large contribution to the average, so that a larger fraction of structures with longer distances can be hidden or at least underestimated (Daura et al., 1999a; Bürgi et al., 2001). Second, the interatomic distances and their fluctuations are not the only quantities which determine NMR relaxation rates and thus, for example, ROESY crosspeak intensities, but there is also an additional contribution from intramolecular motions to the observed signal by their timescales and their orientational correlations. This contribution is usually neglected or implicitly included using the model-free approach by Lipari and Szabo (1982). This is reasonable in the case of large and comparatively stiff molecules such as proteins, where the internal dynamics is considerably faster than the overall tumbling (Lipari and Szabo, 1982).

*To whom correspondence should be addressed. E-mail: wfvgn@igc.phys.chem.ethz.ch

Besides these difficulties, the comparison between distances is not really necessary for small molecules, because the whole timescale of both rotational and internal motions is accessible via molecular dynamics simulations even when using explicit solvent molecules. Thus, the (back)calculation of the relaxation parameters and of the complete 2D-NMR spectra is possible.

In the present study, two 200 ns MD simulations of a heptapeptide of β -amino acids in methanol at 298 K and at 340 K (Daura et al., 1998, 1999b) are analyzed. The correlation times and order parameters for rotational and internal motions are investigated in detail, as well as possible couplings between tumbling and internal dynamics. Theoretical ROESY spectra for both trajectories are computed. The influence of the various aspects of internal dynamics (distance fluctuations, timescales, orientational correlations) on the crosspeak intensities is tested, and the analysis shows that it is possible to obtain additional information by considering not only distance averages but the complete dynamical information. Finally, the theoretical ROESY spectra and the buildup curves for selected crosspeaks are compared with values obtained from experimental spectra, leading to new observations about the folding/unfolding equilibrium in the probe.

Methods

Theoretical background

Because the major goal in this study is to calculate NOESY and ROESY spectra, all formulae presented below are only given for the proton-proton dipolar interaction. However, they can be easily adapted for other relaxation-inducing interactions such as chemical shift anisotropy.

Time correlation functions and spectral densities

NMR relaxation, i.e., the return of a spin system to equilibrium, is determined by the transition probabilities between the energy levels involved. These transition probabilities depend on the fluctuations of the relaxation-inducing Hamiltonian and especially on those frequency components of the fluctuations that correspond to the transition frequency. The Hamiltonian that describes the spin system can be separated into two terms, spin operator functions and spacial functions, the latter containing all temporal fluctuations. For the dipolar interaction of two spins I and

S , the Hamiltonian is given by

$$\begin{aligned} \hat{H}(t) &= \frac{\mu_0}{4\pi} \frac{\gamma^2 \hbar^2}{r^3(t)} \left(\hat{\mathbf{I}}(t) \cdot \hat{\mathbf{S}}(t) \right. \\ &\quad \left. - \frac{3}{r^2(t)} (\hat{\mathbf{I}}(t) \cdot \mathbf{r}(t)) (\hat{\mathbf{S}}(t) \cdot \mathbf{r}(t)) \right) \quad (1) \\ &= \frac{\mu_0}{4\pi} \frac{\gamma^2 \hbar^2}{r^3(t)} \sum_{m=-2}^2 (-1)^m Y_{2,-m}(\theta(t), \\ &\quad \phi(t)) \hat{T}_{2,m}(I, S). \end{aligned}$$

Here $\hat{T}_{2,m}(I, S)$ are the second rank irreducible tensor spin-operators and $Y_{2,-m}(\theta, \phi)$ are the second rank spherical harmonic functions (Tropp, 1980). $\mathbf{r}(t)$ is the internuclear vector and $\theta(t)$ and $\phi(t)$ are its time-dependent polar angles in a laboratory frame of reference. γ is the gyromagnetic ratio of the proton, \hbar is Planck's constant divided by 2π and μ_0 is the permeability of free space. As shown elsewhere (Appendix of Neuhaus and Williamson (1989) or Fischer et al. (1998)) transition probabilities, and thus relaxation rates, are determined by the discrete Fourier-coefficients of time-correlation functions $C^{mn}(\tau)$ of the spherical harmonics involved in Equation 1

$$C^{mn}(\tau) = 4\pi \left\langle \frac{Y_{2,m}(\theta(t), \phi(t)) Y_{2,n}(\theta(t+\tau), \phi(t+\tau))}{r^3(t) r^3(t+\tau)} \right\rangle. \quad (2)$$

Spectral-density functions $J^{mn}(\omega)$ are the Fourier-transforms of these time-correlation functions

$$J^{mn}(\omega) = \int_{-\infty}^{\infty} C^{mn}(\tau) e^{-i\omega\tau} d\tau. \quad (3)$$

In the following only terms of $C^{mn}(\tau)$ with $m = n$ are needed. Furthermore is $C^{mn}(\tau)$ in an isotropic liquid invariant under rotation of the laboratory frame. It can be shown that therefore C^{mn} is independent of m (Brüschweiler and Case, 1994). Summing over m and using the following addition theorem for spherical harmonics

$$\begin{aligned} \sum_{m=-2}^2 Y_{2,m}(\theta(t), \phi(t)) Y_{2,m}^*(\theta(t+\tau), \phi(t+\tau)) &= \\ \frac{5}{4\pi} P_2(\cos \chi_{t,t+\tau}), \end{aligned} \quad (4)$$

where $P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$ is the second-order Legendre polynomial and $\chi_{t,t+\tau}$ is the angle between the inter-spin vector at the two timepoints t and $t + \tau$, gives

the following formula for the time-correlation function that determines relaxation

$$C(\tau) = \frac{1}{5} \sum_{m=-2}^2 C^{mm}(\tau) = \left\langle \frac{P_2(\cos \chi_{t,t+\tau})}{r^3(t) r^3(t+\tau)} \right\rangle. \quad (5)$$

In the following, only this time-correlation function and its Fourier-transform $J(\omega)$ will be used.

An alternative, equivalent description of the interaction Hamiltonian (see Equation 1) is

$$\hat{H} = -\frac{\mu_0}{4\pi} \gamma^2 \hbar^2 \hat{\mathbf{I}} \cdot \mathbf{D}(t) \cdot \hat{\mathbf{S}}, \quad (6)$$

where

$$\mathbf{D}(t) = \frac{1}{r^5} \begin{pmatrix} 3x^2 - r^2 & 3xy & 3xz \\ 3yx & 3y^2 - r^2 & 3yz \\ 3zx & 3zy & 3z^2 - r^2 \end{pmatrix} \quad (7)$$

is the dipole-dipole interaction tensor. This form is often computationally more convenient because it relies on cartesian coordinates instead of spherical harmonics. The correlation functions of the tensor elements (only five are independent because $\mathbf{D}(t)$ is symmetric and traceless) are equivalent to those of the spherical harmonics. Again, in an isotropic liquid all auto- and cross-correlation functions of the matrix elements are either equal (except for some constant scaling factors) or zero (H.J.C. Berendsen, personal communication), and the averaging over all nine matrix-element correlation functions is equal to the averaging in Equation 4. This second approach will be mainly used where models for some internal motions have to be included and a straightforward use of the second Legendre polynomial as in Equation 5 is impossible, e.g., in the case of the united-atom methyl-groups as described below.

Dipolar relaxation in a two-spin system

In a system of two protons i and j the longitudinal relaxation of each proton is given by

$$\frac{d(I_z(t) - I_0)_i}{dt} = -\rho_{ij} (I_z(t) - I_0)_i - \sigma_{ij} (I_z(t) - I_0)_j. \quad (8)$$

Here, $(I_z(t) - I_0)_i$ is the deviation of the 'magnetization' (the z component expectation value of the spin operator of spin i) from its equilibrium value, ρ_{ij} is the direct dipolar relaxation rate constant of spin i by spin j , and σ_{ij} is the cross-relaxation rate. These relaxation rates, which control the magnitude of an NOE enhancement and the intensity of the signals in a NOESY spectrum, are given by the following

formulae (Neuhaus and Williamson, 1989)

$$\rho_{ij} = \frac{1}{10} K^2 \left(3 J(2\omega_0) + \frac{3}{2} J(\omega_0) + \frac{1}{2} J(0) \right) \quad (\text{NOESY}) \quad (9)$$

$$\sigma_{ij} = \frac{1}{10} K^2 \left(3 J(2\omega_0) - \frac{1}{2} J(0) \right), \quad (10)$$

$J(\omega)$ being spectral-density functions of the inter-spin vector, ω_0 the Larmor frequency of the protons, and $K = (\mu_0/4\pi)\hbar\gamma^2$. In the case of a rigid molecule which tumbles isotropically according to Brownian motion, the time-correlation function is a single exponential function with a correlation time τ_c

$$C(\tau) = \frac{1}{r_{ij}^6} e^{-\tau/\tau_c} \quad (11)$$

and

$$J(\omega) = \frac{1}{r_{ij}^6} \frac{2\tau_c}{1 + \omega^2\tau_c^2}, \quad (12)$$

r_{ij} being the interspin distance.

Due to the minus sign in Equation 10, the cross-relaxation rate changes sign for tumbling rates with $\omega_0\tau_c = \sqrt{\frac{5}{4}} \approx 1$, and therefore, the NOESY cross-peak intensity can come close to zero depending on the tumbling rate of the molecule. In this case, rotating frame experiments (ROESY) are performed where this zero-transition does not occur. Then, the measured intensities are determined by the following relaxation rates (K as defined above) (Neuhaus and Williamson, 1989):

$$\rho = \frac{1}{10} K^2 \left(\frac{3}{2} J(2\omega_0) + \frac{9}{4} J(\omega_0) + \frac{5}{4} J(0) \right) \quad (\text{ROESY}), \quad (13)$$

$$\sigma = \frac{1}{10} K^2 \left(\frac{3}{2} J(\omega_0) + J(0) \right) \quad (14)$$

Multispin systems and the relaxation matrix

In the case of a system of n spins, a set of coupled differential equations similar to Equation 8 describes the magnetization of the individual spins:

$$\frac{d(I_z(t) - I_0)_i}{dt} = -\rho_i (I_z(t) - I_0)_i - \sum_{j=1, \neq i}^n \sigma_{ij} (I_z(t) - I_0)_j. \quad (15)$$

Here $\rho_i = \sum_{j=1, \neq i}^n \rho_{ij} + L_{1,i}$ is the total longitudinal relaxation rate of spin i with the 'leakage' $L_{1,i}$, i.e., relaxation due to other relaxation mechanisms, such as chemical shift anisotropy, and intermolecular dipole-dipole relaxation by paramagnetic species

such as dissolved oxygen. This additional contribution is unfortunately hard to estimate and can only approximately be included into the calculations.

The isolated-spin-pair approach, in which a cross-peak intensity is given directly by the cross relaxation rate of the corresponding spin-pair, is only as a first-order approximation applicable. In most cases the so-called 'spin-diffusion' is not negligible, i.e., the coupled system of Equations 15 has to be solved (Macura and Ernst, 1980; Keepers and James, 1984; Boelens et al., 1989). The equations can be rewritten in the matrix form

$$\frac{d\mathbf{M}(t)}{dt} + \mathbf{R}\mathbf{M}(t) = \mathbf{0} \quad (16)$$

with

$$M_{ii}(t) = (I_z(t) - I_0)_i \quad (17)$$

$$M_{ij}(t) = 0 \quad \text{for } i \neq j \quad (18)$$

$$R_{ii} = \rho_i \quad (19)$$

$$R_{ij} = \sigma_{ij}. \quad (20)$$

Here, $\mathbf{M}(t_m)$ is the matrix with the intensities of the NOESY- (or ROESY-)spectrum after a mixing time t_m . $\mathbf{M}(0)$, the spectrum at mixing time 'zero', contains the intensities of the 1D-spectrum along the diagonal.

The matrix Equation 16 may be solved by diagonalization of the relaxation matrix \mathbf{R} , leading to the solution

$$\mathbf{M}(t_m) = \mathbf{X} e^{-\Lambda t_m} \mathbf{X}^{-1} \mathbf{M}(0), \quad (21)$$

where $\Lambda = \mathbf{X}^{-1}\mathbf{R}\mathbf{X}$ contains the eigenvalues of \mathbf{R} as diagonal elements, and \mathbf{X} is the corresponding eigenvector matrix.

Methyl groups – a slight complication

In many MD simulations protons that are bound to an aliphatic carbon atom are, together with that carbon, incorporated into a united atom to save computational effort. In the case of methyl groups this has the effect that the exact positions of the three protons are not known but only their positional average. Thus, models for the methyl rotation have to be used to describe the relaxation of proton pairs where one or both protons are members of methyl groups.

One model for methyl-group rotation is a three-site jump described in (Tropp, 1980). Here, the three possible vectors connecting a nonmethyl proton to the positions of the methyl protons are evaluated (respectively the nine vectors in case of the interaction between two methyl groups) and an averaging over

the spherical harmonics of the N ($N = 3$ or 9) interaction vectors is performed before the correlation function of the spherical harmonics is calculated according to Equation 2. This is more conveniently done in Cartesian space by calculating the five independent elements of matrix \mathbf{D} (see Equation 7) for the N vectors

$$C_{ij}(\tau) = \left\langle \left(\sum_{k=1}^N D_{ij,k}(t) \right) \left(\sum_{k=1}^N D_{ij,k}(t + \tau) \right) \right\rangle$$

$$1 \leq i, j \leq 3; \quad N = 3 \text{ or } 9. \quad (22)$$

Here $N = 3$ or 9 is the number of different interaction vectors and $D_{ij,k}$ is the element ij of \mathbf{D} corresponding to the k th interaction vector. To obtain the optimal statistics over all orientations in space, averaging equivalent to Equations 4 and 5 is performed (the prefactor $\frac{1}{6} = \text{Tr}(r^3 \cdot \mathbf{D})^2$ was chosen to obtain the same scaling as in Equation 5)

$$C(\tau) = \frac{1}{6} \sum_{i,j=1}^3 C_{ij}. \quad (23)$$

In this treatment the correlation time of the methyl-group rotation is not included. In principle this is only correct if this rotation is much faster than the other motions in the system (Edmondson, 1994), which might be problematic for small molecules like peptides, but should not qualitatively change the results.

One term that should still be added is the intramethyl-group relaxation. It gives an additional contribution to the diagonal elements of \mathbf{R} if i is a methyl group. The spectral density functions for interactions between two protons in the same methyl group are (Edmondson, 1992; Woessner, 1962)

$$J(\omega) = \frac{1}{r_m^6} \left(\frac{1}{4} \frac{\tau_c}{1 + \omega^2 \tau_c^2} + \frac{3}{4} \frac{\tau_a}{1 + \omega^2 \tau_a^2} \right), \quad (24)$$

$$\frac{1}{\tau_a} = \frac{1}{\tau_c} + \frac{1}{\tau_m}, \quad (25)$$

where τ_m is the correlation time of the methyl-group rotation and r_m is the distance between the two protons. The values $\tau_m = 0.1$ ns and $r_m = 0.17$ nm were chosen according to (Edmondson, 1992).

An additional complication is that a methyl group is treated as a single site in the relaxation and the magnetization matrix, i.e., in the expression $(I_z - I_0)_i$ the magnetization of the three protons has been summed if i is a methyl group. As a consequence \mathbf{R} becomes nonsymmetric because $\sigma_{ij} = 3\sigma_{ji}$ if i is a methyl group and j is a nonmethyl proton, as can be seen from

Equation 15. This problem can be solved by making \mathbf{R} symmetric before diagonalizing it, as described in (Olejniczak, 1989).

Computation of the spectra

The computation of NOESY or ROESY spectra was done in a three-step procedure in order to make it as flexible as possible. Because the calculation of time-correlation functions is computationally expensive, it was done not for all proton pairs but only for a selected (significant) subset:

1. Those proton pairs were selected which were either close or for which the distance fluctuated significantly:

$$\langle r^{-6} \rangle_{\text{traj}}^{-1/6} \leq \text{MAXDIST} \quad \text{or}$$

$$(\langle r^2 \rangle_{\text{traj}} - \langle r \rangle_{\text{traj}}^2)^{1/2} / \langle r \rangle_{\text{traj}} \geq \text{MINFLUC}$$

The subscript traj indicates an average over the whole MD trajectory. For the β -heptapeptide test-system the criteria MAXDIST = 0.4 nm and MINFLUC = 0.4 were chosen, so that all long-distance NOEs measured experimentally were explicitly calculated.

2. The time-correlation and spectral-density functions of the interproton vectors of the selected pairs were computed according to Equation 5 or 22 and 23.
3. A relaxation matrix for a rigidly tumbling molecule was computed using Equation 12. The values of the spectral densities of the selected pairs were substituted by those that had been calculated explicitly in step 2. After symmetrizing and diagonalizing the relaxation matrix the theoretical spectrum was obtained. For the distances needed in Equation 12, $\langle r^{-6} \rangle^{-1/6}$ -distance averages from a set of structures were used. This set included either the whole trajectory ($\langle r^{-6} \rangle_{\text{traj}}^{-1/6}$) or structures that were representative for a specific conformation ($\langle r^{-6} \rangle_{\text{conf}}^{-1/6}$).

Results and discussion

A detailed description of the MD simulations of the β -heptapeptide in methanol at 298 K and at 340 K is given by Daura et al. (1998, 1999b). At 298 K the system is found mainly ($\approx 95\%$ of the simulation) in the folded conformation, which is a left-handed 3_{14} -helix. In the 340 K simulation the system spends only 36% of the time in the folded state, folding/unfolding

occurring on a timescale of about 10 ns. Although at 340 K the system is more than 50% of the time in conformations that deviate considerably from the NOE-derived model structure, the trajectory-averages of the interproton distances still satisfy quite well the NOE distance restraints (Daura et al., 1999a). Therefore, comparison of average distances from a simulation with experimentally derived distances is not a very reliable criterium to investigate the folding/unfolding equilibrium of a peptide.

A model of the folded conformation of the peptide is displayed in Figure 1. Additionally some vectors that will be frequently referred to later are indicated. From the 200 ns trajectories at 298 K and 340 K structures were saved every ps for the preliminary analysis of the system dynamics. The calculation of ROESY spectra was performed with peptide configurations that were saved every 5 ps. Before, tests had shown that a more frequent saving of conformations (every 2 ps) would not alter the results.

Preliminary analysis of the system dynamics

Before calculating the NMR relaxation parameters of the system, a general analysis of the overall and internal system dynamics was performed for both simulations in order to get an answer to the following questions:

1. In which timescales do internal dynamics and overall tumbling take place?
2. Is the overall tumbling isotropic?
3. To which extent are internal dynamics and overall tumbling correlated?

Normalized time-correlation functions of the motional processes of selected interatomic vectors were calculated according to

$$C(\tau) = \langle r^{-6}(t) \rangle^{-1} \left\langle \frac{P_2(\cos \chi_{t,t+\tau})}{r^3(t) r^3(t+\tau)} \right\rangle \quad (26)$$

both with and without fitting the structures to a (helical) reference configuration. For the fitting a translation of the center of mass to the origin followed by a rotational fit minimizing the root-mean-square positional deviation of backbone atoms was performed. With this method the time correlation functions of all motions in the laboratory frame, $C_{\text{all}}(\tau)$, as well as those of the internal motions in a molecule-fixed frame, $C_{\text{int}}(\tau)$, could be obtained. By calculating the time correlation functions of the three orthonormal row vectors of the rotation matrix that resulted from the fitting, the rotational time correlation function for three directions, $C_{\text{rot}}(\tau)$, could also be obtained.

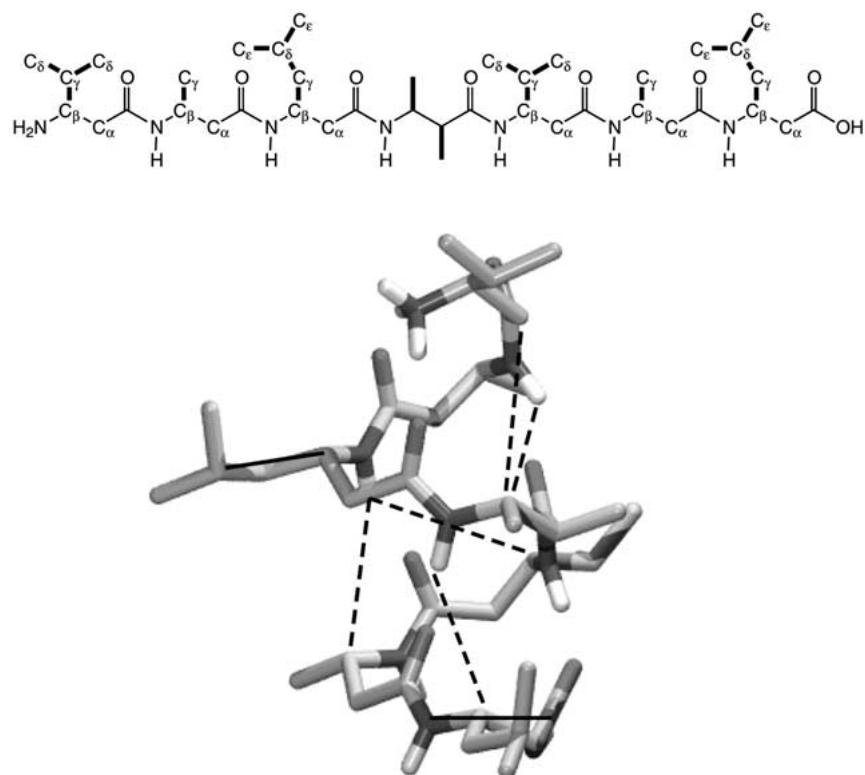


Figure 1. Chemical structure and model helix of the β -heptapeptide with some vectors indicated; dashed lines: some long distance NOEs that are typical for the helical structure (see also Figure 6); solid lines: two vectors (residue 3: $C_\beta - C_\delta$; residue 7: N-C(carbonyl)) for which a preliminary analysis of the occurring motional processes was performed. Note that in the simulation the N-terminus was protonated (NH_3^+).

In the case of the three rotational time correlation functions the correlation times were determined by fitting a single exponential to $C_{\text{rot}}(\tau)$. As in all other cases the decay of the correlation functions was non-exponential, a generalized order-parameter S^2 and an estimate of the correlation time τ_{corr} were determined via

$$S^2 = \lim_{\tau \rightarrow \infty} C(\tau) \quad (27)$$

$$\tau_{\text{corr}} = \frac{1}{C(0) - S^2} \int_0^{t_{\text{int}}} (C(\tau) - S^2) d\tau, \quad (28)$$

where the integration limit t_{int} was chosen as the first timepoint where the correlation function had reached the value of S^2 .

If internal and overall rotational motion were uncorrelated and the overall tumbling corresponded to Brownian motion the following Equation should hold

$$C_{\text{all}}(\tau) \approx C_{\text{rot}}(\tau)C_{\text{int}}(\tau) = e^{-\tau/\tau_c} C_{\text{int}}(\tau), \quad (29)$$

where τ_c is the correlation time of the overall rotation.

Table 1. Rotational correlation times

T (K)	τ_c – three directions (ps)			Average (ps)
298	153	143	158	≈ 150
340	74	76	75	≈ 75

Additionally to the complete internal dynamics, the motions of the individual sidechains were investigated by translating in each case the corresponding β -carbon atom to the origin and performing a rotational fit based on the three atoms directly bonded to the β -carbon atom.

The results for the rotational motion can be seen in Table 1. The overall tumbling is found to be approximately isotropic and has an average correlation time of about 150 ps for the 298 K simulation and 75 ps for the 340 K simulation. The rotational correlation time of the molecule in the simulation at 298 K is probably underestimated. It is known experimentally that the NOESY intensity at 500 MHz is approximately

Table 2. Order parameters S^2 (Equation 27) and correlation times τ (Equation 28) of vectors in the third and in the last residue. The subscript *all* refers to inclusion of all motions, *int* to internal motions (overall translation and rotation removed); the subscript *COM* refers to a translational/rotational fit around the center of mass and therefore represents all internal motions, the subscript *SC* refers to a translational/rotational fit around the C_β -atom and therefore represents the sidechain motions

Vector	T (K)	τ_{all} (ps)	$\tau_{\text{int,COM}}$ (ps)	$\tau_{\text{int,SC}}$ (ps)	$S_{\text{int,COM}}^2$	$S_{\text{int,SC}}^2$
$C_\beta^3 - C_\delta^3$	298	94	126	206	0.57	0.65
$C_\gamma^3 - C_\delta^3$	298	66	140	63	0.30	0.36
$C_\gamma^3 - C_\epsilon^3$	298	61	99	101	0.29	0.32
$C_\delta^3 - C_\epsilon^3$	298	47	63	41	0.21	0.24
$C_\beta^3 - C_\delta^3$	340	44	1077	19	0.19	0.67
$C_\gamma^3 - C_\delta^3$	340	30	270	17	0.09	0.38
$C_\gamma^3 - C_\epsilon^3$	340	25	332	15	0.04	0.32
$C_\delta^3 - C_\epsilon^3$	340	17	212	12	0.03	0.20
$N^7 - C_\delta^7$	298	84	646	287	0.27	0.68
$N^7 - C^7(\text{carbonyl})$	298	105	1229	1882	0.29	0.68
$N^7 - C_\delta^7$	340	31	170	438	0.05	0.59
$N^7 - C^7(\text{carbonyl})$	340	39	190	442	0.04	0.54

zero. According to Equation 10 combined with 12 this zero transition should occur for correlation times of roughly 360 ps. The fact that the rotational diffusion is too fast is not completely unexpected, because it is known that the diffusion constant for the methanol model used is also too large (Walser et al., 2000).

Correlation times of all motions, as well as correlation times and order parameters of the complete internal dynamics and of the motions in the sidechains, are presented in Table 2 for vectors in the third and in the last residue. If no additional types of motions appear at the higher temperature, correlation times are expected to decrease with increasing temperature. This is the case for the last residue, which is never strongly fixed in the helical secondary structure. However, the internal correlation times of vectors in the third residue increase with increasing temperature, which indicates that an additional, comparatively slow component is added to the internal dynamics. This behaviour cannot be observed for the internal motions of the sidechains only, which is to be expected because these sidechain rotations and vibrations should already be expressed at the lower temperature. These observations are confirmed by the order parameters. $S_{\text{int,SC}}^2$, the order parameter of the sidechain motion, does hardly change with temperature, while $S_{\text{int,COM}}^2$, the order parameter of the complete internal motion, decreases strongly with increasing temperature for both residues. In the case of the third residue, both order parameters are very similar at 298 K, indicating that the sidechain

fluctuations are the major component of the internal dynamics in this part of the molecule, whereas at 340 K $S_{\text{int,COM}}^2$ is much smaller than $S_{\text{int,SC}}^2$. The latter observation holds for both temperatures in case of the terminal residue which shows large conformational changes.

These results are illustrated in Figures 2 and 3 where the various time correlation functions of two representative vectors in these residues are shown. For the third residue (Figure 2) the product of the internal time-correlation function with the exponential time correlation function of the tumbling (see Equation 29; dashed lines) matches the time correlation function of all motions (solid lines) for both temperatures. Thus, the overall tumbling and the internal dynamics are not strongly coupled for this residue. Equation 29 is less accurately satisfied in the case of the terminal residue (Figure 3). In all cases $C_{\text{all}}(\tau)$ is not monoexponential, which implies that computing relaxation parameters with a simple Lorentz function as in Equation 12 is not sufficient.

The internal time correlation functions of the residues that take part in helix formation show at 340 K some periodicity with the approximate frequency of the folding and unfolding transitions. This effect is absent at 298 K and in the case of the terminal (not helix forming) residue.

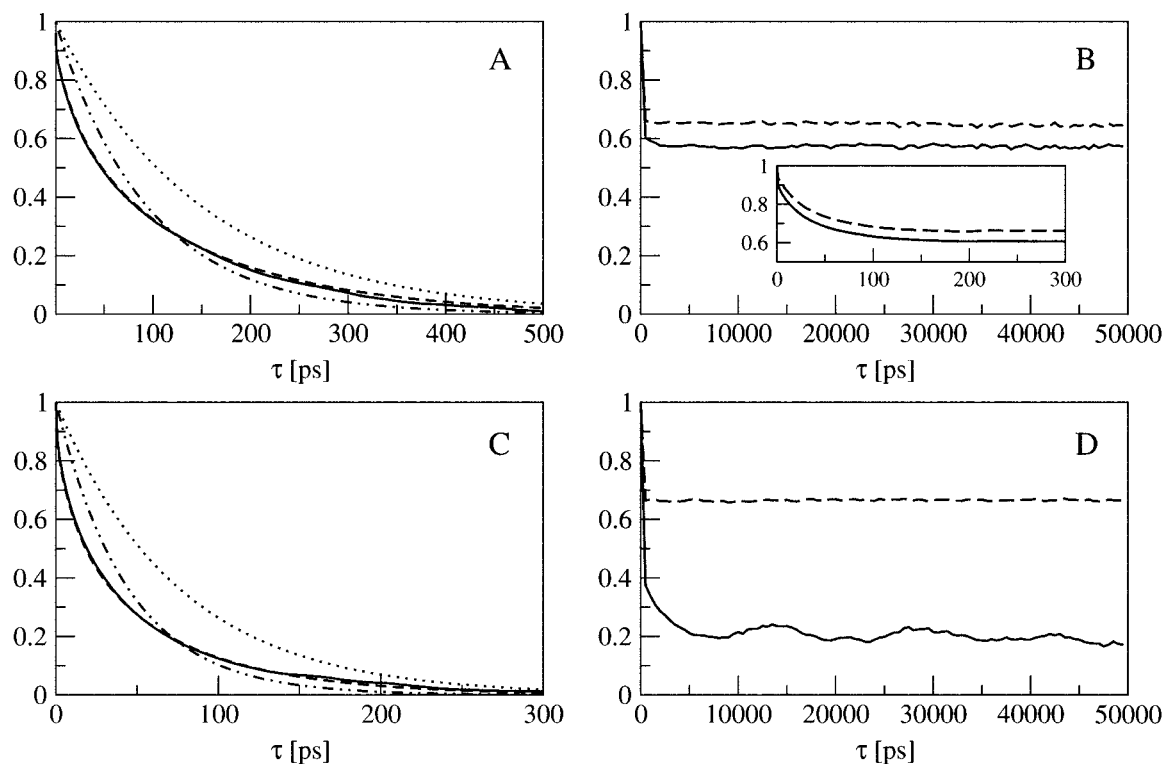


Figure 2. Time correlation functions (TCF) for the vector between the β - and the δ -carbon in the third residue – A and C (A: 298 K; C: 340 K): external coordinate system; TCF of all motions (solid lines), monoexponential TCF using $\tau_{\text{all}} = 94$ ps (A) and 44 ps (C) (dotted/dashed), monoexponential TCF using $\tau_{\text{rot}} = 150$ ps (A) and 75 ps (C) (dotted), internal TCF multiplied with rotational TCF (dashed); B and D (B: 298 K; D: 340 K): internal coordinate system; TCF for all internal motions (solid lines), TCF only for sidechain dynamics (dashed).

Theoretical ROESY-spectra and buildup curves

As already mentioned, there are several aspects of internal flexibility that influence NOESY (or in this case ROESY) cross-relaxation rates. From looking at the functional form of the spectral densities of a tumbling rigid molecule (Equation 12) three major effects of neglecting internal dynamics can be distinguished:

- Internal dynamics lead to fluctuations in the interproton distances. This effect will from now on be referred to as ‘distance fluctuations’, it is in the conventional methods to analyse NOEs approximately taken into account by using $\langle r^{-6} \rangle$ - or $\langle r^{-3} \rangle$ -averaging (Tropp, 1980).
- The overall correlation times of the motions of the interproton vectors contain contributions from both tumbling motions and internal dynamics and are thus different for each proton pair. This effect will from now on be referred to as ‘individual correlation times’.
- Due to the superposition of overall rotations and possibly various types of internal motions,

the overall time correlation function is not a single exponential and thus the spectral density is not a simple Lorentz function as in Equation 12. This effect will from now on be referred to as ‘non-monoexponentiality of the time correlation function’ or as ‘orientational effect’. ‘Orientational’ because it implicitly results from the second Legendre polynomial in Equation 5: this term in the time-correlation function is altered by internal dynamics and makes the time correlation function not monoexponential.

In order to investigate the impact of these influencing factors, various types of theoretical spectra were calculated for both simulations: spectra that correctly use the spectral density functions from the simulations, and therefore correctly incorporate all aspects of internal dynamics and tumbling (Flex₂₉₈ and Flex₃₄₀), and spectra that assume a rigidly tumbling molecule with the tumbling time determined from the simulation and with interproton distances that were computed as $\langle r^{-6} \rangle$ -averages over a set of structures. For the latter, either the structures of the whole simu-

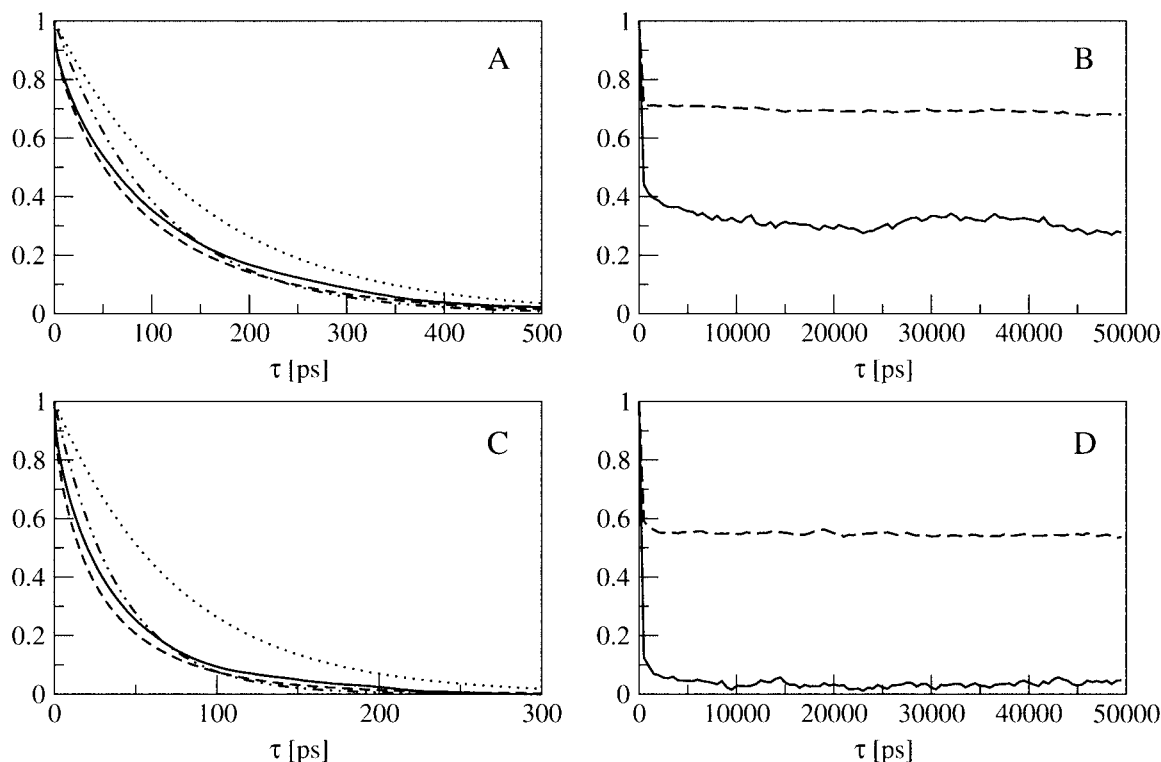


Figure 3. Time correlation functions (TCF) for the vector between the nitrogen and the carbonyl carbon in the last residue – A and C (A: 298 K; C: 340 K): external coordinate system; TCF of all motions (solid lines), monoexponential TCF using $\tau_{\text{all}} = 105$ ps (A) and 39 ps (C) (dotted/dashed), monoexponential TCF using $\tau_{\text{rot}} = 150$ ps (A) and 75 ps (C) (dotted), internal TCF multiplied with rotational TCF (dashed); B and D (B: 298 K; D: 340 K): internal coordinate system; TCF for all internal motions (solid lines), TCF only for sidechain dynamics (dashed).

lation ($\text{Rig}_{298}^{\text{traj}}$ and $\text{Rig}_{340}^{\text{traj}}$) or a subset of structures that are representative for the helical conformation ($\text{Rig}_{298}^{\text{helix}}$ and $\text{Rig}_{340}^{\text{helix}}$) were used. For this set, the conformations from the 340 K simulation were chosen which have a backbone-atom root-mean-square deviation from the NMR model structure smaller than 0.1 nm (Daura et al., 1999b). By looking at the relative or the absolute differences between, for example, $\text{Rig}_{340}^{\text{traj}}$ and $\text{Flex}_{340}^{\text{traj}}$ the effects that do not stem from ‘distance fluctuations’ can be extracted. This allows to test whether there is indeed a substantial neglect of information when comparing only distance averages from simulations and experimentally derived average distances.

The results are illustrated in Figure 4, where difference spectra are plotted as matrices with the 50 proton sites ordered along the axes of the figures according to their sequence in the peptide. This sequence is for each residue i : N^iH , C_β^iH , sidechain protons from $\text{C}_\gamma^i\text{H}$ to maximal $\text{C}_\epsilon^i\text{H}$, $\text{C}_\alpha^i\text{H}_{\text{ax}}$ and $\text{C}_\alpha^i\text{H}_{\text{eq}}$, the subscripts ax and eq indicate the axial and equatorial positions

of the two C_α -protons in the model helix. Panel A shows the relative difference between intensities of all (cross)peaks in the ROESY spectra at 298 K for a rigid (trajectory average) and a flexible molecule, calculated using $(\text{Rig}_{298}^{\text{traj}} - \text{Flex}_{298}^{\text{traj}})/\text{Rig}_{298}^{\text{traj}}$. Panel B shows the same for the 340 K simulation. To make sure that large relative differences between the spectra are really significant, the absolute difference is also given for the simulation at 340 K ($\text{Rig}_{340}^{\text{traj}} - \text{Flex}_{340}^{\text{traj}}$) in panel C. Here the intensities were scaled such that they could be compared with experimentally measured peak volumes. The experimentally determined intensities range from approximately 5 to 200 arbitrary units, which implies that matrix elements with a colour that is different from dark blue in panel C denote a detectable difference in the spectra. As the main focus of this investigation is to study the effect of large conformational changes of the peptide (mainly visible in the behaviour of the backbone), backbone protons are marked with coloured bars along the axes. Additionally some backbone proton pairs are highlighted with

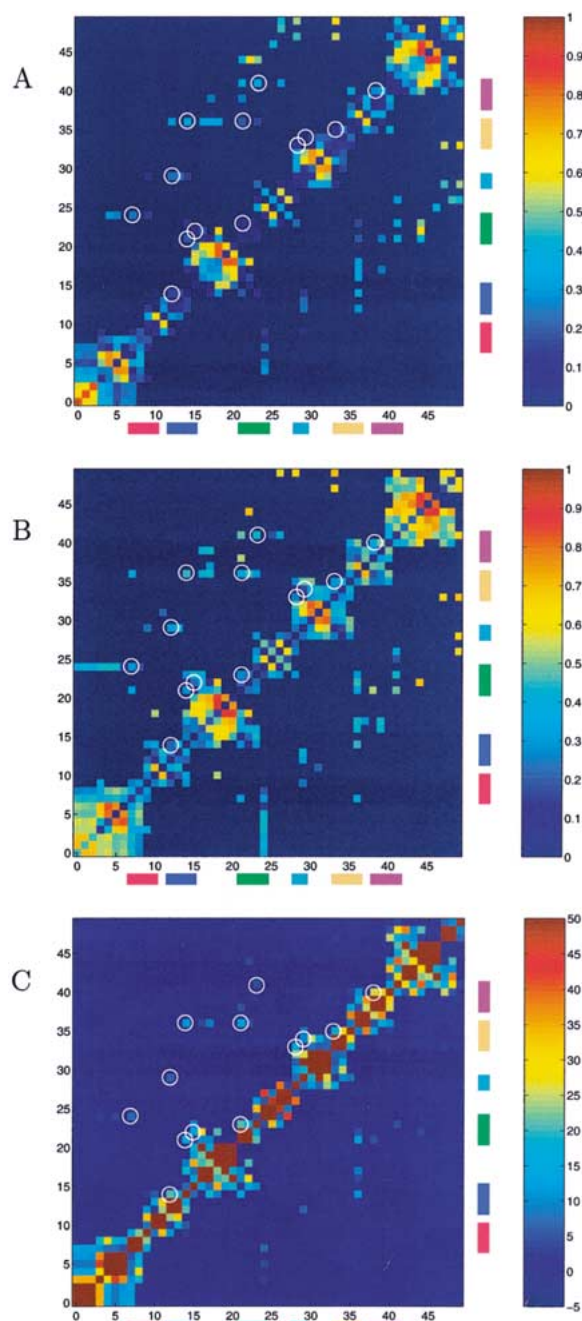


Figure 4. Difference plots for theoretical ROESY spectra of the β -heptapeptide (mixing time: 100 ms) between tumbling rigid ($\langle r^{-6} \rangle$ -average over full trajectory) and flexible molecules; the numbering of the protons follows the sequence in the peptide which is for each residue i : N^iH , C_β^iH , sidechain protons from C_γ^iH to maximal C_ϵ^iH , $C_\alpha^iH_{ax}$, $C_\alpha^iH_{eq}$; coloured bars at the axes indicate backbone protons ($C_\alpha^iH_{ax}$, $C_\alpha^iH_{eq}$, $N^{i+1}H$ and $C_\beta^{i+1}H$; red: $i = 1$, blue: $i = 2$, green: $i = 3$, cyan: $i = 5$, yellow: $i = 6$, magenta: N^5H and C^5H); white rings mark selected pairs of backbone protons – A: 298 K, relative difference; B: 340 K, relative difference; C: 340 K, absolute difference (arbitrary units which correspond to measured peak volumes – detectable intensities range from 5 to 200).

white circles. It can be observed that there is a change in the relative differences upon raising the temperature from 298 K to 340 K, especially for those proton pairs which are in close proximity along the peptide backbone, e.g., $C_\alpha^iH_{ax} - N^{i+1}H$, $N^iH - C_\alpha^iH_{ax}$, $C_\beta^iH - N^{i+1}H$. These are proton pairs where the ‘individual correlation times’ and the ‘orientational effects’ have a strong influence (in addition to the change in the distance average) on the intensity of crosspeaks in case of conformational fluctuations. It can also be seen, that these influences are not equally strong for all proton pairs. This can even lead to an inversion of the relative intensity of two peaks depending on whether the internal dynamics is treated exactly or only approximately via distance averages (see below). Consequently in the case of the simulation at 340 K where strong internal dynamics occur, the explicit computation of the spectral density functions of individual proton pairs appears to be a requirement to reliably (back)calculate ROESY spectra from MD simulations.

To monitor these effects in more detail theoretical ROESY-intensity buildup curves (i.e., crosspeak intensity as function of mixing time) of some selected proton pairs were calculated. The results for both simulation temperatures and for all three above-mentioned cases (Rig^{helix} , Rig^{traj} and $Flex^{traj}$) are shown in Figure 5. Comparing Rig_{298}^{helix} with Rig_{298}^{traj} and Rig_{340}^{helix} with Rig_{340}^{traj} gives an impression how strongly the ‘fluctuating distances’ affect the intensities. Comparing these with $Flex_{298}^{traj}$ and $Flex_{340}^{traj}$ permits assessing all other additional effects included only by explicit calculation of time correlation functions and spectral densities, i.e., ‘individual correlation times’ and ‘non-monoexponentiality of the time correlation function’. As expected, the differences are rather small for the simulation at 298 K, because the peptide is most of the time in the folded state. The situation is different for the simulation at 340 K. The ‘distance fluctuations’ can lead to a decrease or increase in the average distance resulting in an increase or decrease in intensity. This depends on whether the ‘collapse’ of a rigid structure adds shorter distances to the average and thus makes it smaller, or whether the breaking of a rigid structure tears apart proton pairs which are kept close, e.g., by hydrogen bonding, and thus leads to an increase in the average distance. Overall correlation times which include internal motions are usually smaller than the rotational correlation times only, so that the ‘individual correlation times’ lead to a decrease of intensity compared to a ‘rigid spectrum’.

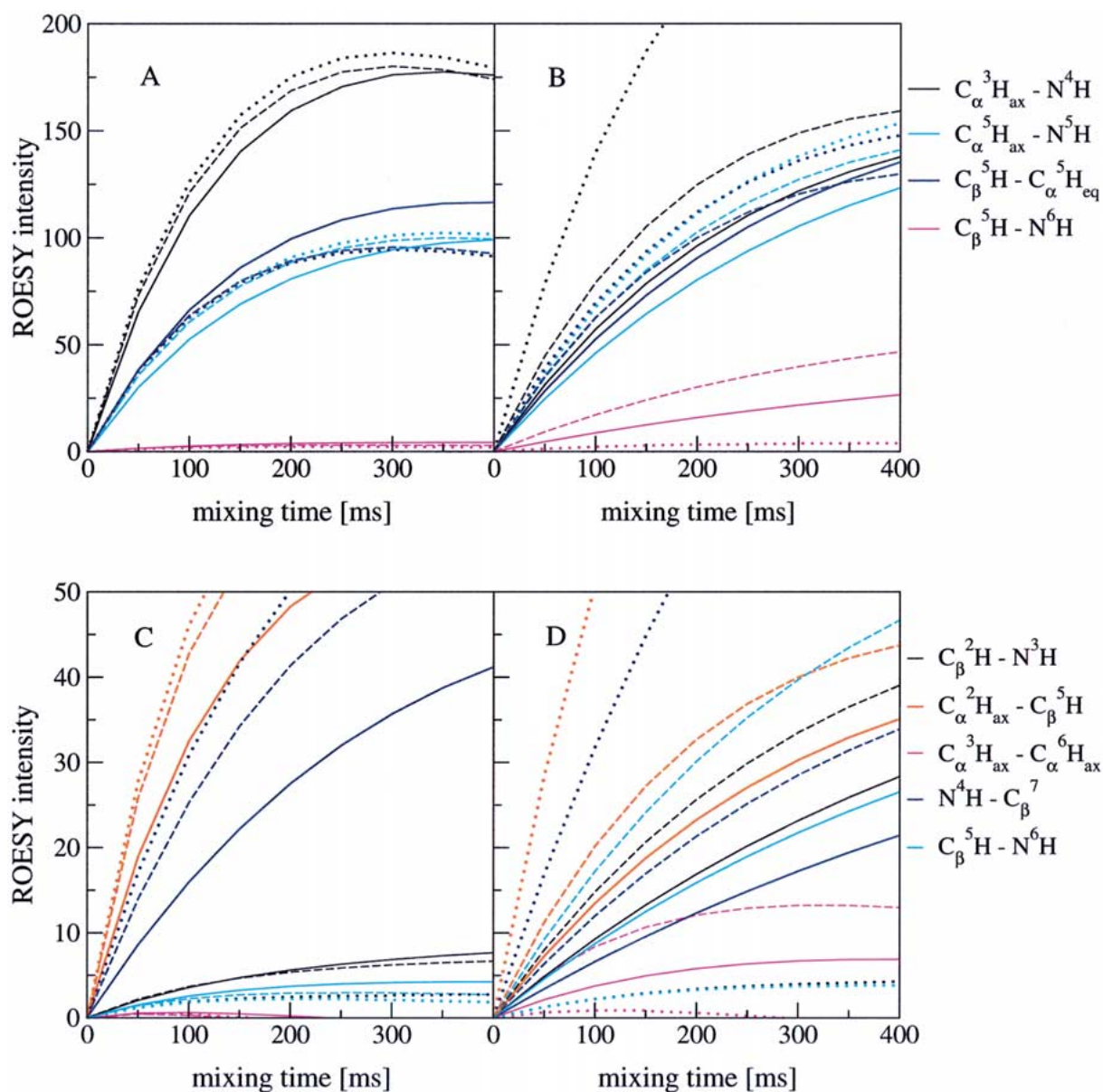


Figure 5. Theoretical buildup curves – left panels (A and C) 298 K ($\tau_{\text{rot}} = 150$ ps), right panels (B and D) 340 K ($\tau_{\text{rot}} = 75$ ps); solid lines: fully flexible molecule ($\text{Flex}_{298}^{\text{traj}}$ and $\text{Flex}_{340}^{\text{traj}}$); dashed lines: rigid molecule, average over all conformations ($\text{Rig}_{298}^{\text{traj}}$ and $\text{Rig}_{340}^{\text{traj}}$); dotted lines: rigid molecule, average over helical structures ($\text{Rig}_{298}^{\text{helix}}$ and $\text{Rig}_{340}^{\text{helix}}$); the proton pairs are indicated in the figure.

This decrease is of course individually strong, and depends on the internal correlation time. The consequences for the relative intensities of a set of peaks in a spectrum can, for example, be seen in Figure 5 by comparing the buildup curves of $\text{C}_{\alpha}^2\text{H}_{\text{ax}} - \text{C}_{\beta}^5\text{H}$, $\text{C}_{\beta}^5\text{H} - \text{N}^6\text{H}$ and $\text{N}^4\text{H} - \text{C}_{\beta}^7\text{H}$.

By treating all internal motions explicitly, even the ‘negative’ information that a signal is not observable

although it should be according to the distances in a rigid structure, or that a peak is observable although the interproton distance in a rigid structure would be too big, can be used to analyse NOESY or ROESY spectra. Yet, it should be always kept in mind that ‘leakage’ contributions to the longitudinal relaxation rate in Equation 15 cannot be calculated quantitatively and may lead to artifacts.

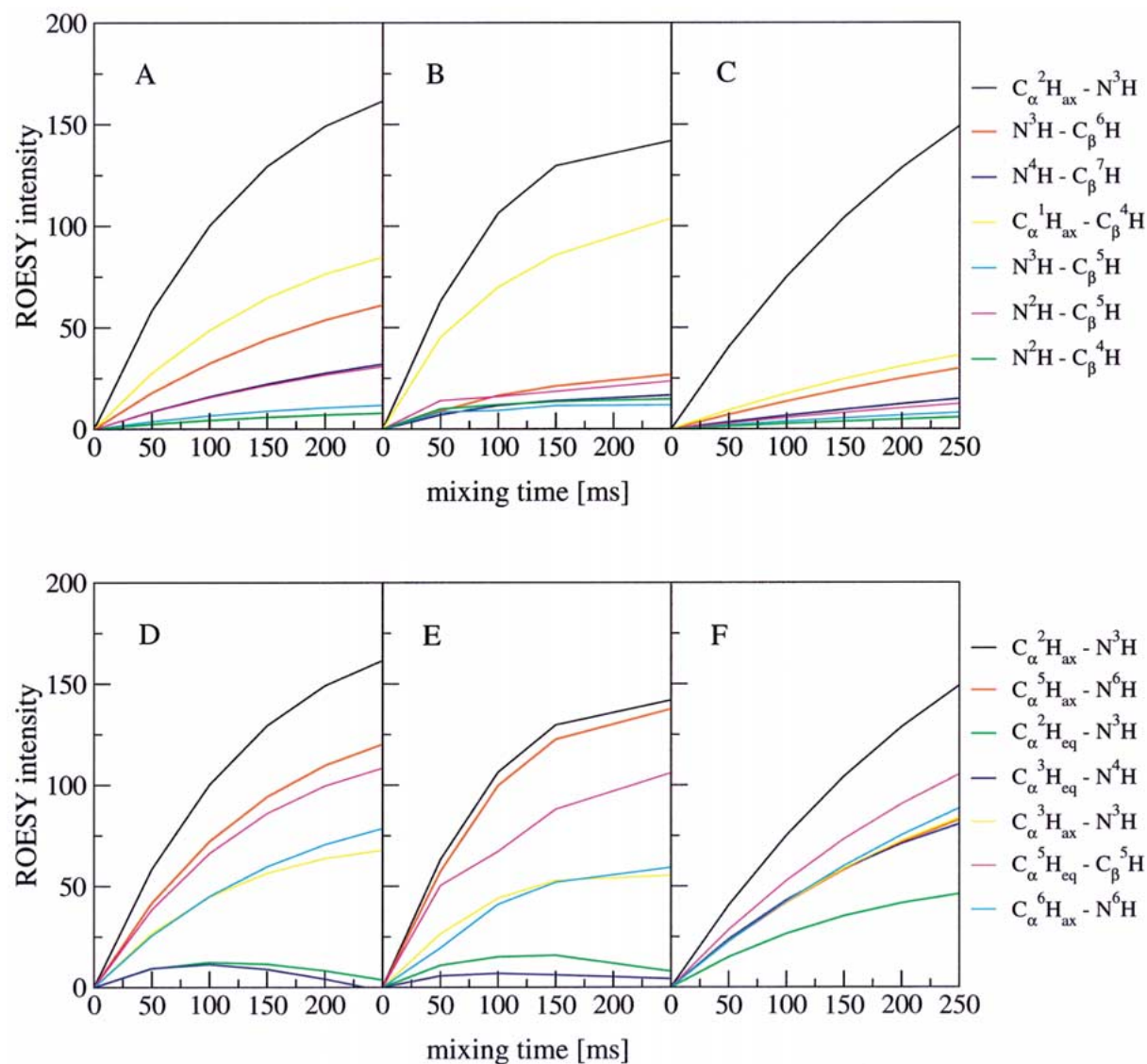


Figure 6. Experimental and theoretical buildup curves – left (A and D): simulation at 298 K (Flex₂₉₈^{traj}), middle (B and E): experimental buildup curves at 300 K, right (C and F): simulation at 340 K (Flex₃₄₀^{traj}) – upper panels (A–C): long distance backbone NOEs; lower panels (D–F): short distance backbone NOEs – the proton pairs are indicated in the figure.

Comparison with experimental data

The theoretically calculated ROESY spectra from the two simulations can now be used to make a more reliable prediction of the ‘real’ folding/unfolding equilibrium of the β -heptapeptide. Experimental ROESY spectra at room temperature with mixing times of 50, 100, 150 and 250 ms are available (Seebach et al., 1996). From these, a set of non-overlapping peaks was selected for which the intensities can be reliably determined by integration. To compare the buildup

curves of those experimental signals with the theoretically calculated ones one scaling factor per simulation temperature was determined by least-square fitting. Some of these buildup curves are shown exemplarily in Figure 6. The agreement between simulation and experiment for the long-distance (‘helical’) NOEs (upper panel) is reasonable but not excellent for both simulations, albeit still better for the simulation at 298 K. However for NOEs that belong to backbone proton pairs which are close in sequence (lower panel)

the agreement with experiment is clearly better for the 298 K simulation compared to that for the 340 K simulation. This can be explained by the fact that the long-distance NOEs are probably mainly sensitive to the distance average and that the effects of correlation times and the shape of the spectral density functions ('non-monoexponentiality of the time correlation function') have a minor impact, whereas they play a more important role for pairs which are separated only by a few bonds.

Comparison of the curvatures of the experimental buildup curves with the ones from the simulation at 298 K again indicates that the simulated correlation times are a bit too short, because an increase in the correlation times of all proton pairs shifts the maxima of the buildup curves towards shorter mixing times.

Besides these small discrepancies the results lead to the conclusion, that the simulation at 298 K gives a more realistic description of the true conformational equilibrium of the peptide at room temperature compared to the simulation at 340 K.

Conclusions

Two 200 ns MD simulations of a β -heptapeptide testsystem at 298 K and 340 K have been used to investigate the influence of internal flexibility of a molecule on NMR relaxation parameters. Various aspects of the system dynamics have been considered (overall tumbling, internal motions, sidechain dynamics). The overall rotation of the molecules is found to be essentially isotropic and the tumbling and internal motions appear not to be coupled. The rotational correlation times are approximately 150 ps and 75 ps at 298 K and 340 K, respectively, which is too short compared to values indicated by NMR experiments at room temperature. In the simulation at 298 K the peptide stays most of the time in a folded (helical) conformation, and the internal dynamics is mainly confined to motions in sidechains and in the terminal residues. In the simulation at 340 K several unfolding events and comparatively slow internal motional processes occur. There are several reasons why this peptide is a good testsystem to assess the advantages of (back)calculating ROESY spectra compared to the conventional method of analysing ROESY spectra by only comparing simulated distance averages with experimentally derived average distances: it shows internal fluctuations which have correlation times that are of the same order of magnitude or lower than the cor-

relation time of the overall tumbling, the correlation times of interatomic vectors are quite diverse, and the corresponding time correlation functions are not simply monoexponential. It could be shown, that internal dynamics have a strong influence on the cross-peak intensities of ROESY spectra. The explicit computation of spectral densities via the time correlation functions of interproton vectors leads to significant differences in the spectra compared to only calculating distance averages and neglecting possible other aspects of internal dynamics, such as 'individual correlation times' and 'orientational effects'. This method makes the use of 'negative' information, i.e., the absence of specific crosspeaks in the spectra, possible. Only the potential influence of 'leakage' contributions from other relaxation mechanisms cannot be computed properly.

For the β -heptapeptide the comparison of theoretical and experimental ROESY buildup curves suggests that the folding/unfolding equilibrium in the experiment lies more on the side of the folded (helical) state. This conclusion could hardly be made from comparing distance averages with ROE-derived distances, as the experimentally derived distances were still in good agreement with distance averages from a simulation with a 50/50 equilibrium distribution. This proves that it is possible to gain more information about the relationship between experiment and simulation by (back)calculating NMR spectra explicitly than by comparing only distance averages.

For future applications this method of performing unrestrained, explicit solvent MD simulations and (back)calculation of NMR relaxation parameters and spectra can be used in cases of even more flexible molecules, where no single configuration that satisfies the NOE derived distances can be found at all. For such systems, where the interpretation of NMR data alone seems to be a desperate venture because of the extreme internal mobility, this method could open new possibilities to analyse structure and dynamics.

Acknowledgements

The authors wish to thank Prof. B. Jaun and co-workers for providing the ROESY spectra of the β -heptapeptide and for giving a first introduction to spectra processing. Special thanks go to Prof. H.J.C. Berendsen for a number of helpful discussions about rotational averaging of tensors and suggestions especially on the treatment of united-atom methyl groups. Financial support was obtained from the Schweiz-

erischer Nationalfonds, project number 21-50929.97, which is gratefully acknowledged.

References

- Boelens, R., Koning, T.M.G., van der Marel, G.A., van Boom, J.H. and Kaptein, R. (1989) *J. Magn. Reson.*, **82**, 290–308.
- Brüschweiler, R. and Case, D. (1994) *Progr. NMR Spectrosc.*, **26**, 27–58.
- Daura, X., Antes, I., van Gunsteren, W.F., Thiel, W. and Mark, A.E. (1999a) *Proteins*, **36**, 542–555.
- Daura, X., Jaun, B., Seebach, D., van Gunsteren, W.F. and Mark, A.E. (1998) *J. Mol. Biol.*, **280**, 925–932.
- Daura, X., van Gunsteren, W.F. and Mark, A.E. (1999b) *Proteins*, **34**, 269–280.
- Edmondson, S.P. (1992) *J. Magn. Reson.*, **98**, 283–298.
- Edmondson, S.P. (1994) *J. Magn. Reson.*, **B103**, 222–233.
- Fischer, M.W.F., Majumdar, A. and Zuiderweg, E.R.P. (1998) *Progr. NMR Spectrosc.*, **33**, 207–272.
- Keepers, J.W. and James, T.L. (1984) *J. Magn. Reson.*, **57**, 404–426.
- Lipari, G. and Szabo, A. (1982) *J. Am. Chem. Soc.*, **104**, 4546–4559.
- Macura, S. and Ernst, R.R. (1980) *Mol. Phys.*, **41**, 95–117.
- Neuhaus, D. and Williamson, M. (1989) *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH, New York, NY.
- Olejniczak, E.T. (1989) *J. Magn. Reson.*, **81**, 392–394.
- Seebach, D., Ciceri, P.E., Overhand, M., Jaun, B., Rigo, D., Oberer, L., Hommel, U., Amstutz, R. and Widmer, H. (1996) *Helv. Chim. Acta*, **79**, 2043–2066.
- Tropp, J. (1980) *J. Chem. Phys.*, **72**, 6035–6043.
- Walser, R., Mark, A.E., van Gunsteren, W.F., Lauterbach, M. and Wipff, G. (2000) *J. Chem. Phys.*, **112**, 10450–10459.
- Woessner, D.E. (1962) *J. Chem. Phys.*, **36**, 1–4.