Journal of Biomolecular NMR **30:** 121–131, 2004.
© 2004 *Kluwer Academic Publishers. Printed in the Netherlands.*

Improved simulation of NOESY spectra by RELAX-JT2 including effects of J-coupling, transverse relaxation and chemical shift anisotropy

Andreas Ried[a], Wolfram Gronwald[a], Jochen M. Trenner[a], Konrad Brunner[a], Klaus-Peter Neidig[b] & Hans Robert Kalbitzer[a,*]

[a]*Department of Biophysics and Physical Biochemistry, University of Regensburg, Postfach, D-93040 Regensburg, Germany;* [b]*Bruker BioSpin GmbH, Software Department, Silberstreifen 4, D-76287 Rheinstetten, Germany*

Received 2 March 2004; Accepted 28 June 2004

Key words: J-coupling, NMR, NOESY, protein structure, relaxation, T_2-simulation

Abstract

RELAX-JT2 is an extension of RELAX, a program for the simulation of ^{1}H 2D NOESY spectra and ^{15}N or ^{13}C edited 3D NOESY-HSQC spectra of biological macromolecules. In addition to the already existing NOE-simulation it allows the proper simulation of line shapes by the integrated calculation of T_2 times and multiplet structures caused by J-couplings. Additionally the effects of relaxation mediated by chemical shift anisotropy are taken into account. The new routines have been implemented in the program AUREMOL, which aims at the automated NMR structure determination of proteins in solution. For a manual or automatic assignment of experimental spectra that is based on the comparison with the corresponding simulated spectra, the additional line shape information now available is a valuable aid. The new features have been successfully tested with the histidine-containing phosphocarrier protein HPr from *Staphylococcus carnosus*.

Abbreviations: HPr – histidine-containing phosphocarrier protein; *S. carnosus* – *Staphylococcus carnosus*

Introduction

The back calculation of multidimensional NOESY spectra from a single trial structure or a set of trial structures represents a powerful tool in macromolecular NMR spectroscopy. In its main applications it is used as a tool for structure validation, for obtaining more reliable distance information, for structure driven manual or automated assignment and structure determination. Validation of NMR structures is often done on the basis of NMR *R*-factors (see e.g., Lefevre et al., 1987; Borgias and James, 1990; Nilges et al., 1991; Bonvin et al., 1991; Clore et al., 1993; Xu et al., 1995; Cullinan et al., 1996; Gronwald et al., 2000) where the experimental NOESY volumes are compared with the volumes obtained from the back calculation of the NOESY spectrum from the trial structure. Another interesting application of back calculated spectra is it use for recognizing the structure that exists predominantly under given conditions in solution. This has been used to prove that an interesting strained conformation of HPr from *E. faecalis* obtained by X-ray crystallography (Jia et al., 1993) cannot be the dominant conformation in solution (Maurer et al., 2001).

For obtaining distance information from NOESY spectra the initial slope approach is only valid when the mixing time is very small where spin diffusion can usually be neglected. As a consequence the complete relaxation matrix approach has been introduced initially by (Keepers and James, 1984) and has been used to delineate spin diffusion pathways in proteins (Shibata and Akasaka, 1990). The full relaxation matrix approach can then be used to obtain reliable

*To whom correspondence should be addressed. E-mail: Hans-Robert.Kalbitzer@biologie.uni-regensburg.de

inter-atomic distances (Boelens et al., 1989; Borgias and James, 1990; Post et al., 1990; van de Ven et al., 1991; Madrid et al., 1991; Kim and Reid, 1992) in an iterative way or to directly enter the simulated NOESY volumes in the target function of the iterative structure calculation (Yip and Case, 1989; Bonvin et al., 1991; Brünger, 1992).

Back calculated NOESY-spectra can be used in the manual assignment procedure by visual comparison with the experimental spectra. However, this procedure has to be applied with care in cases where still strong deviations between the trial structure and the true structure can be assumed, since partially wrong structures may be artificially stabilized by the choice of NOEs fitting the actual imperfect structure. The iterative ARIA method (Linge et al., 2003) for automated NOE assignment is using back-calculated spectra to obtain accurate distance information from the automatically assigned spectra.

The optimal use of back calculated spectra requires that the simulation is as perfect as possible. With regard to NOESY spectra it is clear that the calculation of the NOEs should be optimal. In addition, a real spectrum is determined by multiplet structures (either resolved or not) and the individual line widths (T_2-times) of the signals. The last two properties are neglected in the programs published until now, while for the calculation of NOE-intensities a large number of approaches exist with different levels of complexity.

The most well-known programs for NOE simulations in this context are: CORMA (Keepers and James, 1984), BCKCALC (Banks et al., 1989), IRMA (Boelens et al., 1989), MARDIGRAS (Borgias and James, 1990), MORASS (Post et al., 1990), DINOSAUR (Bonvin et al., 1991), MIDGE (Madrid et al., 1991), NO2DI (van de Ven et al., 1991), a program by Kim and Reid (Kim and Reid, 1992), X-PLOR (Brünger, 1992), BIRDER (Zhu and Reid, 1995), RELAX (Görler and Kalbitzer, 1997; Görler et al., 1999), and SPRIT (Zhu et al., 1998). In practice, the main differences between the various approaches are the motional models available for the description of internal motions and the (possibly anisotropic) overall rotational diffusion.

RELAX-JT2, presented in this paper is based on the program RELAX (Görler and Kalbitzer, 1997; Görler et al., 1999) which is able to treat dipolar relaxation induced by non-isotropic tumbling and internal motion. For every pair of atoms (groups of atoms) an individual model for the internal motion can be specified. Spectral effects resulting from finite relaxation delays are also included in the calculation of the relaxation rates. It allows the simulation of 2D-NOESY and ^{15}N or ^{13}C edited 3D NOESY spectra. Different transfer efficiencies of the INEPT and reverse INEPT steps can be can be taken from the corresponding 2D-HSQC spectrum.

RELAX-JT2 additionally includes relaxation mediated by chemical shift anisotropy (CSA) which may dominate at high fields. To use the full information content of NOE signals it is important to simulate the correct line-shape information. Therefore, the program is able to simulate the multiplet structure due to J-coupling and the line-width of the multiplet components by a T_2-calculation.

The new routines were tested for the simulation of 2D ^{1}H NOESY and 3D ^{1}H-^{15}N NOESY-HSQC spectra of the histidine-containing phosphocarrier protein (HPr). HPr from *Staphylococcus carnosus* is a medium sized protein of 88 residues in size. Its three-dimensional structure was recently solved by NMR-methods (Görler et al., 1999) and shows a well-defined tertiary structure consisting of a four-stranded antiparallel β-sheet and three α-helices. The advantages of the new features will be demonstrated by detailed comparison between the corresponding experimental and simulated 2D spectra.

Materials and methods

Software. The program module RELAX is written in ANSI-C and has been integrated in the most recent version of AUREMOL that is available free of charge from the following website: http://www.auremol.de.

NMR-spectroscopy. Unlabeled HPr-protein from *S. carnosus* was obtained from W. Hengstenberg, Bochum. The sample contained 4.3 mM HPr in 90% H_2O/10% D_2O. In both cases the pH was adjusted to 7.2. The 2D spectrum using a mixing time of 150 ms was acquired at a proton resonance frequency of 800.13 MHz with 1024 * 8192 time domain data points in t_1 and t_2 directions, respectively. A relaxation delay of 2.37 s between the scans was used.

The spectrum was measured at 298 K. The corresponding three-dimensional solution structure of HPr from *S. carnosus* (Görler et al., 1999) used in the calculations was taken from the set of structures submitted to the PDB, accession code 1QR5.

Theoretical considerations

In the following the theoretical basis of the new features introduced in RELAX-JT2 will be recalled. It includes the simulation of J-coupling patterns in NOESY-spectra, the inclusion of chemical shift anisotropy (CSA), and the calculation of the transverse relaxation times. Since normal multidimensional spectra have a low digital resolution and a limited signal-to-noise ratio, only aspects which significantly influence line intensities and peak shapes are taken into account.

Calculation of multiplet patterns in 2D-NOESY and 3D-NOESY-HSQC-spectra

The calculation of J-coupling patterns can become cumbersome and time consuming when strong coupling is considered. Therefore, for the multiplet simulations weak coupling was always assumed. This is in most practical cases a satisfactory approximation and leads to a straightforward calculation of the intensities and positions of the multiplet components (Cavanagh et al., 1996). In the present implementation it has been assumed that heteronuclear J-couplings are suppressed in NOESY experiments by the application of appropriate decoupling schemes. In addition, only two-bond and three-bond homonuclear J-couplings are considered, since the other couplings are too small to have significant effects on the multidimensional spectra recorded with limited spectral resolution. The line splitting between the individual multiplet components is determined by the magnitude of the corresponding coupling constants. Invariant coupling constants in fixed covalent structures were read from a table (Wüthrich, 1976) that is stored as an ASCII file in the global AUREMOL database for invariant coupling constants. Structurally dependent coupling constants $^3J_{ij}$ between spins I_i and I_j were calculated from the given three-dimensional structure by use of a parameterized Karplus Equation 1. In addition, experimentally determined coupling constants can be specified by the user and replace the normally used standard values.

$$^3J_{ij} = A\cos^2\Theta + B\cos\Theta + C. \quad (1)$$

In the parameterized Karplus Equation 1 the dihedral angle Θ is defined as $\Theta = (\phi, \chi_1, \chi_2, \chi_3,...) - 60°$ and the backbone and side-chain dihedral angles ϕ, χ_1, χ_2, χ_3,... are defined according to IUPAC rules. The Karplus parameters A = 6.51, B = −1.76, and C = 1.60 were used for all dihedral angles where the second or third atom defining the angle is a nitrogen e.g. for ϕ backbone angles (Vuister and Bax, 1993). For all other dihedral angles e.g., χ_1 side-chain angles the following values were used: A = 9.5, B = −1.6, and C = 1.8 (deMarco et al., 1978).

Calculation of the NOE-build up in the presence of chemical shift anisotropy

The cross peak volume in NOESY-spectra can be calculated as numerical solution of Equation 2 (see e.g., (Görler and Kalbitzer, 1997))

$$\Delta\mathbf{M}_z(t) = \Delta\mathbf{M}_0(\mathbf{E} - \exp(-t_r\mathbf{R}))\exp(-t\cdot\mathbf{R}) \quad (2)$$

with $\Delta\mathbf{M_0}$ the equilibrium magnetization, $\mathbf{R}$ the relaxation matrix, $\mathbf{E}$ the unity matrix, and t_r the repetition time (the time between the start of the data acquisition and the first 90 degree pulse of the NOESY sequence). The diagonal and outer diagonal elements of the relaxation matrix are the auto-relaxation rates ρ_i and the cross-relaxation rates σ_{ij} of spins I_i and I_j, respectively. The spin lattice relaxation rates ρ_i can be written as the sum of the dipolar relaxation rate ρ_i^D and chemical shift anisotropy (CSA) relaxation rate ρ_i^{CSA}. For two unlike spin 1/2 particles and an axially symmetrical CSA tensor (Luginbühl and Wüthrich, 2002) one obtains:

$$\begin{aligned}\rho_i =& \rho_i^D + \rho_i^{CSA} = \frac{1}{10}\gamma_i^2\gamma_j^2\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2\Big[J^D(\omega_i-\omega_j)\\ &+ 3J^D(\omega_i) + 6J^D(\omega_i+\omega_j)\Big] + \\ &\frac{2}{15}\omega_i^2(\sigma_{||}-\sigma_\perp)^2J^{CSA}(\omega_i).\end{aligned} \quad (3)$$

Here, $\sigma_{||}$ and $\sigma_\perp$ are the components of the chemical shift tensor parallel and perpendicular to the principal axis, $\gamma_{i,j}$ the gyromagnetic ratios and μ_0 the magnetic permeabilty of the vacuum. The various spectral densities $J(\omega)$ are described in more detail by Görler and Kalbitzer (1997). The spectral densities $J^D(\omega)$ and $J^{CSA}(\omega)$ for the dipolar and CSA relaxation and CSA must not be necessarily the same since they may describe different motions. In addition, in the notation used in Görler et al. $J^D(\omega)$ contains a factor r_{ij}^{-6} which is not contained in the corresponding spectral density $J^{CSA}(\omega)$. The cross relaxation rates σ_{ij} are given by

$$\begin{aligned}\sigma_{ij} = \sigma_{ij}^D = &\frac{1}{10}\gamma_i^2\gamma_j^2\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2\\ &\left[-J^D(\omega_i-\omega_j) + 6J^D(\omega_i+\omega_j)\right].\end{aligned} \quad (4)$$

As usual, the terms corresponding to $2I_{i,z}I_{j,z}$ are neglected since their contribution to the relaxation rate

is usually small. For a N-spin system one obtains

$$\rho_i = \rho_i^D + \rho_i^{CSA} = (\sum_{j,j\neq i} \rho_{ij}^D) + \rho_i^{CSA}. \quad (5)$$

The transverse relaxation for two multiplet components $I_i^{(1)}$ and $I_i^{(2)}$ of a spin I_i J-coupled to spin I_j can be described by (Goldman, 1984):

$$\frac{d}{dt}\left\langle I_{i,+}^{(1)}\right\rangle = -i\pi J\left\langle I_{i,+}^{(1)}\right\rangle - (\lambda+\eta)\left\langle I_{i,+}^{(1)}\right\rangle - \mu\left\langle I_{i,+}^{(2)}\right\rangle \quad (6)$$

$$\frac{d}{dt}\left\langle I_{i,+}^{(2)}\right\rangle = -i\pi J\left\langle I_{i,+}^{(2)}\right\rangle - (\lambda-\eta)\left\langle I_{i,+}^{(2)}\right\rangle - \mu\left\langle I_{i,+}^{(1)}\right\rangle \quad (7)$$

with

$$\lambda = \frac{1}{20}\gamma_i^2\gamma_j^2\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2\left[4J^D(0) + J^D(\omega_i-\omega_j) + 3J^D(\omega_i) + 3J^D(\omega_j) + 6J^D(\omega_i+\omega_j)\right] + \frac{1}{15}\omega_i^2(\sigma_{||}-\sigma_\perp)\left[4J^{CSA}(0) + J^{CSA}(\omega_i)\right] \quad (8)$$

$$\eta = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)\gamma_i\gamma_j\hbar\omega_i\left(\sigma_{||}-\sigma_\perp\right)r_{ij}^{-3}\left[4J^{CSA}(0) + 3J^{CSA}(\omega_i)\right] \quad (9)$$

$$\mu = \frac{3}{20}\gamma_i^2\gamma_j^2\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2 J^D(\omega_j). \quad (10)$$

The auto-relaxation rate λ decomposes in a dipolar term given in the first part of Equation 8 and a chemical shift anisotropy term given in the second part of Equation 8. In Equation 9 it is assumed that the symmetry axis of the chemical shift tensor is collinear with the $I_\mathrm{i} - I_\mathrm{j}$ bond vector. The interference term η is neglected in our calculations, since heteronuclear decoupling was assumed and it is strongly scaled down with the distance. In addition, it would lead to different line widths of the individual multiplet components of the split peaks, which would significantly increase the computational time. Due to the cross-relaxation term μ between the individual multiplet components of split peaks (transverse Overhauser effect) the relaxation matrix formalism is necessary. On the other hand the contribution of the cross-relaxation rate to the total-relaxation rate is very small in the presence of indirect coupling between $\mathrm{I_i}$ and $\mathrm{I_j}$ since the splitting of the resonance lines renders the cross relaxation between the split lines ineffective. As a consequence it is possible to neglect the cross-relaxation rate in the case of indirect coupling between $\mathrm{I_i}$ and $\mathrm{I_j}$. If there is no indirect coupling the cross-relaxation rate will be added to the auto-relaxation rate.

The transverse relaxation rate $\mathrm{R_2}$ is:

$$\begin{aligned} R_2^{(1,2)} &= \lambda + \mu \quad \text{if J} = 0 \\ R_2^{(1,2)} &= \lambda \qquad \text{if J} \neq 0. \end{aligned} \quad (11)$$

For resonances with negligible J-coupling one obtains then

$$\frac{1}{T_{2,ij}} = \frac{1}{20}\gamma_i^2\gamma_j^2\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2\left[4J^D(0) + J^D(\omega_i-\omega_j) + 3J^D(\omega_i) + 6J^D(\omega_j) + 6J^D(\omega_i+\omega_j)\right] + \frac{1}{15}\omega_i^2(\sigma_{||}-\sigma_\perp)[4J^{CSA}(0) + J^{CSA}(\omega_i)]. \quad (12)$$

For like spins $\mathrm{I_i}$ and $\mathrm{I_j}$ the transverse relaxation is described by

$$\frac{d}{dt}\langle I_+\rangle = -(\lambda+\mu)\langle I_+\rangle \quad (13)$$

$$\frac{d}{dt}\langle J_+\rangle = -(\lambda+\mu)\langle J_+\rangle \quad (14)$$

with

$$\lambda = \frac{1}{20}\gamma_i^4\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2\left[5J^D(0) + 9J^D(\omega_i) + 6J^D(2\omega_i)\right] + \frac{1}{15}\omega_0^2(\sigma_{||}-\sigma_\perp)[4J^{CSA}(0) + J^{CSA}(\omega_0)] \quad (15)$$

$$\mu_{IS} = \frac{1}{20}\gamma_i^4\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2\left[4J^D(0) + 6J^D(\omega_i)\right]. \quad (16)$$

One then obtains

$$\frac{1}{T_{2,ij}} = \frac{1}{20}\gamma_i^4\hbar^2\left(\frac{\mu_0}{4\pi}\right)^2\left[9J^D(0) + 15J^D(\omega_i) + 6J^D(2\omega_i)\right] + \frac{1}{15}\omega_i^2(\sigma_{||}-\sigma_\perp)\left[4J^{CSA}(0) + J^{CSA}(\omega_i)\right]. \quad (17)$$

The obtained results can be easily extended to a system of N spins. The transverse relaxation rate $1/\mathrm{T_2}$ of a spin I is just the sum of the transverse relaxation rates $1/\mathrm{T}_{2,ij}$

$$\frac{1}{T_{2,i}} = \sum_{\substack{j\neq i \\ \textit{Like}}} \frac{1}{T_{2,ij}} + \sum_{\substack{j\neq i \\ \textit{Unlike}}} \frac{1}{T_{2,ij}}. \quad (18)$$

The longitudinal and transverse relaxation rates depend on the motion of the molecule, which is described by spectral density functions $J(\omega)$. In RELAX separate motional models can be used for given atom pairs to allow for a detailed description of the molecule under investigation. For CSA- relaxation in the present

```
COMPOUND: hpr

section_sequencedefinition
_Residue_seq_code
_Atom_num_code
_Residue_label
_Atom_name
_Atom_type
_Atom_alias
_Atom_equivalent
_Atom_CSA
  1   1 MET   HN H   -      -           8.95
  1   2 MET    N N   -      -           157.00
  1   3 MET   CA C   -      -           40.00
  1   4 MET   HA H   -      -           8.95
  1   5 MET    C C   -      -           40.00
  1   6 MET    O O   -      -           -
  1   7 MET   CB C   -      -           40.00
  1   8 MET  HB2 H   HB    1,9          8.95
  1   9 MET  HB3 H   HB    1,8          8.95
.
.
end_section

section_bond
_Bond_start
_Bond_atom1
_Bond_atom2
_Bond_atom3
_Bond_atom4
  1,1     1,2     -      -      -
  1,2     1,1    1,3     -      -
  1,3     1,2    1,4    1,5    1,7
  1,4     1,3     -      -      -
  1,5     1,3    1,6     -      -
  1,6     1,5     -      -      -
.
.
end_section

section_j-coupling
DEFINE
_Angle_name
_Angle_type
_Karplus_constantA
_Karplus_constantB
_Karplus_constantC
_Karplus_thetainc
phi            C'-N-CA-CB          6.51     -1.76     1.6       +60
chi_1          N-CA-CB-X           9.5      -1.6      1.8       +60
chi_2          CA-CB-CG-X          9.5      -1.6      1.8       +60
chi_2_CYS      CA-CB-SG-HG         9.5      -1.6      1.8       +60
chi_2_SER      CA-CB-OG-HG         9.5      -1.6      1.8       +60
.
.
end_section
```

Figure 1. Example of a compound file. The compound file consists of three sections. In the first section all atoms of a given compound e.g. a protein are defined. Each atom is associated with a residue number, an intra-residual atom number, an atom name, an atom type, an atom alias in case that possibly magnetically equivalent atoms exist e.g., for members of methyl groups, possibly equivalent atoms that are specified by their residue number and intra-residual atom number, and the chemical shift anisotropy. The next section defines the covalent bonds. Each atom is specified by its residue number and intra-residual atom number. The last section contains the information required for the calculation of three-bond J-coupling values. Each dihedral angle is defined by its name, by the corresponding atoms where X can be any heavy atom, the Karplus parameters A, B, and C, and the offset between Θ and the dihedral angles ϕ, χ_1, χ_2, χ_3,.... In the second part of this section (not shown) invariant coupling constants e.g., geminal coupling constants are given.

implementation an isotropic rotational diffusion is always assumed.

Implementation

Data base management. Three-dimensional structural information of the protein of interest is provided to the program by a single pdb file or a set of pdb files. The current IUPAC atom nomenclature is required here. To describe the additional information that is related to the NMR-specific parameters two new file types – the compound file and the meta file – have been developed and defined. The format of these files is related to the NMR-Star format used by the BioMagResBank (Seavey et al., 1991). In a compound file all information of a given compound e.g., a protein is stored that is invariant to the meas-

```
section_sampledefinition
COMPOUNDS:
_Mol_label
_Compound_file
_Concentration_value
_Concentration_value_units
_Isotopic_labeling
NAME   C:\Programme\Auremol\HPr_Daten\hpr.comp    -    -    -
END_COMPOUNDS
PARAMETER:
_Variable_type
_Variable_value
_Variable_value_unit
pH             7.0   -
Temperature    300.00   K
Pressure        1 Bar
END_PARAMETER
end_section


section_compounddefinition 1
NAME: hpr
CLASSDEF:
 DEFINE_CLASSES

 ANIS

 NAME             METHYL
 NUCLEUS 1H
 OCCUPANCY  1
.
.

END_CLASSDEF
SHIFTS:
_Residue_seq_code
_Atom_num_code
_Atom_alias
_Chem_shift_value
_Chem_shift_value_error
_Chem_shift_ambiguity_code
_Atom_class
_Linewidth
 1   1   -   8.210   0.05   1   3   -
 1   2   -   120.000   0.5   1   8   -
 1   3   -   56.660   0.5   1   7   -
.
.
END_SHIFTS
J_COUPL:
_Coupled_atom_1
_Coupled_atom_2
_Coupling_value
 2,1  2,4   10.1
END_J_COUPL
end_section
```

Figure 2. Example of a meta file. Relevant parts of a meta file are displayed. The file consists of two main sections with various subsections. In the first section the sample composition and general sample parameters are defined. For example if protein complexes are investigated each protein is specified as a separate compound. Each compound is associated with the path of its compound file, the concentration of that specific compound in the sample, and the information about isotope labeling. Next pH, temperature, and pressure of the sample are defined. In the next section information about each specific compound is contained. In the sub-section 'CLASSDEF' the various motional models that are applied during the calculations are defined. These definitions are similar to the definitions described in the original RELAX publication (Görler and Kalbitzer, 1997). The 'SHIFTS' subsection associates each atom identified by its residue number and intra-residual number, if applicable with an atom alias identifier, with its chemical shift, with an chemical shift error, with an chemical shift ambiguity code, with a motional model, and if available with an experimentally measured line-width. Please note that only atoms are listed for which chemical shifts are available. In the sub-section 'J_COUPL' experimentally measured J-couplings are stored. Each J-coupling is identified by the coupled atoms and the coupling constant.

urement conditions and the type of performed NMR experiment. The compound file is automatically created from a given protein sequence by AUREMOL. It contains three sections for the definition of atoms, bonds, and J-couplings. Figure 1 shows an example compound file. In the meta file (Figure 2) all information that is dependent on the measurement conditions and on the type of performed experiment(s) is contained, like e.g., the parameters of the applied dynamic models, resonance line assignments, and experimentally determined J-coupling values. It is created semi- or fully-automatically from the corresponding com-

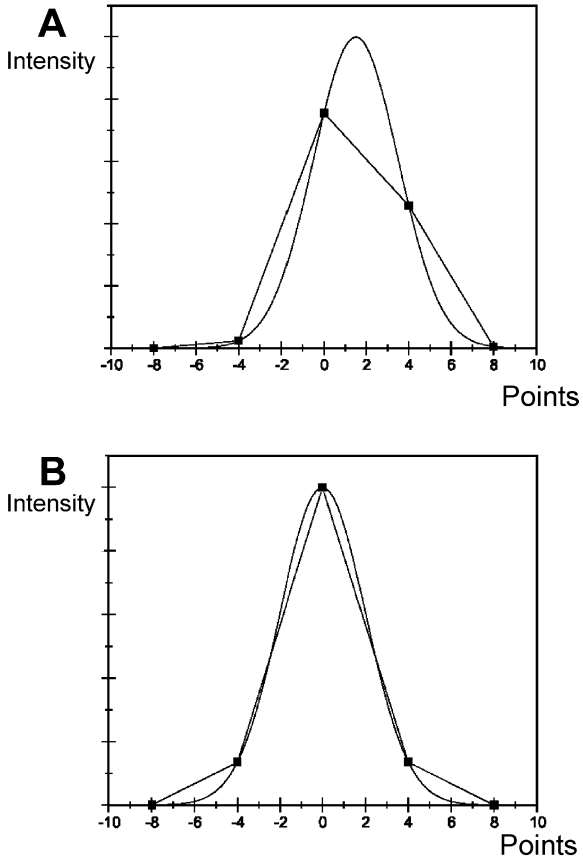

Figure 3. Improved graphical signal representation. (A) The graphical representation of a signal assuming unlimited digital resolution is shown as a smooth line. However, for the real case with finite digital resolution and where the peak maximum does not coincide with a point on the digital grid one obtains distorted line shapes and signal positions as represented by the straight lines connecting the squares. Here each square represents the intensity of one point on the digital grid. (B) An improved line shape representation is obtained by shifting the peak so that its maximum fits nearest point on the digital grid.

pound file. If experimental resonance line assignments are available they can be either read from an ASCII-file or they can be specified within a graphical user interface (GUI). Otherwise the metafile can be automatically filled with database values, e.g., random coil chemical shifts. Also the information about the applied motional models that is stored in the meta file can be edited by the use of the GUI.

Graphical representation of nD-spectra. The frequency domain data were calculated by a superposition of Gaussian or Lorentzian lineshapes (as an option for the user) of suitable line widths. On average a single protein resonance line is split into 16 components due to J-coupling, therefore an average 2D crosspeak would be split into 256 sub-peaks. However, due to the line widths of the individual signals and the limited digital resolution of the multidimensional spectra usually not all of these components are separately visible. As a consequence it is feasible to combine some of the sub peaks of a multiplet before the spectrum is generated, which drastically reduces the required computational effort. Sub-peaks are combined within RELAX in case that at least one of the following two conditions is fulfilled:

$$d_{i,j} < 0.5\Delta\nu_x \tag{19}$$

$$d_{i,j} < 0.5R_D \tag{20}$$

with the distance $d_{i,j}$ measured in [Hz] between two sub-peaks i, j of a multiplet and $\Delta\nu_x$ the corresponding line width measured in [Hz] in a dimension x. Please note that in our calculations within one dimension the same line width is computed for all sub-peaks of a multiplet since interference effects between dipolar relaxation and CSA relaxation that would lead to different line widths between individual multiplet components are neglected. The maximum digital resolution R_D obtained in one of the dimensions is measured in Hz/points. Another problem for the graphical representation of the spectrum arises from the limited digital resolution(s), which may lead to distorted line shapes and peak positions in case that a peak maximum does not coincide with the digital grid. An example is shown in Figure 3. As a consequence within RELAX-JT2 peaks are shifted so that their maxima fit the nearest grid point and much more realistic symmetrical line shapes are obtained.

Results and discussion

The simulation of individual line widths, splittings due to J-couplings and the inclusion of relaxation caused by CSA allows a realistic simulation of multidimensional NOESY spectra. We have tested the new routines for the simulation of 2D ^{1}H NOESY spectra and 3D ^{1}H-^{15}N NOESY-HSQC spectra of the medium sized protein HPr from *S. carnosus*. For the simulation the same parameters e.g. relaxation delay between scans, mixing time, digital resolution etc. that were described for the corresponding experimental spectra were also used in the simulations. The global correlation time $\tau_c = 5.62$ ns of HPr used in the simulations was obtained from relaxation measurements

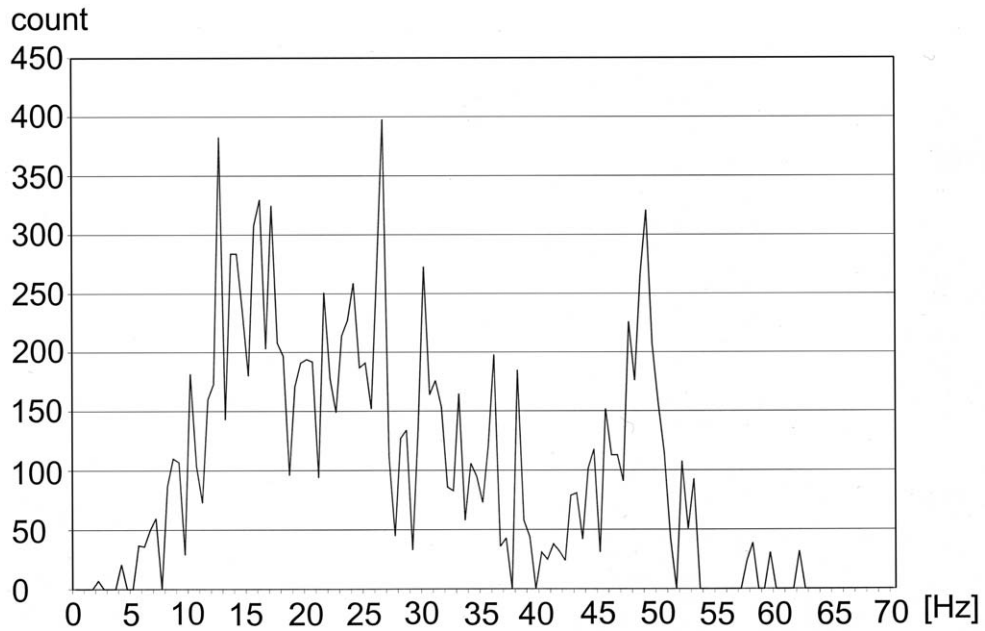

Figure 4. Theoretical distribution of transversal proton relaxation times in HPr from *S. carnosus*. The calculated relaxation rates $1/T_2$ of the proton resonances are depicted. Relaxation times were calculated as described in RESULTS and DISCUSSION. On the x-axis the relaxation rates are shown while on the y-axis the number of occurrences of a specific relaxation rate is displayed. The individual relaxation rates were grouped in steps of 0.5 Hz. Please note that the above distribution was calculated for all signals present in the simulated 2D NOESY spectrum.

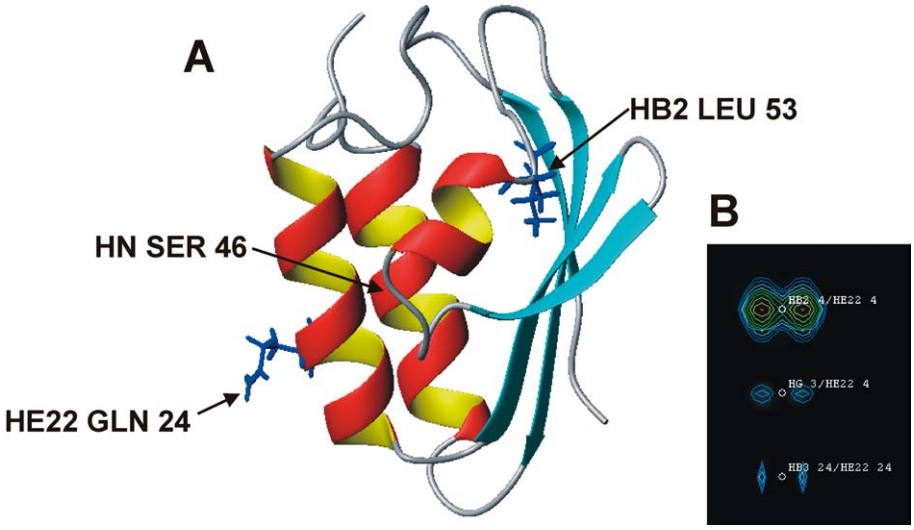

Figure 5. Line width dependence on structural environment. (A) The smallest and largest line widths of 0.658 Hz and 19.729 Hz were calculated for $H^{\varepsilon 22}$ of Gln24 and $H^{\beta 2}$ of Leu53, respectively. The smallest theoretical line width for a backbone amide resonance is 2.899 Hz for Ser46, which is located in a loop region on the surface of the protein. The corresponding side chains are highlighted in blue in a ribbon diagram of HPr. (B) The simulated cross peaks for $H^{\varepsilon 22}$ of Gln24 and $H^{\varepsilon 22}$ of Gln4 in a ^1H 2D NOESY spectrum of HPr are shown. Please note that the signals shown in B correspond only partly to the marked atoms of part A since all signals corresponding to HB2 of Leu 53 appear in heavily overlapped regions of the spectrum. The experimental and the simulated spectrum were artificially broadened by an exponential filter leading to an additional line broadening of 2.4 Hz in the two dimensions.

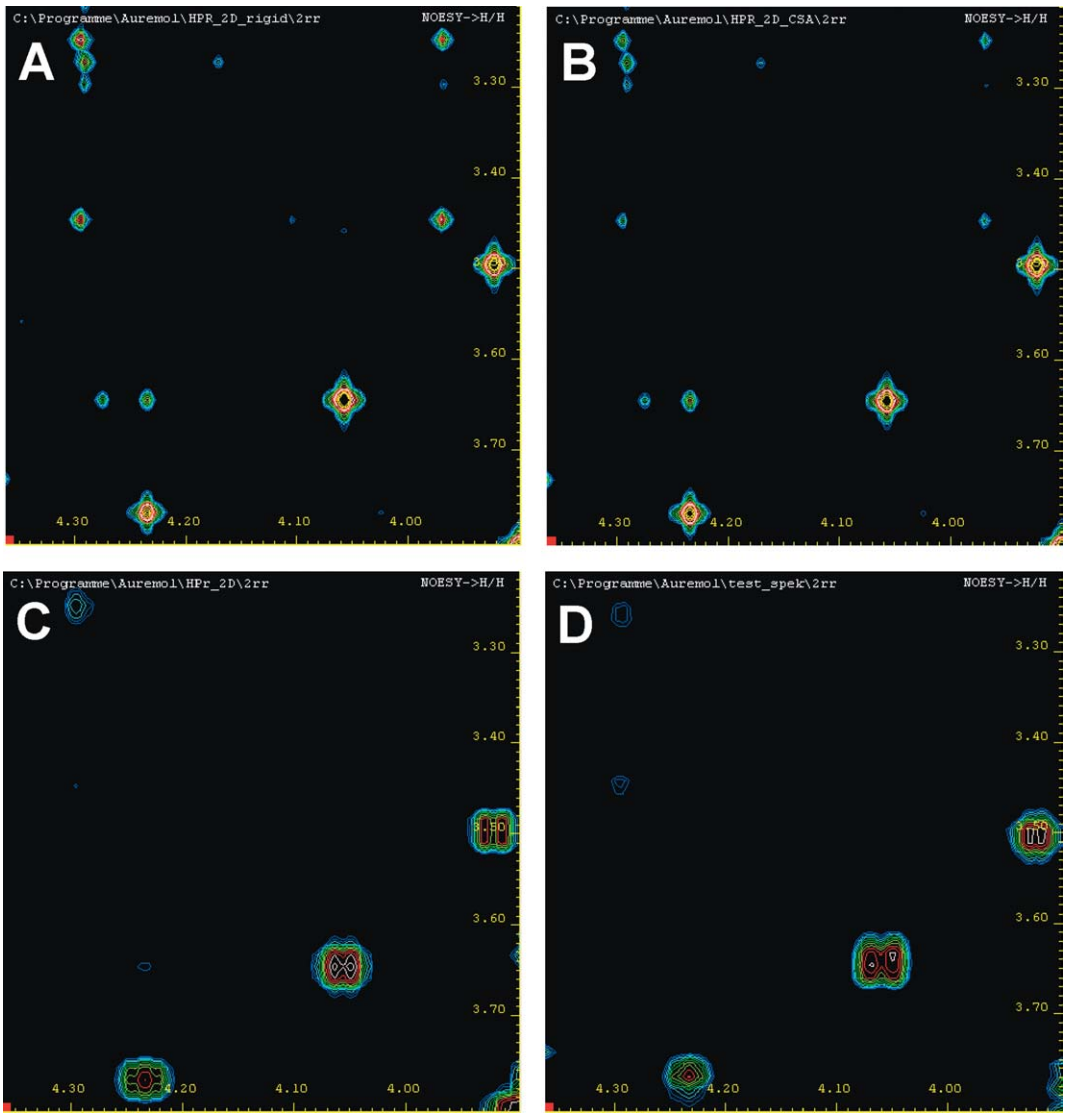

Figure 6. Comparison of experimental and simulated ^{1}H 2D NOESY spectra of HPr. Part of a 2D-NOESY spectrum of HPr from *S. carnosus*. (A) Simulation of the spectrum using a rigid sphere model and assuming a Lorentzian line shape with 8 Hz line width. (B) same as (A) but allowing internal mobility and additional relaxation by chemical shift anisotropy. (C) as (B) but calculating the line shape with J-coupling and individual T_2-times. The experimental spectrum (D) was obtained from 1024 × 8192 time domain data points leading after Fourier transformation to a real frequency domain data set of 1024 × 8192 points. The digital resolution was 10.9 and 1.4 Hz/point in δ_1 and δ_2, respectively. The experimental data were filtered in both dimensions by exponential multiplication with a line broadening of 2.4 Hz. The simulated spectra were calculated with the same digital resolution and a Lorentzian line shape was selected in all cases. An additional line broadening of 2.4 Hz in both directions was assumed for the simulated spectra to take the filtering of the experimental data into account.

performed on uniformly ^{15}N enriched HPr at 298 K (Schubel et al., to be published). From the possible spectral densities as defined in (Görler and Kalbitzer, 1997) LIPARI_1 was selected for all atom pairs not including a methyl group or an aromatic ring. It represents a simplification of the original spectral density defined by Lipari and Szabo (1982a, 1982b). The use of LIPARI_1 is justified in cases where one can assume that the correlation times of the fast internal motions are considerably smaller than the global correlation time. For all atom pairs containing protons from a methyl group a fast-jump model was used for the spectral density where it is assumed that the internal correlation time of the methyl group is much

smaller than the global correlation time. For atom pairs containing members from aromatic rings where it can be assumed that the internal correlation time of the jump motion of the ring is much larger than the global correlation time a slow jump approximation was selected for the spectral density. For all atom pairs containing only backbone atoms an average order parameter S^2 of 0.95 has been experimentally determined (Schubel et al., to be published). For all atom pairs containing side-chain and main-chain atoms an S^2 of 0.80 was used while for side-chain side-chain interactions an S^2 value of 0.65 was assumed. The latter two values were not experimentally determined but taken from the X-PLOR manual (Brünger, 1992). They are in good agreement with the known literature (Buck et al., 1995; Mulder et al., 2001; Skrynnikov et al., 2001). However, it should be noted that the S^2 values determined for the different side-chains of a protein usually vary over a large range. Values for S^2 between 0.2 and 0.9 are not uncommon. As a consequence the description of side-chain mobility is only a relatively rough approximation in this example. In the program it is also possible to automatically correct for deviations of the molecule from spherical shape. However, in case this option is activated, in the current version of RELAX the molecule is treated as a rigid body. Since the three-dimensional structure of HPr can be approximated fairly well with a sphere this option was not used for the tests shown.

In addition to the calculated spectrum RELAX also provides an ASCII text file that indicates for each simulated signal the contributing line widths. Depending on the position in the structure the calculated full proton line widths at half height ($1/\pi T_2$) for HPr vary widely in the range of 0.7 Hz and 19.7 Hz. The average proton line width for all simulated signals is 8.73 Hz for HPr which is a realistic value for a protein of 88 residues in size (Figure 4). The largest line width was obtained for $H^{\beta 2}$ of Leu53, the smallest line width was obtained for $H^{\varepsilon 22}$ of Gln24. As it can be seen in Figure 5A the side chain of Gln24 protrudes from the structure and $H^{\varepsilon 22}$ at the end of this side chain has very few surrounding protons leading to slow relaxation of that atom. Here also a limitation of our model used for the description is apparent; the experimentally observed line width of side chain amide resonances is typically much larger since an additional exchange broadening affects the experimental line width. In our case the calculated line width of the $H^{\varepsilon 22}$–resonance of Gln24 is 0.7 Hz. The smallest line width calculated for a non-exchanging proton is 1.3 Hz for $H^{\varepsilon 1}$ of His15. This is also shown in Figure 4. The opposite is true for $H^{\beta 2}$ of Leu 53 that is located in the core of the protein on top of β-strand three. Here, the high surrounding proton density causes fast relaxation of that nucleus. Figure 5B shows that the line width information can be used for assignment purposes. In this example cross peaks involving $H^{\varepsilon 22}$ of Gln4 and $H^{\varepsilon 22}$ of Gln24 could easily be distinguished even when they would have the same chemical shifts.

Figure 6 shows a comparison between the experimental ^{1}H 2D-NOESY spectrum of HPr and the corresponding simulated spectra displayed on the right hand side. The experimental spectrum (Figure 6D) shows clear variations in line widths and multiplet structures which cannot be simulated by the original version of RELAX using a rigid sphere model (Figure 6A). However assuming internal mobility leads to a better prediction of the intensities (Figure 6B). A satisfactory result is only obtained when line widths and multiplet structures are calculated as it is possible in RELAX-JT2 (Figure 6C).

The realistic line-shape simulation is of special importance for the manual and automated assignment of NOESY spectra. In conventional approaches the assignment is mainly based on chemical shift information and to some degree on peak intensity. With our new routines additional line shape information can be used, which should help to resolve previously ambiguous assignments. Also it is now possible to distinguish between signals that are split due to J-couplings and neighboring signals for example originating from multiple conformers.

In case that a high quality three-dimensional structure of the molecule of interest is available, a particularly interesting possible application of the back calculation of NOESY spectra including the simulation of T_2 times should be the extraction of motional parameters e.g., S^2 values by iteratively adapting these values until the best possible match between experimental and simulated spectra is obtained. In this manner it should be possible to obtain motional information from all atoms that are visible in one particular NOESY spectrum using only this single spectrum. We will pursue this avenue in the future where the motional parameters are directly extracted from the NMR-spectra. In summary we feel that the program is a useful tool for the assignment of 2D and 3D NOESY spectra.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft and the European Commission (SPINE-QLG2-CT-2002-00988). We thank W. Hengstenberg for providing the HPr protein.

References

Banks, K.M., Hare, D.R. and Reid, B.R. (1989) *Biochemistry*, **28**, 6996–7010.

Boelens, R., Koning, M.G., van der Marel, G.A., van Boom, J.H. and Kaptein, R. (1989) *J. Magn. Reson.*, **82**, 290–380.

Bonvin, A.M.J.J., Boelens, R. and Kaptein, R. (1991) *J. Biomol. NMR*, **1**, 305–309.

Borgias, B.A. and James, T.L. (1990) *J. Magn. Reson.*, **87**, 475–487.

Brünger, A.T. (1992) *X-PLOR Version 3.1*, Yale Univ. Press, New Haven/London.

Buck, M., Boyd, J., Redfield, C., MacKenzie, D.A., Jeenes, D.J., Archer, D.B. and Dobson, C.M. (1995) *Biochemistry*, **34**, 4041–4055.

Cavanagh, J., Fairbrother, W.J., Palmer III, A.G. and Skelton, N.J. (1996) *Protein NMR Spectroscopy Principles and Practice*, Academic Press Inc., San Diego.

Clore, G.M., Robien, M.A. and Gronenborn, A. (1993) *J. Mol. Biol.*, **231**, 82–102.

Cullinan, D., Korobka, A., Grollman, A.P., Patel, D.J., Eisenberg, M. and de los Santos, C. (1996) *Biochemistry*, **35**, 13319–13327.

deMarco, A., Llinas, M. and Wüthrich, K. (1978) *Biopolymers*, **17**, 617–636.

Goldman, M. (1984) *J. Magn. Reson.*, **60**, 437–452.

Görler, A. and Kalbitzer, H.R. (1997) *J. Magn. Reson.*, **124**, 177–188.

Görler, A., Gronwald, W., Neidig, K.P. and Kalbitzer, H.R. (1999) *J. Magn. Reson.*, **137**, 39–45.

Görler, A., Hengstenberg, W., Kravanja, M., Beneicke, W., Maurer, T. and Kalbitzer, H.R. (1999) *Appl. Magn. Reson.*, **17**, 465–480.

Gronwald, W., Kirchhöfer, R., Görler, A., Kremer, W., Ganslmeier, B., Neidig, K.P. and Kalbitzer, H.R. (2000) *J. Biomol. NMR*, **17**, 137–151.

Jia, Z., Vandonselaar, M., Quail, J.W. and Delbaere, L.T. (1993) *Nature*, **361**, 94–97.

Keepers, J.W. and James, T.L. (1984) *J. Magn. Reson.*, **57**, 404–426.

Kim, S.-G. and Reid, B.R. (1992) *J. Magn. Reson.*, **100**, 382–390.

Lefevre, J.-F., Lane, A.N. and Jardetzky, O. (1987) *Biochemistry*, **26**, 5076–5090.

Linge, J.P., Habeck, M., Rieping, W. and Nilges, M. (2003) *Bioinformatics*, **19**, 315–316.

Lipari, G. and Szabo, A. (1982a) *J. Am. Chem. Soc.*, **104**, 4546–4559.

Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.*, **104**, 4559–4570.

Luginbühl, P. and Wüthrich, K. (2002) *Prog. NMR Spectrosc.*, **40**, 199–247.

Madrid, M., Llinas, E. and Llinas, M. (1991) *J. Magn. Reson.*, **93**, 329–346.

Maurer, T., Döker, R., Görler, A., Hengstenberg, W. and Kalbitzer, H.R. (2001) *Eur. J. Biochem.*, **268**, 635–644.

Mulder, F.A.A., Skrynnikov, N.R., Hon, B., Dahlquist, F.W. and Kay, L.E. (2001) *J. Am. Chem. Soc.*, **123**, 967–975.

Nilges, M., Habazettl, J., Brünger, A.T. and Holak, T.A. (1991) *J. Mol. Biol.*, **219**, 499–510.

Post, C.B., Meadows, R.P. and Gorenstein, D.G. (1990) *J. Am. Chem. Soc.*, **112**, 6796–6803.

Seavey, B.R., Farr, E.A., Westler, W.M. and Markley, J.L. (1991) *J. Biomol. NMR*, **1**, 217–236.

Shibata, S. and Akasaka, K. (1990) *Magn. Reson. Chem.*, **28**, 129–137.

Skrynnikov, N.R., Mulder, F.A.A., Hon, B., Dahlquist, F.W. and Kay, L.E. (2001) *J. Am. Chem. Soc.*, **123**, 4556–4566.

van de Ven, F.J.M., Blommers, M.J.J., Schouten, R.E. and Hilbers, C.W. (1991) *J. Magn. Reson.*, **94**, 140–151.

Vuister, G.W. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 7772–7777.

Wüthrich, K. (1976) *NMR in Biological Research: Peptides and Proteins*, American Elsevier, North Holland.

Xu, Y., Sugar, I.P. and Krishna, N.R. (1995) *J. Biomol. NMR*, **5**, 37–48.

Yip, P. and Case, D.A. (1989) *J. Magn. Reson.*, **83**, 643–648.

Zhu, L. and Reid, B.R. (1995) *J. Magn. Reson.*, **B106**, 227–235.

Zhu, L., Dyson, H.J. and Wright, P.E. (1998) *J. Biomol. NMR*, **11**, 17–29.