Journal of Magnetic Resonance **155,** 106–118 (2002)
doi:10.1006/jmre.2001.2499, available online at http://www.idealibrary.com on IDEAL®

Complete Relaxation and Conformational Exchange Matrix (CORCEMA) Analysis of Intermolecular Saturation Transfer Effects in Reversibly Forming Ligand–Receptor Complexes

V. Jayalakshmi and N. Rama Krishna[1]

Department of Biochemistry and Molecular Genetics, NMR Core Facility, Comprehensive Cancer Center, University of Alabama at Birmingham, Birmingham, Alabama 35294-2041

E-mail: nrk@uab.edu

Received September 26, 2001; revised December 4, 2001; published online February 8, 2002

A couple of recent applications of intermolecular NOE (INOE) experiments as applied to biomolecular systems involve the (i) saturation transfer difference NMR (STD-NMR) method and (ii) the intermolecular cross-saturation NMR (ICS-NMR) experiment. STD-NMR is a promising tool for rapid screening of a large library of compounds to identify bioactive ligands binding to a target protein. Additionally, it is also useful in mapping the binding epitopes presented by a bioactive ligand to its target protein. In this latter application, the STD-NMR technique is essentially similar to the ICS-NMR experiment, which is used to map protein–protein or protein–nucleic acid contact surfaces in complexes. In this work, we present a complete relaxation and conformational exchange matrix (CORCEMA) theory (H. N. B. Moseley *et al., J. Magn. Reson. B* 108, 243–261 (1995)) applicable for these two closely related experiments. As in our previous work, we show that when exchange is fast on the relaxation rate scale, a simplified CORCEMA theory can be formulated using a generalized average relaxation rate matrix. Its range of validity is established by comparing its predictions with those of the exact CORCEMA theory which is valid for all exchange rates. Using some ideal model systems we have analyzed the factors that influence the ligand proton intensity changes when the resonances from some protons on the receptor protein are saturated. The results show that the intensity changes in the ligand signals in an intermolecular NOE experiment are very much dependent upon: (1) the saturation time, (2) the location of the saturated receptor protons with respect to the ligand protons, (3) the conformation of the ligand–receptor interface, (4) the rotational correlation times for the molecular species, (5) the kinetics of the reversibly forming complex, and (6) the ligand/receptor ratio. As an example of a typical application of the STD-NMR experiment we have also simulated the STD effects for a hypothetical trisaccharide bound to a protein. The CORCEMA theory for INOE and the associated algorithm are useful in a quantitative interpretation of the intensity changes in the ligand in both the STD-NMR and ICS-NMR, provided the identity of the receptor protons experiencing direct RF saturation is known. The formalism presented here is likely to be useful in the design of bioactive ligands to a specific target protein and in the quantitative mapping of binding epitopes and interfaces between molecules in complexes. © 2002 Elsevier Science (USA)

Key Words: **STD-NMR; CORCEMA; relaxation rate matrix; exchange matrix; intermolecular cross-saturation; ligand–protein complexes; intermolecular NOE; epitope mapping; protein contact surfaces; drug design; screening libraries.**

INTRODUCTION

Intermolecular NOE (INOE) experiments in which the NMR signals from one molecule are perturbed and the changes due to intermolecular dipolar cross-relaxation in the signals from a different species of a molecule are observed have been described a long time ago (*1–3*). This intermolecular dipolar interaction could be modulated by either the translational diffusion (*3, 4*), or by rotational diffusion (*2, 4*) in the case of molecular complexes. NOEs between a protein and its solvent molecules (*3, 5*) and the intermolecular transferred NOESY (*6–9A*) are special examples of the INOE experiment.

A couple of recent interesting applications of intermolecular NOE (INOE) experiments to biomolecular systems involve: (i) the saturation transfer difference NMR (STD-NMR) method (*10, 11*) for screening compound libraries and for mapping the binding epitopes of the ligand, and (ii) the intermolecular cross-saturation NMR (ICS-NMR) experiment for mapping the protein–protein contact surface (*12*) or protein–RNA contact surface in a complex (*13*). In both cases, the evolution of proton magnetization is modulated by coupled dipolar relaxation networks of each molecular species in their free state and within the complex, together with a coupled evolution from molecular exchange between free and bound states.

In the "traditional" STD-NMR experiment, one generally works with a solution containing an excess of a single ligand or a library of ligands in the presence of a macromolecule (*10, 11*). These measurements work best under weak binding conditions with dissociation constants (K_d) in the range $\sim 10^{-3}$

[1] To whom correspondence should be addressed.

1090-7807/02 $35.00

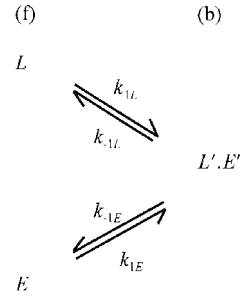
to $\sim 10^{-7}$ M. The NMR spectra of the ligand with and without irradiation of a region within the broad resonance envelope of the macromolecule are collected and a difference spectrum is obtained. In the case of a library of compounds, only those ligands that specifically bind to the target macromolecule will experience intensity changes and show up in the difference NMR spectrum, while the resonances from the remaining nonassociating ligands get canceled out in the difference spectrum. In the 2D-NMR implementation, the 2D-TOCSY spectra are obtained with and without irradiation of the resonances of the macromolecule and a difference is obtained. In the STD-NMR experiment only those protons within the ligand which are at the interface with the protein are generally expected to exhibit significant intensity changes and hence the ligand surface epitopes are easily identified.

In the intermolecular cross-saturation (ICS) experiment (*12*) to identify protein–protein contact interfaces, the protein of interest is uniformly labeled with ^{15}N and ^{2}H. The perdeuteration eliminates all aliphatic hydrogens from this protein, thus enabling a selective saturation of the aliphatic protons on the receptor protein. Perdeuteration also minimizes the deleterious effects on ICS due to leakage relaxation. The [^{1}H–^{15}N] correlation NMR spectra of this labeled protein are then recorded with and without a wide but selective saturation of the aliphatic proton resonances of the unlabeled large protein within the complex. The amide protons from those amino acids within the labeled protein which are at the protein–protein interface will experience greater saturation transfer and hence their correlation peaks will exhibit significantly reduced intensity changes, thus facilitating an easy identification of residues at the interface. The initial slopes of ICS intensities are generally utilized to gain structural information. Deuterium labeling is not essential to map the interface in RNA/protein complexes (*13*). Protein–protein and protein–nucleic acid complexes generally form tight complexes due to large contact surfaces and the associated interactions. Thus, the ICS experiments have been performed usually on such tight complexes. However, as will be apparent from this work, the ICS experiments are likely to work even on reversibly forming weak complexes also.

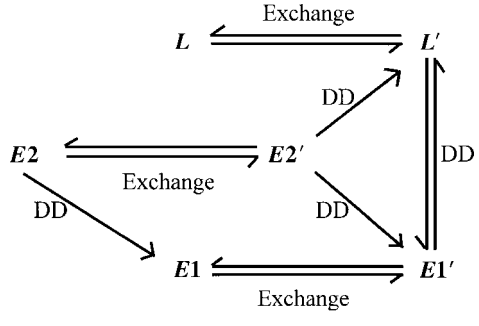
In this work, we present a complete relaxation and conformational exchange matrix (CORCEMA) theory (*6, 7*) that presents a unified formulation for both these essentially identical experiments, although much of the discussion will utilize the STD-NMR experiment as an illustration. The theory is easily applicable to the ICS-NMR experiment by properly adjusting the rotational correlation times for the interacting molecules. We also have identified and analyzed some of the factors that influence the ligand proton intensity changes in the STD-NMR experiments using some hypothetical model systems, and also simulate a typical STD-NMR application involving a hypothetical trisaccharide–protein complex. Whereas the protein–ligand interactions can sometimes influence intraligand transferred NOEs through indirect effects (*6–9, 20*), these interactions are directly responsible for the intermolecular transferred NOEs (*8*) and for the STD (*10, 11*), and ICS effects (*12, 13*).

THEORY

As in our previous work (*6*) we will assume a two-state kinetic model that involves the molecular species in their free (L, E) and bound (L', E') states according to the following scheme (I):



In this scheme, (f) and (b) refer to the free and bound states respectively, k_{1L} and k_{1E} are the on-rates for the ligand (L) and the receptor (E) respectively, and k_{-1L} and k_{-1E} are the corresponding off-rates. We have also assumed, for simplicity, that each molecular species in the above scheme undergoes simple isotropic rotational diffusion. Internal motions (in particular the methyl groups) are incorporated (*7, 8*) by "model-free" calculations for spectral density in terms of order parameters, internal correlation times, and overall correlation times (*14–16*). We also assume that the receptor protons can be subdivided into two classes, ***E*1** (and ***E*1′**) and ***E*2** (and ***E*2′**), where ***E*2** and ***E*2′** stand for protons that experience RF irradiation directly (e.g., aromatic ring protons, or some select CH_3 protons from some specified residues in the case of an STD-NMR, or aliphatic protons in the ICS-NMR). The primes indicate protons within the complex. ***E*1** and ***E*1′** stand for all the remaining protons in the receptor in its free and bound states respectively; some of these may experience considerable saturation indirectly through spin diffusion due to their spatial proximity to the irradiated ***E*2** and ***E*2′** protons or only a weak saturation if the spin diffusion is not efficient. Thus, in our model, the saturation originates from ***E*2** and ***E*2′** and spreads to other protons according to the following scheme (II) (DD stands for exchange through dipole–dipole relaxation):



Under these conditions, the expression for the observable magnetization in an INOE experiment in which the $E2$ and $E2'$ protons are *instantaneously* saturated by RF irradiation is given by

$$\mathbf{I}(t) = \mathbf{I}_0 + [\mathbf{1} - \exp\{-(\mathbf{R}+\mathbf{K})t\}](\mathbf{R}+\mathbf{K})^{-1}\mathbf{Q}, \qquad [1]$$

where t is the time period for which the protons remain saturated. This expression is a solution of the nonhomogeneous set of differentiatial equations obtained from standard equations of motions for the magnetizations coupled by dipolar and chemical exchange processes after setting the magnetizations of saturated receptor protons (***E2*** and ***E2′***) to zero. In this paper, all quantities in bold letters (such as $\mathbf{I}$, $\mathbf{R}$, $\mathbf{I}_{E1}$, $\mathbf{k}_{1L}$, $\mathbf{Q}$, etc.) stand for groups of protons and are represented by matrices in the equations. Even though in this work the expressions are given for the analysis of STD-1DNMR measurements, they are easily adapted for 2D-NMR measurements as well. In the above equation,

$$\mathbf{I}(t) = \begin{vmatrix} \mathbf{I}_{\mathrm{f}} \\ \mathbf{I}_{\mathrm{b}} \end{vmatrix} = \begin{vmatrix} \mathbf{I}_L \\ \mathbf{I}_{E1} \\ \mathbf{I}_{L'} \\ \mathbf{I}_{E1'} \end{vmatrix}, \qquad [2]$$

where $\mathbf{I}_{\mathrm{f}}$ and $\mathbf{I}_{\mathrm{b}}$ are column matrices of intensities of the protons from L and $E1$ in their free and bound states, respectively. $\mathbf{I}_L$, $\mathbf{I}_{E1}$, $\mathbf{I}_{L'}$, and $\mathbf{I}_{E1'}$ are also column matrices representing the magnetizations of protons in ***L***, ***E1***, ***L′***, and ***E1′*** in Scheme II above. $\mathbf{I}_0$ is the corresponding thermal equilibrium matrix, where the intensities are simply proportional to the concentrations of the respective species, multiplied by column vectors of appropriate dimensions containing ones. $\mathbf{1}$ is a column vector consisting of 1's as elements. The term $\mathbf{R}+\mathbf{K}$ is the dynamic matrix (*6, 17*). $\mathbf{R}$ is the generalized relaxation rate matrix composed of rate matrices for the free ($\mathbf{R}_{\mathrm{f}}$) and bound ($\mathbf{R}_{\mathrm{b}}$) states and defined by Eq. [9] of Ref. (*6*) (with ***E*** and ***E′*** replaced by ***E1*** and ***E1′***, respectively). The generalized kinetic matrix $\mathbf{K}$ is defined by Eq. [10] of Ref. (*6*).

The elements $\mathbf{k}_{1L}$, $\mathbf{k}_{-1L}$, $\mathbf{k}_{1E}$, and $\mathbf{k}_{-1E}$ of the $\mathbf{K}$ matrix are defined as $\mathbf{k}_{1L} = k_{\mathrm{on}}[E]\,\mathbf{1}_L$, $\mathbf{k}_{-1L} = k_{\mathrm{off}}\mathbf{1}_L$; $\mathbf{k}_{1E} = k_{\mathrm{on}}[L]\,\mathbf{1}_{E1}$, and $\mathbf{k}_{-1E} = k_{\mathrm{off}}\mathbf{1}_{E1}$. $\mathbf{1}_L$ and $\mathbf{1}_{E1}$ are unity square matrices of dimensions indicated by the respective subscripts (*6*). k_{on} and k_{off} are respectively the on- and off-rate constants. The matrix $\mathbf{Q}$ is defined as

$$\mathbf{Q} = \begin{vmatrix} \mathbf{Q}_{\mathrm{f}} \\ \mathbf{Q}_{\mathrm{b}} \end{vmatrix} = \begin{vmatrix} \mathbf{0} \\ \mathbf{R}_{E1E2}\mathbf{I}_{E20} \\ \mathbf{R}_{L'E2'}\mathbf{I}_{E2'0} \\ \mathbf{R}_{E1'E2'}\mathbf{I}_{E2'0} \end{vmatrix}, \qquad [3]$$

where $\mathbf{Q}_{\mathrm{f}}$ and $\mathbf{Q}_{\mathrm{b}}$ are submatrices corresponding to the free and bound states, respectively. Note that $\mathbf{R}_{E1E2}$, $\mathbf{R}_{L'E2'}$, $\mathbf{R}_{L'E1'}$, etc., can be rectangular matrices in general if the number of hydrogens in each species (included in the calculation) is different.

For steady state conditions (i.e., long times), the solution of Eq. [1] is given by

$$\mathbf{I}(\infty) = \mathbf{I}_0 + (\mathbf{R}+\mathbf{K})^{-1}\mathbf{Q}. \qquad [4]$$

When the exchange is fast on the relaxation rate scale, it can be shown using the matrix algebra procedures for multistate kinetics of biomolecules developed by us previously (*6, 7, 19*) that an approximate solution to Eq. [1] is obtained (assuming fast exchange on the chemical shift scale),

$$\begin{vmatrix} \mathbf{I}_L + \mathbf{I}_{L'} \\ \mathbf{I}_{E1} + \mathbf{I}_{E1'} \end{vmatrix} = \begin{vmatrix} \mathbf{I}_{L0} + \mathbf{I}_{L'0} \\ \mathbf{I}_{E10} + \mathbf{I}_{E1'0} \end{vmatrix} + [\mathbf{1} - \exp(-\langle\mathbf{R}\rangle t)]\langle\mathbf{R}\rangle^{-1} \times \begin{vmatrix} \mathbf{R}_{L'E2'}\mathbf{I}_{E2'0} \\ \mathbf{R}_{E1E2}\mathbf{I}_{E20} + \mathbf{R}_{E1'E2'}\mathbf{I}_{E2'0} \end{vmatrix}, \qquad [5]$$

where $\langle\mathbf{R}\rangle$ is the generalized average relaxation rate matrix and is expressed formally as a sum of the products of the relaxation rate matrix and the corresponding population matrix for the free and the bound states in Scheme I. It is expressed as (*6, 7, 20*)

$$\langle\mathbf{R}\rangle = \mathbf{R}_{\mathrm{f}}\mathbf{P}_{\mathrm{f}} + \mathbf{R}_{\mathrm{b}}\mathbf{P}_{\mathrm{b}} = \begin{vmatrix} p_L\mathbf{R}_L + p_{L'}\mathbf{R}_{L'} & p_{E'}\mathbf{R}_{L'E1'} \\ p_{L'}\mathbf{R}_{E1'L'} & p_E\mathbf{R}_{E1} + p_{E'}\mathbf{R}_{E1'} \end{vmatrix}. \qquad [6]$$

The relaxation rate matrices $\mathbf{R}_{\mathrm{f}}$ and $\mathbf{R}_{\mathrm{b}}$ and population matrices $\mathbf{P}_{\mathrm{f}}$ and $\mathbf{P}_{\mathrm{b}}$ for the free and bound states, respectively, in Scheme I are given by Eqs. [9] and [16] in Moseley *et al.* (*6*). This $\langle\mathbf{R}\rangle$ matrix takes into account the ligand–protein cross-relaxation involving the protons $E1'$ in the complex, while the term $\mathbf{R}_{L'E2'}$ in the $\mathbf{Q}$ matrix takes into account the cross-relaxation with the saturated protons. The quantities p_L (=$[L]/[L_{\mathrm{T}}]$) and $p_{L'}$ (=$[L']/[L_{\mathrm{T}}]$) are normalized fractional populations (*6*) for the ligand in its free and bound states (i.e., $p_L + p_{L'} = 1$). Similarly, p_E (=$[E]/[E_{\mathrm{T}}]$) and $p_{E'}$ (=$[E']/[E_{\mathrm{T}}]$) are the fractional populations for the protein in its free and bound states (with $p_E + p_{E'} = 1$). These populations are determined from a knowledge of L_{T}, E_{T}, and K_{eq} and using Eq. [52] in Ref. (*7*). From Eq. [1] it is seen that for very small times after instantaneous saturation, the *initial* signal is given by

$$\mathbf{I}(t) \approx \mathbf{I}_0 + \mathbf{Q}t - (\mathbf{R}+\mathbf{K})\,\mathbf{Q}t^2 + \mathbf{H}(\mathbf{R}, \mathbf{K}, t), \qquad [7]$$

where $\mathbf{H}$ stands for higher order terms in the expansion. If the exchange is fast on both chemical shift and relaxation rate scales,

a similar expression is obtained from Eq. [5]:

$$\left|\begin{array}{c}(\mathbf{I}_L+\mathbf{I}_{L'})\\(\mathbf{I}_{E1}+\mathbf{I}_{E1'})\end{array}\right| \approx \left|\begin{array}{c}(\mathbf{I}_{L0}+\mathbf{I}_{L'0})\\(\mathbf{I}_{E10}+\mathbf{I}_{E1'0})\end{array}\right| + (1-\langle\mathbf{R}\rangle t)t \times \left|\begin{array}{c}\mathbf{R}_{L'E2'}\mathbf{I}_{E2'0}\\(\mathbf{R}_{E1E2}\mathbf{I}_{E20}+\mathbf{R}_{E1'E2'}\mathbf{I}_{E2'0})\end{array}\right| + \mathbf{H}(\langle\mathbf{R}\rangle, t). \quad [8]$$

Thus, the initial portion of the ligand signal in an STD-NMR experiment (irrespective of the exchange rate) is sensitive to a *direct* transfer of saturation from the saturated protein proton to the ligand proton in the complex, as reflected by the $\mathbf{R}_{L'E2'}\mathbf{I}_{E2'0}$ term. For longer times, the STD-NMR intensity spectrum displays additional effects from indirect saturation transfer from the $\mathbf{E1}'$ set of protons in the bound state, as well as exchange-mediated effects (e.g., from the $\mathbf{R}_{E1E2}\mathbf{I}_{E20}$ term in the $\mathbf{Q}$ matrix).

The dynamic matrix $\mathbf{R}+\mathbf{K}$ is asymmetric, and can be put in a symmetric form for convenient numerical diagonalization purposes using a symmetrization matrix $\mathbf{S}$ (*6*). Thus, if $\mathbf{\Lambda}$ is the diagonalized form of $\mathbf{R}+\mathbf{K}$ matrix,

$$\mathbf{\Lambda} = \mathbf{T}^{-1}\,\mathbf{S}^{-1}(\mathbf{R}+\mathbf{K})\,\mathbf{S}\,\mathbf{T}, \quad [9]$$

where $\mathbf{T}$ is the matrix that diagonalizes the symmetric matrix $\mathbf{S}^{-1}(\mathbf{R}+\mathbf{K})\,\mathbf{S}$. Thus,

$$\mathbf{I}(t) = \mathbf{I}_0 + [\mathbf{1} - \mathbf{ST}\exp\{-\mathbf{\Lambda}t\}\mathbf{T}^{-1}\,\mathbf{S}^{-1}](\mathbf{ST}\,\mathbf{\Lambda}^{-1}\,\mathbf{T}^{-1}\,\mathbf{S}^{-1})\,\mathbf{Q}. \quad [10]$$

From Eq. [1] or [10] it is easy to compute the quantities like $\mathbf{I}_{Lk}$ and $\mathbf{I}_{L'k}$ which are the magnetizations of the k'-th proton of the ligand in its free and bound states respectively. If the exchange is fast on the chemical shift scale, $\{(\mathbf{I}_{Lk}+\mathbf{I}_{L'k}) - (\mathbf{I}_{L0k}+\mathbf{I}_{L'0k})\}$ gives the absolute STD-NMR signal for proton k in the ligand, and $[(\mathbf{I}_{Lk}+\mathbf{I}_{L'k}) - (\mathbf{I}_{L0k}+\mathbf{I}_{L'0k})]/(\mathbf{I}_{L0k}+\mathbf{I}_{L'0k})$ gives the corresponding "fractional" intensity change.

METHODS

Figure 1 shows the CORCEMA-STD protocol that employs Eq. [1] or [10] (a choice is provided in the protocol) to compute the absolute $(I-I_0)$ and the fractional $\{(I-I_0)/I_0\}$ STD-1DNMR intensity spectra. A program employing this protocol using MATLAB (The Mathworks, Inc., Natick, MA) was written for use on a personal computer. Our experience using the examples in this paper shows that MATLAB can directly handle asymmetric matrices of the type $\mathbf{R}+\mathbf{K}$ without a need for symmetrization, and that the results obtained with Eqs. [1] and [10] are identical. The starting point of the program is reading the protein data bank (PDB) coordinates for the free ligand, free protein, and the ligand–protein complex either for existing models from the PDB or from proposed and/or modified models. This program also permits the conformations of the interacting

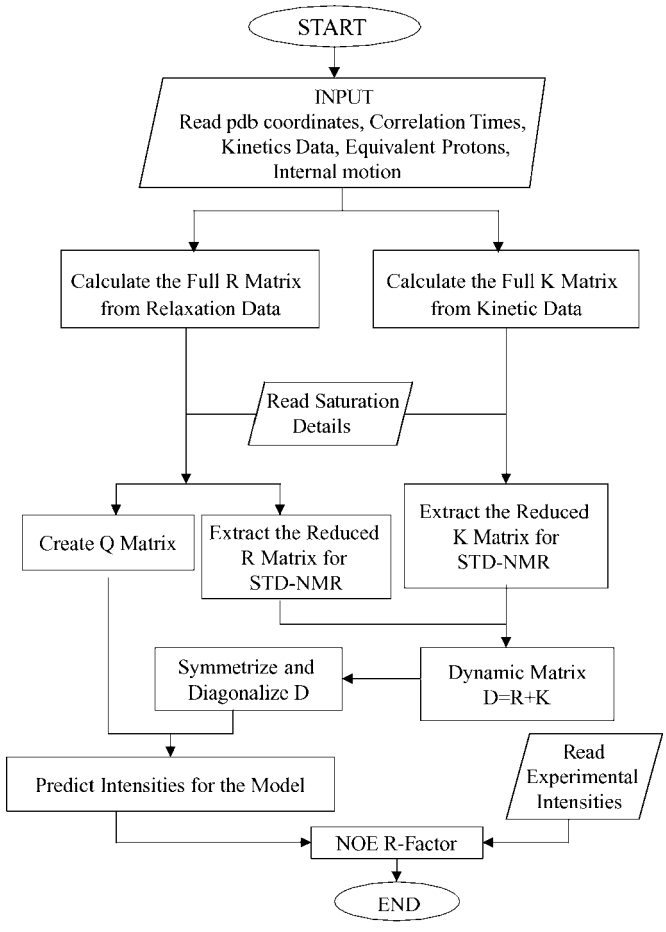


FIG. 1. CORCEMA Protocol for calculating STD-NMR intensities.

molecules to be different, if necessary, in the free and bound states. Other input parameters consist of the total number of the ligand protons (N), and the protein protons (M) near the active site to be included in the CORCEMA calculations, the rotational correlation times, the kinetic data (i.e., equilibrium constant $K_{\rm eq}$ $(=1/K_{\rm d}$, where $K_{\rm d}$ is the dissociation constant) for the complex together with estimates of the $k_{\rm on}$ or $k_{\rm off}$ rates), and an estimate for the leakage relaxation rate of the individual protons in the free and bound states (usually kept at 0.2 to 0.3 s^{-1} to account for nonspecific relaxation due to dissolved paramagnetic oxygen (*8*)). Additionally, in setting up the $\mathbf{R}$ matrix, flags are set to identify situations that require a consideration of internal motions. Specifically, the methyl group protons (from Leu, Val, Thr, Ala, and Ile) and the aromatic ring protons (from Tyr and Phe with two-fold symmetry) are identified. For methyl groups where the internal rotational correlation times can be considerably shorter than the overall rotational correlation time, a "model-free" calculation of spectral densities with order parameters (S^2) and internal ($\tau_{\rm m}$) and overall ($\tau_{\rm r}$) correlation times was used to calculate spectral densities involving intramethyl and methyl-to-other proton dipolar interactions, as in Refs. (*7, 8, 14–16*). For intramethyl proton–proton interactions, S^2 was

set to 0.25. For methyl-to-other proton interactions, the S^2 value was a read-in value, but generally kept in the range of 0.6 to 0.9 (*8, 15, 21*). For Tyr and Phe, it was assumed that the ring-flip rate was much longer than the rotational correlation time; a simple $\langle 1/r^6 \rangle$ average was used for the interaction of external protons with the equivalent protons in these two aromatic residues.

The program first sets up sparse matrices for **R** and **K**. From a knowledge of the identity of the saturated receptor protons, the **Q** matrix consisting of cross-relaxation elements involving $\boldsymbol{E2}'$ (and $\boldsymbol{E2}$) and the reduced **R** and **K** matrices corresponding to $\boldsymbol{L}$, $\boldsymbol{L}'$, $\boldsymbol{E1}$, and $\boldsymbol{E1}'$ protons are created. The dimensions of the reduced **R** and **K** matrices (in Eq. [1]) are $2(N + M - J)$, where J is the number of $\boldsymbol{E2}$ protons. The dimension of the **Q** matrix is $2(N + M - J) \times 1$. To speed up the computations considerably on a PC, the program also gives an option of using an arbitrary cutoff distance (r_{cutoff}, usually 10 Å) in computing the **R** matrix; i.e., if the distance between any pair of intra- or intermolecular hydrogens is $\geq$10 Å, the corresponding element in the relaxation rate matrix is set automatically to 0 without further calculation. The concentrations of the species are calculated using standard equations (e.g., see the equations on p. 254 of Ref. (*6*)). The elements of $\mathbf{I}_0$ are set to be identical to the concentrations of the molecular species. In the final step, the "predicted" STD changes are calculated and compared with the experimental values using the NOE R-Factor (*18, 22*).

We first did some simulations using two hypothetical models consisting of 6 ligand protons and 5 protein protons with the configurations shown in Figs. 2A (the symmetric model) and 2B (the asymmetric model). For simulating a typical experimental situation, we also took the PDB coordinates (ID# 2 kmb) for the complex of sLex/MBP (mannose-binding protein) (*23*). From this data set, we retained three sugars (Gal, MAG (α-methyl-*N*-acetyl-D-glucosamine), and Fuc), and 8 residues (D184, E185, N187, H189, G190, N205, D206, and I207) in the MBP binding pocket. Figure 3 shows the stereo picture of the model used in the simulations. To simulate results of STD-NMR measurements usually performed in D_2O, all exchangeable hydrogens (OH and NH) were excluded in the calculation. Additionally, for the sake of simplicity, we have also assumed that the conformations of the trisaccharide and the protein in the free and bound forms in Scheme I are identical. However, as mentioned at the beginning of this section, the program can easily accommodate situations where the conformations are different. Unless otherwise specified, a uniform leakage factor of 0.3 s^{-1} was assumed for all the protons in their free and bound states. This term is added to all the diagonal elements of the **R** matrix. The k_{on} rate was assumed to be 10^9 s^{-1} M^{-1}. The spectrometer frequency was set at 600 MHz. In all the simulations a ligand/protein ratio of 10 : 1 was assumed. A somewhat longer correlation time of 10^{-7} s was assumed for the protein and the complex, primarily to ensure

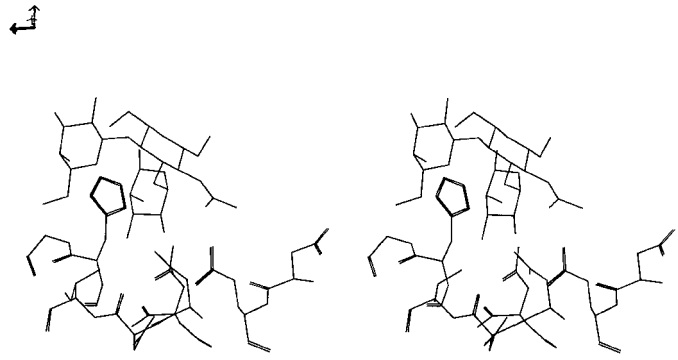


FIG. 3. Stereo view of a hypothetical trisaccharide, Gal–MAG–Fuc bound to the mannose-binding protein (MBP). The hydrogens were omitted for clarity. The model was generated by modifying the PDB file 2kmb for sLex/MBP (*23*).

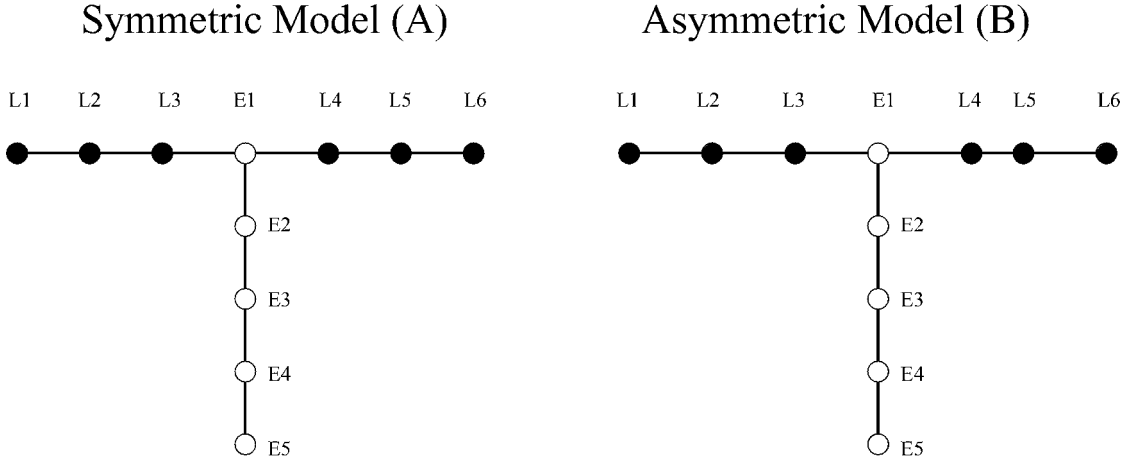


FIG. 2. (A) Symmetric model of ligand–protein complex. The filled circles represent the ligand protons while the open circles represent the protein protons. All the distances between the nearest-neighbor protons in the complex are assumed to be 3 Å. For the free ligand, the $L3$–$L4$ distance is 6 Å. (B) The asymmetric model, with the $L4$–$L5$ distance set at 1.8 Å. The saturated protein protons $E2$ ($E2'$) or $E5$ ($E5'$) correspond to $\boldsymbol{E2}$ ($\boldsymbol{E2}'$) in Scheme II. The remaining protons are represented by $\boldsymbol{E1}$ ($\boldsymbol{E1}'$) in Scheme II.

efficient spin diffusion. The ligand correlation time was fixed at 0.2966 ns corresponding to null NOE at 600 MHz.

(1) Experiments with Finite Delays

Equation [1] was derived assuming that the nuclear spin magnetizations are at thermal equilibrium values prior to the start of the presaturation. In practice, due to time constraints on the instrument, this condition may not be usually realized and the nuclear spin magnetization can be generally in a quasi-equilibrium state prior to presaturation. If $(t_d + t)$ is the delay between two consequtive 90° observe pulses, where t is the presaturation period and t_d is the time delay before presaturation (this includes the data acquisition time for the previous pulse), then the appropriate expressions for STD and for control NMR spectra are given by

$$\mathbf{I}(t) = \mathbf{I}_0 + [\mathbf{1} - \exp\{-(\mathbf{R} + \mathbf{K})t\}](\mathbf{R} + \mathbf{K})^{-1}\mathbf{Q} + \exp\{-(\mathbf{R} + \mathbf{K})t\}[\mathbf{I}(0)_r - \mathbf{I}_0], \quad [11]$$

where

$$\mathbf{I}(0)_r = \{[\mathbf{1} - \exp\{-(\mathbf{R}_F + \mathbf{K}_F)t_d\}]\mathbf{I}_{0F}\}_r$$

and

$$\mathbf{I}^{\text{control}} = \{[\mathbf{1} - \exp\{-(\mathbf{R}_F + \mathbf{K}_F)(t_d + t)\}]\mathbf{I}_{0F}\}_r. \quad [12]$$

In the above expressions, the subscript F refers to the full relaxation and exchange matrices that include the $E2$ and $E2'$ protons since their magnetizations do not experience RF saturation during t_d in the STD and during $t_d + t$ in the control experiment, and hence experience coupled recovery during these periods. The subscript "r" refers to the reduced matrix containing elements for $\boldsymbol{L}$, $\boldsymbol{L}'$, $\boldsymbol{E1}$, and $\boldsymbol{E1}'$ extracted from the full matrix. We have implemented the above expressions as an option in the CORCEMA—STD Program.

(2) Noninstantaneous Saturation

Even though the assumption of *instantaneous* saturation of the irradiated proton signal is generally easily met under standard experimental conditions by adjusting the bandwidth and strength of the irradiating RF field, a provision is also provided in the protocol to treat in a phenomenological way those situations where the saturated proton(s) does not experience instantaneous saturation, but experiences saturation while undergoing Torrey oscillations (*24, 25*).

RESULTS AND DISCUSSION

(1) Effect of Saturation Time

Even if a given set of protein proton signals ($\boldsymbol{E2}$ and $\boldsymbol{E2}'$ protons) are instantaneouly saturated, the saturation will take a finite time to spread to other protein protons ($\boldsymbol{E1}$ and $\boldsymbol{E1}'$ sets of protons) and the bound ligand protons ($\boldsymbol{L}'$) through dipolar networks, and through chemical exchange from the bound to the free ligand ($\boldsymbol{L}$) protons. Figures 4A and 4B show the ligand STD-NMR fractional intensity changes plotted as a function of saturation time (in Eq. [1]) for the symmetric and asymmetric

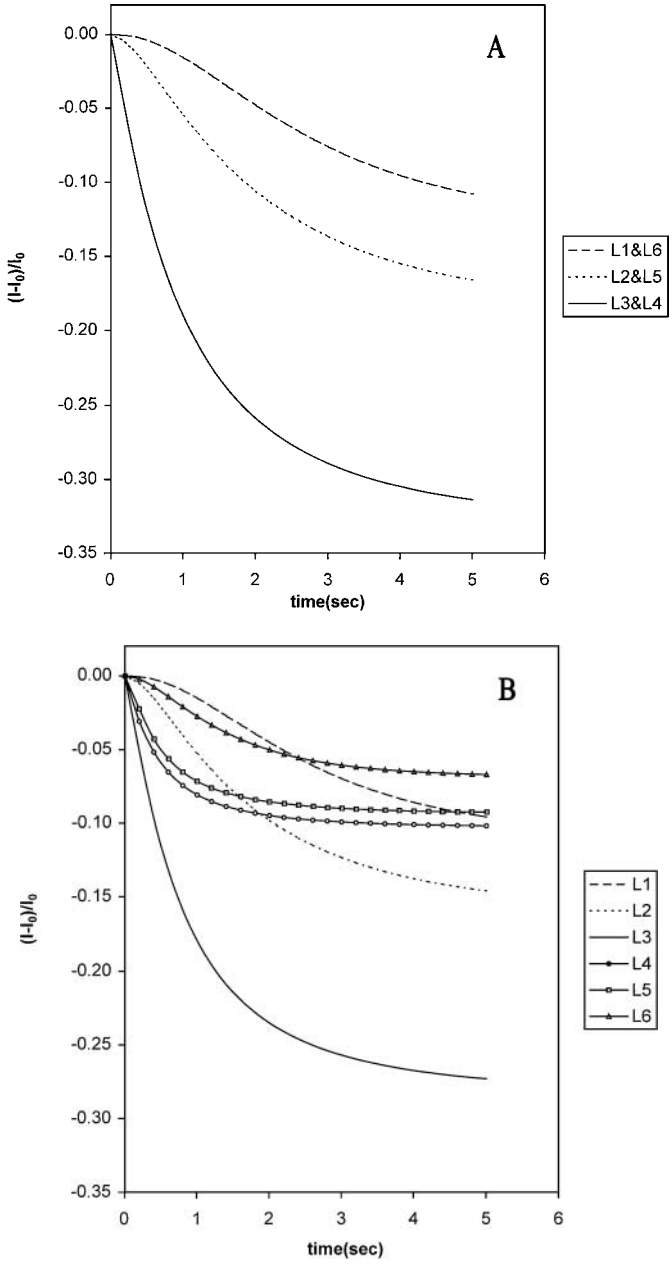


FIG. 4. (A) STD curves for the symmetric ligand–protein complex in Fig. 2A. (B) STD curves for the asymmetric ligand–protein complex in Fig. 2B. The $E2'$ and $E2$ protrons are saturated. A spectrometer frequency of 600 MHz and the free ligand correlation time of 2.966×10^{-10} s corresponding to null NOE at 600 MHz were assumed. The protein correlation time was 10^{-7} s. $L_t/E_t = 10:1$. Leakage rate $= 0.3\ s^{-1}$.

ligand/protein complexes, respectively in Fig. 2. The saturation spreads from $E2'$ to $E1'$ in the bound state and from $E2'$ and $E1'$ to the bound ligand protons, $L1'$ to $L6'$. Similarly, there is also a saturation transfer from the free $E1$ protons to the bound ligand protons through an exchange-mediated transfer (Scheme II).

The time-dependent intensity changes in the initial regions (0 to 1 s in Fig. 4) clearly reflect the spatial proximity of the ligand protons to the enzyme protons in the bound state. For example, in Fig. 4A for the symmetric model, it is the $L3$ and $L4$ protons that show the largest initial slopes, as well as the largest STD effects. The farther protons, $L2$ and $L5$ and similarly $L1$ and $L6$, first show a STD-lag behavior for short times, followed by smaller steady state STD intensity changes, reflecting the finite time for spread of saturation predominantly along a straight line in conformity with the assumed linear arrangement of spins assumed in the model (i.e., $L3 \rightarrow L2 \rightarrow L1$ and $L4 \rightarrow L5 \rightarrow L6$).

(2) Effect of Relative Location of Saturated Proton(s) at the Active Site and Exchange-Mediated Leakage

In performing the STD-NMR experiments, it is a common practice to set the irradiating RF field at one specific location generally devoid of ligand proton resonances, such as the $+1$ to -1 ppm range corresponding to the methyl proton resonances, or the 6.5 to 8 ppm range corresponding to the aromatic ring protons. Thus, it is fair to assume that one or more protons ($E2$ and $E2'$) with resonances corresponding to the general vicinity of irradiating RF field experience instantaneous saturation first. Figure 4 shows the STD effects when the $E2'$ and $E2$ protons are saturated, while Fig. 5 shows the STD effects when the $E5'$ and $E5$ protons are saturated. The lag in intensity change for the $L3$ and $L4$ protons in Fig. 5A is due to the finite time the saturation takes to spread from $E5'$ to $E2'$ and $E1'$ protons. The dramatic difference in the magnitudes of the STD effects between Figs. 4 and 5 is self-evident. This is a reflection of the fact that even for long correlation times of 10^{-7} s, with all the protein protons separated by 3 Å each (and a leakage of 0.3 s^{-1}), the saturation due to spin diffusion is not 100% efficient in the complex. This is shown in Fig. 6 where the STD effects for the $E4$ to $E1$ protons (in the asymmetric model, Fig. 2B) are plotted as a function of the time of irradiation of $E5'$ and $E5$. Even though the protein protons reach steady state much faster than the ligand protons, it is evident that the $E1'$ experiences only about 12% saturation, significantly less than 100%. This significantly poor saturation is in large part due to the large *exchange-mediated leakage* experienced by the protons in the binding pocket ($E2'$ and $E1'$) from their strong dipolar relaxation coupling (see Eq. [6]) with the bound ligand protons, which themselves are reversibly and rapidly exchanging into the large pool of free ligand protons that have not yet experienced saturation. If the distances of $L3'$ and $L4'$ from $E2'$ and $E1'$ are large, say 6 Å, the $E1'$ experiences a considerably larger degree of saturation ($\sim$62%). Thus, the commonly made assumption that the selective saturation of a protein resonance leads to a very efficient and complete or near-complete saturation of the protons within the binding pocket through spin diffusion is not always justified for reversible formation of ligand–protein complexes, and some caution is needed in interpreting STD-NMR results based on such an assumption. In addition to this exchange-mediated leakage, in practice, many large proteins are likely to contain domains or regions with varying degrees of proton density, loss of exchangeable hydrogens when dissolved in D_2O, and varying degrees of segmental mobility. These additional factors will also interfere with the efficiency

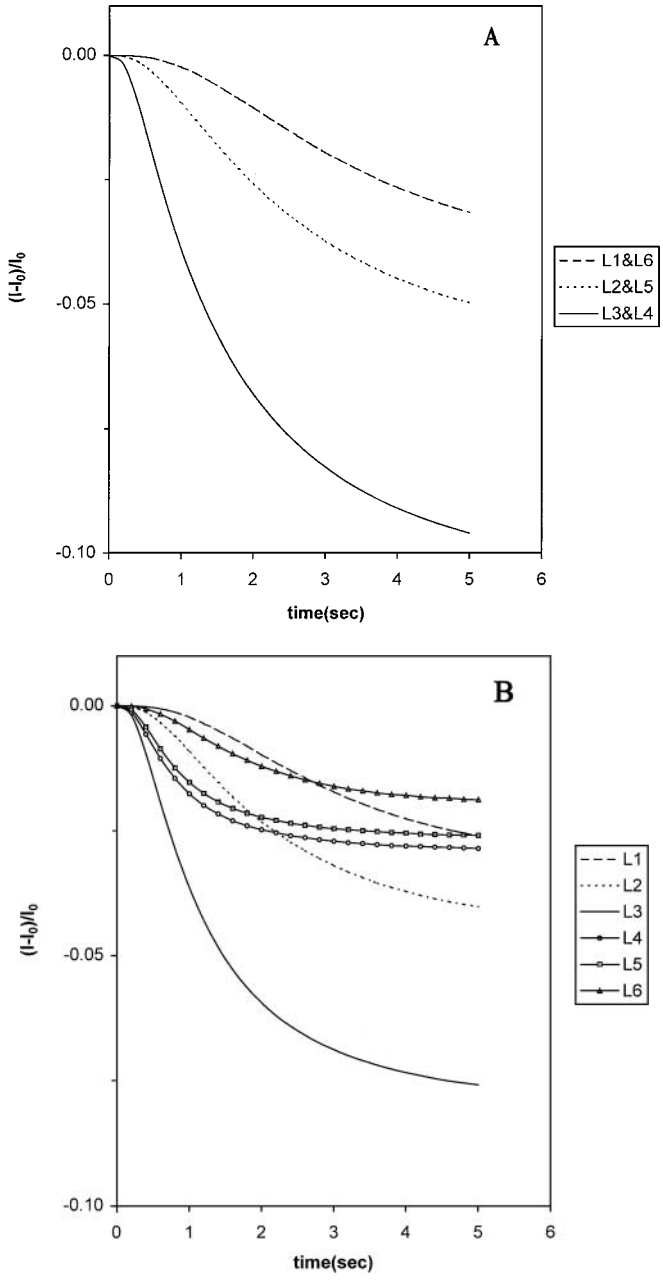


FIG. 5. Same as described in the legend to Fig. 4, but with $E5'$ and $E5$ protons saturated.

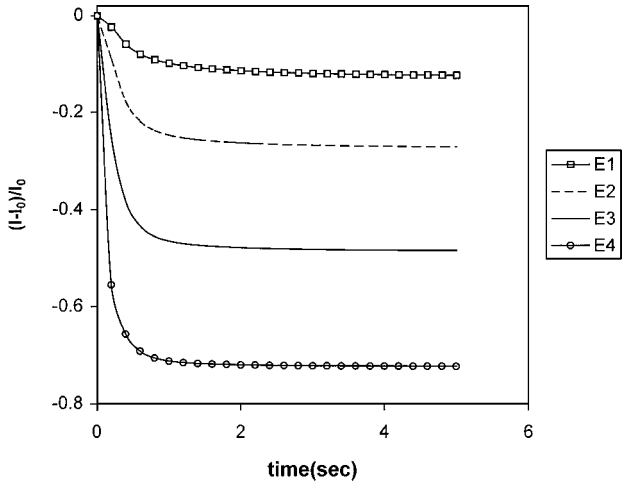


FIG. 6. STD curves for the protein protons $E1$ to $E4$ (free + bound) when the $E5'$ and $E5$ in the asymmetric model are saturated. $K_{eq} = 10^6$ M^{-1}. If the uniform leakage rate is set to zero, the $E1$ ($E1'$) saturation increases to only about 18%. If the $L3$–$E1$ and $L4$–$E1$ distances are increased to 6 Å, the $E1$ and $E1'$ saturation increases to 62% (with leakage rate 0.3) and 83% (with leakage rate 0). The saturation increases to 98% for $L3$–$E1$ and $L4$–$E1$ distances of 10 Å (and 0 leakage).

of spin diffusion and thus the spread of saturation through the entire protein.

(3) Effect of Ligand Conformation

To illustrate the sensitivity of the STD-NMR effects to differences in the bound-ligand conformation, the predicted STD-NMR (fractional changes) results for the two models in Figs. 2A and 2B are shown in Figs. 4A and 4B, respectively. It can be seen that whereas the initial time-dependent regions of the STD curves reflect more or less the spatial proximity of the ligand protons to the protein protons in the complex (with ligand protons at shorter distances showing larger initial slopes for the intensity changes), the steady state values show markedly different results. For example, in Fig. 4B, the $L3$ and $L4$ protons initially (for times less than 1 s) show the largest slopes, reflecting their closeness to the saturated $E2'$ proton (the fact that the slopes are slightly different for $L3$ and $L4$ in the figure suggests that higher order terms are contributing even at 50 ms). The initial slopes for the remaining protons $L1$, $L2$, $L5$, and $L6$ similarly reflect approximately their spatial proximity to the saturated $E2'$ proton of the protein. However, for long times and the steady state (or near steady state) conditions at 5 s, it can be seen that $L3$ and $L4$ have substantially different STD values with $L4$ showing a significantly smaller effect, *even though these two protons are equidistant* from the $E2'$ proton. This is a simple consequence of the differences in the relaxation rates for these $L3$ and $L4$ protons due to differences in their "local environments" (e.g., in the asymmetric model, the $L4$–$L5$ distance is shorter than the $L3$–$L2$ distance). The largest STD effect is shown by the $L3$ proton, followed by the $L2$ proton, even though this proton is farther than the $L4$ proton. These observations suggest that some degree of caution needs to be exercised when one is tempted to do a "qualitative interpretation" of STD effects by ascribing the spatial proximity of the ligand protons to protein protons simply from the magnitudes of the steady state STD effects. A rigorous CORCEMA analysis of STD effects will be needed to make a quantitative interpretation. These conclusions also apply to the ICS-NMR experiments to map the protein–receptor contact surface.

(4) Effect of Free Ligand Correlation Time

Figure 7 shows the calculated STD effects for the asymmetric model assuming 10^{-7} s as the protein correlation time, and with the free ligand correlation times of 10^{-11} and 6×10^{-10} s.

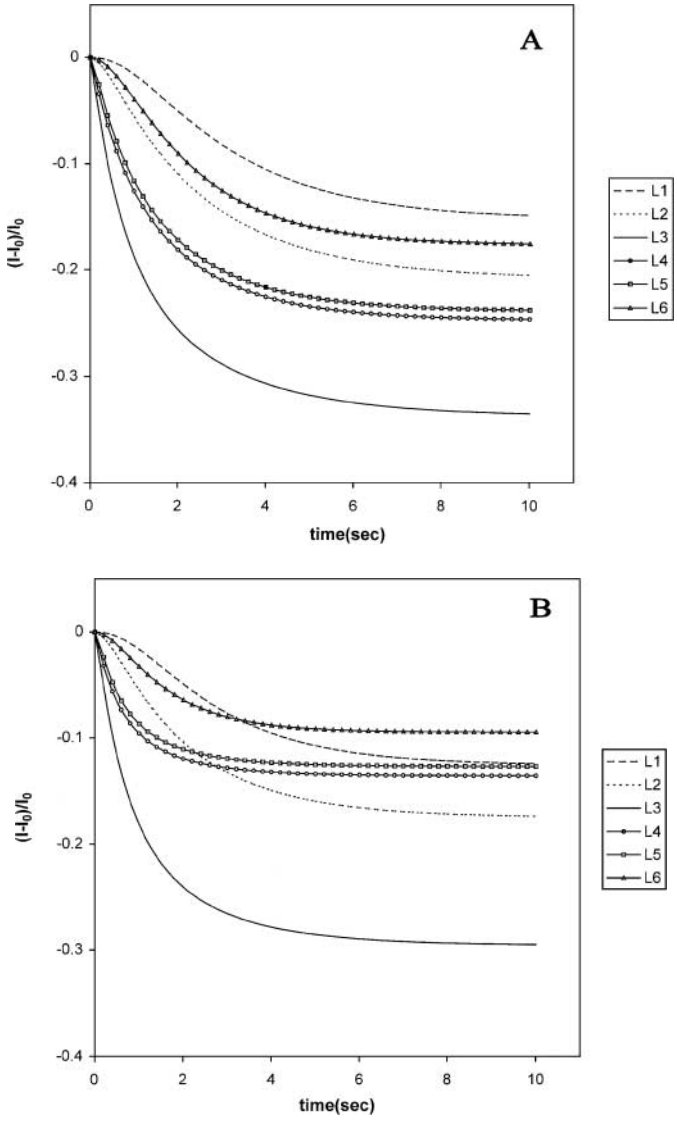


FIG. 7. Effect of free ligand correlation time on STD effects for the asymmetric model. (A) Correlation time = 10^{-11} s. (B) Correlation time = 6×10^{-10} s.

(Figure 4B shows the result for free ligand correlation time of 2.966×10^{-10} s corresponding to null NOE.) It is obvious that even though the initial time-dependent portions (for <200 ms) are identical, the steady state values and the order of these effects are significantly different. This result also underscores the importance of a need for quantitative interpretation of the STD intensities for rigorous structural conclusions.

(5) Effect of K_{eq}

Figures 8A and 8B show the calculated fractional intensity variations for the free ligand and the bound ligand protons, respectively, for the asymmetric model plotted as a function of the equilibrium constant, K_{eq} ($=1/K_d$, where K_d is the dissociation constant) in the range 10^0 to $10^{14}\,M^{-1}$ when the $E5$ (and $E5'$) protons are saturated. Since under normal experimental situations, the exchange is generally fast on the chemical shift scale for $K_{eq} \lesssim 10^7\ M^{-1}$, we have also shown in Fig. 8C the fractional intensity changes in the total ligand signal (i.e., $[(I_L + I'_L) - (I_{L0} + I'_{L0})]/(I_{L0} + I'_{L0})$).

These figures illustrate some interesting properties of the STD-NMR experiment which reflect the rather complex dependence of STD intensities or relaxation, populations, and exchange rates. The range of $K_{eq} \lesssim 10^7$ generally corresponds to

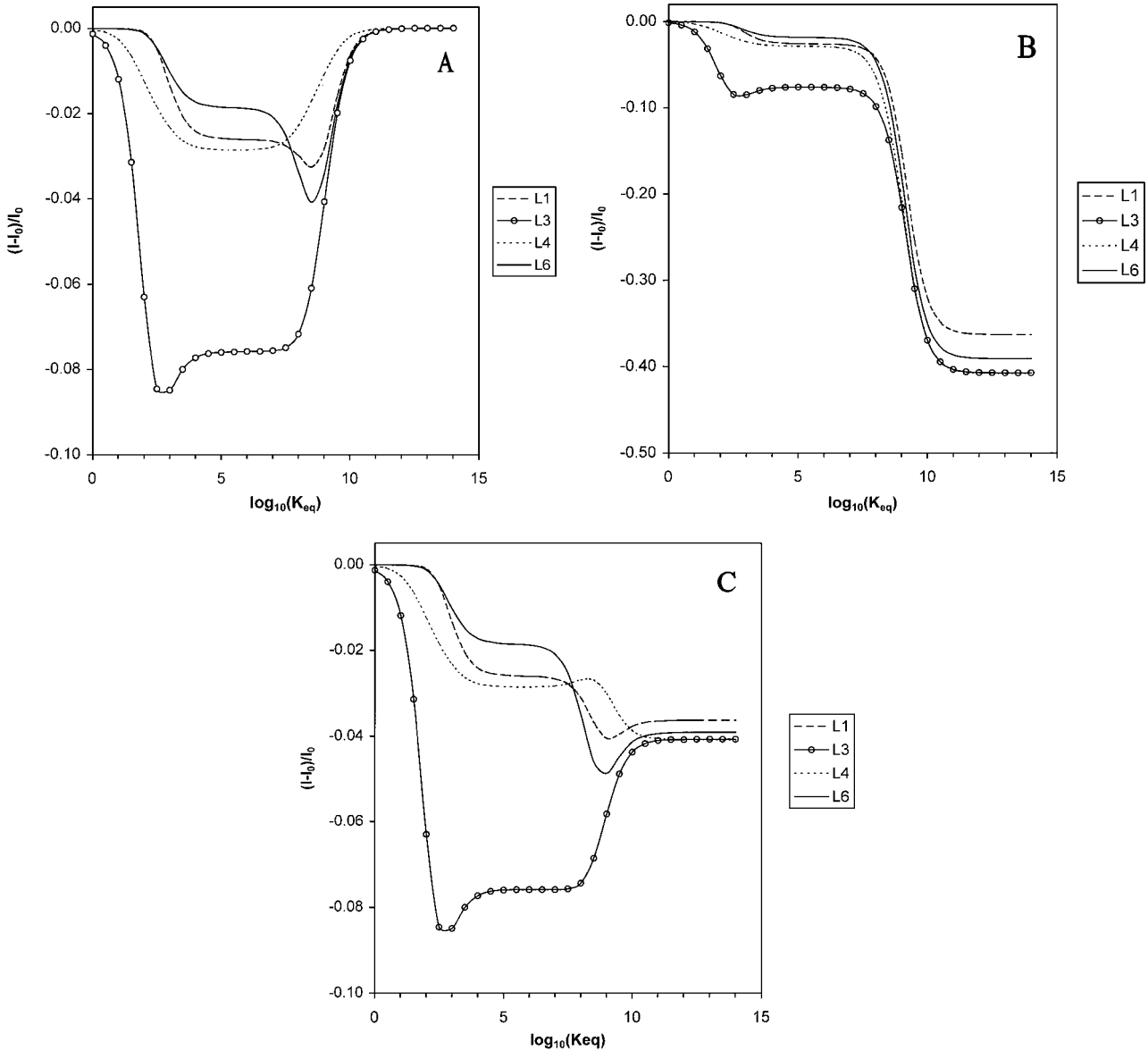


FIG. 8. Dependence of the STD values on K_{eq} for the asymmetric ligand–protein complex in Fig. 2B. A saturation time of 5 s was assumed. (A) Free ligand STD curves; (B) bound ligand STD curves, and (C) STD curves for the total ligand signal (free plus bound). The curves for $L2$ and $L5$ are not shown due to crowding, but are similar to $L4$ in their behavior.

fast exchange on the relaxation rate scale (and generally on the chemical shift scale as well), and hence the STD effects in this range are governed by the generalized average relaxation rate matrix given in Eq. [6], together with variations in the fractional populations of the free and bound forms of the protein and ligand. For the k_{on} ($=10^9$ s^{-1} M^{-1}) value assumed in this model, the populations of the free and bound molecules (ligand and protein) remain more or less constant for $K_{eq} \gtrsim 10^4$ M^{-1}, but undergo rapid variations between K_{eq} values of 4 and 2 (when we assume a 10 : 1 ratio for the ligand/protein ratio). Thus, not surprisingly the STD curves also remain relatively flat in the range of K_{eq} values $\sim 10^4$ to $\sim 10^7$. For $K_{eq} \gtrsim 10^7$ M^{-1}, with the forward and reverse rates k_{1L} and k_{-1L} becoming comparable to the longitudinal relaxation rates, the STD-NMR behavior is no longer described by the average relaxation rate matrix theory, but by the exact theory in Eq. [1].

For K_{eq} values $\sim 10^4$ to $\lesssim 10^7$ in the fast exchange regime, relaxation behaviors of the ligand and protein protons are very strongly coupled to each other as shown by Eq. [6] due to the terms $p_{E'}\mathbf{R}_{L'E1'}$ and the term $p_{L'}\mathbf{R}_{E1'L'}$ (note that $p_{E'} \approx 1$ in this range for large ligand/protein ratios). As a result, the $E1'$ proton experiences less saturation (see Fig. 6). As K_{eq} is further decreased from $\sim 10^4$, the population of bound protein begins to decrease while that of free protein increases. This results in increased saturation of the free $E2$ and $E1$ protons, which is transferred by exchange to $E2'$ and $E1'$ and then to $L3'$. As a result, $L3'$ (and $L3$) show slightly enhanced saturation at $K_{eq} \sim 10^{2.8}$ M^{-1} compared to the plateau STD values for $10^4 \lesssim K_{eq} \lesssim 10^7$ M^{-1}. This is because the small decrease in the fraction of bound ligand population is more than compensated by the larger rotational correlation time and the increased saturation experienced by $E1'$ and $E2'$ protons. For further decrease of K_{eq}, however, the population of bound complex becomes too small to contribute saturation transfer. This results in a diminished STD for $L3'$ and all the other ligand protons. If the $E2'$ ($E2$) proton is saturated, the $E1'$ ($E1$) proton experiences significantly larger saturation, resulting in a large STD value for the plateau region. As a result the dip in STD at $K_{eq} \sim 10^{2.8}$ M^{-1} will be much less apparent.

In Figs. 8A and 8C we observe an interesting dip in STD for the free $L6$ proton as well as the $L1$ proton of the ligand in its free state near $K_{eq} \sim 10^{8.5}$. At first sight, this might look surprising since the $L1$ and $L6$ protons are farthest from the protein protons in the bound state, and yet show the largest STD effects. In the bound state, the $L1'$ and $L6'$ protons show significant saturation transfer at $K_{eq} \sim 10^{8.5}$ (see Fig. 8B), nearly comparable to their respective neighboring protons (i.e., $L2'$ and $L5'$). However, at this K_{eq} value, the life time of the ligand in the free state is relatively long (2.85 s). It is much longer than the longitudinal relaxation times of the $L4$ (0.48 s) and $L5$ (0.46 s) protons, but comparable to that of the $L6$ (2.62 s) and $L1$ (2.64 s) protons. Thus, the larger STD effect observed for $L6$ and $L1$ is simply a reflection of the slower (or smaller) longitudinal relaxation (or recovery from saturation) rates of these two protons in the free state (since each one has only one neighboring proton at 3 Å in this model). In contrast, the $L2$ and $L5$ protons show smaller STD effects since their relaxation rates are faster due to their interaction with two neighboring protons (in particular, the $L5$–$L4$ distance was assumed to be 1.8 Å, resulting in an even faster recovery for the $L5$ proton in its free state). At $K_{eq} \sim 10^{8.5}$ M^{-1} the exchange is generally slow on the chemical shift scale.

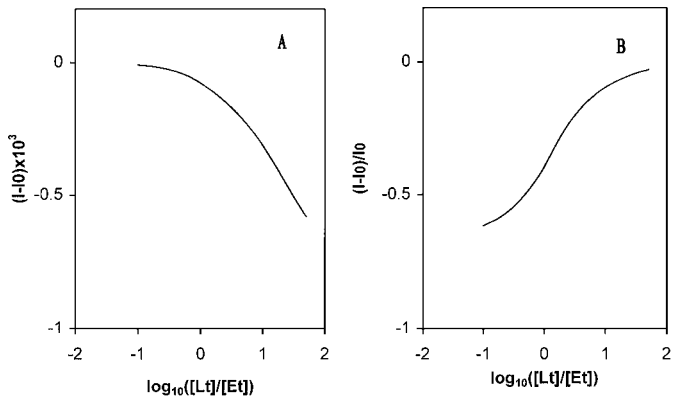


FIG. 9. Effect of ligand/protein ratio on (A) the absolute intensity change (plotted as intensity difference, $I - I_0$), and (B) as a fractional intensity change $(I - I_0)/I_0$ of $L3$ (and $L4$) protons. The symmetric model was used for the calculation. A saturation time of 5 s was assumed. $K_{eq} = 10^6$ M^{-1}.

(6) Effect of L_t/E_t Ratio

In Fig. 9 we have expressed the STD-NMR changes as a function of the L_T/E_T ratio for a couple of hydrogens in the symmetric model in Fig. 2A. In Fig. 9A, the STD is plotted as the absolute intensity change ($[I_L + I'_L] - [I_{L0} + I'_{L0}]$). As intuitively expected, and experimentally observed before (*11*), the absolute intensities increase as the L_T/E_T ratio increases, corresponding to increasing amounts of the concentration of the complex. This will be followed by a plateau when all the protein exists only in a bound state. On the other hand, when the STD is measured as a *fractional change* in intensity (i.e., $([I_L + I_{L'}] - [I_{L0} + I_{L'0}])/[I_{L0} + I_{L'}]$), the STD decreases as the L_T/E_T ratio increases. This is because the free ligand concentration also increases rapidly, thus in effect diluting the amount of STD intensity experienced by the bound ligands, which undergo rapid exchange with the excess of free ligand molecules which have not yet experienced the saturation.

(7) Comparison of Exact and Approximate CORCEMA Theories

Equation [1] (or its equivalent Eq. [10]) gives the calculated intensities in the STD-NMR experiment using an exact CORCEMA treatment applicable over the entire range of exchange rates while Eq. [5] gives the CORCEMA calculated intensities using the "average relaxation rate matrix theory" (*6, 7, 20*) that is valid strictly only when the exchange rates are

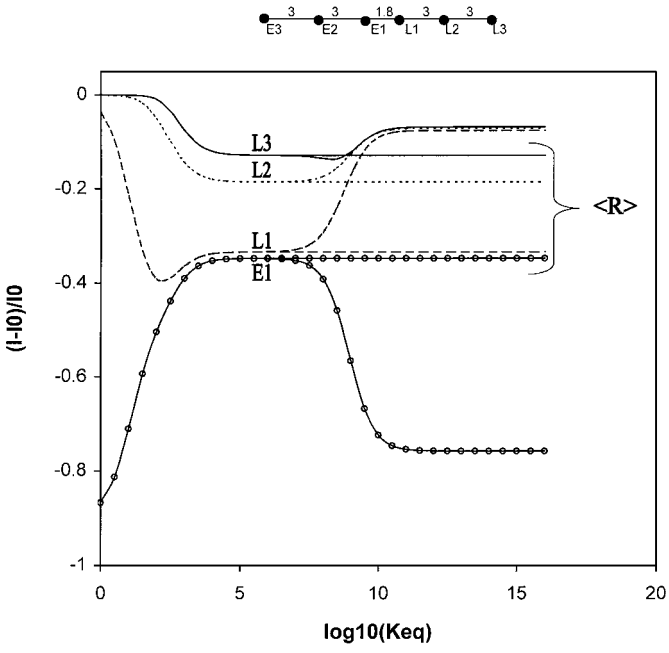


FIG. 10. Comparison of CORCEMA predictions using the generalized average relaxation rate matrix theory (Eq. [5]) with the exact calculation (Eq. [1]). A hypothetical protein–ligand complex with three protons each in a straight line arrangement as shown was assumed. The distances are in ångstroms. $E3$ protons (free and bound) were saturated for 10 s. The correlation times are 10^{-7} s for the complex and the free protein, and 0.2966 ns for the ligand. The STD values for the $L1$, $L2$, and $L3$ and $E1$ (free plus bound ligand signals) are plotted as a function of K_{eq}. The predictions of Eq. [5] are indicated by $\langle \mathbf{R} \rangle$ next to the curves. It is seen that the average relaxation rate matrix theory and the exact calculations give identical results up to $K_{\mathrm{eq}} \lesssim 10^7$.

much larger than the relaxation rates. It is reasonable to ask under what conditions the predictions of Eq. [5] begin to deviate from that of Eq. [1] (or Eq. [10]). We have performed simulations for a hypothetical linear configuration of a protein–ligand complex consisting of three protons in the protein ($E3$, $E2$, and $E1$) and three protons in the ligand ($L1$, $L2$, and $L3$). All the protons were assumed to be in a straight line arrangement given by $E3$–$E2$–$E1$: $L1$–$L2$–$L3$, with the $E1$ proton being closest to $L1$. The $E1$–$L1$ distance was assumed to be 1.8 Å. The distances between the remaining neighboring proton pairs (i.e., $E1$–$E2$, $E2$–$E3$, $L1$–$L2$, and $L2$–$L3$) was assumed to be 3 Å. The $E3$ proton in its free and bound states (i.e., $E3$ and $E3'$) was assumed to be saturated instantaneously. Figure 10 shows the results of calculated STD curves for $L1$, $L2$, and $L3$ and $E1$ protons under the assumption of fast exchange on the chemical shift scale (i.e., the intensities are plotted as a sum of free and bound ligand intensity changes). It is seen that for $K_{\mathrm{eq}} \lesssim 10^7$ (and an assumed value of $k_{\mathrm{on}} = 10^9\ \mathrm{s}^{-1}\,\mathrm{M}^{-1}$), the approximate CORCEMA theory using an "average relaxation rate matrix" (Eq. [5]) and the exact CORCEMA treatment (Eq. [1]) give *identical* results, thus validating this theory. The results begin to diverge for K_{eq} values larger than 10^7 where the fast exchange approximation (compared to relaxation rates) begins to breakdown, thus giving an idea of the range of validity of the average relaxation rate matrix theory.

(8) Trisaccharide/MBP Complex: A Hypothetical Example

Figure 11A shows the calculated STD-NMR spectra as a function of the saturation-on time for the trisaccaharide/MBP model shown in Fig. 3. We have selected some hypothetical but reasonable values for the $k_{\mathrm{on}}(=10^9\ \mathrm{s}^{-1}\,\mathrm{M}^{-1})$, the $[L_{\mathrm{T}}]/[E_{\mathrm{T}}]$ ratio $(=10:1)$, the rotational correlation times for the free ligand $(=2.966 \times 10^{-10}$ s), and for the protein and its complex $(10^{-7}$ s), and a proton frequency of 600 MHz. To account for

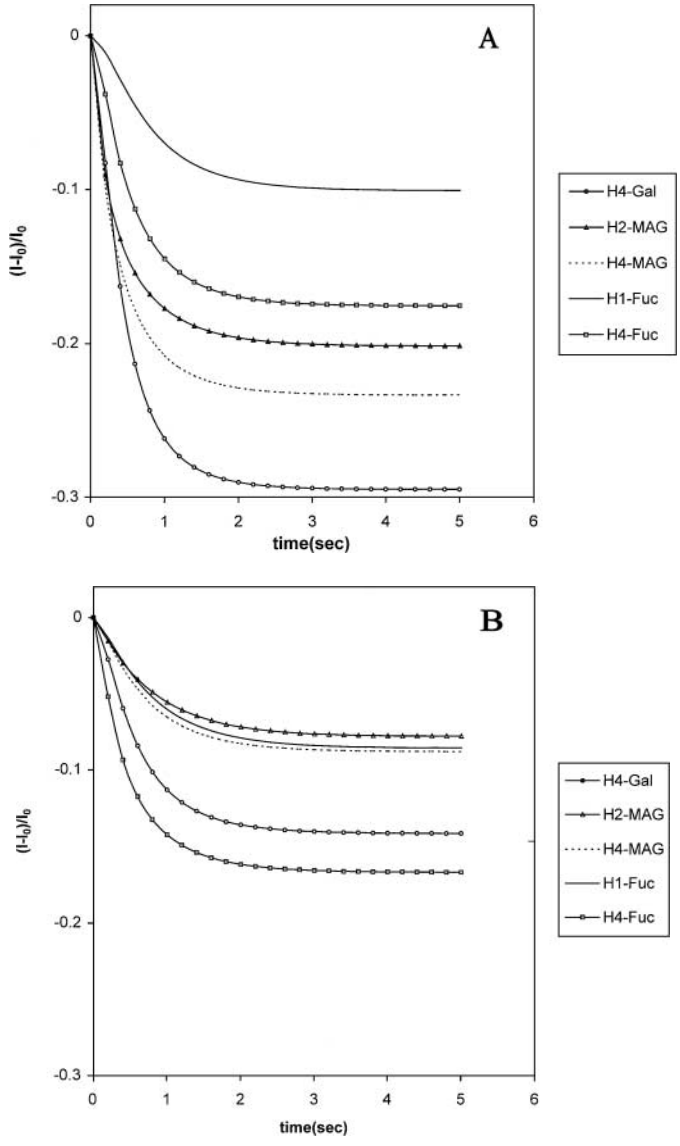


FIG. 11. (A) The calculated STD curves for Gal-H4, MAG-H2, MAG-H4, Fuc-H1, and Fuc-H4 as a function of time when the C_2H and C_4H protons on the His189 ring are saturated. The His189 side chain orientation is the same as in the crystal structure ($\chi_1 = -164.1°$). (B) The His189 side chain torsion angle, χ_1, was changed to 52.1°.

dipolar interactions with methyl groups, S^2 was set at 0.85 and an internal correlation time of 10^{-11} s. We have also assumed that the His189 ring protons were instantaneously saturated. From this figure, it is seen that in general that those protons that are close to the saturated protein protons show the largest steady state STD values, but the order of magnitudes of STD does not necessarily reflect the order of proximity of ligand protons to the His protons in the complex. As pointed out earlier using the asymmetric 6-proton model, the steady state STD values are significantly influenced by the intraligand dipolar relaxation behavior of different proton pairs; i.e., those protons, such as the geminal protons in a ligand which can have a very strong intradipolar relaxation, are likely to show a smaller steady state STD effect even when they are closer to a saturated protein proton than a relatively isolated ligand proton, which is a bit farther. In Fig. 11B, we show the STD-NMR curves predicted if the χ_1 of His has been changed to 52.1° from its value of −164.1° in the crystallographic structure. A comparison of the data in Figs. 11A and 11B shows that the relative order and magnitudes of STD effect in Fig 11B are completely different from those in Fig 11A. These calculations, together with the data in Figs. 4A and 4B, demonstrate the sensitivity of the STD-NMR effects to the conformation of the residues in the active site.

CONCLUSIONS

The STD and ICS experiments have been successfully employed in literature, but their application has remained essentially qualitative. In this work we have explored the possibility of gaining quantitative structural information from these two experiments. We have presented a CORCEMA theory useful for a quantitative analysis of STD-NMR effects in reversibly forming ligand–protein complexes. The theory is continuously applicable over a rather wide range of binding conditions, from very weak complexes (e.g., K_d much larger than 10^{-3} M) to very tight complexes (K_d much less 10^{-14} M). We have also presented an average relaxation rate matrix CORCEMA theory (*6, 7, 20*) to describe STD-NMR effects when the exchange rates are much faster than relaxation rates, and have established its range of validity. Several factors that can influence the STD effects have been considered in detail, and their effects have been illustrated using some hypothetical models of ligand–protein complexes. Simulations were also performed for a more realistic hypothetical model (involving a trisaccharide–protein complex) representative of some of the applications in literature. The sensitivity of STD-NMR to changes in the bound conformation of residues in the binding pocket (both ligand and protein residues) has been demonstrated. The current CORCEMA-STD protocol can be used to predict the STD values for a given ligand-receptor model and a given set of experimental parameters such as binding constant, correlation times, and spectrometer frequency. As such, it can be utilized directly in iterative structure-refinement programs that employ back-calculation of STD spectra as a component, in a manner analogous to iterative structure-refinement protocols that employ a back-calculation of NOESY spectra (e.g., *22, 26*). Thus, CORCEMA theory for STD-NMR is likely to be useful in structure-based drug design efforts and in a quantitative identification of the binding epitopes of bioactive ligands, provided the identity of receptor protons experiencing direct RF saturation is known. It is desirable to have an independent estimate of values of as many parameters as possible, such as the binding constants and rotational correlation times, to minimize the number of variables in an iterative optimization-based protocol. If a reasonably good model for the protein–bound conformation of a ligand is available by independent methods (e.g., through quantitative transferred NOESY), such information will augment efforts at defining the protein–ligand interface using CORCEMA-STD. It should also be useful in a quantitative data interpretation in intermolecular cross-saturation experiments to map protein–protein and protein–nucleic acid contact surfaces in reversibly forming complexes.

ACKNOWLEDGMENTS

This work is supported in part by the NCI grants 1RO1 CA-84177 and P30 CA-13148.

REFERENCES

1. R. Kaiser, Intermolecular nuclear Overhauser effect in liquid solutions, *J. Chem. Phys.* **42,** 1838–1839 (1965).
2. P. Balaram, A. A. Bothner-By, and E. Breslow, Localization of tyrosine at the binding site of neurophysin II by negative nuclear Overhauser effects, *J. Am. Chem. Soc.* **94,** 4017–4018 (1972).
3. N. R. Krishna and S. L. Gordon, Intermolecular nuclear Overhauser studies of coupled spin systems in high resolution NMR, *J. Chem. Phys.* **58,** 5687–5696 (1973).
4. A. Abragam, "The Principles of Nuclear Magnetism," Chap. VIII, Claredon Press, Oxford (1961).
5. G. Otting, Hydration studies of biological macromolecules by intermolecular water-solute NOEs, *in* "Biological Magnetic Resonance," Vol. 17, Structure Computation and Dynamics in Protein NMR," pp. 485–544, Kluwer Academic/Plenum Press, New York (1999).
6. H. N. B. Moseley, E. V. Curto, and N. R. Krishna, Complete relaxation and conformational exchange matrix (CORCEMA) analysis of NOESY spectra of interacting systems: Two-dimensional transferred NOESY, *J. Magn. Reson. B* **108,** 243–261 (1995).
7. N. R. Krishna and H. N. B. Moseley, Complete relaxation and conformational exchange matrix (CORCEMA) analysis of NOESY spectra of reversibly forming ligand-receptor complexes, *in* "Biological Magnetic Resonance, Vol. 17, Structure Computation and Dynamics in Protein NMR," pp. 223–310, Kluwer Academic/Plenum Press, New York (1999).
8. H. N. B. Moseley, W. Lee, C. H. Arrowsmith, and N. R. Krishna, Quantitative determination of conformational, dynamic, and kinetic parameters of a ligand-protein/DNA complex from a complete relaxation and conformational exchange matrix analysis of intermolecular transferred NOESY, *Biochemistry* **36,** 5293–5299 (1997).
9. (*A*) F. Ni, Recent developments in transferred NOE methods, *Prog. NMR Spectrosc.* **26,** 517–606 (1994); (*B*) R. E. London, M. E. Perlman, and D. G. Davis, Relaxation-matrix analysis of the transferred nuclear Overhauser

effect for finite exchange rates, *J. Magn. Reson.* **97,** 79–98 (1992); (*C*) J. Zheng and C. B. Post, Variation in protein indirect relaxation effects in one- and two-dimensional exchange-transferred Overhauser experiments, *J. Phys. Chem.* **100,** 2675–2680 (1996); (*D*) L. Y. Lian, L. Barsukov, M. J. Sutcliffe, K. H. Sze, and G. C. K. Roberts, Protein-ligand interactions: Exchange processes and determination of ligand conformation and protein-ligand contacts, *Methods Enzymol.* **239,** 657–700 (1994).

10. (*A*) M. Mayer and B. Meyer, Characterization of ligand binding by saturation transfer difference NMR spectroscopy, *Agnew. Chem. Int. Ed. Engl.* **38,** 1784–1788 (1999); (*B*) M. Mayer and B. Meyer, Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor, *J. Am. Chem. Soc.* **123,** 6108–6117 (2001).
11. (*A*) M. Vogtherr, and T. Peters, Application of NMR based binding assays to identify key hydroxy groups for intermolecular recognition, *J. Am. Chem. Soc.* **122,** 6093–6099 (2000); (*B*) H. Maaheimo, P. Kosma, L. Brade, H. Brade, and T. Peters, Mapping the binding of synthetic disaccharides representing epitopes of chlamydial lipopolysaccharide to antibodies with NMR, *Biochemistry* **39,** 12,778–12,788 (2000); (*C*) T. Peters, T. Biet, and L. Herfurth, NMR methods for screening the binding of ligands to proteins—Identification and characterization of bioactive ligands, *in* "Biological Magnetic Resonance, Vol. 22, Protein NMR for the Millennium," Kluwer Academic/Plenum Press, New York, in press (2002).
12. H. Takahashi, T. Nakanish, K. Kami, Y. Arata, and I. Shimada, A novel NMR method for determining the interfaces of large protein-protein complexes, *Nat. Struct. Biol.* **7,** 220–223 (2000).
13. A. Ramos, G. Kelly, D. Hollingworth, A. Pastore, and T. Frenkiel, Mapping of the Interfaces of Protein-Nucleic Acid Complexes Using Cross-saturation, *J. Am. Chem. Soc.* **122,** 11,311–11,314 (2000).
14. G. Lipari and A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. I. Theory and range of validity, *J. Am. Chem. Soc.* **104,** 4546–4559 (1982).
15. M. J. Dellwo and A. J. Wand, The influence of methyl rotor dynamics on hydrogen relaxation networks: Derivation of spectral densities in model-free form, *J. Am. Chem. Soc.* **115,** 1886–1893 (1993).
16. J. D. Baleja, R. T. Pon, and B. D. Sykes, Solution structure of phage λ half-operator DNA by use of NMR, restrained molecular dynamics, and NOE-based redefinement, *Biochemistry* **29,** 4828–4839 (1990).
17. R. R. Ernst, G. Bodenhausen, and A. Wokaun, "Principles of Nuclear Magnetic Resonance in One and Two Dimensions," Oxford Science, Oxford (1987).
18. N. R. Krishna, D. G. Agresti, J. D. Glickson, and R. Walter, Solution conformation of peptides by the intramolecular nuclear Overhauser effect experiment, *Biophys. J.* **24,** 791–814 (1978).
19. N. R. Krishna, G. Goldstein, and J. D. Glickson, A general multistate model for the analysis of hydrogen-exchange kinetics, *Biopolymers* **19,** 2003–2020 (1980).
20. P. L. Jackson, H. N. B. Moseley, and N. R. Krishna, Relative effects of protein-mediated and ligand-mediated spin-diffusion pathways on transferred NOESY, and implications on the accuracy of the bound ligand conformation, *J. Magn. Reson. B* **107,** 289–292 (1995).
21. R. Brüschweiler and D. A. Case, Characterization of biomolecular structure and dynamics by NMR cross relaxation, *Prog. NMR Spectrosc.* **26,** 27–58 (1994).
22. Y. Xu, I. P. Sugar, and N. R. Krishna, A variable target intensity-restrained global optimization (VARTIGO) procedure for determining three-dimensional structures of polypeptides from NOESY data: Application to gramicidin-S, *J. Biomol. NMR* **5,** 37–48 (1995).
23. K. K. Ng and W. I. Weis, Structure of a selectin-like mutant of mannose-binding protein complexed with sialylated and sulfated Lewis(x) oligosaccharides, *Biochemistry* **36,** 979–988 (1997).
24. H. C. Torrey, Transient nutations in nuclear magnetic resonance, *Phys. Rev.* **76,** 1059–1068 (1949).
25. G. Wagner and K. Wuthrich, Truncated driven nuclear Overhauser effect (TOE). A new technique for studies of selective ^{1}H-^{1}H Overhauser effects in the presence of spin diffusion, *J. Magn. Reson.* **33,** 675–680 (1979).
26. B. A. Borgias and T. L. James, MARDIGRAS-A procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure, *J. Magn. Reson.* **87,** 475–487 (1990).