Journal of Biomolecular NMR, **22: PLEASE RUN AGAIN** , 2002.
KLUWER/ESCOM

NMRKIN: Simulating line shapes from two-dimensional spectra of proteins upon ligand binding

Ulrich L. Günther[a,*] & Brian Schaffhausen[b]
[a]*Biozentrum N230, Institut für Biophysikalische Chemie, J.W. Goethe Universität, Frankfurt, Marie-Curie-Str. 9, 60439 Frankfurt, Germany;* [b]*Department of Biochemistry, School of Medicine, Tufts University, 136 Harrison Avenue, Boston, MA 02111, U.S.A.*

Received 20 August 2001; Accepted 3 January 2002

Key words: kinetics, line shapes, protein-ligand interactions

Abstract

The analysis of the shape of signals in NMR spectra is a powerful tool to study exchange and reaction kinetics. Line shapes in two-dimensional spectra of proteins recorded for titrations with ligands provide information about binding rates observed at individual residues. Here we describe a fast method to simulate a series of line shapes derived from two-dimensional spectra of a protein during a ligand titration. This procedure, which takes the mutual effects of two dimensions into account, has been implemented in MATLAB as an add-on to NMRLab (Günther et al., 2000). In addition, more complex kinetic models, including sequential and parallel reactions, were simulated to demonstrate common features of more complex line shapes which could be encountered in protein-ligand interactions. As an example of this method, we describe its application to line shapes obtained for a titration of the p85 N-SH2 domain of PI3-kinase with a peptide derived from polyomavirus middle T antigen (MT).

Introduction

Line shape analysis of signals in NMR spectra is a well-established method to study the kinetics of processes occurring on a micro- to millisecond time scale (Rao, 1989; Sandström, 1982; Lian and Roberts, 1993). Unlike many other methods for kinetic investigation, NMR spectroscopy can provide kinetic information on individual residues of a protein. In some cases it is also possible to obtain equilibrium constants. Increasing use of NMR to analyze protein-ligand interactions is opening new opportunities for line shape analysis. However with very few exceptions, line shapes have been derived from one-dimensional NMR spectra. The degeneracy of chemical shifts in proteins restricts this type of analysis to the study of isolated protons. Line shape analysis has been used in this way to investigate protein folding by monitoring line broadening of the aromatic signals of a protein (Huang and Oas, 1995; Burton et al., 1998). Protein-ligand interactions have also been frequently studied by monitoring the line broadening of the NMR signals of small ligands when they are bound to large proteins (Fejzo et al., 1999; Hoyt et al., 1994). Although analysis of line shapes has most commonly been carried out on one-dimensional spectra, lines can also be obtained from two-dimensional NMR spectra. This elegant approach has been applied by Balbach to study protein-folding (Balbach et al., 1996). In some instances it has been used to analyze protein-ligand interactions (Craven et al., 1996; Johnson et al., 1998). In studying the kinetics of protein-ligand interactions, the variation of line shapes as a function of the ligand concentration provides information for individual residues. Here, the variation of line shapes as a function of the ligand concentration provides information on the kinetics of the interaction. Another great advantage of NMR line shape analysis is that it can potentially provide qualitative insight into the mechanism of reactions. For example, Hensman showed different kinetic mechanisms for two different residues

*Author to whom correspondence should be addressed. E-mail: ugunt@bpc.uni-frankfurt.de

of the N-SH2 domain of PI3-kinase (Hensmann et al., 1994).

Our goal was to develop a readily applicable routine for the simulation of individual dimensions of two-dimensional spectra. Such analysis of line shapes of signals derived from ^{15}N-HSQC spectra in either the proton or nitrogen dimension allows calculation of off-rates for the protein-ligand interaction observed at individual amino acid residues of the protein. A detailed theoretical treatment of line shapes in two-dimensional HSQC spectra including the treatment of non-equilibrium chemical exchange has recently been published (Helgstrand et al., 2000). However, it is impractical to simulate the complete two-dimensional signal for series of line shapes in HSQC spectra. Therefore we decided to simulate cross-sections for both spectral dimensions. Because line broadening occurs in both spectral dimensions, the effect of line broadening in the other dimension must be considered in the line shape simulation. This was achieved by simulating simultaneously the cross-section of the signal in both spectral dimensions using identical exchange parameters (**P**, and k_{off}) but considering the different peak separations ($\Delta\nu_{\text{N}}$ and $\Delta\nu_{\text{H}}$) in each dimension. The relative intensities of lines in the second dimension are used to scale the intensity of lines in the first dimension and vice versa. This simple procedure allows fast simulations of series of line shapes derived from HSQC spectra. In these calculations we did not use a quantum-mechanical treatment of the HSQC experiment. This is justified for exchange rates which are significantly larger than the overall correlation time of the protein τ_c and in the absence of spin-spin couplings on the same order as exchange rates. In a classical treatment all relaxation rates are summarized in the line width of the signal. If the line width does not change significantly between the free and the complexed state of the protein the error arising from a classical treatment is small.

The method described here was implemented in MATLAB with an interface to the NMRLab (Günther et al., 2000) processing package.

Theoretical concepts

Dynamic processes with reaction rates on a μs to ms time scale that occur within the course of a single free induction decay (FID) affect the line shape of the corresponding signal (see Binsch, 1968; Rao, 1989; Sandström, 1982; Lian and Roberts, 1993 for reviews). This phenomenon was first studied by Anderson assuming Markovian transitions (Anderson, 1954). Qualitatively, in a two-state system the NMR signal experiences its most extreme broadening at the coalescence point where the exchange rate is $k_{\text{coal.}} = \pi\ \Delta\nu/\sqrt{2}$ ($\Delta\nu$ is the separation between lines). For faster rates a single signal is observed at an intermediate chemical shift between that of the two exchanging nuclei (fast exchange on an NMR timescale). Slower processes will yield two distinguishable signals for the exchanging nuclei.

A treatment of two-site chemical exchange based on the Bloch equations (Bloch, 1946) was first derived by Gutowsky et al. (1953):

$$M = \frac{i\omega_1 M_0 \left[(\tau_A + \tau_B) + \tau_A\tau_B\,(\alpha_A p_B + \alpha_B p_A)\right]}{(1 + \alpha_A\tau_A)\,(1 + \alpha_B\tau_B) - 1} \tag{1}$$

with $\alpha_A = T_2^{-1} - i\,(\Delta\omega + \delta\omega/2)$ and $\alpha_B = T_2^{-1} - i\,(\Delta\omega - \delta\omega/2)$, $\tau_A = 1/k_A = \tau/p_B$ and $\tau_B = 1/k_B = \tau/p_A$ employing a reduced life time $\tau = \tau_A\tau_B/\tau_A + \tau_B$. Here $\Delta\omega = 2\pi\Delta\nu$, $\Delta\nu$ is the separation of lines in Hz; $\delta\omega/2$ is the width of the signal at half height.

McConnell outlined a more common set of equations which describe the line shape of an NMR signal that are subject to chemical exchange (McConnell, 1958).

To study exchange between multiple sites or reactions involving more complex kinetic mechanisms it is preferable to write the Bloch equations modified for chemical exchange under steady-state conditions in a matrix notation (Binsch, 1968):

$$\mathbf{AM} = iC\mathbf{P}, \tag{2}$$

where **M** is a column vector containing the magnetizations M_i of the nuclei which are subject to chemical exchange. **P** is a column vector with the populations (mole fractions) of the reaction components and $C = \gamma B_1 M_0$. **A** is a square matrix that contains the

$$\mathbf{A} = 2\pi i\,(\mathbf{I}\nu - \mathbf{W}_0) + \mathbf{R}_2 + \mathbf{K}, \tag{3}$$

where **I** = the unity matrix, ν = a variable frequency defining the frequency range, $\mathbf{W}_0$ = a diagonal matrix with the larmor frequencies ν_i, $\mathbf{R}_2$ = a diagonal matrix with transverse relaxation rates, $R_2 = 1/T_2 = \pi \cdot \text{LW}$ (line width at half height) and **K** = a matrix containing the rate constants.

The total magnetization is the sum of the individual magnetization components:

$$M_{\text{tot}} = \mathbf{I}^T\mathbf{M} = iC1^T\mathbf{A}^{-1}\mathbf{P}. \tag{4}$$

For exchange between two sites we obtain

$$M_{\text{tot}} = -iC[1,1]\mathbf{A}^{-1}\begin{bmatrix} p_1 \\ p_2 \end{bmatrix}.$$

The analytical solution of this equation is the Gutowsky Equation (1).

However, it is more convenient to calculate line shapes directly from Equation 4 . This requires the inversion of the matrix **A** for each point of the spectrum. For iterative line shape fitting, this procedure is too time consuming. For this reason we have calculated the corresponding frequency according to Equation 5:

$$F(t) = \mathbf{I}\exp\left[(-\mathbf{K} + 2\pi i\mathbf{W}_0 - \mathbf{R}_2)\,t\right]\mathbf{P}. \qquad (5)$$

This equation was first derived by Anderson (1954) employing Markov processes for the transition between different states. In Equation 5 **K**, $\mathbf{W_0}$, $\mathbf{R_2}$ and **P** are the same matrices as in Equations 3 and 4. After Fourier transformation the resulting spectrum is identical to that obtained using Equation 4 which had been derived from the Bloch equations. This procedure is much faster and has the intrinsic advantage that typical processing steps such as zero filling or apodization functions can be incorporated into the line shape simulation. The calculation of the matrix exponential in Equation 5 is critical because the matrix elements can become very large causing significant numerical errors. This problem is particularly serious when **K** contains large exchange rates. We therefore use Padé rational polynomials to approximate matrix exponentials (Moler and Loan, 1979).

Line shape analysis was traditionally used to study two-site exchange processes ($\mathrm{A}\underset{k_2}{\overset{k_1}{\rightleftharpoons}}\mathrm{B}$) in small organic molecules using one-dimensional spectra. Exchange between two states is described by the mole fractions (= the population p_i) of one of the two states ($p_A = 1 - p_B$), the line widths of the two signals and their frequency separation ($\Delta\nu_{AB}$). A popular simplification is the use of reduced rates or lifetimes $\tau = \tau_1\tau_2/(\tau_1 + \tau_2) = 1/(k_1 + k_2)$. However this simplification may cause erroneous line shapes when the two rates are different, i.e. when the equilibrium constant $K_{eq} = \frac{k_1}{k_2}$ is far from unity. The location of the equilibrium is usually determined by thermodynamic parameters and can often be shifted by changing temperature. For two-site exchange processes the matrix **K** adopts the form

$$\mathbf{K} = \begin{bmatrix} k_1 & -k_2 \\ -k_2 & k_1 \end{bmatrix}. \qquad (6)$$

Similar principles can be employed to study the kinetics of reactions. Here the fate of the line shape depends on the relative concentration of the reaction partners. Sudmeier et al. (1980) have described theoretical line shapes for second order reactions of the type

$$A + X \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} B. \qquad (7)$$

When line shapes for A are obtained at different concentrations of the component X, they can be used to study the kinetics of this reaction. The location of the equilibrium can be conveniently adjusted by the addition of the component X. In equilibrium $\mathrm{d[A]}/\mathrm{d}t = \mathrm{d[B]}/\mathrm{d}t = 0$ and therefore $k_{\text{on}} \cdot [X]p_A = k_{\text{off}}p_B = k_{\text{off}}(1 - p_A)$. Because the equilibrium concentration of X is usually unknown a rate $k'_{\text{on}} = k_{\text{on}}[X]$ is used. $\tau_{\text{off}} = k_{\text{off}}^{-1}$ describes the life time of the state B. The exchange matrix has the same form as for a simple two-site exchange when k_1 is substituted by $k_{\text{on}}[X]$:

$$\mathbf{K} = \begin{bmatrix} k_{\text{on}}[X] & -k_{\text{off}} \\ -k_{\text{off}} & k_{\text{on}}[X] \end{bmatrix} = \begin{bmatrix} k'_{\text{on}} & -k_{\text{off}} \\ -k_{\text{off}} & k'_{\text{on}} \end{bmatrix}.$$

For a given off-rate and a given set of populations p_A the effective on-rate $k_{\text{on}}[X]$ can be calculated for each step. When line shapes are simulated for N different concentrations of X the minimum number of parameters to be adjusted is $N + 1$: N population values p_A and the off-rate (k_{off}). Figure 1 shows a typical example for line shapes simulated for a frequency separation of 150 Hz and an off-rate k_{off} of 1000 Hz. When the simulations are to be compared to experimental data, these $N + 1$ values can be adjusted manually comparing the observed data and the simulated curves. If a least-squares procedure is used the difference between the observed data and the simulated curves is minimized by adjusting k_{off} and the populations of p_A.

In these kinds of analyses solvent exchange may complicate line shapes. In ^{15}N-edited spectra of the free state of the protein will occasionally show lines at a lower intensity due to slow exchange of amide protons with protons from water. This can be handled by adding a third reaction partner in slow exchange. In our routines we simply add an outside equilibrium constant K_{ex} as a correction factor for the population of the free protein. This treatment is equivalent to adding an extra reaction component in slow exchange except that no chemical shift must be assumed for an additional reaction partner.

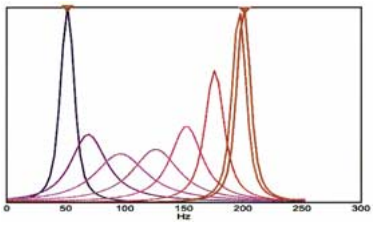

Figure 1. Simulation of line shapes for subsequent steps of a reaction of the type $P + L \rightleftharpoons PL$ using Equation 5 with a subsequent Fourier transformation. Populations (mole fractions) of P were calculated for a dissociation constant K_D of 10^{-6}, a protein concentration of 1 mM and a titration in equal steps until an equimolar amount of ligand was added (population of P: $\mathbf{P}_P = [0.99, 0.83, 0.67, 0.50, 0.34, 0.17, 0.03, 0.01]^T$). The signals of P and PL were separated by 150 Hz and the off-rate was 1000 s^{-1}. Peak centers of P and PL are marked by a red '▽'. The line width of P and PL was 15 Hz. The steps in the titration go from blue to red.

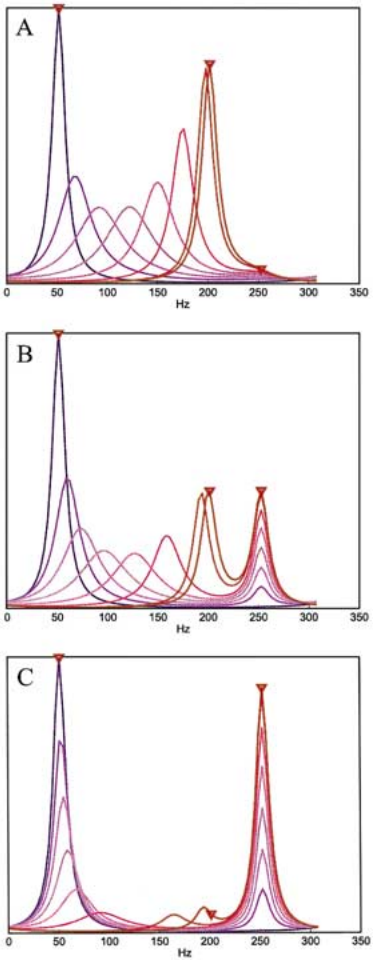

Figure 2. Simulation of line shapes for a two-step mechanism $P + L \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} PL^* \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} PL$. The off-rate for the binding step k_{21} is 1000 s^{-1}, the rate for the exchange between PL and PL* is 10 s^{-1}. The equilibrium constant $K = [PL] / [PL^*]$ is 10 in A, 1 in B and 0.1 in C. Frequency separations were 150 Hz between P and PL* and 50 Hz between PL* and PL, peak centers of P, PL* and PL are marked by a red '▽'. In all three cases the populations of P were $\mathbf{P}_P = [0.99, 0.83, 0.67, 0.50, 0.34, 0.17, 0.03, 0.01]^T$. The population of PL* was calculated by $\mathbf{P}_{PL^*} = (1 - \mathbf{P}_P)(K + 1)$ and $\mathbf{P}_{PL} = 1 - \mathbf{P}_P - \mathbf{P}_{PL^*}$. The line width of P, PL* and PL was 15 Hz.

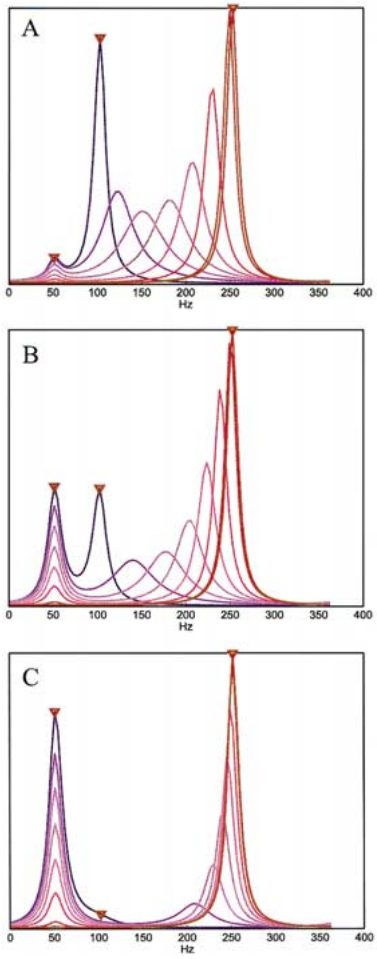

Figure 3. Simulation of line shapes for a two-step mechanism $P \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} P^* + L \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} PL$. The rates k_{12} and k_{21} for the exchange between P and P* were 10 s^{-1}, the off-rate for the actual binding step was 1000 s^{-1}. The equilibrium rate $K = \frac{[P^*]}{[P]}$ was 10.0 in A, 1.0 in B and 0.1 in C. Frequency separations were 50 Hz between P and P * and 150 Hz between P * and PL, peak centers of P, P* and PL are marked by a red '▽'. Populations for PL were $\mathbf{P}_{PL} = [0.001, 0.17, 0.33, 0.50, 0.67, 0.83, 0.97, 0.99]^T$ for PL. Populations of P were calculated by $\mathbf{P}_P = (1 - \mathbf{P}_P) / (K + 1)$ and $\mathbf{P}_{P^*} = 1 - \mathbf{P}_{PL} - \mathbf{P}_P$. The line width of P, P* and PL was 15 Hz.

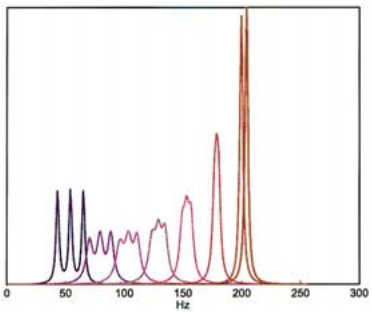

Figure 4. Simulation of a parallel reaction in which three forms of a protein (P_{1a}, P_{1b} and P_{1c}) bind a ligand L leading to the same product PL. Off-rates were 1000 s^{-1} for all three reaction pathways.

As will be shown elsewhere more complex mechanisms may be encountered in protein-ligand interactions (Günther et al., 200; Weyrauch and Günther, 2002). Although line shapes may become too complex for a rigorous simulation it is worthwhile to consider the effect of more complex reactions on NMR line shapes. For this reason we have included the possibility of simulating line shapes for sequential and for parallel reactions in this software. Two types of sequential reactions were included:

$$\mathrm{P} + \mathrm{L} \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} \mathrm{PL}^* \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} \mathrm{PL} \tag{8}$$

and

$$\mathrm{P} \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} \mathrm{P}^* + \mathrm{L} \underset{k_{32}}{\overset{k_{23}}{\rightleftharpoons}} \mathrm{PL}. \tag{9}$$

Here P is always the free protein, PL the protein-ligand complex. Reaction (8) can be regarded as second order binding followed by an exchange step between an intermediate PL^* and the final product PL. The first step of this reaction will cause successive chemical shift changes for increasing concentrations of L for fast exchange between P and PL^*. The second step is determined by an equilibrium constant $K = \frac{[\mathrm{PL}]}{[\mathrm{PL}*]}$ which depends solely on the reaction conditions.

The exchange matrix for reaction (8) is

$$\mathbf{K} = \begin{bmatrix} k'_{12} & -k_{21} & 0 \\ -k'_{21} & k_{12} + k_{23} & -k_{32} \\ 0 & -k_{23} & k_{32} \end{bmatrix}$$

assuming that P cannot be converted to PL directly. The population vector is $\mathbf{P} = [p(\mathrm{P}),\ p(\mathrm{PL}^*),\ p(\mathrm{PL})]$ with $p(\mathrm{P}) + p(\mathrm{PL}^*) + p(\mathrm{PL}) = 1$, $K = k_{23}/k_{32} = p(\mathrm{PL})/p(\mathrm{PL}^*)$. The reaction step is treated the same way as for a simple one-step reaction, i.e., $k_{12} = k_{21}p(\mathrm{PL}^*)/p(\mathrm{P})$. It is convenient to calculate $p(\mathrm{PL}^*)$ from $p(\mathrm{P})$ and $\mathbf{K}$ according to $p(\mathrm{PL}^*) = 1 - p(\mathrm{P})/K + 1$. Thus a series of N line shapes is defined by a population vector $p(\mathrm{P})$ with N elements, k_{off} and the equilibrium constant K. Figure 2 shows typical curves obtained for this mechanism. In all three examples the off-rate for the first step of the reaction was 1000 s^{-1}, for the second step the on-rate was set to 10 s^{-1} and the corresponding off-rate was calculated from the equilibrium constant K which is 10, 1.0 and 0.1 in Figures 2A, 2B and 2C, respectively. The location of the lines of P, PL and PL^* is indicated by a '$\bigtriangledown$'. Figure 2B shows that low intermediate concentrations can cause line shapes which resemble those usually attributed to slow exchange.

A similar **K**-matrix can be derived for reaction (9):

$$\mathbf{K} = \begin{bmatrix} k_{12} & -k_{21} & 0 \\ -k_{21} & k_{12} + k'_{23} & -k_{32} \\ 0 & -k'_{23} & k_{32} \end{bmatrix}.$$

Here the kinetic mechanism is defined by an equivalent set of rules. Again the sum of the populations of all components must be one. An equilibrium constant is defined as $K = k_{12}/k_{21} = p(\mathrm{P})/p(\mathrm{P}^*)$ and $k_{23} = k_{32} \cdot p(\mathrm{PL})/p(\mathrm{P}^*)$. Again a convenient simplification may be applied to reduce the number of adjustable parameters by $p(\mathrm{P}) = 1 - p(\mathrm{PL})/K + 1$. Line shapes for this reaction type are displayed in Figure 3. Here, the equilibrium constant K between free protein and the intermediate P* was again varied between 0.1 and 10.0. The rate for the exchange between P and P^* is 10 s^{-1}, the off-rate for the actual binding step is 1000 s^{-1}. When the equilibrium constant $K_{\mathrm{eq}} = \mathrm{P}^*/\mathrm{P}$ is large we observe binding curves originating from P^* where P^* is in an equilibrium with P and thus its intensity is slightly reduced. Figure 3b shows the case where equal amounts of P and P^* are present in all steps of the reaction and exchange between P and P^* is slow. Because only P^* binds the ligand L, the titration starts at the chemical shift of P^*. In Figure 3c the equilibrium is shifted towards P, i.e., there is always only a low amount of P^* present. If P^*(rather than P) binds the ligand L, the resulting line shapes again resemble the case of slow off-rates. The relevance of this type of reaction for biomolecules will be discussed elsewhere (Günther et al., 2002).

NMRKIN also contains a parallel mechanism for reactions of the type

$$\begin{aligned} \mathrm{P_{1a}} + \mathrm{L} &\rightleftharpoons \\ \mathrm{P_{1b}} + \mathrm{L} &\rightleftharpoons \mathrm{PL} \\ \mathrm{P_{1c}} + \mathrm{L} &\rightleftharpoons \end{aligned}$$

In this case $p(\mathrm{P_{1a}}) + p(\mathrm{P_{1b}}) + p(\mathrm{P_{1c}}) + p(\mathrm{PL}) = 1$ and the exchange matrix is

$$\mathbf{K} = \begin{bmatrix} k_{\mathrm{a12}} & & & -k_{\mathrm{a21}} \\ & k_{\mathrm{b12}} & & -k_{\mathrm{b21}} \\ & & k_{\mathrm{c12}} & -k_{\mathrm{c21}} \\ -k_{\mathrm{a12}} & -k_{\mathrm{b12}} & -k_{\mathrm{c12}} & k_{\mathrm{a21}} + k_{\mathrm{b21}} + k_{\mathrm{c21}} \end{bmatrix}.$$

It has been assumed that $\mathrm{P_{1a}}$, $\mathrm{P_{1b}}$ and $\mathrm{P_{1c}}$ cannot exchange with each other. The result of a simulation for $p(\mathrm{P_{1a}}) = p(\mathrm{P_{1b}}) = p(\mathrm{P_{1c}})$ is shown in Figure 4.

Implementation of line shape calculation in NMRKIN

The matrix notation used in this work was directly implemented within the MATLAB language. Our routines use vectorized programming as far as possible. However, the calculation of the matrix exponential must be performed for each point on the time axis. When the program is started a setup file is requested which determines all important parameters and determines the mechanism to be used. This includes a selection for one or two-dimensional data fitting. Our routines can be used for theoretical simulations or to simulate experimental data. Using a non-linear least squares optimization algorithm (Dennis, 1977) line shapes can be fitted within a few seconds on a standard personal computer. The error of the least-squares fit depends on the type of line shapes and the exchange regime. The best results are obtained for systems in fast exchange with visible intermediate line broadening. When lines are not broadened in intermediate steps because exchange processes are very fast, only a lower limit can be determined from NMR data.

Input data may be a MATLAB matrix or in ASCII format and can readily be generated from NMRLab or any other NMR processing software capable of writing ASCII files. Parameter adjustment can be performed manually or by a least squares routine available in the MATLAB optimization toolbox.

The MATLAB routines used in NMRKIN are listed in Table 1. The calculation is started using a master routine `fitftex` which calls a number of subroutines specified in the parameter file. Several sample parameter files including those used to simulate curves in this work are provided. The parameter file determines the mechanism used and contains the kinetic parameters. The mechanisms that have been implemented are listed in Table 1. For `m21` which is an implementation of a reaction with second order binding

of the type $P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL$ automatic data fitting of experimental line shapes is available. To avoid negative values for populations and rates in least squares optimizations, squares of populations and rates are used as penalty functions when negative values arise. Least squares optimizations converge only when reasonable starting parameters are available. Manual adjustment of parameters is often advisable, especially when peak shapes or the baseline are not ideal.

Additional kinetic mechanisms can be easily implemented following the examples provided in NMRKIN. The implementation of new kinetic models is very straightforward. The required routines are internally split in two sections: An inner section with the **K**-matrix and an outer section implementing the specific kinetic relationships for a given mechanism.

The NMRKIN suite of MATLAB routines along with sample data and fitting results are available on the NMRLab WEB page (www.nmrlab.net).

Results and discussion

The simulation method was applied to data for the interaction of the N-terminal SH2 domain (N-SH2) of the p85 regulatory subunit of PI3-kinase with a peptide (EEEpYMPME-NH_2) derived from polyoma virus middle T antigen. The interaction of the two proteins is known to be important for tumor formation. We have previously analyzed the interactions between the MT N-SH2 interaction using genetics (Yoakim et al., 1992, 1994) and NMR (Günther et al., 1996, 2001; Weber et al., 2000).

Figure 5 shows the ^{1}H cross-sections of signals of the K346 residue of the N-SH2 in a titration with MT (dotted line) together with simulated line shapes calculates as described by Equation 5 with subsequent Fourier transformation (straight line). Optimal values for the populations **P** of the protein in each of the 7 steps of the titration and the off-rate were obtained by a least squares optimization. The contribution of the exchange broadening in the ^{15}N dimension on the ^{1}H-signal was considered during the simulation although its contribution was small because the ^{15}N chemical shift change was only 20 Hz. The final off-

Figure 6. A and B: Experimental signals (lines with ◊ from blue to red) and simulated line shapes (solid line) for the S339 residue in a p85 N-SH2 - MT titration (A: ^{1}H and B: ^{15}N). A: Cross-sections for the ^{1}H-dimension of an HSQC-experiment, B: cross-sections for the ^{15}N-dimension of the experiment. Simulations were performed and optimal kinetic parameters were determined as described in the caption of Figure 5. A and B: Simulation parameters were: $k_{off} = 1000\ s^{-1}$, $K_{ex} = \frac{P}{P^*} = 0.89$. C: Simulation of the ^{1}H cross-section as a pure one-dimensional signal without considering the effect of line broadening on the attached ^{15}N-nucleus ($k_{off} = 250$ Hz). Simulations were carried out as described in Figure 5. D: Simulation of the ^{1}H cross-section as a pure one-dimensional signal without considering the effect of line broadening on the attached ^{15}N-nucleus using $k_{off} = 250$ Hz. In A–D peak centers of P and PL are marked by a red '▽'.

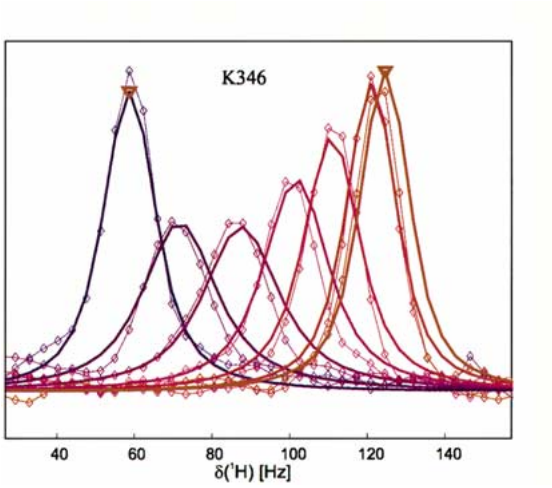

Figure 5.

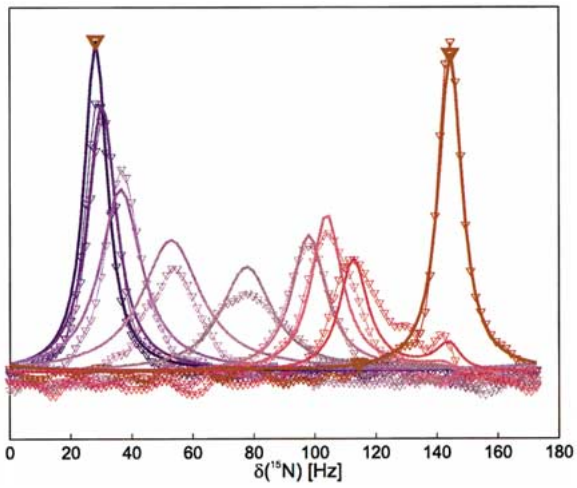

Figure 7.

Figure 5. ^{1}H cross-section of the signal of residue K346 of the p85 N-SH2 in a titration with MT peptide (EEEpYMPME-NH_2) (lines with ◇ from blue to red) and simulated lines (straight lines). Experimental lines were derived from a HSQC spectrum recorded on an AMX500 spectrometer employing mlev decoupling (Bax and Davis, 1985) during acquisition in the fast dimension and decoupling in the incremented dimension by applying a ^{15}N 180° pulse in the middle of the evolution period. Water suppression was achieved employing a Messerle pulse (Messerle et al., 1989) and otherwise as reported in (Günther et al., 1996). Optimal simulation parameters were obtained by fitting signals employing Equation 5 with a subsequent Fourier transformation for an exchange matrix for a two-state exchange model (6) to the experimental signal. The effect of the second dimension on one-dimensional cross-sections was considered by scaling intensities as described in the text. The signal was automatically fitted employing a least squares minimization optimizing the population **P** for each of the 7 spectra and the off-rate. The calculated off rate was 700 s^{-1}. Peak centers of P and PL are marked by a red '▽'.

Figure 7. Cross-sections in the ^{15}N-dimension of HSQC spectra for residue I381 in a titration of the P395S mutant of the p85 N-SH2 with MT. Experimental signals (lines with ◇ from blue to red) and simulated line shapes (solid line) assuming a mechanism as described by Equation 8 with $k_{off} = k_{21} = 1000\ s^{-1}$ for the first step of the reaction and $k_{off} = 10\ s^{-1}$ for k_{32}; $p(P) = [0.99, 0.97, 0.88, 0.7, 0.3, 0.12, 0.08, 0.01, 0.01, 0.01]^T$, $p(PL^*) = [0.01, 0.03, 0.11, 0.3, 0.4, 0.5, 0.56, 0.41, 0.01, 0.01]^T$. Peak centers of P and PL are marked by a red '▽'.

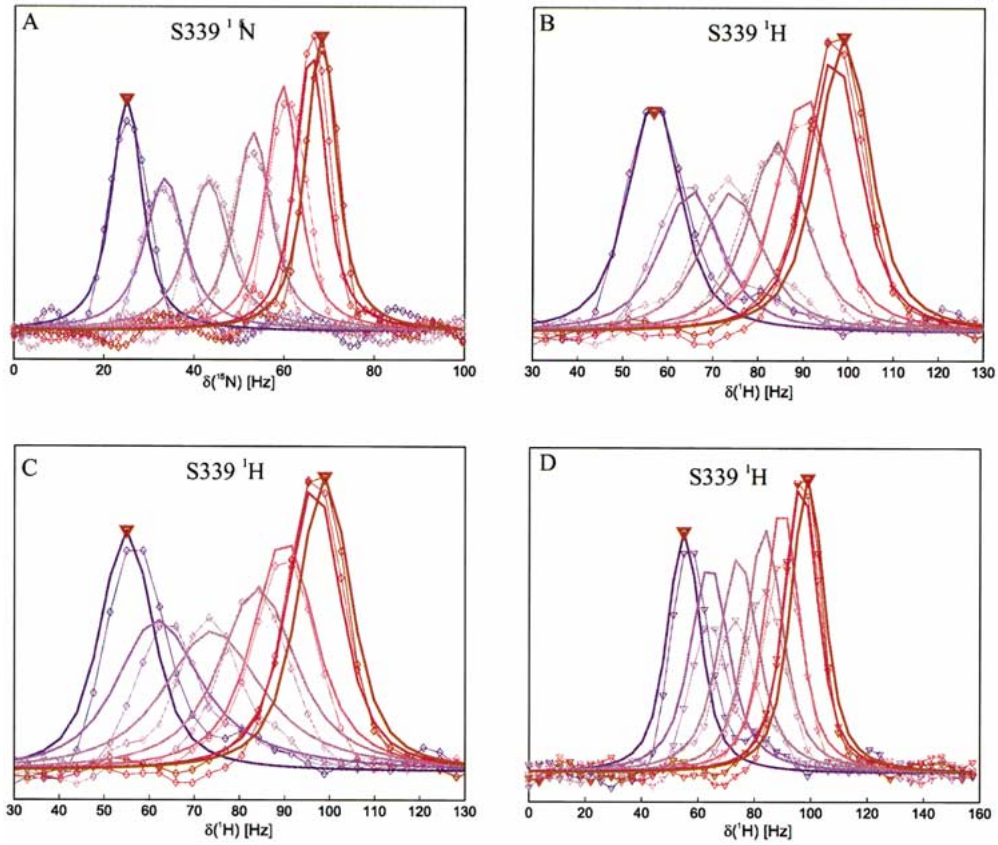

Table 1. Routines in NMRKIN

nmrkin.m	Calling routine
mklspar.m	Read and check parameters
testpar.m	Sample parameter file
redat.m	Read experimental data
m11.m	Mechanism file for two-site exchange (A $\rightleftharpoons$ B)
m21.m	Mechanism file for A+X $\rightleftharpoons$ B
m2111.m	Mechanism file for A+X $\rightleftharpoons$ A* $\rightleftharpoons$B
m1121.m	Mechanism file for A $\rightleftharpoons$ A*+X $\rightleftharpoons$B
m321.m	Mechanism file for parallel reaction
wdwf.m	Calculate window function[a]

[a]Apodization routine from NMRLab.

rate was $\sim$700 s^{-1}, corresponding to a life time of $\sim$1.4 ms. The unsymmetric course of the series of line shapes and the observed broadening of the line shapes depending on the concentration of the ligand was well represented by the simulation. The error of such simulations depends not only on the quality of the experimental lines but also on the exchange regime. For very fast exchange or close to the coalescence point the error will be much higher because line shapes become relatively insensitive for exchange rates.

Figure 6 shows another example, S339 of the N-SH2 upon titration with the same MT peptide, where the chemical shift changed in both dimensions of the HSQC spectrum. Cross-sections for each dimension were simulated. The least-squares optimization was allowed to alter k_{off} and the population of one reaction component for each of the 7 lines. The simulation also required the consideration of an external exchange process for the free protein ($K_{ex} = 0.89$) because the intensity of the first line was smaller than the intensity of the last line of the spectral series. The final off-rate for each dimension was 1000 s^{-1}.

Simulation without the consideration of the second dimension yielded erroneous line shapes and incorrect kinetic constants (Figures 6C and 6D). When the ^{1}H line shapes of S339 were simulated without considering the second dimension the best fit that could be obtained had a k_{off} value of 250 s^{-1} (Figure 6C). A simulation at a k_{off} of 1000 s^{-1} (Figure 6D) was obviously inadequate.

Figure 7 depicts a more complex type of line shape observed for the titration of residue I381 of the P395S mutant (Yoakim et al., 1994) of the p85 N-SH2. The experimental line shapes ($\Diamond$) represent a cross-sections in the ^{15}N dimension; the proton dimension was not affected in this titration. The data can be simulated using the mechanism of Equation 8. The simulation reflects the principal features of these line shapes reasonably well including the late peak just above 140 Hz, but the kinetic mechanism underlying these line shapes is clearly even more complex. Signals towards the end of the titration show multiple shoulders. In addition, the ratio of populations of PL* and PL is not constant. Complex line shapes in titrations of proteins with small ligands will be described in detail elsewhere (Günther et al., 2001; Weyrauch and Günther, 2001).

In summary we provide simple routines that can be used to simulate one-dimensional line shapes derived from two-dimensional NMR spectra and that allow the derivation of off rates. Using line shapes from two-dimensional NMR spectra of ^{15}N-labeled proteins has an inherent advantage that the effect of binding on two adjacent nuclei can be used.

A fully satisfactory simulation of experimental line shapes is only possible for the simple case of a one-step reaction. However, simulations of line shapes for more complex mechanisms may still be useful for a qualitative discussion of possible mechanisms for protein-ligand interactions.

Acknowledgements

We thank H. Rüterjans for kindly giving U.G access to his laboratory. This work was supported by a grant of the Deutsche Forschungsgemeinschaft, by the NMR large scale facility Frankfurt, by a NATO collaborative research grant and by grants from the N.I.H. to B.S.

References

Anderson, P. (1954) *J. Phys. Soc. Jpn.*, **9**, 316–339.
Balbach, J., Forge, V., Lau, W., van Nuland, N., Brew, K. and Dobson, C. (1996) *Science*, **274**, 1161–1163.
Bax, A. and Davis, D. (1985) *J. Magn. Reson.*, 65355–65360.
Binsch, G. (1968) *Top. Stereochem.*, **3**, 97–191.
Bloch, F. (1946) *Phys. Rev.*, **70**, 460–474.
Burton, R., Busby, R. and Oas, T. (1998) *J. Biomol. NMR*, **11**, 335–360.
Craven, C., Whitehead, B., Jones, S., Thulin, E., Blackburn, M. and Waltho, J. (1996) *Biochemistry*, **35**, 10287–10299.
Dennis, J. (1977) In *State of the Art in Numerical Analysis*, Jacobs, D. (Ed.), Academic Press, New York, pp. 269–312.
Fejzo, J., Lepre, C., Peng, J., Bemis, G., Ajay, Murcko, M. and Moore, J. (1999) *Chem. Biol.*, **6**, 755–769.
Günther, U., Liu, Y., Sanford, D., Bachovchin, W. and Schaffhausen, B. (1996) *Biochemistry*, **35**, 15570–15581.
Günther, U., Liu, Y. and Schaffhausen, B. (2002) submitted.

Günther, U., Ludwig, C. and Ruterjans, H. (2000) *J. Magn. Reson.*, **145**, 101–108.
Gutowsky, H., McCall, D. and Slichter, C. (1953) *J. Chem. Phys.*, **21**, 279–292.
Helgstrand, M., Hard, T. and Allard, P. (2000) *J. Biomol. NMR*, **18**, 49–63.
Hensmann, M., Booker, G.W., Panayotou, G., Boyd, J., Linacre, J., Waterfield, M. and Campbell, I.D. (1994) *Prot. Sci.*, **3**, 1020–1030.
Hoyt, D., Harkins, R., Debanne, M., O'Connor-McCourt, M. and Sykes, B. (1994) *Biochemistry*, **33**, 15283–15292.
Huang, G. and Oas, T. (1995) *Proc. Nat. Acad. Sci. USA*, **92**, 6878–6882.
Johnson, P., Creagh, L., Brun, E., Joe, K., Tomme, P., Haynes, C. and McIntosh, L. (1998) *Biochemistry*, **37**, 12772–12781.
Lian, L.-Y. and Roberts, J. (1993) In *NMR of Macromolecules. A Practical Approach*, Oxford University Press, Oxford, pp. 153–182.
McConnell, H. (1958) *J. Chem. Phys.*, **28**, 430–431.
Messerle, B.A., Wider, G., Otting, G., Weber, C. and Wŵhrich, K. (1989) *J. Magn. Reson.*, **85**, 608–613.
Moler, C.B. and Loan, C.F.V. (1979) *SIAM Rev.*, **20**, 801–836.
Rao, B. (1989) *Meth. Enzymol.*, **176**, 279–311.
Sandstrom, J. (1982) *Dynamic NMR Spectroscopy*, Academic Press, New York.
Sudmeier, J.L., Evelhoch, J.L. and Jonsson, N. (1980) *J. Magn. Reson.*, **40**, 377–390.
Weber, T., Schaffhausen, B., Liu, Y. and Günther, U. (2000) *Biochemistry*, **39**, 15860–15869.
Weyrauch, B. and Günther, U. (2002) In preparation.
Yoakim, M., Hou, W., Liu, Y., Carpenter, C.L., Kapeller, R. and Schaffilausen, B. (1992) *J. Virol.*, **66**, 5485–5491.
Yoakim, M., Hou, W., Songyang, Z., Liu, Y., Cantley, L. and Schaffhausen, B. (1994) *Mol. Cell. Biol.*, **14**, 5929–5938.