continous shift in the structure toward the one calculated without dipolar couplings when an increasing fraction of the dipolar couplings is taken out of the calculation. Considering that it frequently will be impossible to get as complete a set of dipolar couplings as was used for this dodecamer, this is a somewhat disappointing result. However, preliminary results obtained from novel experiments for quantitative measurement of ^{1}H–^{1}H couplings indicate that a smaller number of heteronuclear dipolar couplings may suffice if quantitative interproton dipolar couplings are available. Even in cases where such information cannot be retrieved, one bond–heteronuclear dipolar couplings will provide tight restraints for orienting regions of regular helical structure relative to one another.[131–133]

[9] Nuclear Magnetic Resonance Methods for High Molecular Weight Proteins: A Study Involving a Complex of Maltose Binding Protein and β-Cyclodextrin

By LEWIS E. KAY

Introduction

The development of multidimensional, multinuclear nuclear magnetic resonance (NMR) spectroscopy that started well over a decade ago has significantly increased the scope of molecules that are amenable to solution NMR studies.[1] With uniform ^{15}N,^{13}C labeling, many proteins in the 15–20 kDa molecular mass regime can now be investigated in a routine manner. The use of deuteration in concert with complete ^{15}N,^{13}C labeling has increased the size limits still further, facilitating structural studies of proteins on the order of 30–40 kDa.[2,3]

More recently, Wüthrich's group has developed a very elegant class of experiments that makes use of cross-correlated spin relaxation between dipolar and chemical shift anisotropy (CSA) relaxation mechanisms.[4,5] The signal derived from spins where the two competing relaxation fields partially cancel is selected, resulting in increased sensitivity and resolution in important classes of experiments. This so-called TROSY (Transverse Relaxation Optimized Spectroscopy) promises to increase the size of proteins amenable to solution NMR

[1] A. Bax, *Curr. Opin. Struct. Biol.* **4,** 738 (1994).
[2] K. H. Gardner and L. E. Kay, *Annu. Rev. Biophys. Biomol. Struct.* **27,** 357 (1998).
[3] B. T. Farmer and R. A. Venters, *in* "Biological Magnetic Resonance" (N. R. Krishna and L. J. Berliner, eds.), Vol. 16, p. 75. Kluwer Academic/Plenum Publishers, New York, 1998.
[4] K. Pervushin, R. Riek, G. Wider, and K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **94,** 12366 (1997).
[5] K. Pervushin, R. Riek, G. Wider, and K. Wüthrich, *J. Am. Chem. Soc.* **120,** 6394 (1998).

0076-6879/00 $35.00

study still further. Indeed, applications of this exciting methodology have already emerged.[6-8]

Approaches for measuring dipolar couplings in solution have also been described.[9,10] Many important anisotropic interactions, including dipolar couplings between pairs of spins, are averaged out in the isotropic solvent environment typically used in solution NMR studies. Information provided by such interactions has been exploited for a long time in the solid state where it forms the basis of structural studies.[11,12] Recently, however, it has become possible to reintroduce these interactions by addition of bicelles[9,13] or phage particles[14,15] that orient in a magnetic field. The small degree of alignment imparted to the solute particles leads to measurable dipolar couplings between pairs of interacting spins, providing important probes of structure.

The goal of the present chapter is not to describe in detail any one of these advances. Rather, the aim is to provide a picture of how all of these tools were used in concert to study the solution structural properties of maltose binding protein (MBP), a 370 residue single polypeptide chain, in complex with the cyclic heptasaccharide, β-cyclodextrin. At each stage a brief description of the methodology employed will be presented, but the interested reader is referred to the original literature, including a number of reviews,[2,3,16-18] for more details.

Labeling Strategy Used for MBP

The first triple-resonance experiment which exploited the line narrowing of ^{13}C and $^{1}H^{N}$ resonances afforded by high levels of deuteration was published by Bax and co-workers in 1993.[19] Subsequently Yamazaki *et al.* published a suite of

[6] M. Pellecchia, P. Sebbel, U. Herrmanns, K. Wüthrich, and R. Glockshuber, *Nat. Struct. Biol.* **6,** 336 (1999).

[7] Y. X. Wang, J. Jacob, F. Cordier, P. Wingfield, S. J. Stahl, S. Lee-Huang, D. Torchia, S. Grzesiek, and A. Bax, *J. Biomol. NMR* **14,** 181 (1999).

[8] K. Pervushin, A. Ono, C. Fernandez, T. Szyperski, M. Kainosho, and K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **95,** 14147 (1998).

[9] N. Tjandra and A. Bax, *Science* **278,** 1111 (1997).

[10] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, and J. H. Prestegard, *Proc. Natl. Acad. Sci. USA* **92,** 9279 (1995).

[11] R. R. Ketchem, W. Hu, and T. A. Cross, *Science* **261,** 1457 (1993).

[12] R. G. Griffin, *Nat. Struct. Biol. NMR Supplement* **5,** 508 (1998).

[13] M. Ottiger and A. Bax, *J. Biomol. NMR* **12,** 361 (1998).

[14] M. R. Hansen, L. Mueller, and A. Pardi, *Nat. Struct. Biol.* **5,** 1065 (1998).

[15] G. M. Clore, M. R. Starich, and A. M. Gronenborn, *J. Am. Chem. Soc.* **120,** 10571 (1998).

[16] K. W. Wüthrich, *Nat. Struct. Biol. NMR Suppl.* **5,** 492 (1998).

[17] G. Wider and K. Wüthrich, *Curr. Opin. Struct. Biol.* **9,** 594 (1999).

[18] J. H. Prestegard, *Nat. Struct. Biol. NMR Suppl.* **5,** 517 (1998).

[19] S. Grzesiek, J. Anglister, H. Ren, and A. Bax, *J. Am. Chem. Soc.* **115,** 4369 (1993).

triple resonance experiments for the assignment of backbone ^{15}N, $^{13}C^{\alpha}$, $^{1}H^{N}$ and side-chain $^{13}C^{\beta}$ chemical shifts,[20,21] which was later extended by Shan *et al.* in studies of a 64 kDa complex of *trp* repressor and operator DNA.[22] Other groups have also contributed significantly to the array of pulse schemes that can be used in the study of highly deuterated, ^{15}N,^{13}C-labeled proteins.[23–30] It became very clear early on, however, that the elimination of all nonexchangeable protons, although beneficial for the assignment of backbone and side-chain heavy atoms, would pose serious challenges to structural studies of proteins relying on traditional nuclear overhauser effect (NOE) approaches. With this problem in mind a protocol was developed for the production of Val, Leu, Ile (δ1 only) methyl protonated, ^{15}N,^{13}C,^{2}H-labeled proteins and applied to the specific case of MBP.[31] In this approach, protein is expressed using [^{13}C,^{2}H]glucose and $^{15}NH_4Cl$ as the carbon and nitrogen sources, respectively, in a D_2O based medium, with the (commercially available) methyl protonated, ^{2}H,^{13}C-labeled metabolites, [3,3-$^{2}H_2$] α-ketobutyrate and [3-^{2}H] α-ketoisovalerate added approximately 1 hr prior to induction. Figure 1 illustrates the metabolites that are used during biosynthesis and the resultant labeling pattern in the amino acids that are produced directly from the metabolites. A detailed mass-spectroscopic and NMR analysis (Fig. 2) has established that the only NMR visible isotopomers produced are of the $^{13}CH_3$ variety and that Val, Leu, and Ile (δ1) methyl groups are protonated at approximately 90% when 75 mg/liter of α-ketoisovalerate and 50 mg/liter of α-ketobutyrate are employed.[31] The excellent dispersion available for the Ile methyl groups, in particular, is evident from Fig. 2 and is critical for the rapid assignment of NOE correlations involving these groups (see below).

[20] T. Yamazaki, W. Lee, M. Revington, D. L. Mattiello, F. W. Dahlquist, C. H. Arrowsmith, and L. E. Kay, *J. Am. Chem. Soc.* **116,** 6464 (1994).

[21] T. Yamazaki, W. Lee, C. H. Arrowsmith, D. R. Muhandiram, and L. E. Kay, *J. Am. Chem. Soc.* **116,** 11655 (1994).

[22] X. Shan, K. H. Gardner, D. R. Muhandiram, N. S. Rao, C. H. Arrowsmith, and L. E. Kay, *J. Am. Chem. Soc.* **118,** 6570 (1996).

[23] B. T. Farmer and R. Venters, *J. Am. Chem. Soc.* **117,** 4187 (1995).

[24] B. T. Farmer and R. A. Venters, *J. Biomol. NMR* **7,** 59 (1996).

[25] M. Shirakawa, M. Walchli, M. Shimizu, and Y. Kyogoku, *J. Biomol. NMR* **5,** 323 (1995).

[26] H. Matsuo, E. Kupce, H. Li, and G. Wagner, *J. Magn. Reson. Series B* **111,** 194 (1996).

[27] H. Matsuo, H. Li, and G. Wagner, *J. Magn. Reson. Series B* **110,** 112 (1996).

[28] D. Nietlispach, R. T. Clowes, R. W. Broadhurst, Y. Ito, J. Keeler, M. Kelly, J. Ashurst, H. Oschkinat, P. J. Domaille, and E. D. Laue, *J. Am. Chem. Soc.* **118,** 407 (1996).

[29] Y. Lin and G. Wagner, *J. Biomol. NMR* **15,** 227 (1999).

[30] M. Hennig, D. Ott, P. Schulte, R. Lowe, J. Krebs, T. Vorherr, W. Bermel, H. Schwalbe, and C. Griesinger, *J. Am. Chem. Soc.* **119,** 5055 (1997).

[31] N. K. Goto, K. H. Gardner, G. A. Mueller, R. C. Willis, and L. E. Kay, *J. Biomol. NMR* **13,** 369 (1999).

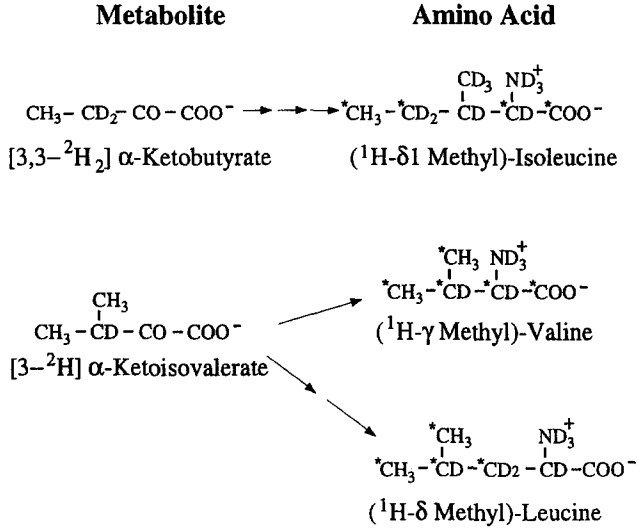


FIG. 1. Chemical structures of the metabolites supplementing the $^{2}H_2O/^{13}C,^{2}H$ glucose/$^{15}NH_4Cl$ growth media used to produce Val, Leu, Ile ($\delta 1$) methyl protonated, $^{15}N,^{13}C,^{2}H$-labeled proteins. The labeling pattern on each of Val, Leu, and Ile is shown, with carbons derived from the metabolites indicated by an asterisk (*). The metabolites [3,3-$^{2}H_2$]-^{13}C α-ketobutyrate and ^{13}C, protonated α-ketoisovalerate are available commercially. [3-^{2}H]-α-ketoisovalerate is easily produced from the protonated form by incubation in D_2O at pH 12.5, 45° for 2–3 hr. The reaction is monitored by NMR and goes to completion with no detectable production of other products. Reproduced with permission from Goto *et al.*[31]

It is worth mentioning at this point that a new labeling approach is presently under development by Otomo *et al.*[32–34] and Cowburn, Muir, and co-workers[35] that involves $^{15}N,^{13}C$ labeling of specific peptide segments in a protein, while the remaining fragments remain unlabeled. For application to high molecular weight proteins where signal overlap can be critical this has the advantage in that only those regions that are labeled are observed via edited experiments (or eliminated via filtered schemes). Yamazaki and collaborators have applied this intein methodology to studies of MBP.[32] In a demonstration of the power of the method, two separate inteins have been used to ligate unlabeled segments to an ^{15}N-labeled middle fragment of the molecule.

[32] T. Otomo, N. Ito, Y. Kyogoku, and T. Yamazaki, *Biochemistry* **38,** 16040 (1999).
[33] T. Otomo, K. Teruya, K. Uegaki, T. Yamazaki, and Y. Kyogoku, *J. Biomol. NMR* **14,** 105 (1999).
[34] T. Yamazaki, T. Otomo, N. Oda, Y. Kyogoku, K. Uegaki, K. Ito, Y. Ishino, and H. Nakamura, *J. Am. Chem. Soc.* **120,** 5591 (1998).
[35] R. Xu, B. Ayers, D. Cowburn, and T. W. Muir, *Proc. Natl. Acad. Sci. USA* **96,** 388 (1999).

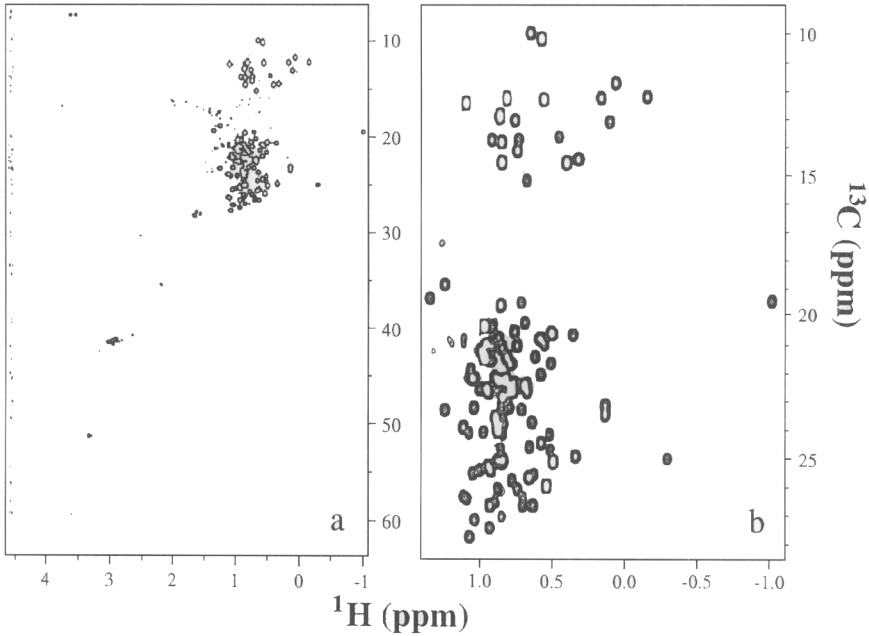


FIG. 2. ^{1}H–^{13}C constant-time HSQC spectrum of a 1.4 m*M* complex of Val, Leu, Ile ($\delta 1$) methyl protonated, ^{15}N,^{13}C,^{2}H-labeled MBP, and β-cyclodextrin recorded at 37° on a Varian Unity+ 500 MHz spectrometer. (a) Aliphatic region of the spectrum indicating the selectivity of labeling. Small levels of protonation at C^{γ} positions of Pro/Arg, $C^{\gamma 2}$ methyls of Ile, and C^{β} positions of Asp/Ser are observed with intensities less than 10% of the expected methyl peaks. (b) Methyl region of the HSQC, illustrating that CH_2D and CHD_2 isotopomers are not produced. Reproduced with permission from Goto *et al.*[31]

Backbone and Side-Chain Assignments of MBP

The high levels of deuteration achieved by the ^{15}N,^{13}C,^{2}H (^{1}H-methyl)-labeling method described above is particularly desirable for backbone assignment. On the basis of CT-HNCA, CT-HN(CO)CA, CT-HN(CA)CB and CT-HN(COCA)CB experiments[20–22] $\geq$95% of the backbone ^{15}N, $^{13}C^{\alpha}$, $^{1}H^{N}$, and side-chain $^{13}C^{\beta}$ chemical shifts could be assigned from spectra recorded at 600 MHz on a 0.7 m*M* sample of MBP at 37°. At this temperature the overall rotational correlation time as established by ^{15}N spin relaxation experiments[36] is 17 ns, and although TROSY-based schemes would be preferred, excellent sensitivity was nevertheless obtained using non-TROSY versions of the experiments listed above. Figure 3 illustrates the quality of the data obtained for MBP and, in particular, shows the high resolution

[36] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay, *Biochemistry* **33,** 5984 (1994).

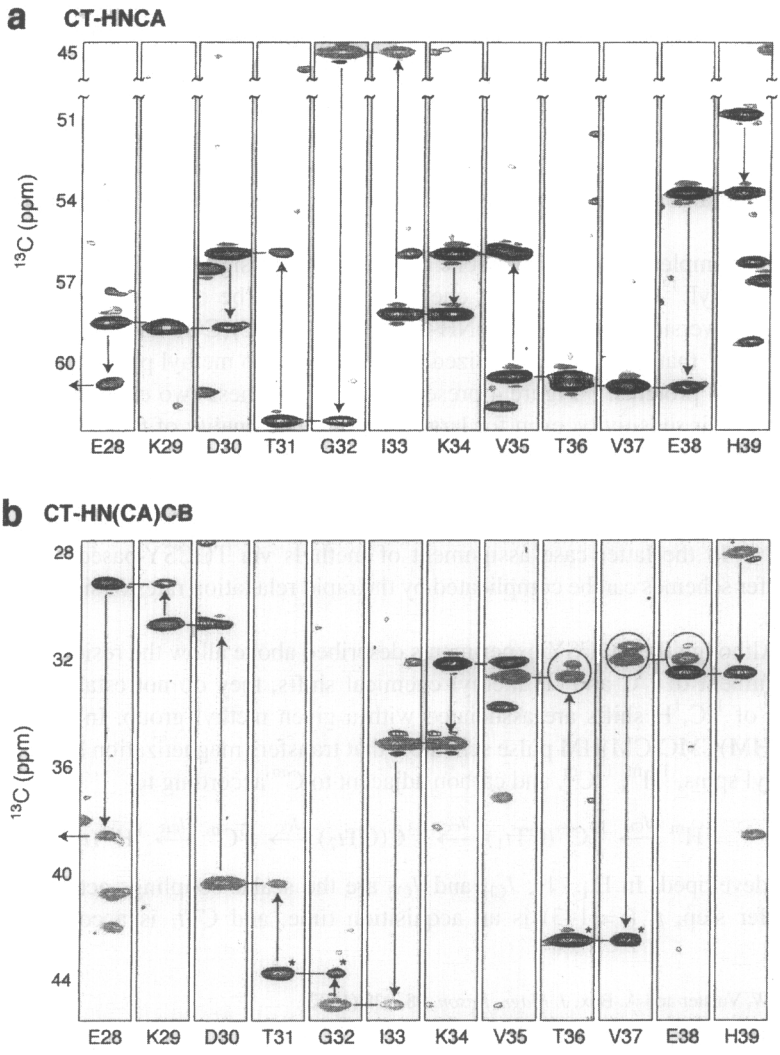


FIG. 3. Slices from CT-HNCA (a) and CT-HN(CA)CB (b) spectra recorded on a 0.7 m*M* sample of Val, Leu, Ile (δ1) methyl protonated, ^{15}N, ^{13}C, ^{2}H-labeled MBP and β-cyclodextrin at 37° (600 MHz). Each slice is taken at the F_2 frequency (^{15}N) of the amino acid indicated. Negative peaks arise from Gly in (a) and from residues with C^{β} carbons one-bond coupled to an odd number of aliphatic carbons or Gly in (b). The weak upfield shoulders in the carbon dimension for Val residues (circled) are the result of ~10% CD_3 isotopomer associated with Val that is produced by the labeling scheme of Fig. 1. Peaks marked with * are aliased in the carbon dimension. Reproduced with permission from Gardner *et al.*[45]

that constant-time ^{13}C spectroscopy[37,38] affords, which is absolutely critical in 3D NMR applications to high molecular weight proteins such as MBP. Of note is that correlations in the 3D CT-HN(CA)CB [and CT-HN(COCA)CB] that derive from residues with $^{13}C^{\beta}$ carbons one-bond coupled to an odd number of aliphatic ^{13}C spins are of opposite phase relative to correlations involving $^{13}C^{\beta}$ carbons coupled to an even number of ^{13}C spins.[22] This feature is particularly useful in resolving ambiguities that arise from degeneracies in pairs of $^{15}N,^{1}H^{N}$ chemical shifts.

On completion of the backbone chemical shift assignment, Val, Leu, and Ile ($\delta 1$) methyl $^{13}C,^{1}H$ ($^{13}C^{m},^{1}H^{m}$) chemical shifts can be obtained using slightly modified versions of (H)C(CO)NH-TOCSY and H(C)(CO)NH-TOCSY experiments[39–41] that have been optimized for application to methyl protonated, highly deuterated proteins.[42] Figure 4 presents strips from these two experiments illustrating their sensitivity, even for large proteins. The quality of these types of experiments points to an advantage associated with selective methyl labeling of sidechains relative to other labeling schemes such as those that result in complete protonation of side chains of selected amino acids, such as Val, Leu, and/or Ile.[43,44] In the latter case assignment of methyls via TOCSY-based coherence transfer schemes can be complicated by the rapid relaxation rates of sidechain ^{13}C spins.

Although the TOCSY experiments described above allow the residue specific assignment of ^{13}C and ^{1}H methyl chemical shifts, they do not establish which pairs of $^{13}C,^{1}H$ shifts are associated with a given methyl group. In this regard the (HM)CMC(CM)HM pulse scheme[45] that transfers magnetization between the methyl spins, $^{1}H^{m}$, $^{13}C^{m}$, and carbon adjacent to C^{m} according to

$$^{1}H^{m} \xrightarrow{J_{CH}} {}^{13}C^{m}(CTt_1) \xrightarrow{J_{CC}} {}^{13}C(CTt_2) \xrightarrow{J_{CC}} {}^{13}C^{m} \xrightarrow{J_{CH}} {}^{1}H^{m}(t_3) \qquad (1)$$

was developed. In Eq. (1), J_{CH} and J_{CC} are the scalar couplings active in each transfer step, t_i ($i = 1$–3) is an acquisition time, and CTt_i is a constant-time

[37] G. W. Vuister and A. Bax, *J. Magn. Reson.* **98,** 428 (1992).
[38] J. Santoro and G. C. King, *J. Magn. Reson.* **97,** 202 (1992).
[39] S. Grzesiek, J. Anglister, and A. Bax, *J. Magn. Reson. Series B* **101,** 114 (1993).
[40] T. M. Logan, E. T. Olejniczak, R. Xu, and S. W. Fesik, *FEBS Lett.* **314,** 413 (1992).
[41] G. T. Montelione, B. A. Lyons, S. D. Emerson, and M. Tashiro, *J. Am. Chem. Soc.* **114,** 10974 (1992).
[42] K. H. Gardner, R. Konrat, M. K. Rosen, and L. E. Kay, *J. Biomol. NMR* **8,** 351 (1996).
[43] W. J. Metzler, M. Wittekind, V. Goldfarb, L. Mueller, and B. T. Farmer, *J. Am. Chem. Soc.* **118,** 6800 (1996).
[44] B. O. Smith, Y. Ito, A. Raine, S. Teichmann, L. Ben-Tovim, D. Nietlispach, R. W. Broadhurst, T. Terada, M. Kelly, K. Oschkinat, T. Shibata, S. Yokoyama, and E. D. Laue, *J. Biomol. NMR* **8,** 360 (1996).
[45] K. H. Gardner, X. Zhang, K. Gehring, and L. E. Kay, *J. Am. Chem. Soc.* **120,** 11738 (1998).

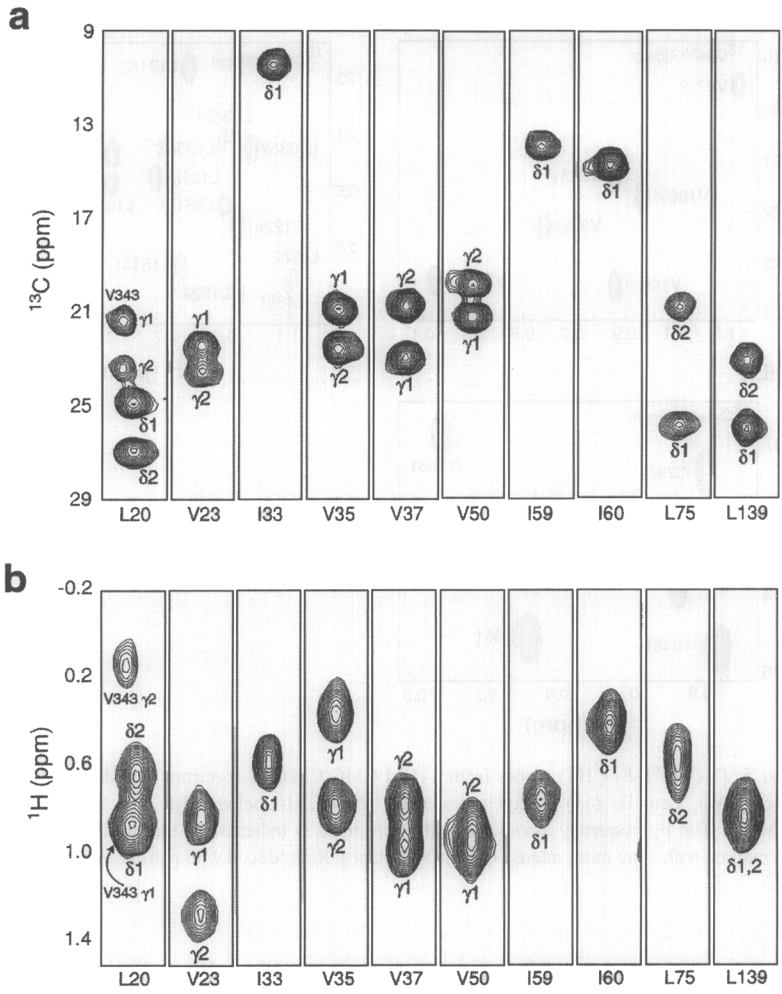


FIG. 4. Strips from (H)C(CO)NH-TOCSY (a) and H(C)(CO)NH-TOCSY (b) spectra recorded on a 0.9 m*M* sample of Val, Leu, Ile ($\delta 1$) methyl protonated, ^{15}N,^{13}C,^{2}H-labeled MBP/β-cyclodextrin at 37° (600 MHz). Each strip is labeled according to the identity of the methyl-containing residue. Reproduced with permission from Gardner *et al.*[45]

indirect detection period. Data sets are obtained with correlations at (ω_C^m, ω_C, ω_H^m), as illustrated in Fig. 5. It is also possible to separate correlations from Val and from Leu/Ile using a form of difference spectroscopy described in detail by Gardner *et al.*[45]

In total 119 of the 122 methyl groups in MBP were assigned using the methodology described above. Two of the three remaining methyl groups lie in regions

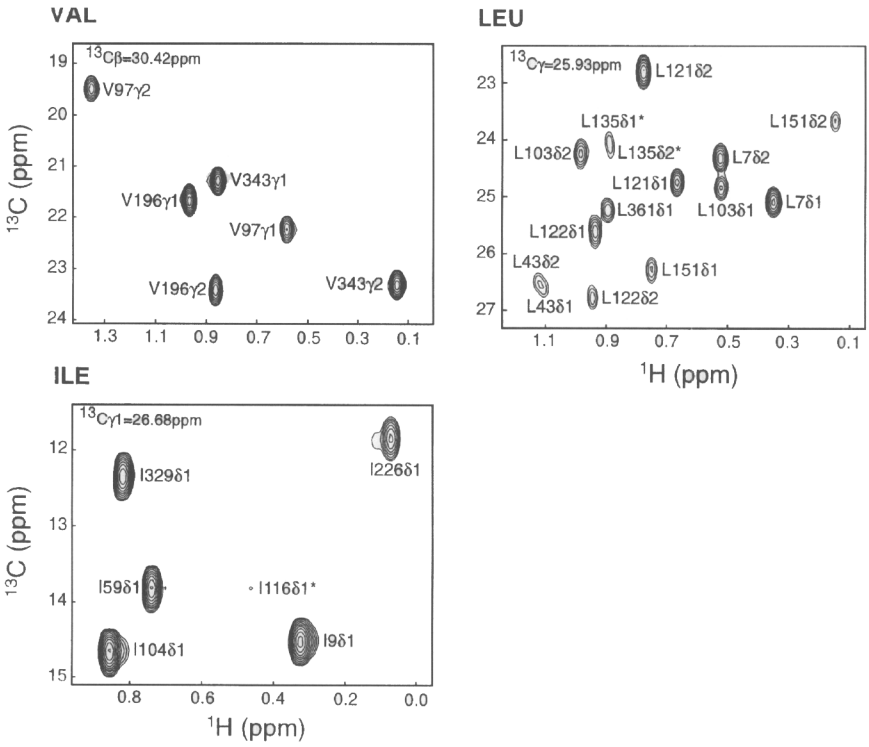
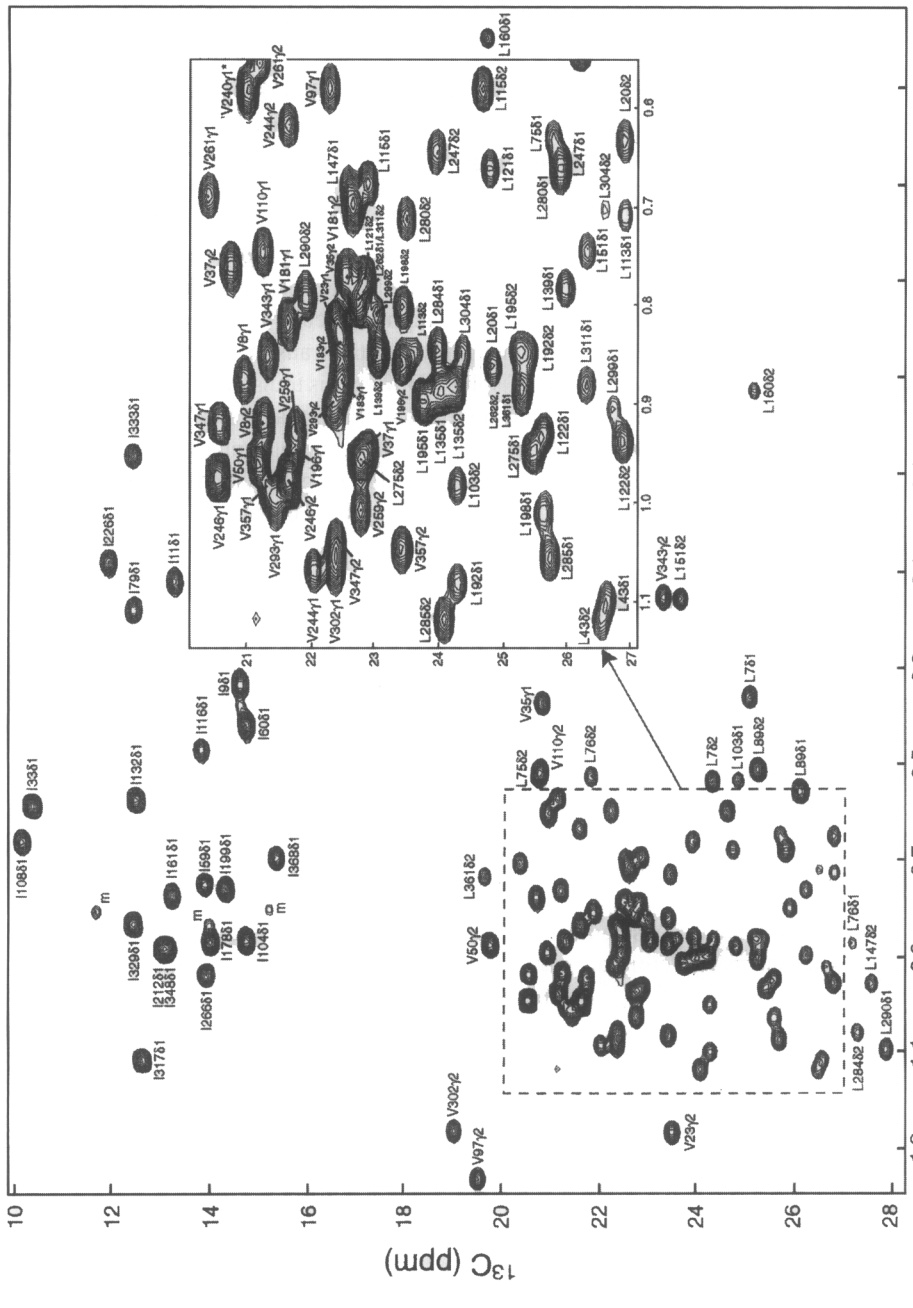


FIG. 5. F_1 ($^{13}C^m$)–F_3 ($^1H^m$) planes from a (HM)CMC(CM)HM spectrum recorded on a 0.9 m*M* sample of Val, Leu, Ile ($\delta 1$) methyl protonated, ^{15}N,^{13}C,^2H-labeled MBP/β-cyclodextrin at 37° (600 MHz). The F_2 frequency associated with each plane is indicated in the top left-hand corner. Peaks marked with * are more intense on adjacent planes. Reproduced with permission from Gardner *et al.*[45]

close to the β-cyclodextrin binding site that are significantly exchange broadened. Figure 6 shows the ^{13}C,^1H correlation map of Val, Leu, Ile ($\delta 1$ only) methyl-protonated, ^{15}N, ^{13}C, ^2H-labeled MBP, with stereospecific assignments, obtained by the method of Neri *et al.*[46] labeled on the diagram.

[46] D. Neri, T. Szyperski, G. Otting, H. Senn, and K. Wüthrich, *Biochemistry* **28,** 7510 (1989).

FIG. 6. CT ^{13}C-^1H HSQC of Val, Leu, Ile ($\delta 1$) methyl protonated, ^{15}N,^{13}C,^2H-labeled MBP/β-cyclodextrin at 37° (600 MHz), with assignments. Peaks marked with an "m" are from a small fraction of maltose-bound MPB. Reproduced with permission from Gardner *et al.*[45]


^{13}C (ppm)
^{1}H (ppm)

TROSY-Based Triple Resonance 4D Spectroscopy of MBP at Low Temperature

As described in the introduction, the development of TROSY spectroscopy by Pervushin, Wüthrich, and co-workers[4,5] is an extremely important advance that promises to significantly increase the range of proteins that can be studied using the ^{15}N, ^{13}C, ^{2}H labeling approaches discussed above. Descriptions of the TROSY principle have appeared in reviews by Wüthrich[16] and by Wider and Wüthrich[17] as well as in the original papers by Pervushin *et al.*[4,5] In what follows, therefore, only a brief discussion of the method is presented followed by a number of 4D experiments that are useful for assignment of high molecular weight proteins.

Many important classes of experiments involve the transfer of magnetization from $^{1}H^{N}$–^{15}N spin pairs to adjacent ^{13}C spins.[1] During this transfer in non-TROSY triple resonance experiments $^{1}H^{N}$ spins are decoupled so that the effective signal decay of ^{15}N magnetization is given by the relaxation rate of inphase magnetization, N_{TR}.[47,48] For reasonably sized molecules at high magnetic fields the components of ^{15}N magnetization coupled to $^{1}H^{N}$ spins in the "up" and "down" states, $N_{TR}(1 + 2H_z^N)$ and $N_{TR}(1 - 2H_z^N)$, decay at significantly different rates resulting from interference between $^{1}H^{N}$–^{15}N dipolar and ^{15}N chemical shift anisotropy (CSA) relaxation interactions.[4,49] Sensitivity and resolution in these experiments can therefore be improved by ensuring that these two components do not interchange during the course of the pulse scheme and by selecting signal derived from the slowly relaxing component exclusively.[4] It is noteworthy that interference between $^{1}H^{N}$–^{15}N dipolar and $^{1}H^{N}$ CSA relaxation operates during acquisition so that maximum gains can be achieved by choosing the pathway where magnetization is transferred from the slowly relaxing ^{15}N component to the slowly relaxing $^{1}H^{N}$ signal. Triple resonance 3D experiments that incorporate the TROSY principle have been published by Salzmann *et al.*,[50] Yang and Kay[51] (see also supplemental material of this reference), and Palmer and co-workers.[52]

As described above, assignment of MBP was achieved using a series of CT-based experiments that correlate either interresidue or both intra- and interresidue $^{1}H^{N}$,^{15}N,$^{13}C^{\alpha}$ or $^{1}H^{N}$,^{15}N,$^{13}C^{\beta}$ chemical shifts. The CT-$^{13}C^{\beta}$ experiments are particularly important since many potential ambiguities can be eliminated based on the $^{13}C^{\beta}$ chemical shift and the sign of the cross peaks[22] (see above). Unfortunately

[47] S. Grzesiek and A. Bax, *J. Magn. Reson.* **96,** 432 (1992).

[48] B. T. Farmer, R. A. Venters, L. D. Spicer, M. G. Wittekind, and L. Mueller, *J. Biomol. NMR* **2,** 195 (1992).

[49] M. Goldman, *J. Magn. Reson.* **60,** 437 (1984).

[50] M. Salzmann, K. Pervushin, G. Wider, H. Senn, and K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **95,** 13585 (1998).

[51] D. Yang and L. E. Kay, *J. Am. Chem. Soc.* **121,** 2571 (1999).

[52] J. P. Loria, M. Rance, and A. G. Palmer, *J. Magn. Reson.* **141,** 180 (1999).

these experiments are likely to fail in applications to proteins with correlation times in excess of 30–40 ns, even when TROSY is employed, since these schemes make use of extended periods during which transverse ^{13}C magnetization is present. Although it is possible to record the $^{13}C^{\beta}$ shift in non-CT mode, the lack of resolution in cases of application to high molecular weight, single polypeptide chains is likely to be limiting. With this in mind we have developed a set of 4D TROSY-based pulse schemes for backbone assignment that offer significantly improved resolution over their 3D counterparts.[51,53] As a test of the methodology the experiments have been evaluated by recording data sets on a 1.4 m*M* sample of methyl-protonated, ^{15}N, ^{13}C, ^{2}H-labeled MBP/β-cyclodextrin at 5°, where the complex tumbles with an overall correlation time of 46 ns.

The 4D experiments, 4D HNCOCA,[51] 4D HNCACO,[51] and 4D HNCO$_{i-1}$CA$_i$[53] are based on 3D HN(CO)CA[54] and HN(CA)CO[55] pulse schemes and the magnetization transfer pathways in each experiment can be described succinctly by 4D HNCOCA:

$$\begin{aligned} {}^{1}H^{N}_{(i)} \xrightarrow{J_{HN}} {}^{15}N_{(i)} \xrightarrow{J_{NC'}} {}^{13}C'_{(i-1)}[CTt_1] \xrightarrow{J_{C'C\alpha}} {}^{13}C^{\alpha}_{(i-1)}[t_2] \xrightarrow{J_{C'C\alpha}} {}^{13}C'_{(i-1)} \\ \xrightarrow{J_{NC'}} {}^{15}N_{(i)}[CTt_3] \xrightarrow{J_{HN}} {}^{1}H^{N}_{(i)}[t_4] \end{aligned} \tag{2}$$

4D HNCACO:

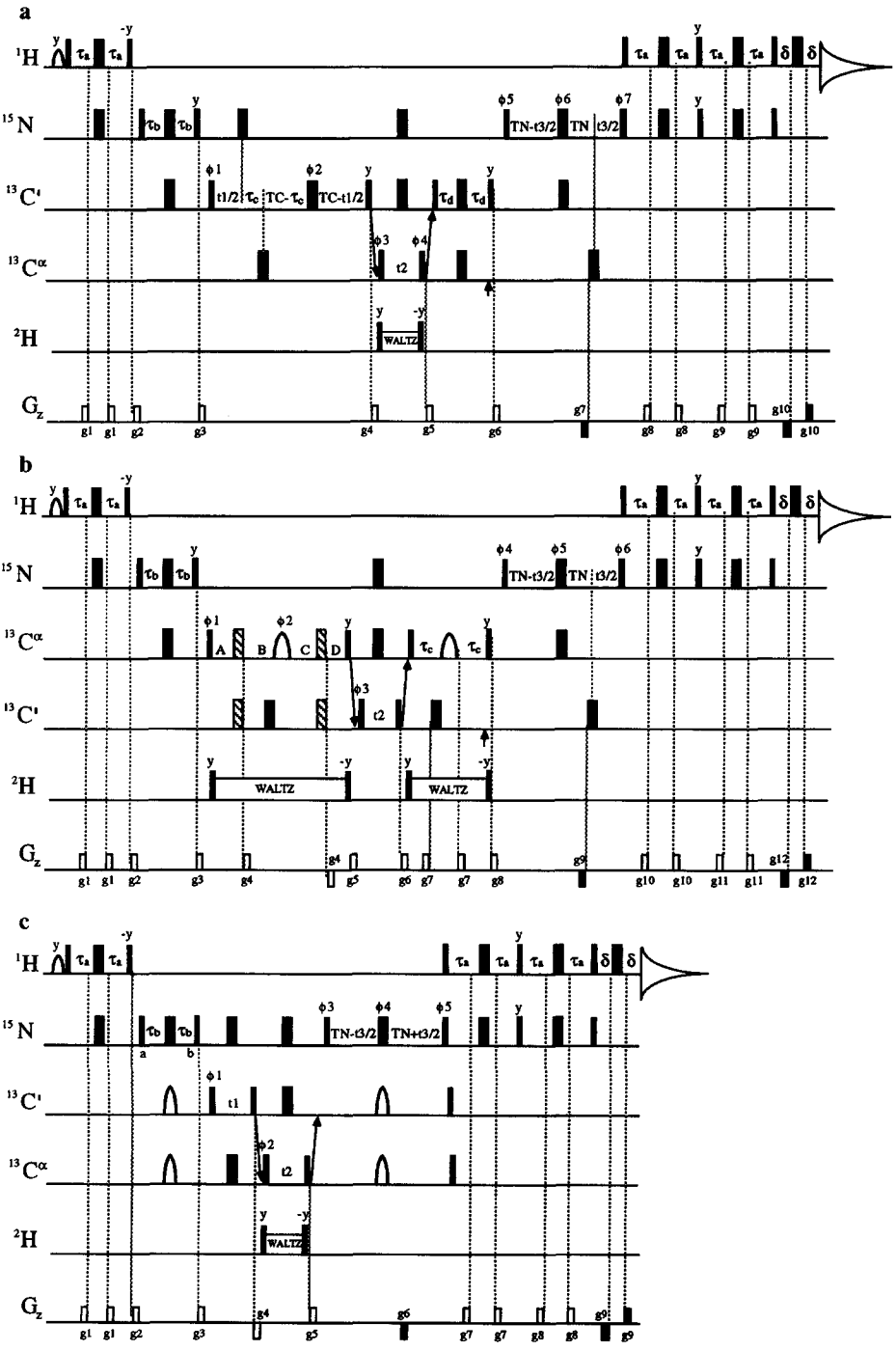
$$\begin{aligned} {}^{1}H^{N}_{(i)} \xrightarrow{J_{HN}} {}^{15}N_{(i)} \xrightarrow{J_{NC\alpha}} {}^{13}C^{\alpha}_{(i-1,i)}[CTt_1] \xrightarrow{J_{C'C\alpha}} {}^{13}C'_{(i-1,i)}[t_2] \xrightarrow{J_{C'C\alpha}} {}^{13}C^{\alpha}_{(i-1,i)} \\ \xrightarrow{J_{NC\alpha}} {}^{15}N_{(i)}[CTt_3] \xrightarrow{J_{HN}} {}^{1}H^{N}_{(i)}[t_4] \end{aligned} \tag{3}$$

4D HNCO$_{i-1}$CA$_i$:

$$\begin{aligned} {}^{1}H^{N}_{(i)} \xrightarrow{J_{HN}} {}^{15}N_{(i)} \xrightarrow{J_{NC'}, J_{NC\alpha}} {}^{13}C'_{(i-1)}[t_1], {}^{13}C^{\alpha}_{(i-1,i)}[t_2] \xrightarrow{J_{NC'}, J_{NC\alpha}} \\ {}^{15}N[CTt_3] \xrightarrow{J_{HN}} {}^{1}H^{N}_{(i)}[t_4] \end{aligned} \tag{4}$$

Figure 7 illustrates the pulse schemes of the 4D experiments discussed above. Unlike the approach described by Pervushin and co-workers where the fast relaxing component of magnetization is selected against,[4] both fast and slowly relaxing components are preserved in the experiments of Fig. 7. However, for applications to proteins tumbling with correlation times $\geq$30 ns at fields of 600 MHz or larger, the fast-relaxing component decays completely during the course of the delays in the pulse scheme so that only the desired TROSY component is observed.[56] For the case of MBP at 5° the enhanced sensitivity gradient implementations of TROSY illustrated in Fig. 7 offer approximately 20% improved sensitivity relative to experiments that actively suppress the anti-TROSY contributions to the signal.[56]

[53] R. Konrat, D. Yang, and L. E. Kay, *J. Biomol. NMR* **15,** 309 (1999).
[54] A. Bax and M. Ikura, *J. Biomol. NMR* **1,** 99 (1991).
[55] R. T. Clubb, V. Thanabal, and G. Wagner, *J. Magn. Reson.* **97,** 213 (1992).
[56] D. Yang and L. E. Kay, *J. Biomol. NMR* **13,** 3 (1999).


a
b
c
WALTZ
TN-t3/2
TN t3/2
TC-t1/2
TN+t3/2

Representative $^{13}C^{\alpha}$–$^{13}C'$ and ^{15}N–$^{1}H^{N}$ planes from the 4D HNCACO and 4D HNCOCA data sets of MBP at 5° are shown in Fig. 8, illustrating the assignment strategy that is used. Starting from slice A in the Figure, the $^{13}C^{\alpha}$,$^{13}C'$ coordinates of Ala-168 are obtained, from which the $^{1}H^{N}$ and ^{15}N shifts of Phe-169 are assigned from the 4D HNCOCA (slice B). The $^{1}H^{N}$,^{15}N shifts in turn allow assignment of the $^{13}C^{\alpha}/^{13}C'$ carbons of Phe-169 from the 4D HNCACO. This approach is then continued for successive residues.

Although there is essentially no overlap of cross peaks in each of the 4D data sets of MBP, there is, of course, overlap in 2D ^{15}N–$^{1}H^{N}$ and $^{13}C^{\alpha}$–$^{13}C'$ correlation maps. Such overlap complicates assignment based on the 4D HNCOCA and 4D HNCACO experiments, as illustrated schematically in Fig. 9 (see color plate). For example, in the first stage of the assignment procedure the $^{13}C^{\alpha}/^{13}C'$ shifts obtained in the 4D HNCACO are used to assign the $^{1}H^{N}$,^{15}N pair of the successive residue using the 4D HNCOCA. In cases where the $^{1}H^{N}$,^{15}N shifts are degenerate (denoted by the green spins in Fig. 9) it is not possible to unambiguously select the intraresidue $^{13}C^{\alpha}/^{13}C'$ shifts from the 4D HNCACO. The 4D $HNCO_{i-1}CA_i$ resolves this ambiguity by correlating pairs of chemical shifts [$^{13}C'_{(i-1)}$ and $^{13}C^{\alpha}_{(i)}$] that bridge the degenerate $^{1}H^{N}$,^{15}N spin pair. Figure 10 illustrates how the 4D $HNCO_{i-1}CA_i$ is used in concert with the 4D HNCACO and 4D HNCOCA experiments to resolve ambiguities in the assignment processes arising from degenerate pairs of $^{1}H^{N}$,^{15}N chemical shifts. Leu-43/Leu-135 (Fig. 10a) and Ala-324/Asp-363 (Fig. 10b) have degenerate amide proton and nitrogen chemical shifts. Hence it is not possible to assign the $^{13}C^{\alpha}/^{13}C'$ chemical shifts of Leu-43, for example, from the 4D HNCACO. However, because the $^{13}C'$ and $^{13}C^{\alpha}$ chemical shifts of Lys-42 and Leu- 43, respectively, are distinct from the corresponding shifts of Ala-134 and Leu-135, it is possible to assign the $^{13}C^{\alpha}/^{13}C'$ shifts of Leu-43 (and Leu-135) from the $HNCO_{i-1}CA_i$. A similar situation occurs in the case of Ala-324/Asp-363, illustrated in Fig. 10(b).

An additional bottleneck in the above assignment procedure can result due to degeneracies of $^{13}C^{\alpha}/^{13}C'$ pairs of chemical shifts. In this case step 1 of the assignment illustrated in Fig. 9 would not be possible since several candidate sequential ^{15}N,$^{1}H^{N}$ pairs of shifts would emerge from the 4D HNCOCA. For these cases a 4D $^{1}H^{N}$–$^{1}H^{N}$ NOESY data set[57,58] is extremely useful since correlations between

[57] R. A. Venters, W. J. Metzler, L. D. Spicer, L. Mueller, and B. T. Farmer, *J. Am. Chem. Soc.* **117,** 9592 (1995).

[58] S. Grzesiek, P. Wingfield, S. Stahl, J. Kaufman, and A. Bax, *J. Am. Chem. Soc.* **117,** 9594 (1995).

FIG. 7. Pulse schemes of the 4D TROSY HNCOCA (a), 4D TROSY HNCACO (b), and 4D TROSY $HNCO_{i-1}CA_i$ (c) used to obtain backbone chemical shifts of high molecular weight ^{15}N,^{13}C,^{2}H-labeled proteins. Details can be found in Yang and Kay[51] (a,b) and Konrat *et al.*[53] (c).

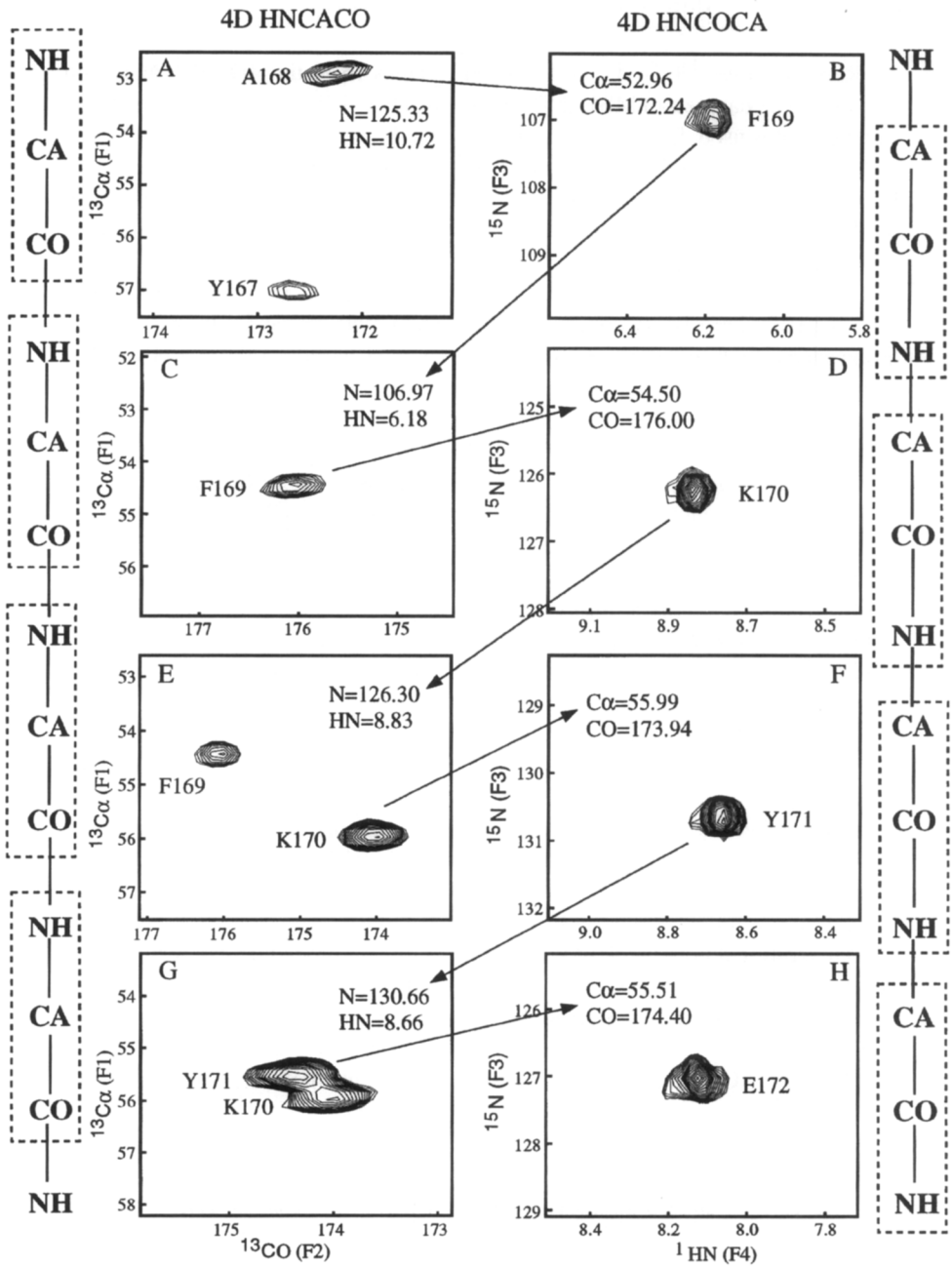


FIG. 8. Representative planes from the 4D HNCACO and 4D HNCOCA recorded at 600 MHz on a 1.4 m*M* sample of Val, Leu, Ile ($\delta 1$) methyl protonated, ^{15}N,^{13}C, ^{2}H-labeled MBP/β-cyclodextrin, 5° (correlation time of 46 ns). Cross peaks are labeled on the basis of the $^{13}C^{\alpha}$,$^{13}C'$ or ^{15}N,$^{1}H^{N}$ shifts in the 4D HNCACO and 4D HNCOCA, respectively. Reproduced with permission from Yang and Kay.[51]

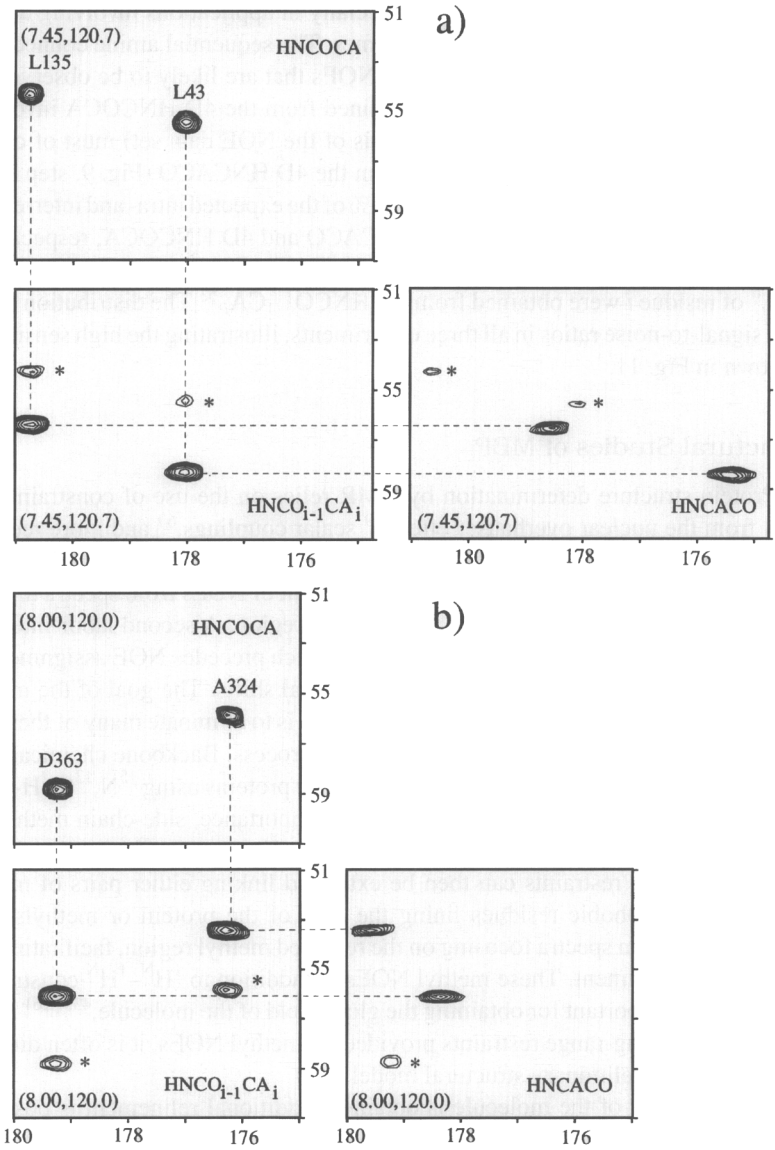


FIG. 10. Selected slices from the 4D HNCOCA, 4D HNCO$_{i-1}$CA$_i$, and 4D HNCACO data sets at ($^1H^N$,^{15}N) chemical shifts of (7.45 ppm, 120.7 ppm) (a) and (8.00 ppm, 120.0 ppm) (b) illustrating the utility of the 4D HNCO$_{i-1}$CA$_i$ in resolving ambiguities in chemical shift assignments resulting from degenerate $^1H^N$,^{15}N spin pairs. Reproduced with permission from Konrat *et al.*[53]

sequential amides are often obtained, especially in applications involving deuterated proteins and relatively long mixing times. The sequential amide connectivity can be chosen from the several $^{1}H^{N}$–$^{1}H^{N}$ NOEs that are likely to be observed for each amide since the $^{13}C^{\alpha}/^{13}C'$ shifts obtained from the 4D HNCOCA involving the sequential amide (assigned on the basis of the NOE data set) must of course match the intraresidue $^{13}C^{\alpha}/^{13}C'$ shifts from the 4D HNCACO (Fig. 9, step 1).

As a final important point, more than 95% of the expected intra- and interresidue correlations were observed in the 4D HNCACO and 4D HNCOCA, respectively, whereas more than 95% of the correlations linking the C′ of residue $i-1$ with the C^{α} of residue i were obtained from the $HNCO_{i-1}CA_{i}$.[53] The distribution of the peak signal-to-noise ratios in all three experiments, illustrating the high sensitivity, is shown in Fig. 11.

Structural Studies of MBP

Protein structure determination by NMR relies on the use of constraints derived from the nuclear overhauser effect,[59] scalar couplings,[60] and more recently dipolar couplings.[9,10] It is clear that the time-consuming step in the entire structure determination procedure lies in the assignment of NOEs from spectra that can be very significantly overlapped in the aliphatic region. A second labor-intensive step for applications involving large proteins, which precedes NOE assignment, is the assignment of ^{1}H and ^{13}C side-chain chemical shifts. The goal of the methyl protonation, ^{2}H-labeling scheme described above is to eliminate many of the time-consuming steps in the structure determination process. Backbone chemical shift assignment can be readily accomplished for many proteins using ^{15}N,^{13}C,^{2}H-triple resonance-based experiments, and, of specific importance, side-chain methyl assignments for Val, Leu, and Ile ($\delta 1$) can be obtained in a straightforward and rapid manner. Distance restraints can then be extracted linking either pairs of methyl groups of hydrophobic residues lining the core of the protein or methyls with amide groups from spectra focusing on the resolved methyl region, facilitating unambiguous assignment. These methyl NOEs, in addition to $^{1}H^{N}$–$^{1}H^{N}$ constraints, are extremely important for obtaining the global fold of the molecule.[43,44,61] In the absence of the long-range restraints provided by methyl NOEs, it is often difficult to build even a preliminary structural model.[61]

Once the fold of the molecule is obtained, additional refinement is possible, although for studies of large proteins it is not clear how successful traditional approaches based on assignment of large numbers of NOEs will be since in many cases cross-peak overlap is likely to be prohibitive. In this regard a method described by Fesik and co-workers for assignment of NOEs to phenylalanine residues

[59] K. Wüthrich, "NMR of Proteins and Nucleic Acids." John Wiley & Sons, New York, 1986.

[60] A. Bax, G. W. Vuister, S. Grzesiek, F. Delaglio, A. C. Wang, R. Tschudin, and G. Zhu, *Methods Enzymol.* **239,** 79 (1994).

[61] K. H. Gardner, M. K. Rosen, and L. E. Kay, *Biochemistry* **36,** 1389 (1997).

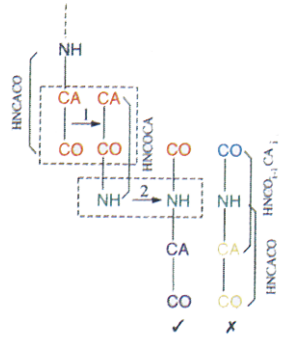


FIG. 9. Schematic diagram illustrating the ambiguities that can arise in chemical shift assignment resulting from degeneracies in $^{1}H^{N}$, ^{15}N shifts using the 4D HNCACO and 4D HNCOCA exclusively. Like spins with degenerate chemical shifts have the same color. In step 1 of the assignment the $^{13}C^{\alpha}$,$^{13}C'$ shifts of residue i from the 4D HNCACO are linked with the $^{13}C^{\alpha}$,$^{13}C'$ shifts from the 4D HNCOCA to obtain the $^{1}H^{N}$, ^{15}N shifts of residue $i+1$. Because this spin pair is not unique, it is not possible to continue assignment using the 4D HNCACO. The 4D HNCO$_{i-1}$CA$_i$ removes the ambiguity by correlating chemical shifts that bridge the degenerate $^{1}H^{N}$, ^{15}N pair. Reproduced with permission from Konrat *et al.*[53]

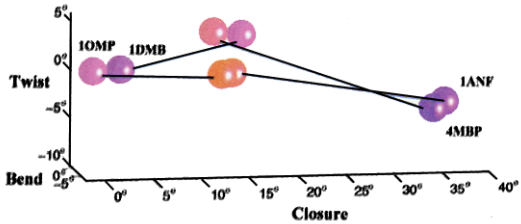


FIG. 14. Domain orientation for the N- and C-domains of MBP obtained using dipolar couplings and four different starting X-ray structures. Spheres with PDB accession numbers correspond to the starting X-Ray structures with the resulting solution structures from each X-ray structure indicated by the lines. All structures are defined using a set of rotations about three orthogonal axes denoted by closure, twist, bending in that order, relative to the reference structure, 1OMP. The closure, twist, and bending axes have polar angles (109°, 124°), (159°, 279°), and (82°, 212°), respectively, in the 1OMP frame. The coloring scheme is related to the quality of the fit of experimental dipolar coupling values to those predicted from the structure with red corresponding to the best fit. Modified from Skrynnikov *et al.*[74]

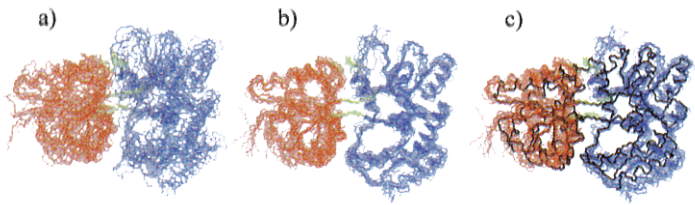


FIG. 20. The 10 lowest energy structures generated using (a) 1943 NOE, 555 dihedral and 48 hydrogen bonding restraints or (b) restraints in (a) with orientational restraints derived from dipolar couplings as defined in text for 188 residues. In (c) the backbone C^{α} trace from the X-ray structure 1DMB[70] is superimposed on the structures for comparison. Reprinted with permission from Mueller *et al.*[83]

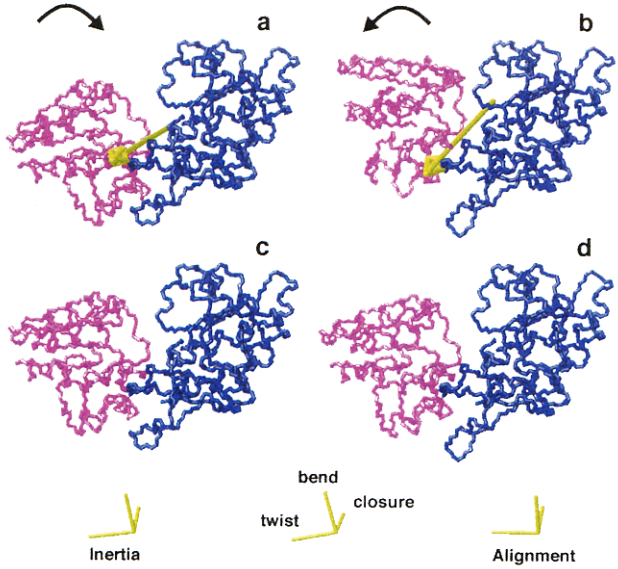


FIG. 15. Backbone representation of the MBP structures before (a,b) and after (c,d) rotations. In a and b the X-ray structures, 1OMP and 1ANF, are shown, while the corresponding solution structures (MBP with β-cyclodextrin) are indicated in c and d. The hinge axes about which rotation occurs on transforming from the X-ray to solution structures are indicated by the yellow arrows. Also shown are the inertial, alignment, and closure,twist,bend frames relative to the coordinate system of 1OMP. Reprinted with permission from Skrynnikov *et al.*[74]

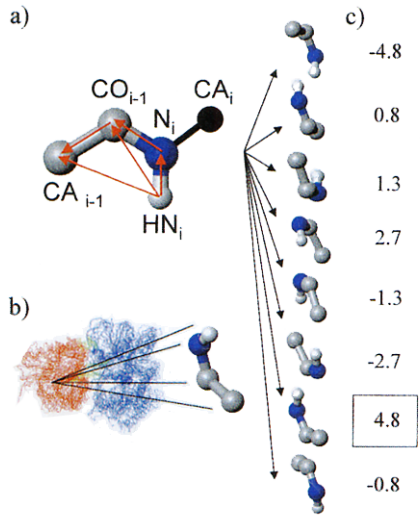


FIG. 19. Summary of the protocol used to choose between the eight possible orientations of the peptide plane established on the basis of dipolar coupling data. See text for details. Reprinted with permission from Mueller *et al.*[83]

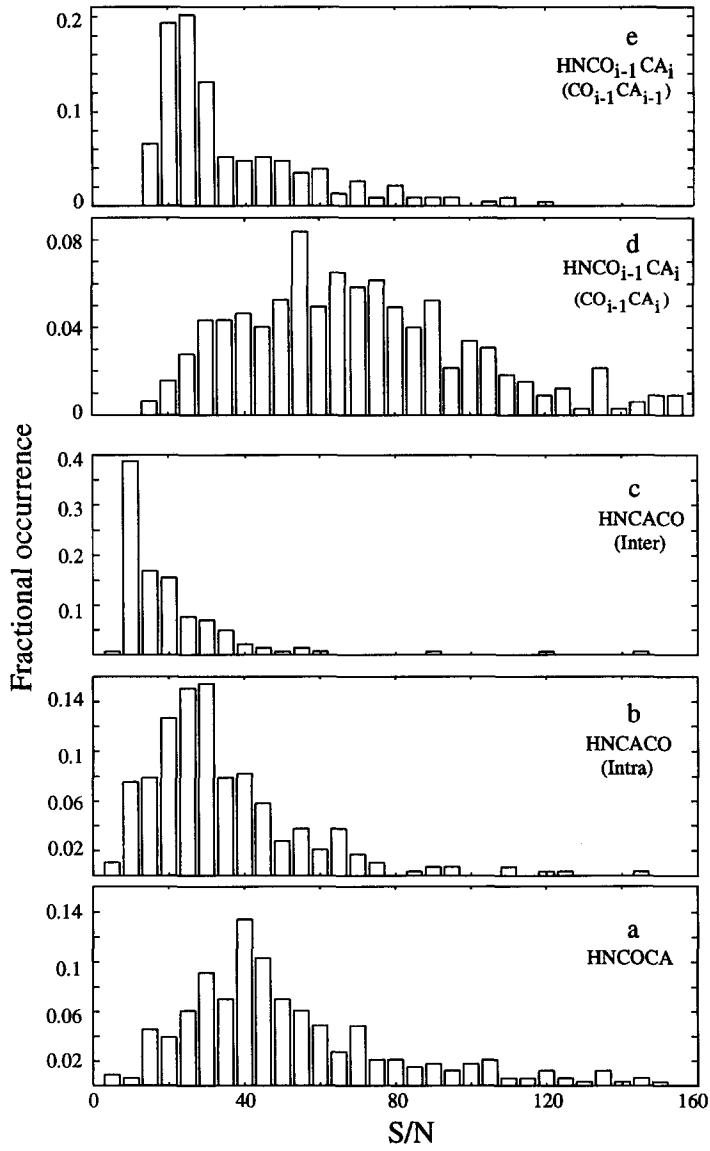


FIG. 11. Histograms showing fractional occurrence vs signal to noise (*S/N*) of the 4D HNCOCA (a), 4D HNCACO (b, intraresidue; c interresidue) and 4D HNCO$_{i-1}$CA$_i$ (d, correlations linking $^1H_i^N$, $^{15}N_i$, $^{13}C_i^\alpha$, $^{13}C'_{i-1}$; (e) correlations linking $^1H_i^N$, $^{15}N_i$, $^{13}C_{i-1}^\alpha$, $^{13}C_{i-1}$) recorded on a 1.4 m*M* sample of Val, Leu, Ile (δ1) methyl protonated, ^{15}N, ^{13}C, ^2H labeled MBP/β-cyclodextrin, 5°, 600 MHz. Note that the HNCO$_{i-1}$CA$_i$ data set was recorded with a different probe and the *S/N* values are similar for this experiment and the HNCOCA.

in proteins is likely to be very useful.[62] This group has develop a simple method for the production of ϵ-^{13}C-labeled Phe, which can be added to the growth medium in a manner similar to the precursors used for methyl labeling. Using a combination of ^{13}C-editing and ^{12}C-filtering experiments,[63,64] NOEs to the H^{ϵ} of phenylalanine residues can be assigned in proteins, providing additional important distance restraints.

A major advance in the field of biomolecular NMR in the past several years has been the development of anisotropic media for the alignment of macromolecules in solution with only a very modest decrease in spectral quality.[9,14,15] This has allowed measurement of residual dipolar couplings between pairs of NMR active spins.[9] Such information is particularly useful for large proteins and protein complexes where the number of NOE restraints available for structural studies is dramatically reduced, due to either deuteration or overlap of NOEs involving aliphatic side-chain protons or both.

It is readily shown that the dipolar splitting, D_{IM}, resulting from the proximity of two spin 1/2 nuclei I and M is given by

$$D_{IM} = -\gamma_I \gamma_M h/\left(4\pi^2 r_{IM}^3\right)\langle 3\cos^2\theta - 1\rangle \tag{5}$$

where γ_i is the gyromagnetic ratio of spin i, r_{IM} is the distance between spins I, M, and θ is the angle that the IM internuclear vector makes with respect to the magnetic field.[65] The angular brackets, $\langle\rangle$, describe the averaging that occurs because of overall molecular tumbling and internal dynamics. In the absence of alignment, $D_{IM} = 0$ since $\langle 3\cos^2\theta - 1\rangle = 0$. Equation (5) shows that dipolar couplings contain both structural and dynamical information. In order to separate these two contributions the angular part of Eq. (5) can be recast according to

$$\langle 3\cos^2\theta - 1\rangle = (16\pi/5)^{0.5}\Sigma_p Y_{2p}(\theta', \phi')\langle D_{p0}(\alpha, \beta, \gamma)\rangle \tag{6}$$

where Y_{2p} is a second-order spherical harmonic, with θ', ϕ' the polar angles of the IM vector in the molecular alignment frame, and $D(\alpha, \beta, \gamma)$ is the Wigner rotation matrix describing the transformation from the alignment frame to the laboratory frame. In Eq. (6) α, β are the polar angles describing the orientation of the z axis of the laboratory frame (i.e., the magnetic field) in the alignment frame. Expanding the sum given by the right-hand side of Eq. (6) and transforming into the principal alignment frame gives

$$D_{IM} = -\gamma_I \gamma_M h/\left(4\pi^2 r_{IM}^3\right) A_a S\{(3\cos^2\theta'' - 1) + 1.5R\sin^2\theta''\cos 2\phi''\} \tag{7}$$

[62] H. Wang, D. A. Janowick, J. M. Schkeryantz, X. Liu, and S. Fesik, *J. Am. Chem. Soc.* **121,** 1611 (1999).

[63] M. Ikura and A. Bax, *J. Am. Chem. Soc.* **114,** 2433 (1992).

[64] G. Otting, H. Senn, G. Wagner, and K. Wüthrich, *J. Magn. Reson.* **70,** 500 (1986).

[65] R. R. Ernst, G. Bodenhausen, and A. Wokaun, "Principles of Nuclear Magnetic Resonance in One and Two Dimensions." Oxford University Press, Oxford, 1987.

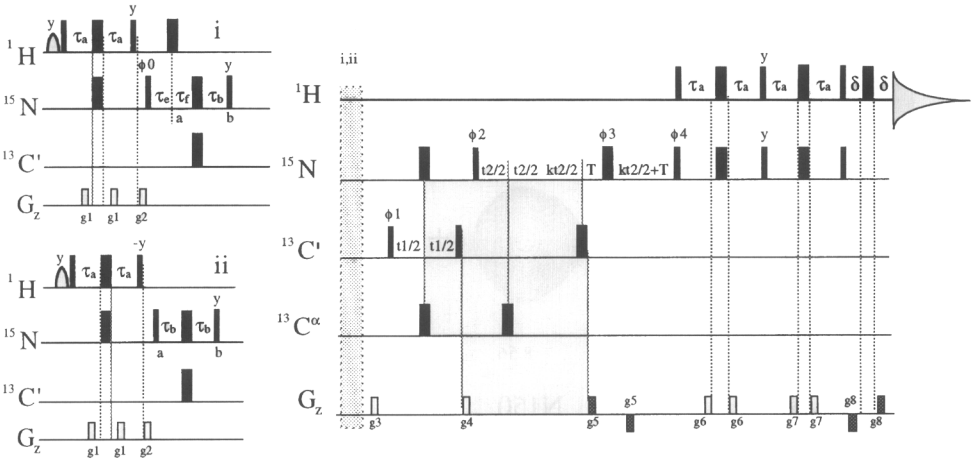


FIG. 12. E.COSY-based TROSY HNCO pulse scheme for measurement of $^1J_{NC'}$ and $^2J_{HN,C'}$ couplings. Details of the pulse sequence are reported by Yang *et al.*[67]

where θ'', ϕ'' are the polar angles of the IM vector in the principal alignment frame of the molecule, S is an order parameter that reflects isotropic averaging due to fast local dynamics, and A_a, R are the axial and rhombic components of the alignment tensor, respectively.[9] Note that A_a and R are proportional to the averages $\langle 3\cos^2\alpha' - 1\rangle$, $\langle \sin^2\alpha' \cos 2\beta'\rangle$, respectively, where α', β' are the polar angles of the magnetic field in the principal alignment frame. Equation (7) contains a clear separation of dynamics, described by the parameters A_a, R, S, from structure, given by the polar angles, θ'', ϕ''.

Bax and co-workers have described a simple 2D experiment for the simultaneous measurement of $^1H^N$–^{15}N, ^{15}N–$^{13}C'$, and $^1H^N$–$^{13}C'$ dipolar couplings in ^{15}N, ^{13}C, ^2H labeled proteins.[66] For molecules the size of MBP (370 residues) it is necessary to develop 3D analogs of this experiment that offer improved resolution. Yang *et al.* have described a suite of TROSY-based triple resonance HNCO pulse sequences for measurement of one-bond $^1H^N$–^{15}N, ^{15}N–$^{13}C'$, $^{13}C'$–$^{13}C^\alpha$, two-bond $^1H^N$–$^{13}C'$, and three-bond $^1H^N$–$^{13}C^\alpha$ dipolar couplings in highly deuterated, ^{15}N, ^{13}C labeled proteins.[67] Not surprisingly, when used in concert these couplings provide particularly valuable structural restraints, illustrated by Wang *et al.*[66]

Figure 12 shows the TROSY scheme that we have used to measure one-bond ^{15}N–$^{13}C'$ and two-bond $^{13}C'$–$^1H^N$ dipolar couplings in MBP, obtained from peak splittings in the frequency domains. The pulse sequence makes use of the E.COSY

[66] Y. X. Wang, J. L. Marquardt, P. Wingfield, S. J. Stahl, S. Lee-Huang, D. A. Torchia, and A. Bax, *J. Am. Chem. Soc.* **120,** 7385 (1998).

[67] D. Yang, R. A. Venters, G. A. Mueller, W. Y. Choy, and L. E. Kay, *J. Biomol. NMR* **14,** 333 (1999).

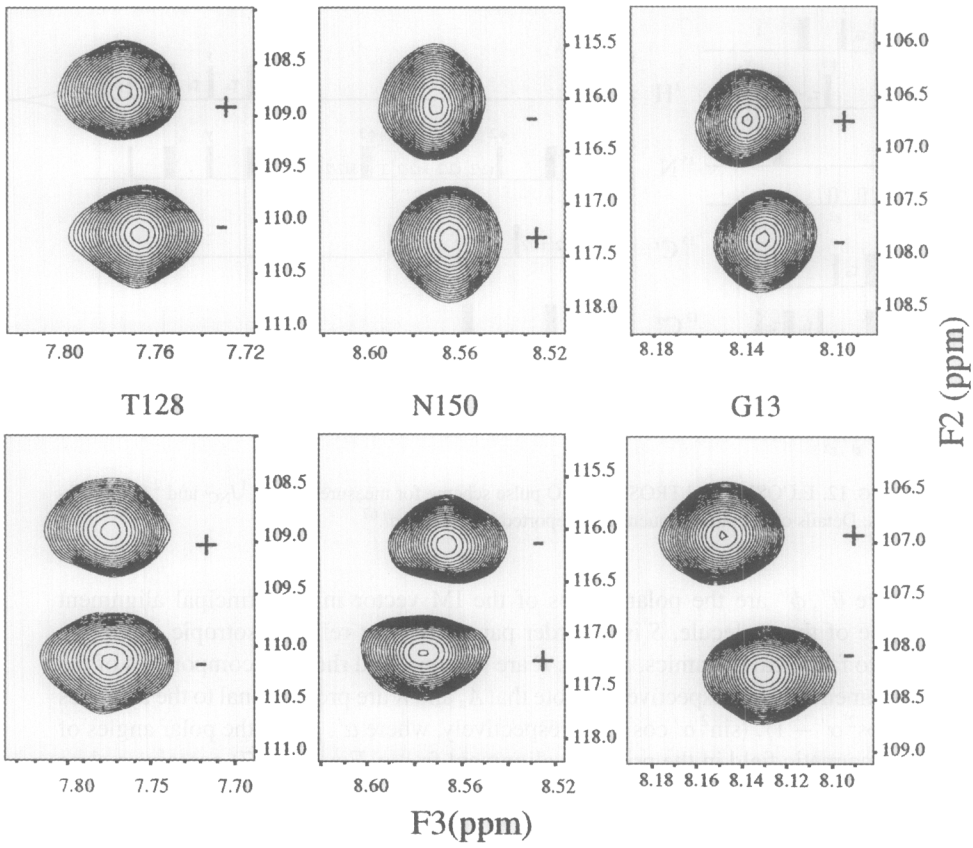


FIG. 13. Selected regions from spectra recorded on a 1.0 m*M* sample of Val, Leu, Ile ($\delta 1$) methyl protonated, ^{15}N, ^{13}C, ^{2}H-labeled MBP/β-cyclodextrin, 37°, 600 MHz without (upper panels) and with (lower panels) Pf1 phage (19 mg/ml). The splittings in F2 and F3 provide a measure of $^{1}J_{NC'}$ and $^{2}J_{HN,C'}$ couplings, respectively. The phases of the multiplet components are indicated by +/−. The components for Asn-150 are of opposite phase relative to the other peaks due to aliasing in the $^{13}C'$ dimension where the first time point is set to half the dwell. Note that the one-bond ^{15}N–$^{13}C'$ and the two-bond $^{1}H^{N}$–$^{13}C'$ scalar couplings are negative and positive, respectively. Reprinted with permission from Yang *et al.*[67]

principle,[68] with the $^{13}C'$ spin state preserved between the ^{15}N and $^{1}H^{N}$ evolution periods, t_2 and t_3, respectively. Because the ^{15}N–$^{13}C'$ coupling is small ($^{1}J_{N,C'} \sim -15$ Hz) accordion spectroscopy[69] has been employed in the ^{15}N dimension so that

[68] C. Griesinger, O. W. Sorensen, and R. R. Ernst, *J. Chem. Phys.* **85,** 6387 (1986).
[69] G. Bodenhausen and R. R. Ernst, *J. Magn. Reson.* **45,** 367 (1981).

cross peaks corresponding to the two $^{13}C'$ spin states are split by $|(1+\kappa)^1 J_{NC'}|$ and $|^2 J_{HN,C'}|$ in F_2 and F_3, respectively, where J is the sum of dipolar and scalar couplings and κ is a scaling factor that can be adjusted. Selected regions from spectra recorded on MBP using this sequence are shown in Fig. 13. The top panel in Fig. 13 illustrates correlations from a spectrum recorded in isotropic media, while the bottom traces show the corresponding regions in the case of alignment. Couplings can change quite significantly, as shown for Asn-150, where the sign of $^2 J_{HN,C'}$ is reversed on alignment. By measuring spectra with and without orientation and taking the difference in splittings obtained in each case, dipolar couplings are readily extracted. Using this sequence and others reported by Yang *et al.*[67] dipolar couplings were obtained for 280 ^{15}N–$^1H^N$, 262 ^{15}N–$^{13}C'$, 276 $^{13}C^\alpha$–$^{13}C'$, 262 $^{13}C'$–$^1H^N$, and 276 $^{13}C^\alpha$–$^1H^N$ internuclear vectors.

Orienting Domains in MBP: A Study Based on Combined NMR and X-Ray Data

MBP comprises two domains that are comparable in size connected by an antiparallel β sheet (two strands) as well as a single α helix.[70] X-ray studies have demonstrated that the relative orientation of the two domains depends on the type of ligand bound and that the transformation from one structure to another can be described in terms of simple hinge rotations, with the structure of each domain preserved[71]. On the basis of backbone chemical shifts, hydrogen exchange data, and preliminary structures calculated from NOE data exclusively (see below), the structures of each domain in solution and in the crystal state are essentially the same.[45] Hence, rigid body domain rotations can be used to obtain the solution orientation of the domains from starting X-ray coordinates.

An alternative method to X-ray crystallography for obtaining the relative orientation of domains in multidomain proteins[72,73] is extremely important. The crystalline state conformation of a protein can be affected by crystal packing forces that may be of the same order as interactions between domains. Thus, it might not be unexpected to find examples where high-resolution X-ray structures show significant differences relative to their solution state counterparts. It is far less likely, however, that the intradomain structures will vary appreciably between solution and crystal forms. An approach based on using dipolar couplings to reorient domains starting from X-ray coordinates provides a rapid method to establish average solution structures of multidomain proteins.

[70] A. J. Sharff, L. E. Rodseth, and F. A. Quiocho, *Biochemistry* **32,** 10553 (1993).
[71] A. J. Sharff, L. E. Rodseth, J. C. Spurlino, and F. A. Quiocho, *Biochemistry* **31,** 10657 (1992).
[72] J. A. Losonczi, M. Andrec, M. W. Fischer, and J. H. Prestegard, *J. Magn. Reson.* **138,** 334 (1999).
[73] M. W. Fischer, J. A. Losonczi, J. L. Weaver, and J. H. Prestegard, *Biochemistry* **38,** 9013 (1999).

Four different sets of X-ray coordinates were used to determine the average solution state conformation of MBP bound to β-cyclodextrin.[74] X-ray sets of two open conformations of the molecule (PDB accession numbers 1OMP[75] and 1DMB[70]) and two closed conformations (1ANF[76] and 4MBP[76]) where chosen. In the analysis, all structures were transformed into the 1OMP coordinate frame by superposition of C-domains and the N-domains of each structure were rotated to minimize the residual between calculated and experimental dipolar coupling values.[74] A total of eight fitting parameters are required, corresponding to three structural degrees of freedom describing the orientation of the hinge axis about which rotation occurs and the amplitude of rotation, ω, as well as five alignment parameters, A_a, R, and the three angles describing the orientation of the alignment frame [see Eq. (7)] in the chosen coordinate frame.

Figure 14 (see color plate) shows a schematic representation of results obtained from application of the conformational search algorithm described above to MBP. Each structure is represented by a sphere, color coded to indicate the χ^2 value corresponding to the level of agreement between predicted and measured dipolar coupling values. The PDB accession code is written above each of the starting X-ray conformers with solid lines connecting the X-ray structure to the corresponding dipolar coupling directed conformation. Structures are defined relative to the reference 1OMP structure via a set of three rotations about orthogonal axes, corresponding to closure, twist, and bending. The twist axis connects the centers of mass of each of the two domains, the closure axis lies in the plane formed by the twist and hinge axes, and the bending axis is formed by the cross product of vectors along the twist and closure axes. In Fig. 15 (see color plate) backbone structures of MBP before (a and b) and after (c and d) rotation are shown. Despite the fact that distinctly different starting X-ray structures are selected (1OMP and 1ANF in a and b) the resulting structures in c and d are very similar. The closure, twist, bend axes and the inertial and alignment frame for solution structure c are indicated. Note that the closure axis is nearly perpendicular to the long axis of the alignment frame. Thus the measured dipolar couplings are very sensitive to the degree of closure, with a considerably smaller degree of sensitivity to twist [via the term proportional to R in Eq. (7)].

The average solution structure of β-cyclodextrin loaded MBP is related to its crystal counterpart 1DMB via $11 \pm 1°$ closure, $1 \pm 3°$ twist, and $1 \pm 2°$ bending. The errors are estimated from the positions of the four solution conformers generated from the different starting X-ray structures. Although the extent of closure, twist, and bending is completely defined from the dipolar coupling data, the position of the hinge axis is not. The solution structures indicated in Fig. 15 were built

[74] N. R. Skrynnikov, N. K. Goto, D. Yang, W. Y. Choy, J. R. Tolman, G. A. Mueller, and L. E. Kay, *J. Mol. Biol.* **295,** 1265 (2000).

[75] A. J. Sharff, L. E. Rodseth, J. C. Spurlino, and F. A. Quiocho, *Biochemistry* **31,** 10657 (1992).

[76] F. A. Quiocho, J. C. Spurlino, and L. E. Rodseth, *Biochemistry* **31,** 10657 (1997).

using the pivot obtained by superimposing the C-domains of the crystal structures 1OMP and 1ANF and subsequently best fitting one N-domain onto the other. Because dipolar couplings are invariant when domains are rotated by 180° about any of the three alignment axes, there are four sets of hinge rotations that are consistent with the experimental input. In general, however, only one hinge rotation is of moderate amplitude, while the three others can be discarded because they are much larger.[74,77]

The analysis described above has assumed that there is a single conformation of the protein in solution. However, in the case of a two-domain protein such as MBP, where the domains reorient upon ligand binding, it is certainly possible that some level of flexibility exists. We have considered two possible models of domain motion. In the first, it is assumed that there is a rapid interconversion between open (o) and closed (c) conformations, 1OMP and 1ANF, respectively. In this case each dipolar coupling, D_{IM}, is given by $D_{IM} = p_o D_{IM}(o) + (1 - p_o) D_{IM}(c)$, where p_o is the probability of the open state and $D_{IM}(o)$, $D_{IM}(c)$ are dipolar couplings in the open and closed states, respectively. Alignment parameters for each of the open and closed states and p_o are determined from an 11-parameter fit. A best solution of $p_o A_a(o)/(1 - p_o)A_a(c) = 1.67$ is obtained, where $A_a(o)$ and $A_a(c)$ are the A_a values for the open and closed states, respectively [Eq. (7)]. Since the open and closed states differ by a closure of approximately 35°, this analysis suggests an average solution structure that is 13° more closed than 1OMP, assuming equal degrees of alignment in open and closed states. This is in complete agreement with results obtained from an analysis of the data based on the assumption of a single solution conformation, illustrated in Fig. 15.

A second motional model is one in which each of the domains of MBP rotates equally about a common hinge axis. In this case it can be shown that although the alignment parameters are affected by the motion, the average orientation of domains is not. Because the two domains of MBP are approximately equal in size, this model is likely to be quite realistic. In the less likely scenario that amplitudes of fluctuation of each domain differ, it can be shown that both alignment parameters and the average relative orientation of the domains will be affected. For example, in the static model, a hinge axis was determined about which the N-domain of MBP was rotated to generate the solution structure. Assuming further that the N-terminal domain fluctuates about this axis by $\pm 20°$ while the C-terminal domain is static, the average solution conformation is characterized by $12 \pm 2°$ closure, $-3 \pm 6°$ twist, and $-2 \pm 3°$ bending relative to 1DMB. These numbers are only slightly different from those determined using a model that assumes a static structure ($11 \pm 1°$, $1 \pm 3°$, and $1 \pm 2°$ closure, twist, and bend, respectively) and illustrate that for moderate differences in domain dynamics it is still possible to obtain the average solution structure.

[77] R. Brüschweiler, X. Liao, and P. E. Wright, *Science* **268,** 886 (1995).

A Global Fold of MBP Based on a Limited Set of NOEs and Dipolar Couplings

As described above, dipolar couplings can be used in concert with methyl–methyl, methyl–$^1H^N$, and $^1H^N$–$^1H^N$ NOEs to generate global folds of proteins. As with any structure determination relying on distance constraints, it is important that as many NOEs as possible be assigned. However, because only a limited subset of ^1H–^1H distances can be obtained in highly deuterated, methyl protonated molecules assignment is even more critical. With this in mind a number of experiments have been developed for measuring methyl NOEs in ^{15}N, ^{13}C, ^2H methyl-protonated proteins. These methods exploit the narrow linewidths of CH_3 groups and the significantly improved ^{13}C relaxation times of nonmethyl carbon spins where protons are replaced by deuterons.

In the first of the NOE experiments, (HM)CMCB(CMHM)-NOESY,[78] the ^{13}C chemical shift of the carbon adjacent to the methyl is recorded in constant-time mode[37,38] to improve resolution. The flow of magnetization can be described according to

$$^1H^m \xrightarrow{J_{CH}} {}^{13}C^m(CTt_1) \xrightarrow{J_{CC}} {}^{13}C^i(CTt_2) \xrightarrow{J_{CC}} {}^{13}C^m \xrightarrow{J_{CH}} {}^1H^m \xrightarrow{NOE} {}^1H(t_3) \quad (8)$$

Proton magnetization originating on the methyl group is transferred to the methyl carbon, where chemical shift is recorded during a constant-time evolution period of duration $1/2J_{CC}$. During this interval evolution proceeds due to the one-bond $^{13}C^m$–$^{13}C^i$ coupling so that at the conclusion of the CT delay, transverse $^{13}C^i$ magnetization is generated. As described in detail by Zwahlen *et al.*,[78] it is possible to design this second CT period so that evolution of magnetization from all one-bond ^{13}C–$^{13}C^i$ couplings is refocused. The duration of the second CT interval is therefore not restricted to multiples of $1/J_{CC}$. Subsequently magnetization is transferred back to $^1H^m$ and then to proximal proton spins during an NOE period. Thus, correlations are observed at $(\omega_C^m, \omega_C^i, \omega_H^m)$, corresponding to diagonal peaks, and at $(\omega_C^m, \omega_C^i, \omega_H)$ (cross peaks). Figure 16 shows slices from the 100 ms mixing time NOESY recorded on Val, Leu, Ile ($\delta 1$) methyl-protonated, ^{15}N, ^{13}C, ^2H MBP/β-cyclodextrin showing NOEs involving Val-97 $H^{\gamma 1}/H^{\gamma 2}$, extracted at the $^{13}C^\beta$ shift of Val-97. The NOEs in this figure correlate the methyls of five residues and involve amino acids located on different β strands of two opposing faces of a β sandwich.

In a second NOE experiment,[79] methyl–methyl and methyl–$^1H^N$ distances are obtained using a simple sequence that relays magnetization according to

$$^1H_i^m \xrightarrow{J_{CH}} {}^{13}C_i^m(CTt_1) \xrightarrow{J_{CH}} {}^1H_i^m \xrightarrow{NOE} {}^1H_j^m \xrightarrow{J_{CH}} {}^{13}C_j^m(t_2) \xrightarrow{J_{CH}} {}^1H_j^m(t_3) \quad (9)$$

[78] C. Zwahlen, S. J. F. Vincent, K. H. Gardner, and L. E. Kay, *J. Am. Chem. Soc.* **120,** 4825 (1998).

[79] C. Zwahlen, K. H. Gardner, S. P. Sarma, D. A. Horita, R. A. Byrd, and L. E. Kay, *J. Am. Chem. Soc.* **120,** 7617 (1998).

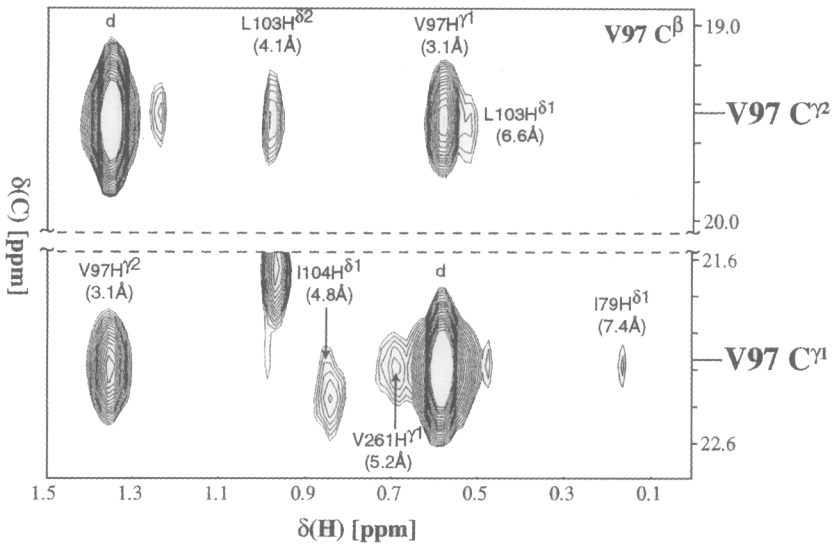


FIG. 16. Methyl–methyl NOEs involving Val 97 $H^{\gamma 1,\gamma 2}$ observed in a 100 ms (HM)CMCB (CMHM)-NOESY recorded on a sample of 1.4 m*M* Val, Leu, Ile ($\delta 1$) methyl protonated, ^{15}N,^{13}C,^{2}H-labeled MBP/β-cyclodextrin at 37°, 600 MHz. The diagonal peaks are identified by "d," with the destination methyl group indicated above each of the cross peaks, along with the distance between pseudoatoms located at the average coordinates of each set of methyl protons. Reprinted with permission from Zwahlen *et al.*[78]

where $^{1}H_i^m$ and $^{1}H_j^m$ are proximal methyl protons. In this case the CT t_1 period is set to $1/J_{CC}$. Fourier transformation of the resultant data set yields cross peaks at $(\omega_{Ci}^m, \omega_{Cj}^m, \omega_{Hj}^m)$, $i \neq j$, and diagonal peaks at $(\omega_{Ci}^m, \omega_{Ci}^m, \omega_{Hi}^m)$. It is also possible to simultaneously record NOEs between proximal methyls and amide protons, as described by Zwahlen *et al.*[79] Figure 17 illustrates F_1–F_3 strips from a 100 ms NOESY spectrum recorded on MBP, with CT acquisition in both ^{13}C dimensions.

In addition to the data sets described above, 4D ^{15}N–^{15}N[57,58] and ^{15}N–^{13}C edited[80] NOESYs were recorded on MBP. On the basis of the NOEs obtained from analysis of all spectra, 555 ϕ, ψ dihedral angle restraints generated to a large extent from the program TALOS,[81] and 48 hydrogen bonding restraints (derived from hydrogen exchange data[45] and included only for residues in regular secondary structure), preliminary structures were produced from a simulated annealing X-PLOR protocol and used in the assignment of further NOEs. A total of 348 methyl–methyl, 769 methyl–$^{1}H^N$, and 826 $^{1}H^N$–$^{1}H^N$ NOEs were assigned

[80] D. R. Muhandiram, G. Y. Xu, and L. E. Kay, *J. Biomol. NMR* **3,** 463 (1993).
[81] G. Cornilescu, F. Delaglio, and A. Bax, *J. Biomol. NMR.* **13,** 289 (1999).

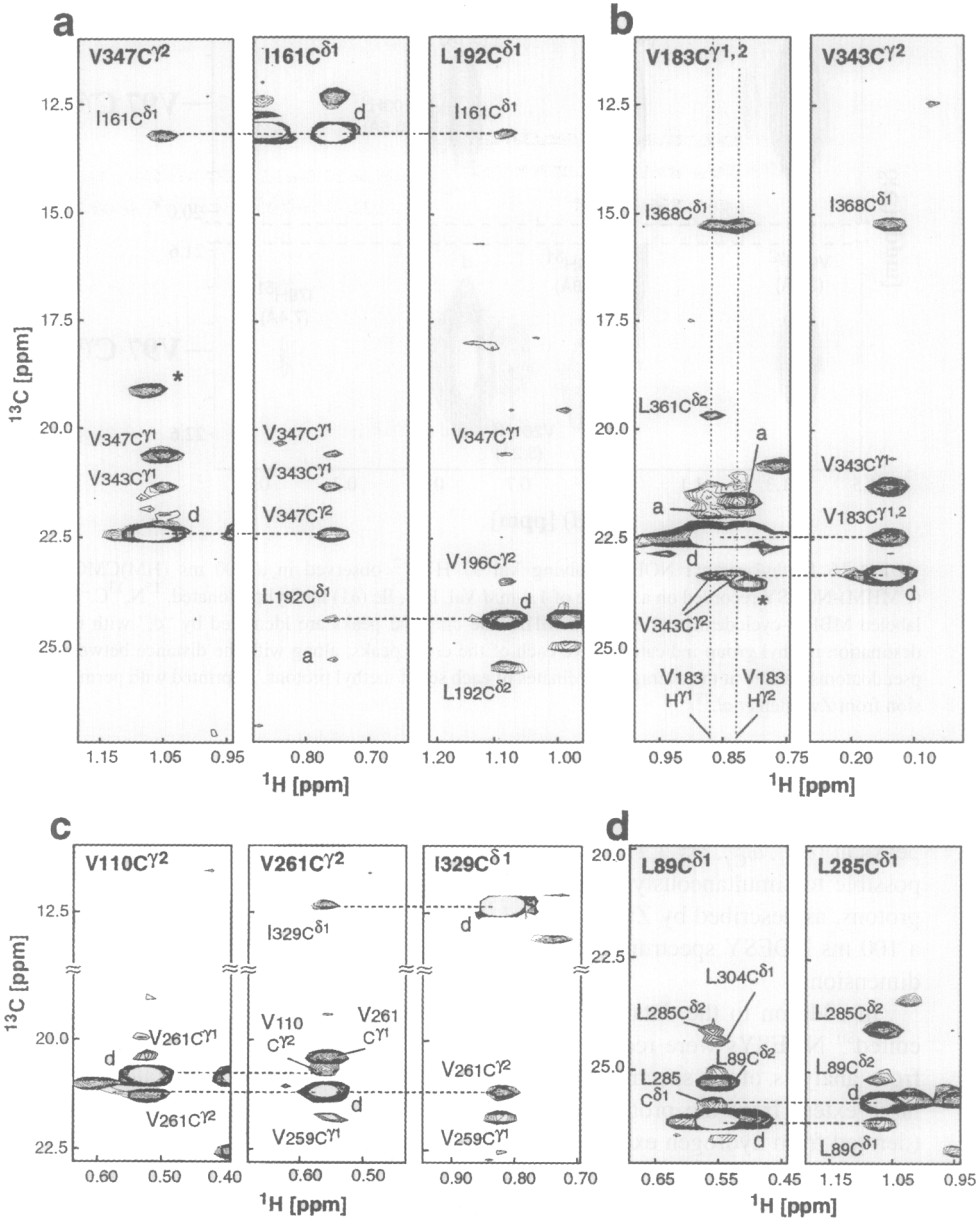


FIG. 17. Strips from the 3D $^{13}C,^{13}C$ methyl NOESY spectrum recorded on a sample of 1.4 m*M* Val, Leu, Ile ($\delta 1$) methyl protonated, $^{15}N,^{13}C,^{2}H$-labeled MBP/β-cyclodextrin at 37°, 600 MHz, 100 ms mixing time. Diagonal peaks are marked with "d," ambiguous peaks with "a," and cross peaks with greater intensity on an adjacent slice are indicated by an asterisk. Symmetry related correlations are connected via dashed horizontal lines. Reprinted with permission from Zwahlen *et al.*[79]

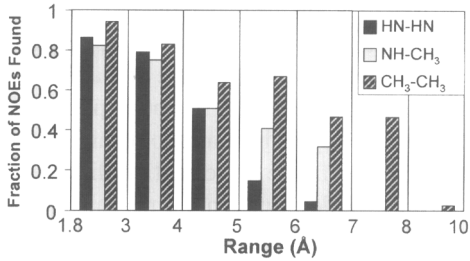


FIG. 18. Fraction of NOEs observed in spectra recorded on a sample of 1.4 m*M* Val, Leu, Ile ($\delta 1$) methyl protonated, ^{15}N,^{13}C,^{2}H-labeled MBP/β-cyclodextrin at 37° as a function of the distance range calculated on the basis of the X-ray structure IDMB.[70]

using this procedure. A summary of the NOEs obtained relative to those expected on the basis of the X-ray structure of β-cyclodextrin loaded MBP, 1DMB, is presented in Fig. 18. Structures with a pairwise global backbone rmsd of 5.7 Å were obtained at this stage, with rmsd values of 2.4 and 3.9 Å for the N and C domains, respectively.

In addition to calculating structures using only NOE, dihedral angle, and hydrogen bonding restraints, structures were also generated by including all of the above restraints in a protocol that directly refines against one-bond ^{15}N–$^{1}H^{N}$, ^{15}N–$^{13}C'$, $^{13}C^{\alpha}$–$^{13}C'$, two-bond $^{13}C'$–$^{1}H^{N}$, and three-bond $^{13}C^{\alpha}$–$^{1}H^{N}$ dipolar couplings. Despite numerous attempts, we were unsuccessful in obtaining structures with low numbers of violations in the dipolar coupling restraints. Prestegard and coworkers have also noted that for a two-domain lectin where the density of NOEs is low, direct refinement against dipolar couplings did not produce a set of structures with the same relative domain orientation.[73] These convergence problems likely result from the fact that for a given dipolar coupling there are an infinity of orientations for the dipole vector that are consistent with the measured coupling value. In total the couplings lead to an extremely complex energy surface and the global minimum is very difficult to find. Clearly structures with a low density of NOEs (in the case of MBP there are ~ 5 NOEs/residue) will suffer the most, since the NOEs are insufficient to "guide" the dipolar vectors to their correct orientation. Notably, in cases where >10 restraints/residue are obtained, direct refinement against dipolar couplings can lead to converged structures with improved precision.[73,82]

The lack of success in using direct refinement against dipolar couplings to improve the calculated structures of MBP led to the development of a different protocol,[83] illustrated in Fig. 19 (see color plate). In some respects the approach is

[82] G. M. Clore, M. R. Starich, C. A. Bewley, M. Cai, and J. Kuszewski, *J. Am. Chem. Soc.* **121,** 6513 (1999).

[83] G. A. Mueller, W. Y. Choy, R. A. Venters, D. Yang, J. D. Forman-Kay, and L. E. Kay, *J. Mol. Biol.* **300,** 197 (2000).

not unlike the method used to orient the two domains of MBP starting from crystal structures,[74] described above. In that case the structure of each of the two domains was fixed and only the relative orientation of the domains was allowed to vary. In the present case the fixed structural unit is a peptide plane, indicated in Fig. 19a. For each plane in the protein the goal becomes to obtain the Euler angles (α,β,γ), which transform from the initial frame to the peptide alignment frame. Assuming a rigid structure, each of the peptide alignment frames must coincide with the global alignment coordinate system so that this procedure effectively rotates each peptide into its proper orientation in the overall structure.

In principle there are five degrees of freedom associated with the choice of a suitable alignment frame, including the three Euler angles described above and two order parameters A_a and R [see Eq. (7)]. Following the approach of Clore *et al.*,[84] the distribution of measured dipolar couplings is used to extract A_a and R, leaving three parameters $[(\alpha,\beta,\gamma)]$ to be determined for each plane. These Euler angles can be obtained by a grid search procedure that establishes the orientation of the peptide alignment frame by minimizing the difference between experimental dipolar couplings and those predicted from a "trial" alignment frame. At least three of the five dipolar couplings illustrated in Fig. 19 are required for this approach; in the case of MBP three or more couplings could be obtained for 240 residues.

Unfortunately, the orientation of a peptide plane in its alignment frame, as defined by the dipolar coupling data, is not unique. Rotation of the plane by 180° about any one of the three alignment axes produces a structure that is consistent with the dipolar data. Moreover, for each of the four orientations it is possible to rotate the plane about its normal by 180° to generate additional orientations that satisfy the dipolar data. This is illustrated in Fig. 19b where the eight possible orientations of the peptide plane bridging residues Phe-149 and Asn-150 in MBP are shown. In order to choose the proper orientation from the eight possibilities, we make use of the preliminary NMR structures derived on the basis of NOEs and dihedral angles, shown in Fig. 19. An average structure is calculated from the initial structures and the measured dipolar couplings are used to define an overall molecular alignment frame. The average structure is subsequently rotated into this alignment frame, and the orientation of the peptide plane from the average structure is compared with the corresponding eight orientations generated from the starting peptide plane in Fig. 19a. The level of agreement between the orientation of a plane isolated from the average structure and the eight possible planes derived from the dipolar coupling data exclusively is assessed by evaluating the dot product of the five dipolar vectors highlighted in red in Fig. 19a for each of the eight structures with the corresponding vectors from the average structure. The plane giving the largest dot product is used to provide restraints for the dipolar

[84] G. M. Clore, A. M. Gronenborn, and A. Bax, *J. Magn. Reson.* **113,** 216 (1998).

vectors in subsequent structural refinements (providing the product is >3.5) using a new module in CNS/X-PLOR that has been written for this purpose.[83] The dot products obtained for each of the planes are indicated in Fig. 19c; a product of 4.8 is obtained for the best orientation in this example. Details of the methodology are described in Mueller *et al.*[83]

Figure 20 (see color plate) illustrates the improvements in the quality of structures that are obtained by including dipolar couplings in the refinement procedure using the approach described above. In all cases the ensemble of 10 lowest energy structures is shown, with residues 6–370 superimposed. Structures obtained from refinement using only NOEs, dihedral angle and hydrogen bonding restraints are shown in Fig. 20 (a), while structures generated by including orientational restraints generated from dipolar couplings for 188 residues are illustrated in Fig. 20b. In Fig. 20c the X-ray derived structure of MBP with β-cyclodextrin[70] (1DMB) is superimposed on the NMR structures for comparison.

Concluding Remarks

The present review describes the application of recently developed NMR methods to study the solution structure of maltose binding protein in complex with β-cyclodextrin. The size of the protein and the corresponding spectral complexity have led to the development of new labeling methodology and new NMR pulse sequences for both chemical shift and NOE assignment. In addition, a new protocol for incorporation of dipolar coupling data in cases where only limited NOE constraints are available has been described. The methodology outlined will facilitate the rapid determination of global folds of high molecular weight proteins, and these structures can be subsequently refined using further experimental or data base information.

Acknowledgments

L.E.K. is extremely grateful to all of his coworkers who have contributed to the maltose binding protein project over the past several years. Funding for this research has been provided by the Medical Research Council of Canada. L.E.K. is a foreign investigator of the Howard Hughes Medical Research Institute.