SIDECHAIN ROTATIONAL ISOMERIZATION IN PROTEINS

Dynamic Simulation with Solvent Surroundings

INDIRA GHOSH AND J. ANDREW MCCAMMON
Department of Chemistry, University of Houston-University Park, Houston, Texas 77004

ABSTRACT Molecular dynamics simulations are used to study the rotational isomerization of the tyrosine 35 ring in bovine pancreatic trypsin inhibitor immersed in liquid water. Inclusion of the solvent surroundings improves the agreement with experimental results significantly, although the theoretical free energy barrier (13 kcal/mol at 300K) is still ~3 kcal/mol below that found by nuclear magnetic resonance studies. This remaining discrepancy will probably be eliminated in future calculations by the use of a more accurate model for the hydrogen atoms on the tyrosine ring. An important finding in the present work is that frictional effects due to solvent damping appear to be small for the tyrosine 35 ring, which is largely but not completely buried in the protein surface.

INTRODUCTION

Sidechain reorientations are an essential feature of many of the activities of proteins, including ligand binding and catalysis of substrate reactions (1–4). As a relatively simple model of such reorientations and other local activated processes in biopolymers, the rotational isomerization of tyrosine rings in the bovine pancreatic trypsin inhibitor (BPTI) has been the subject of particularly intense experimental and theoretical study (4–14, 33). As discussed below, the rotation of rings that are buried in BPTI is coupled to larger-scale structural fluctuations in the surrounding protein (4, 12, 13, 15). In principle, such collective fluctuations may be altered by changes in the solvent surroundings of the protein and thereby mediate solvent effects on local dynamical processes inside the protein. Two different types of solvent effects that can be imagined are specific effects that depend on solvent composition (e.g., differences in solvent association with protein surface groups that lead to shifts in protein conformation) and viscosity effects that alter the dynamics of the collective fluctuations and thereby the rate of local transitions in the interior (4, 12, 16–20). Here we report the first results from a molecular dynamics study of ring rotations in BPTI in water. The questions that arise in this study arise also in connection with sidechain conformational transitions in other proteins (17, 18, 20), ligand penetration into myoglobin (16, 19, 21), and other cases in which local transitions in the interior of a protein may be coupled to collective motions (4).

The most detailed theoretical results presently available for ring rotation in BPTI pertain to tyrosine 35, which is the most deeply buried and has the smallest rate of ring rotation of the four tyrosine rings in the protein. These results were obtained by the activated trajectory method, in which one uses molecular dynamics simulation methods first to compute the free energy barrier (and thus the ideal transition state theory rate) for an activated process, and then to determine the detailed reaction mechanism and corrections to the ideal rate by analysis of trajectories initiated at the peak of the barrier (4, 10, 12). The results obtained to date have added considerably to our understanding of the nature of local conformational changes in proteins, particularly in showing how such changes can be controlled or "gated" by larger scale fluctuations in the structure of a protein. In the specific case of tyrosine 35, the ring rotation is preceded by displacement of a section of backbone that covers one face of the ring (12).

The theoretical studies described above included a number of simplifying features designed to reduce computational requirements (10). The most important of these was the neglect of explicit solvent surroundings of the protein. The solvent surroundings are expected to influence both the structure and dynamics of the protein surface. Because the section of backbone that gates the ring rotation and a small part of the tyrosine 35 ring itself are exposed at the protein surface, solvent effects are expected to be significant. In the previous work, the charges of the protein atoms were set to zero to reflect the dielectric screening properties of water. This clearly represents an overcorrection, however, and does not allow for such structural influences as solvent hydrogen bonding. In fact, the difference between the calculated and experimental free energy barriers (10 kcal/mol and 16 kcal/mol, respectively) was attributed to distortions of the protein surface arising from the absence of detailed solvation effects (10). Frictional effects due to solvent damping were also absent in the calculations. Finally, only atoms within 7.7 Å of the tyrosine ring centroid (a total of 94 atoms) were allowed to move during these earlier simulations.

In the calculations described here, all of the above approximations have been relaxed. The system now consists of BPTI with partial charges on the atoms, immersed in a detailed model of liquid water. As before, only a finite part of the system is allowed to move in most of the simulations, but this part is five times larger than in the previous work. We show that the agreement with experimental results for the kinetics of rotation is significantly improved by use of the more detailed model. We also show that the explicit effect of solvent damping is small, a result of interest in connection with this particular process and also more generally. In a subsequent paper, we plan to examine the mechanistic details of the transitions and the effects of variations in solvent viscosity.

METHODS

The approach used in this work is largely the same as that described by Northrup et al. (10). We therefore provide only an outline of the overall procedure for the sake of clarity in the subsequent sections, and describe new elements of this procedure in greater detail.

The rate constant k for an activated transition can be written as (4)

$$k = (1/2)\,\kappa < |\dot{\xi}| > \rho(\xi^{\neq}) \,/\, \int \rho(\xi)\, d\xi, \qquad (1)$$

where ξ is a reaction coordinate that describes the progress of the transition from its initial to its final state through a transition state where $\xi = \xi^{\neq}$, κ is a transmission coefficient (discussed below), $< |\dot{\xi}| >$ is the average absolute value of the velocity $d\xi/dt$ evaluated at $\xi^{\neq}$, $\rho(\xi)d\xi$ is the relative frequency of occurrence of configurations with values of the reaction coordinate in the range ξ, $\xi + d\xi$, and the integral is over the range of ξ corresponding to the initial configurations of the system. The ideal transition state theory ($\kappa = 1$) corresponds to the assumption that any trajectory at the transition state that is moving toward the product state will continue in that direction and lead to a stable product. In a dense system such as a protein, crossings of moderate free energy barriers are likely to be interrupted so that $\kappa < 1$. As before, the reaction coordinate used to describe the rotation of the tyrosine 35 ring is $\xi = \chi^2_{35} - \chi_v$, where χ^2_{35} is the familiar dihedral angle $C^{\alpha}_{35} - C^{\beta}_{35} - C^{\gamma}_{35} - C^{\delta 1}_{35}$ and χ_v is the "virtual" dihedral angle $C^{\beta}_{35} - C^{\gamma}_{35} - C^{\delta 2}_{35} - N_{36}$.

The first step in the calculation of k consists of the determination of $\rho(\xi)$ by an umbrella sampling procedure (4, 10). A sequence of molecular dynamics calculations is carried out with "window potentials" added to the usual potential energy function of the system to concentrate the sampling in a sequence of overlapping regions of "windows" of ξ. The data from these simulations are analyzed as described previously to provide $\rho(\xi)$, the corresponding potential of mean force or free energy

$$W(\xi) = -k_B T \ln \rho(\xi), \qquad (2)$$

and the mean potential energy $<V(\xi)>$. The calculation differs from the earlier one primarily in the molecular model that is used, although there are also differences in the simulation and analysis procedures.

The initial structure for the system was obtained by superimposing the atomic coordinates of the 1.5-Å resolution x-ray structure of BPTI (with 60 associated water molecules) (22, 23) and those of a large volume of bulk water in a typical configuration from a dynamic simulation of pure water. Bulk water molecules closer than 2.3 Å to any nonhydrogen atom in the crystal strucutre were deleted, as was one bulk water molecule that was left in a poor location for forming multiple hydrogen bonds in the protein interior. Finally, all water molecules more than 27 Å from C^{γ}_{35} were deleted.

Energy relaxation and molecular dynamics calculations were then carried out using the GROMOS suite of programs (24–26). This includes a model for water molecules (SPC water) that is compatible with the protein model. The SHAKE algorithm was used to constrain all bond lengths and the internal geometries of the water molecules (4, 27). The Verlet leapfrog method was used for the dynamics (4, 28). Nonbonded interactions were cut off at a distance of 8.0 Å. The waters exterior to the protein were partly equilibrated by 100 steps of steepest descent energy refinement followed by alternation of Maxwellian velocity reassignment (at 295K) and dynamical propagation (time step = 2 fs) for 0.2-ps intervals for a total period of 1.0 ps. The dynamics of the external solvent was then continued for 4.5 ps with Maxwellian velocity reassignment at 295K every 0.9 ps. The protein was held fixed to avoid the development of any structural distortions during this partial equilibration of the solvent surroundings.

The final list of mobile atoms was then generated and these atoms were further equilibrated in the field of the remaining fixed atoms (which are either protein atoms in the x-ray structure or water atoms in the partly relaxed solvent). All atoms closer than 10 Å to the C^{γ} of tyrosine 35 were initially placed in the mobile atom list. The list was then expanded by adding the remaining atoms of any residue or water molecule that had one or more atoms in the initial list. The final list comprised 306 protein atoms and 162 water atoms. The next phase of dynamic equilibration of the mobile atoms consisted of Maxwellian velocity reassignments at 295K followed by 1.5-ps intervals of dynamics for a total period of 7.5 ps. The masses of all hydrogen atoms were then increased to 10 amu (to improve sampling efficiency by slowing hydrogen libration during the calculation of equilibrium properties) (29), the time step was increased from 2 to 4.5 fs, and the dynamic equilibration was continued for an additional 4.5 ps but with Maxwellian velocity assignments at 300K. The equilibration was completed by coupling the system to a heat bath at 300K with a relaxation time of 0.1 ps (30) and continuing the dynamics for 10 ps.

At this point the umbrella sampling calculations were carried out to determine $\rho(\xi)$, $W(\xi)$, and $<V(\xi)>$. The procedure is essentially the same as previously except that a single type of window potential was used (harmonic in ξ), rather than separate window potentials in χ^2_{35} and χ_v; the dynamic equilibration for each window was longer than before, 7 ps vs. 3 ps; the simulation for each window was slightly shorter than before, 12 ps vs. 17 ps; and more windows were used than before, 21 windows vs. 9.

The second step in the calculation of k consists of the determination of $\langle|\dot{\xi}|\rangle$ and κ. The former quantity can be obtained easily from the velocities of the atoms at a representative set of occurrences of the transition state structures. The transmission coefficient, however, depends on the history of trajectories that pass through the transition state region. To compute these quantities, a 43-ps simulation was performed with the window potential $U(\xi) = (K/2)(\xi - \xi^{\neq})^2$, $K = 100$ kcal mol^{-1} rad^{-2} and $\xi^{\neq} = 90°$, to generate a representative set of transition state coordinate and velocity sets. This simulation was done with normal hydrogen masses (so that correct dynamic or nonequilibrium results are obtained), and was preceded by a 3-ps equilibration with Maxwellian reassignments at 300K every 0.5 ps, to allow the system to adjust to the window potential and the normal masses. 190 representative coordinate and velocity sets were selected from this simulation. Each set was separated by at least 0.2 ps from the others to ensure dynamical independence and had $89° < \xi < 91°$ to avoid artifacts due to the subsequent removal of the window potential.

The transmission coefficient was calculated by the absorbing boundary method recently introduced by Berne and Straub (31, 32). Each of the 190 coordinate and velocity sets obtained above was dynamically propagated in the absence of the window potential. A number N_B of these trajectories had $\dot{\xi} > 0$ initially; these were terminated when ξ reached 300° (at which point it is assumed that the system has been trapped in the product state), or when ξ returned to 90°. The number $N_B(t)$ of these trajectories that were not absorbed at $\xi = 90°$ by time t was recorded for times up to 1.2 ps. Corresponding numbers N_A and $N_A(t)$ were determined for trajectories that started in the reactant direction and had not returned to $\xi = 90°$ by time t. The transmission coefficient κ was then

obtained from the plateau region of the function

$$\kappa(t) = T_A(t)T_B(t)/[T_A(t) + T_B(t) - T_A(t)T_B(t)] \quad (3)$$

where

$$T_A(t) = N_A(t)/N_A \text{ and } T_B(t) = N_B(t)/N_B.$$

RESULTS AND DISCUSSION

The results for the potential of mean force $W(\xi)$ and the mean potential energy $\langle V(\xi) \rangle$ are presented in Fig. 1. By separate analysis, the data from the first and second halves of the 21 different windows in the umbrella sampling calculation gave results for $W(\xi)$ and $\langle V(\xi) \rangle$ that were generally within 0.5 and 1.5 kcal/mol of those shown, respectively. Thus, the statistical uncertainty in these results appears to be reasonably small.

The potential of mean force and mean potential energy differ somewhat from those obtained in earlier work (10). The free energy barrier is 30% higher than before, ~13 kcal/mol vs. 10 kcal/mol, and the location of the peak has shifted from ~120° to ~90°. As discussed before (10), the internal energy of the system $E(\xi)$ varies with ξ in the same manner as $\langle V(\xi) \rangle$, so that variations in $W(\xi)$ can be decomposed into energy and entropy components as

$$\Delta W(\xi) = \Delta E(\xi) - T\Delta S(\xi) = \Delta \langle V(\xi) \rangle - T\Delta S(\xi). \quad (4)$$

The present results agree with the previous results in showing that the free energy barrier is dominated by the internal energy contribution. There may be a larger entropic contribution to the barrier ($T\Delta S \simeq$ 2–4 kcal/mol vs. ~ 1 kcal/mol) in the present case, but the statistical uncertainty in $\langle V(\xi) \rangle$ limits the reliability of this result.

The results of the absorbing barrier calculation are

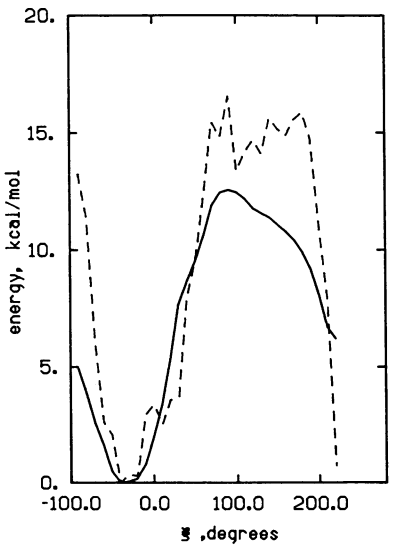

FIGURE 1 Potential of mean force $W(\xi)$ (——) and average potential energy $\langle V(\xi) \rangle$ (---) as functions of the tyrosine 35 ring-rotation reaction coordinate ξ.

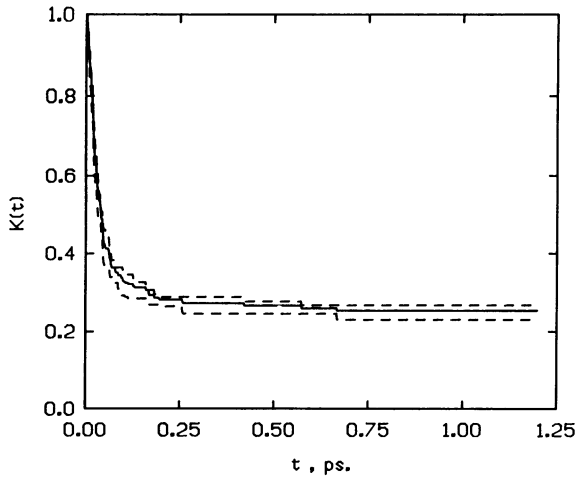

FIGURE 2 Values of the apparent transmission coefficient $\kappa(t)$ as a function of trajectory length t. Results are given for the full set of 190 trajectories (——) and for the first and second half of the set (---).

shown in Fig. 2. The transmission coefficient estimated from the plateau region of the curve is $\kappa = 0.25$. This is remarkably similar to the value obtained in the previous calculation that did not include explicit solvent surroundings, $\kappa = 0.22$. Although the ring rotation occurs within the surface region of the protein, direct solvent viscosity effects appear to be less important than internal friction effects (due to collisions among atoms in the protein) for the solvent model considered here.

The mean crossing velocity obtained in the present work ($\langle|\dot{\xi}|\rangle = 9.7 \times 10^{12}$ rad/s) is essentially identical to that obtained previously (1.0×10^{13} rad/s). Combining this velocity with the results displayed in Figs. 1 and 2, one obtains from Eq. 1 the rate constant for ring rotation at 300K, $k = 540\ s^{-1}$. This result is closer to the experimental rate constant based on nuclear magnetic resonance studies ($6\ s^{-1}$, reference 6; $1–10\ s^{-1}$, reference 33) than the result of the previous calculation that did not include the solvent surroundings explicitly ($8.9 \times 10^4\ s^{-1}$, reference 10), but it is still too large by two orders of magnitude. The insensitivity of the transmission coefficient to the refinement of the model to date suggests that the free energy barrier obtained here is lower than the true barrier by ~3 kcal/mol. This discrepancy could be due to approximations in the potential function used. In particular, hydrogen atoms that cannot participate in hydrogen bonding are represented implicitly in the present model by adjustment of the nonbonded interaction parameters of the heavy atoms to which they are attached (25). Replacement of this "extended atom" description by one in which all hydrogens are represented explicitly would increase the van der Waals radius of the tyrosine ring by ~0.35 Å in the direction from the rotation axis to C^δ or C^ϵ. A somewhat larger barrier would be expected for the all-hydrogen model, because slightly larger distortions of the protein matrix would be required to allow rotation of the ring.

CONCLUDING REMARKS

Improved agreement with the experimental rate constant for tyrosine 35 ring rotational isomerization in BPTI has been obtained by using a more detailed theoretical model that includes solvent surroundings. The primary effect of the improvements in the model has been thermodynamic rather than explicitly dynamic in character. This result is consistent with the recent experimental finding that the chemical nature of a solvent has a larger effect on tyrosine ring rotation rates in cytochrome *c* than does the viscosity of the solvent, at least for viscosities in the range 0.8–2.0 cp (20). This result can perhaps be understood in terms of variations in solvent exposure of protein groups during the collective "gate opening" step that precedes ring rotation (12). The situation for these ring rotation processes appears to be different from that for ligand binding to myoglobin, where experimental studies indicate that ligand displacement transitions in the protein respond to the viscosity of the surrounding solvent (16). We plan to investigate the mechanistic details in the simulations described here and in other simulations with a high viscosity solvent to provide a more complete picture of the physical nature of the solvent effects.

We thank Professor W. van Gunsteren for providing the GROMOS programs, Professor Jan Hermans for a helpful discussion on the potential function, and Dr. Omar Karim for a helpful discussion on transmission coefficient calculations.

Indira Ghosh is a Fulbright Scholar and Robert A. Welch Foundation Postdoctoral Fellow. J. Andrew McCammon is a Camille and Henry Dreyfus Teacher-Scholar. This work was supported in part by grants from the National Science Foundation and the Texas Advanced Technology Research Program.

Received for publication 6 August 1986 and in final form 1 December 1986.

REFERENCES

1. Schulz, G. E., and R. H. Schirmer. 1979. Principles of Protein Structure. Springer-Verlag New York, Inc., New York.
2. Creighton, T. E. 1983. Proteins. The Freeman, Irvington-on-Hudson, NY.
3. Fersht, A. 1985. Enzyme Structure and Mechanism. 2nd ed. The Freeman, Irvington-on-Hudson, NY.
4. McCammon, J. A., and S. C. Harvey. 1987. Dynamics of Proteins and Nucleic Acids. Cambridge, London.
5. Gelin, B. R., and M. Karplus. 1975. Sidechain torsional potentials and motion of amino acids in proteins: bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA.* 72:2002–2006.
6. Wagner, G., A. DeMarco, and K. Wuthrich. 1976. Dynamics of the aromatic amino acid residues in the globular conformation of the basic pancreatic trypsin inhibitor (BPTI) I. ^{1}H NMR studies. *Biophys. Struct. Mech.* 2:139–158.
7. McCammon, J. A., and M. Karplus. 1979. Dynamics of activated processes in globular proteins. *Proc. Natl. Acad. Sci. USA.* 76:3585–3589.
8. McCammon, J. A., and M. Karplus. 1980. Dynamics of tyrosine ring rotations in a globular protein. *Biopolymers.* 19:1375–1405.
9. Karplus, M., and J. A. McCammon. 1981. Pressure dependence of aromatic ring rotations in proteins: a collisional interpretation. *Fed. Eur. Biochem. Soc. Lett.* 131:34–36.
10. Northrup, S. H., M. R. Pear, C. Y. Lee, J. A. McCammon, and M. Karplus. 1982. Dynamical theory of activated processes in globular proteins. *Proc. Natl. Acad. Sci. USA.* 79:4035–4039.
11. Wagner, G. 1983. Characterization of the distribution of internal motion in the basic pancreatic trypsin inhibitor using a large number of internal NMR probes. *Quart. Rev. Biophys.* 16:1–57.
12. McCammon, J. A., C. Y. Lee, and S. H. Northrup. 1983. Side-chain rotational isomerization in proteins: a mechanism involving gating and transient packing defects. *J. Am. Chem. Soc.* 105:2232–2237.
13. McCammon, J. A. 1984. Protein dynamics. *Rep. Prog. Phys.* 47:1–46.
14. Levy, R. M., and J. W. Keepers. 1987. Computer simulations of protein dynamics: theory and experiment. *Comm. Mol. Cell. Biophys.* In press.
15. Northrup, S. H., and J. A. McCammon. 1984. Gated reactions. *J. Am. Chem. Soc.* 106:930–934.
16. Beece, D., L. Eisenstein, H. Frauenfelder, D. Good, M. C. Marden, L. Reinisch, A. H. Reynolds, L. B. Sorensen, and K. T. Yue. 1980. Solvent viscosity and protein dynamics. *Biochemistry.* 19:5147–5157.
17. Kinsey, R. A., A. Kintanar, and E. Oldfield. 1981. Dynamics of amino acid side chains in membrane proteins by high field solid state deuterium nuclear magnetic resonance spectroscopy. *J. Biol. Chem.* 256:9028–9036.
18. Lecomte, J. T. J., and M. Llinas. 1984. ^{1}H NMR spectral patterns of rapidly flipping tyrosyl rings: a study of crambin in organic solvents. *J. Amer. Chem. Soc.* 106:2741–2748.
19. Frauenfelder, H., and P. G. Wolynes. 1985. Rate theories and puzzles of hemeprotein kinetics. *Science (Wash. DC).* 229:337–345.
20. Ehrenberg, A. 1986. Viscosity and glycerol effects on dynamics of cytochrome c. *In* Structure, Dynamics and Function of Biomolecules. A. Ehrenberg, A. Graslund, R. Rigler, and L. Nilsson, editors. Springer-Verlag, Berlin.
21. Case, D., and J. A. McCammon. 1987. Dynamical simulations of oxygen binding to myoglobin. *Ann. NY Acad. Sci.* In press.
22. Deisenhofer, J., and W. Steigemann. 1975. Crystallographic refinement of the structure of bovine pancreatic trypsin inhibitor at 1.5A resolution. *Acta Crystallogr.* B31:238–250.
23. Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.
24. van Gunsteren, W. F., H. J. C. Berendsen, J. Hermans, W. G. J. Hol, and J. P. M. Postma. 1983. Computer simulation of the dynamics of hydrated protein crystals and its comparison with x-ray data. *Proc. Natl. Acad. Sci. USA.* 80:4315–4319.
25. Hermans, J., H. J. C. Berendsen, W. F. van Gunsteren, and J. P. M. Postma. 1984. A consistent empirical potential for water-protein interactions. *Biopolymers.* 23:1513–1518.
26. Berendsen, H. J. C., W. F. van Gunsteren, H. R. J. Zwinderman, and R. G. Geurtsen. 1987. Simulations of proteins in water. *Ann. NY Acad. Sci.* In press.
27. Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.* 23:327–343.
28. Verlet, L. 1967. Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* 159:98–103.
29. Wood, D. W. 1979. Computer simulation of water and aqueous solutions. *In* Water: A Comprehensive Treatise. F. Franks, editor. Plenum Publishing Corp., New York.

30. Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. 1984. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* 81:3684–3690.
31. Berne, B. J. 1985. Molecular dynamics and Monte Carlo simulations of rare events. *In* Multiple Time Scales. J. U. Brackbill and B. I. Cohen, editors. Academic Press, Inc., New York.
32. Straub, J. E., and B. J. Berne. 1985. A rapid method for determining rate constants by molecular dynamics. *J. Chem. Phys.* 83:1138–1139.
33. Davis, D. G., and A. Bax. 1985. Separation of chemical exchange and cross-relaxation effects in two-dimensional NMR spectroscopy. *J. Mag. Res.* 64:533–535.