157

# Multidimensional triple resonance NMR spectroscopy of isotopically uniformly enriched proteins: a powerful new strategy for structure determination

Ad Bax, Mitsuhiko Ikura, Lewis E. Kay, Gaetano Barbato and Silvia Spera

*Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, Maryland 20892, USA*

*Abstract.* A procedure is described that affords complete $^{1}H$, $^{13}C$ and $^{15}N$ resonance assignment in proteins of up to about 25 kDa. The new approach requires uniform isotopic enrichment of the protein with $^{13}C$ and $^{15}N$ and correlates resonances of adjacent nuclei using the relatively large and well-resolved one-bond J couplings. Spectral overlap, a common problem in the application of two-dimensional NMR, is removed by increasing the dimensionality of the new methods to three or four, without increasing the number of observed resonances. With complete $^{1}H$, $^{13}C$ and $^{15}N$ resonance assignments available, the nuclear Overhauser effect (NOE)-based interproton distance constraints can be extracted in a very straightforward manner from four-dimensional NOE spectra.

*1991 Protein conformation. Wiley, Chichester (Ciba Foundation Symposium 161) p 108–135*

Assignment of all $^{1}H$ resonances of a protein is a prerequisite for determining its conformation in solution by NMR spectroscopy. This resonance assignment process is typically broken up in two parts: sequential assignment of the backbone protons and assignment of all the side chain proton resonances. Once assignments are available, short interproton distances (< 5 Å) can be measured using the nuclear Overhauser effect (NOE). Provided a sufficient number of interproton distances can be determined, sophisticated computer algorithms can be used for determining the ensemble of protein conformations compatible with these distances. Until recently, resonance assignment was accomplished by means of homonuclear $^{1}H$ two-dimensional experiments that identify intraresidue through-bond $^{3}J(NH,H\alpha)$ connectivity and sequential inter-residue through-space (NOE) connectivity (for reviews, see Wüthrich 1986, Kaptein et al 1988, Clore & Gronenborn 1989, Bax 1989). Use of the inter-residue NOE is essential





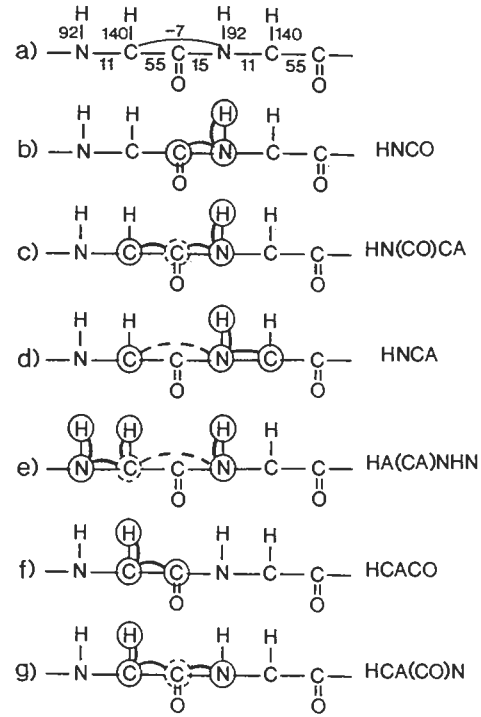
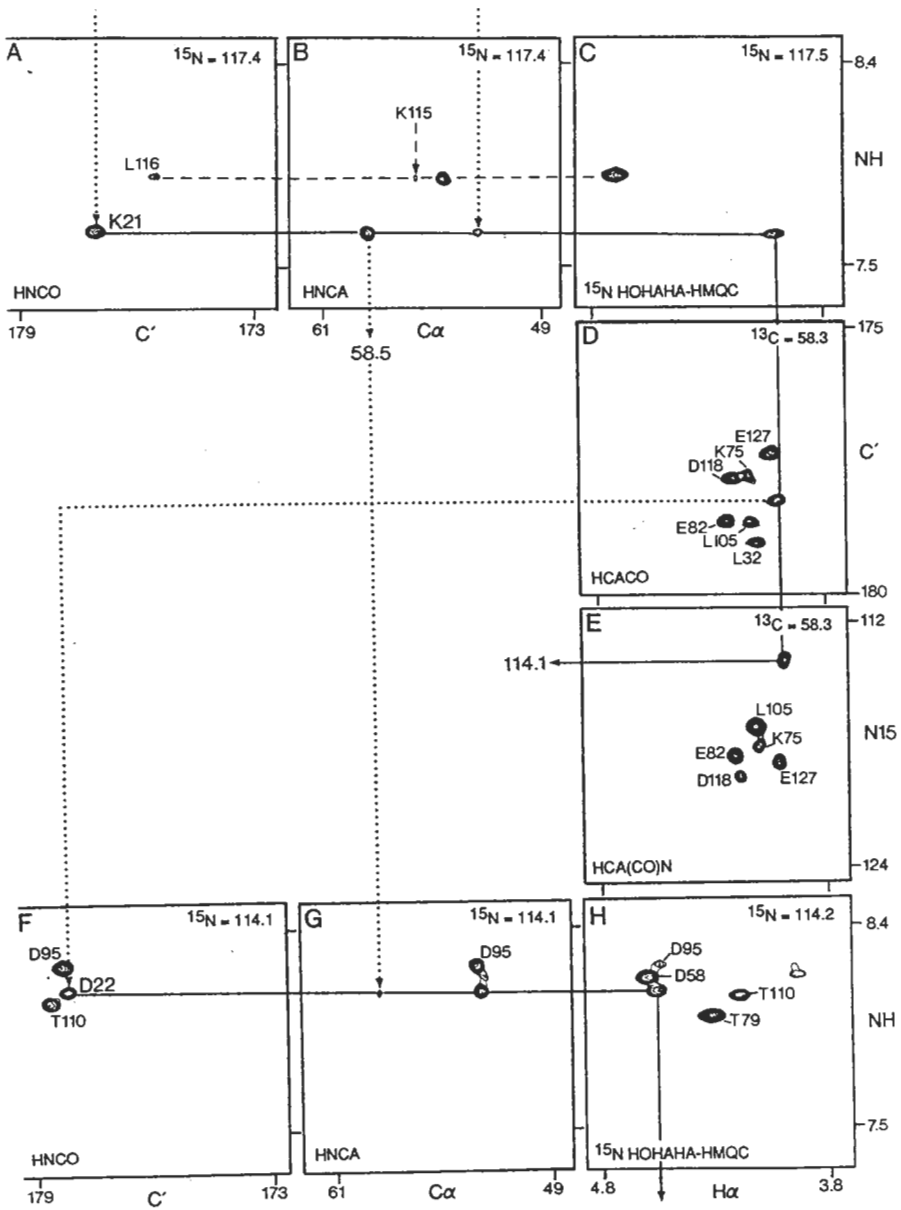



FIG. 1. Backbone atoms of two adjacent amino acids. (a) Size of relevant J couplings, in Hz; (b–g) connectivity diagrams indicating the correlations between different backbone atoms that can be observed with the pulse sequences shown on the right (CA, α-carbon). Circles mark the correlated resonances, with the heavy curved lines indicating the connectivity pathways. The dashed lines indicate transfer via the two-bond $J_{NC\alpha}$ coupling; all other transfers occur via one-bond couplings. Dashed circles mark atoms that participate as 'relay' nuclei in the magnetization transfer pathway; the resonance frequencies of these nuclei are not observed.

in this approach because of the absence of a significant $^{1}H$–$^{1}H$ J coupling between protons of adjacent amino acids. The strength of short range inter-residue NOE interactions depends strongly on the local conformation. In addition, many of these protons can also exhibit long-range NOE interactions, making unambiguous identification of sequential NOEs even more difficult.

In recent years, the sequential assignment procedure has been applied successfully to a large number of small proteins. For larger proteins (molecular mass > 10 kDa), or for proteins with a very narrow chemical shift distribution of the backbone proton resonances, the standard sequential assignment procedure may not yield unambiguous answers because of very exentsive overlap in critical regions of the $^{1}H$ two-dimensional NMR spectra. This overlap






problem can be alleviated dramatically by the recording of isotope-edited two-dimensional NMR experiments on proteins in which specific amino acids are labeled with $^{15}$N, $^{13}$C, $^{2}$H, or a combination thereof. Although this latter approach is extremely powerful, it also is very labour-intensive because of the large number of different protein preparations and NMR spectra that are needed (McIntosh et al 1987, Senn et al 1987, LeMaster & Richards 1988, Torchia et al 1989).

Uniform enrichment of the protein with $^{15}$N permits dispersion of all NOE and J interactions involving amide protons into a third frequency dimension, the $^{15}$N chemical shift (Fesik & Zuiderweg 1988, Marion et al 1989a,b). This approach dramatically reduces the spectral overlap problem, but for proteins larger than about 15 kDa this method on its own is frequently insufficient for obtaining complete resonance assignments. The potential for obtaining sequential assignment information through the use of J connectivities between backbone atoms has long been recognized. Several groups have developed two-dimensional NMR methods relying on these J connectivities that can aid in making the necessary proton resonance assignments (Oh et al 1988, Niemczura et al 1989). Recently, we have extended this idea by correlating not just the heavy backbone atoms, or the protons, to their directly attached $^{13}$C or $^{15}$N nucleus, but by the development of techniques that simultaneously correlate selected heteronuclear and proton chemical shifts in three or four orthogonal frequency dimensions (Ikura et al 1990a, Kay et al 1990a,b,c). The logic behind this new strategy is outlined below.

## Sequential assignment of backbone atoms

Although the assignment process for the backbone atoms requires knowledge of the type of amino acid for at least some residues, the assignments of backbone and side chain atoms are treated separately. Figure 1 shows the backbone nuclei for two amino acids of a polypeptide and indicates the types of connectivities that can be observed in six separate three-dimensional NMR experiments. Nuclei encircled in each of the diagrams of Fig. 1 are the nuclei whose chemical shifts

FIG. 2. Selected regions of slices from five separate three-dimensional NMR experiments recorded for the protein calmodulin. These regions illustrate the J correlation between Lys-21 and Asp-22. Solid and dotted lines trace the connectivity patterns for these two residues. Broken lines correspond to parts of the connectivity patterns observed for other residues. Slices A, B and C are taken at the Lys-21 $^{15}$N chemical shift. Slices D and E are taken at the Lys-21 $C_\alpha$ shift, observed in B. Slices F, G and H are taken at the $^{15}$N frequency of Asp-22, as measured in E. The analysis of the connectivity patterns is discussed in the text. No baseline correction or any other cosmetic procedures were used for any of the three-dimensional spectra. From Ikura et al (1990a).

are being correlated along orthogonal axes of the three-dimensional NMR spectrum. Nuclei with dashed circles around them participate in the magnetization transfer pathway but their chemical shift is not correlated to the other nuclei. Broken lines indicate correlations that occur via two-bond J coupling between the nitrogen of one residue and the $\alpha$-carbon of the preceding residue. The intensity of these two-bond correlations rapidly decreases when the molecular tumbling time, $\tau_c$, exceeds 10 ns. All other correlations occur via the relatively large one-bond J couplings between directly bound nuclei. Fig. 1d shows, for example, that the resonance frequencies of intraresidue amide hydrogen (HN), N and $C_\alpha$ nuclei can be correlated with one another (HNCA spectrum; CA, $\alpha$-carbon). In addition, the dashed line in this diagram indicates that the inter-residue connectivity between HN, N and the $\alpha$-carbon of the preceding residue can also be observed (provided that $\tau_c$ is less than about 10 ns). Panels B and G of Fig. 2 illustrate two cross sections taken from the HNCA spectrum of calmodulin, and show the correlations between the $C_\alpha$ and HN resonances for residues with a $^{15}N$ chemical shift close to 117.4 p.p.m. (Fig. 2B) and close to 114.1 p.p.m. (Fig. 2G). Similarly, the other types of connectivities outlined in Fig. 1 can be observed in other panels of Fig. 2.

Using Fig. 2 we shall briefly outline how the assignment procedure could be executed. The HNCO spectrum (Fig. 2A) correlates the frequency of the Asp-20 carbonyl with the amide $^1H$ and $^{15}N$ chemical shifts of Lys-21. These amide shifts are correlated with the Lys-21 $H_\alpha$ shift in the cross section of the three-dimensional spectrum shown in Fig. 2C. Note that the three-dimensional spectrum of which a slice is shown in Fig. 2C (and in Fig. 2H) was recorded with the older homonuclear Hartmann–Hahn heteronuclear multiple quantum correlation (HOHAHA-HMQC) technique (Marion et al 1989b). The more recent H(CA)NHN technique (Fig. 1e) is a more efficient alternative for obtaining this type of spectrum (Kay et al 1991). Once the $H_\alpha$ and $C_\alpha$ chemical shifts of Lys-21 are known, one can immediately find the carbonyl (C′) chemical shift of this residue by inspection of the HCACO spectrum, which correlates $H_\alpha$, $C_\alpha$ and C′ chemical shifts. The slice of the HCACO spectrum taken at the $C_\alpha$ shift of Lys-21 shows a correlation at the Lys-21 $H_\alpha$ shift to a C′ resonance at 178.3 p.p.m. In addition, the HCA(CO)N spectrum shows that this $H_\alpha$–$C_\alpha$ pair correlates with a 114.1 p.p.m. $^{15}N$ shift for the next residue. Inspection of a slice taken from the HNCO spectrum, perpendicular to the $^{15}N$ axis at an $^{15}N$ chemical shift of 114.1 p.p.m. (Fig. 2F), shows a correlation to the same C′ resonance as observed in Fig. 2D, thus identifying the HN shift of Asp-22. The HNCA spectrum shows the $C_\alpha$ correlation that corresponds to this amide (Fig. 2G), and also shows a weak correlation to the $\alpha$-carbon of Lys-21, confirming the assignment. In practice, for calmodulin, this procedure can be used to link unambiguously as many as 10 to 15 residues. Ambiguities occur, however, if two or more residues have identical amide $^1H$ shifts and identical amide $^{15}N$ shifts, or, as is more common, if they have identical $H_\alpha$ and $C_\alpha$ chemical shifts. At this stage, some knowledge of the amino acid side chains is required to anchor the chain of residues in the correct position along the polypeptide backbone. Of course, the procedure outlined above can be easily automated and most of the assignment process is done by simple computer programs.



## Assignment of the side chain resonances

Assignment of the side chain resonances is conventionally done using $^1H$–$^1H$ J correlation techniques. For proteins larger than about 10 kDa two-dimensional methods based on this J correlation start suffering from severe spectral overlap. In principle, one should expect that isotopic enrichment with $^{13}C$ and spreading the $^1H$–$^1H$ J correlation into an orthogonal frequency dimension (corresponding to the $^{13}C$ chemical shift) would solve this problem. However, incorporation of $^{13}C$ into the protein causes a large increase in the $^1H$ line-width (due to the $^{13}C$–$^1H$ dipolar interaction) which makes homonuclear $^1H$–$^1H$ J correlation techniques ineffective for larger $^{13}C$-enriched proteins.

A more efficient pathway to connect the side chain proton resonances utilizes one-bond couplings, in a similar manner to the assignment procedure described above for the protein backbone nuclei (Fesik et al 1990, Kay et al 1990b, Bax et al 1990). Figure 3 shows the size of the relevant J couplings. Very efficient magnetization transfer can be obtained in three steps: first from a proton to the $^{13}C$ nucleus to which it is directly attached, second from this $^{13}C$ to a second $^{13}C$ spin in the same side chain (possibly via intermediate $^{13}C$ nuclei) and finally back to the proton attached to this second $^{13}C$. This type of technique, most easily executed as a three-dimensional experiment, allows for very straightforward assignment of both $^1H$ and $^{13}C$ resonances of the amino acid side chains.

Figure 4 shows a cross-section from a so-called HCCH-TOCSY (total correlation spectroscopy) spectrum, recorded for the protein calmodulin. This slice shows $^1H$–$^1H$ correlations for residues for which at least one $^{13}C$ resonates at 66.8, 43.0 or 19.2 p.p.m. Correlated resonances are connected by horizontal lines. The $^{13}C$ chemical shifts and the side chain patterns observed are usually indicative of the type of amino acid involved. For example, the top trace in Fig. 3 connects two methyl groups at 0.7 and 0.8 p.p.m. with a proton at 2.4 p.p.m. and a proton at 4.5 p.p.m.; this is a clear signature of a valine residue. For the vast majority

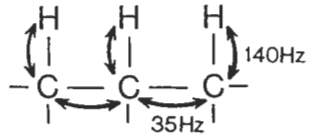



FIG. 3. Connectivity diagram for the so-called HCCH-type experiments (Kay et al 1990b, Fesik et al 1990, Bax et al 1990) which correlate side chain resonances utilizing one-bond $J_{CH}$ and $J_{CC}$ couplings. Arrows mark the magnetization transfer pathways.

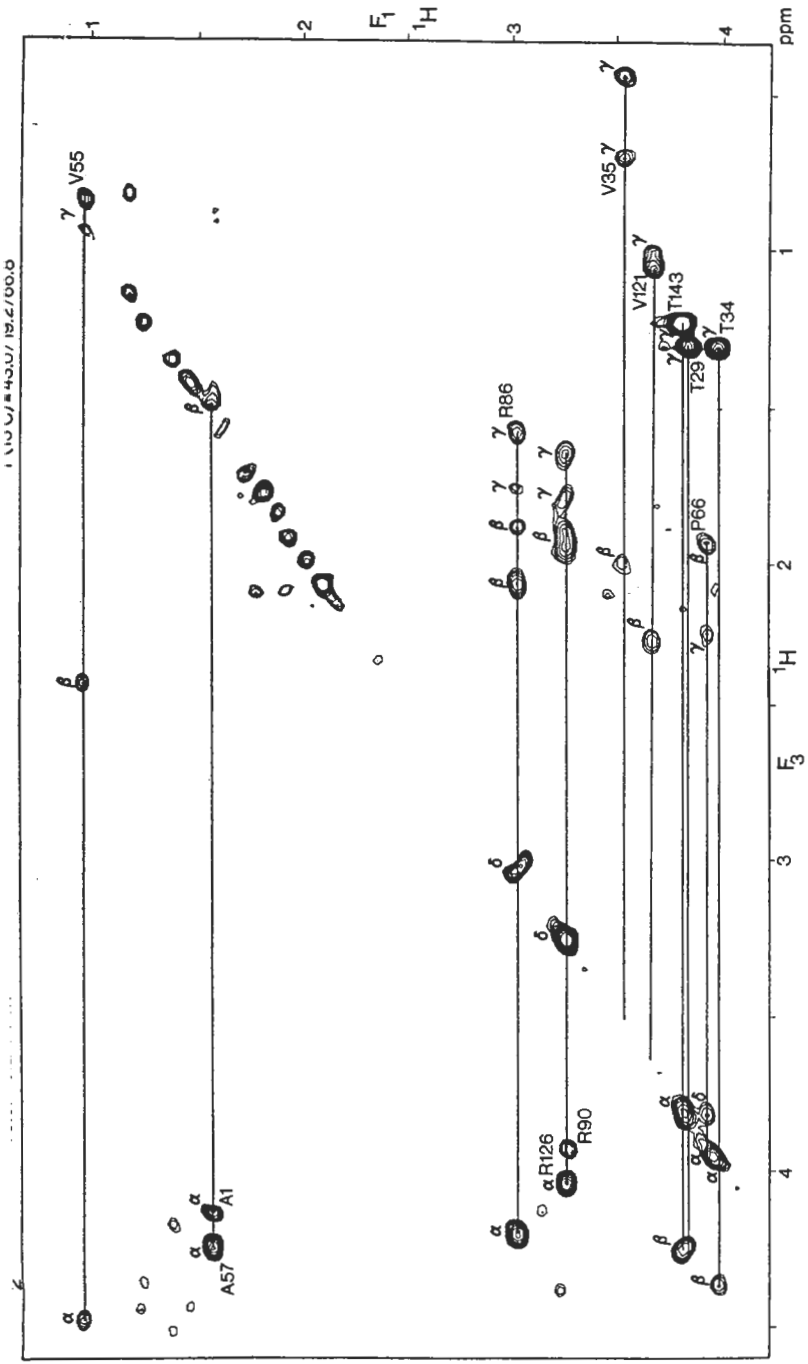



FIG. 4. Example of a slice taken from the three-dimensional HCCH-TOCSY spectrum (Bax et al 1990) of calmodulin. The diagonal resonances correspond to protons attached to carbons that resonate at 66.8, 43.0 or 19.2 p.p.m. A carbon isotropic mixing duration of 24 ms was used in this experiment, which is sufficiently long to yield connectivities between all side chain resonances. For example, the $C_\delta$ methylene protons of Arg-126 show correlations to $C_\gamma$, $C_\beta$ and $C_\alpha$ protons.

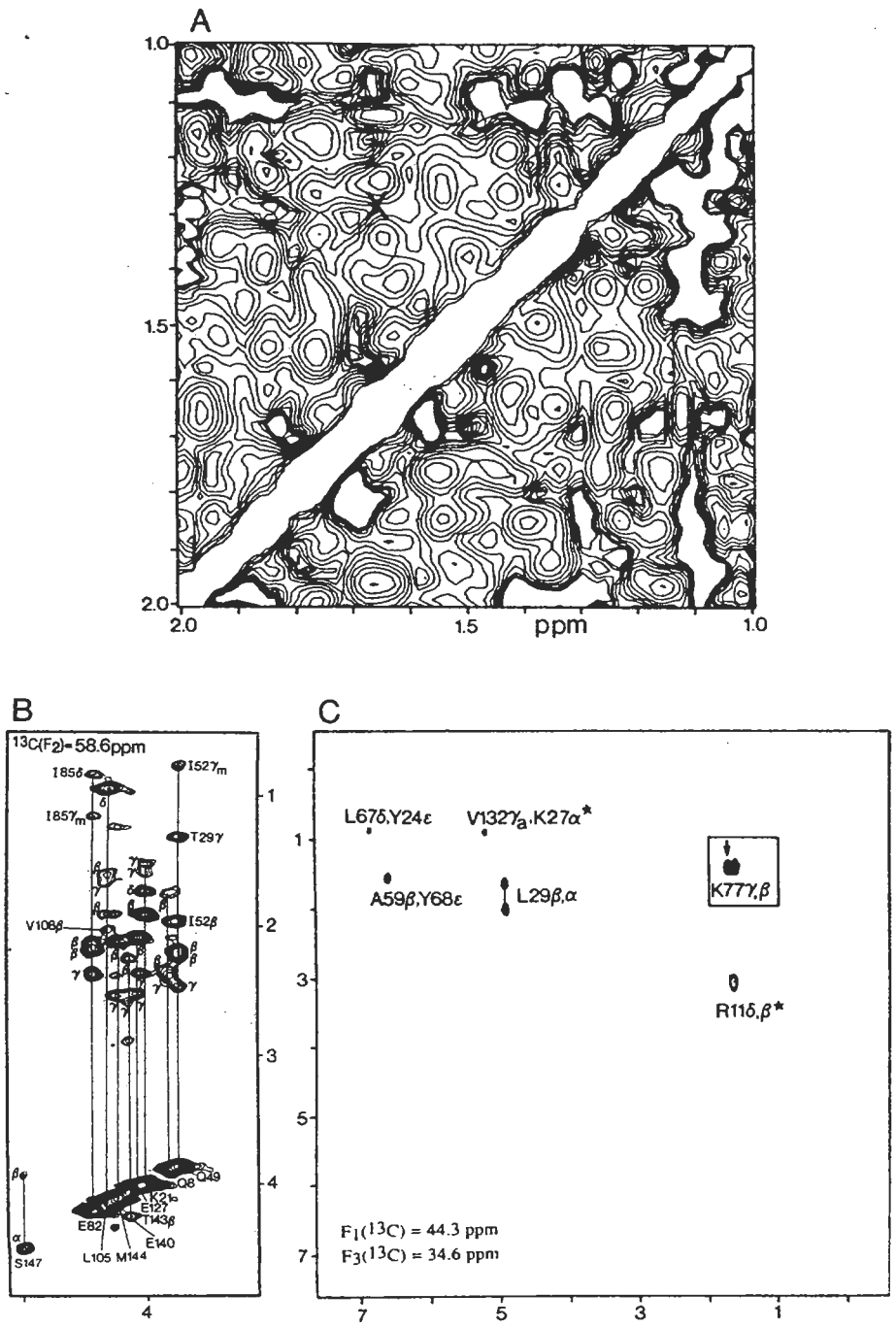


of $H_\alpha$–$C_\alpha$ pairs this technique affords identification of the type of residue. Together with the amino acid linking method described above, this procedure provides a far overdetermined set of information for making complete and unambiguous spectral assignments of virtually all $^1H$, $^{13}C$ and $^{15}N$ nuclei in the protein.

## Obtaining the $^1H$–$^1H$ distance constraints

Measurement of interproton distances using the nuclear Overhauser effect (NOE) constitutes the heart of protein structure determination. NOE measurement is conventionally done using two-dimensional NOESY (nuclear Overhauser effect spectroscopy) experiments, where a short interproton distance gives rise to a resonance in the two-dimensional spectrum with the chemical shift frequency of one of the protons along the $F_1$ axis of the spectrum, with the chemical shift of the second proton being the coordinate in the orthogonal frequency dimension ($F_2$). If two protons, A and B, have identical $^1H$ chemical shifts, but only one of these is proximate to a third proton, C, it is impossible to decide from the NOESY spectrum alone whether the observed interaction is between A and B, or between A and C. This type of ambiguity occurs commonly and can often be solved by spreading the two-dimensional NOESY spectrum into a third dimension. For example, if both A and B are amide protons, $^{15}N$ enrichment of the protein makes feasible a three-dimensional $^{15}N$-separated NOESY experiment which disperses the regular two-dimensional NOESY spectrum into an orthogonal frequency dimension, the $^{15}N$ chemical shift (Fesik & Zuiderweg 1988, Marion et al 1989a, Messerle et al 1989). Further resolution of such a three-dimensional spectrum can be obtained by spreading the frequency of the protons that interact with the amide proton into yet another frequency dimension, the $^{13}C$ chemical shift, in a four-dimensional NMR experiment (Kay et al 1990c).

Interactions between carbon-attached protons are actually more numerous than those involving amide protons, but, even for small proteins, many of these cannot be resolved sufficiently in two-dimensional NOESY spectra to allow their use for distance constraint measurement. The increase in spectral simplicity afforded by increasing the spectral dimensionality is illustrated in Fig. 5. Figure 5A shows a small region of the two-dimensional NOESY spectrum of the protein interleukin 1β, recorded at the highest available magnetic field strength (14.1 tesla, corresponding to a $^1H$ frequency of 600 MHz) for maximum spectral dispersion. Clearly, no individual cross-peaks can be identified in this region of the NOESY spectrum. As was the case with overlapping amide protons, significant alleviation of the overlap problem can be obtained by spreading the spectrum into a third dimension corresponding to the chemical shift of the $^{13}C$ nucleus directly attached to one of the two interacting protons (Ikura et al 1990b, Zuiderweg et al 1990). A small region of a NOESY






slice of such a three-dimensional spectrum (Fig. 5B), taken at a $^{13}$C frequency of 58.6 p.p.m., illustrates that virtually all of the NOE interactions involving $H_\alpha$ protons can be identified in such a spectrum. However, spectral regions such as the one shown in Fig. 5A remain insufficiently resolved in the corresponding three-dimensional spectrum (data not shown) to permit identification of the NOE interactions. Only when the spectrum is dispersed in a four-dimensional manner, to separate the chemical shifts of interacting protons according to both their $^1$H and attached $^{13}$C chemical shifts, can the pertinent interactions be identified (Fig. 5C). As demonstrated clearly by Clore et al (1991), spectral overlap in such four-dimensional spectra is a rarity, and a tremendous number of NOE interactions can be identified unambigously in a straightforward manner because complete $^1$H and $^{13}$C assignments are already available from the new assignment procedure, discussed earlier.

## Conclusions

Uniform isotopic labeling of proteins permits a whole array of sophisticated NMR experiments to be performed that yield spectral assignments and identifiable interproton NOE distance constraints in a relatively straightforward manner. Depending on the level of protein expression available, the cost of the quantities of $^{13}$C and $^{15}$N isotopes needed for this approach can vary from as little as US$1000 to many times this amount. In addition, the approach requires 'high-tech' NMR spectrometers that can generate the complex sequences of radiofrequency pulses needed, and sophisticated software is needed for the analysis for the spectral data. It may be expected that once the suitable hardware and software required for data analysis becomes available, structure determination by NMR spectroscopy for proteins of up to about 20 kDa will become a relatively fast process, requiring approximately four weeks of measurement time and possibly as little as a few months for the subsequent analysis of NOE distance constraints and calculation of the protein structure.

FIG. 5. (A) A small region of the NOESY spectrum of interleukin 1β; (B) a small region of a slice of the $^{13}$C-separated three-dimensional NOESY spectrum of calmodulin and (C) a slice taken from the four-dimensional $^{13}$C/$^{13}$C-separated NOESY spectrum of interleukin 1β. Spectrum A contains all pairwise NOE interactions. Spectrum B shows NOE interactions for pairs of protons where one of the two is attached to a carbon with a 58.6 p.p.m. chemical shift. Spectrum C shows interactions between pairs of protons only where one proton is attached to a carbon with a 44.3 p.p.m. shift and the second proton is attached to a carbon with a 34.6 p.p.m. shift. Panels A and C are adapted from Clore et al (1991).

*knowledgements*

thank Marius Clore, Angela Gronenborn and Dennis Torchia for stimulating ussions. This work was supported by the Intramural AIDS Antiviral Program of office of the Director of the National Institutes of Health. L. E. K. acknowledges entennial Fellowship from the Medical Research Council of Canada and the Alberta itage Trust Foundation.

## ferences

A 1989 Two-dimensional NMR and protein structure. Annu Rev Biochem 58:223–256
A, Clore GM, Gronenborn AM, 1990 $^{1}H$–$^{1}H$ correlation via isotropic mixing of $^{13}C$ magnetization: a new three-dimensional approach for assigning $^{1}H$ and $^{13}C$ pectra of isotopically $^{13}C$-enriched proteins. J Magn Reson 88:425–431
re GM, Gronenborn AM 1989 Determination of three-dimensional structures of roteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. RC Crit Rev Biochem Mol Biol 24:479–564
re GM, Kay LE, Bax A, Gronenborn AM 1991 Four-dimensional $^{13}C/^{13}C$-edited uclear Overhauser enhancement spectroscopy of a protein in solution: application o interleukin 1β. Biochemistry 30:12–18
ik SW, Zuiderweg ERP 1988 Heteronuclear three-dimensional NMR spectroscopy. strategy for the simplication of homonuclear two-dimensional NMR spectra. J Magn eson 78:588–593
ik SW, Eaton HL, Olejniczak ET, Zuiderweg ERP, Dahlquist FW 1990 2D and 3D MR spectroscopy employing $^{13}C$–$^{13}C$ magnetization transfer by isotropic mixing. pin system identification in large proteins. J Am Chem Soc 112:886–888
ra M, Kay LE, Bax A 1990a A novel approach for sequential assignment of $^{1}H$, $^{13}C$, nd $^{15}N$ spectra of larger proteins: heteronuclear triple-resonance NMR spectroscopy. pplication to calmodulin. Biochemistry 29:4659–4667
ra M, Kay LE, Tschudin R, Bax A 1990b Three-dimensional NOESY-HMQC pectroscopy of a $^{13}C$ labeled protein. J Magn Reson 96:204–209
tein R, Boelens R, Scheek RM, van Gunsteren WF 1988 Protein structures from MR. Biochemistry 27:5389–5395
LE, Ikura M, Tschudin R, Bax A 1990a Three-dimensional triple resonance pectroscopy of isotopically enriched proteins. J Magn Reson 89:496–514
LE, Ikura M, Bax A 1990b Proton–proton correlation via carbon–carbon coupling: three-dimensional NMR approach for the assignment of aliphatic resonances in roteins labeled with carbon-13. J Am Chem Soc 112:888–889
LE, Clore GM, Bax A, Gronenborn AM 1990c Four-dimensional heteronuclear iple-resonance NMR spectroscopy of interleukin-1β in solution. Science (Wash DC) 49:411–414
LE, Ikura M, Bax A 1991 The design and optimization of complex NMR experiments: pplication to a triple resonance pulse scheme correlating Hα, NH and $^{15}N$ chemical ifts in $^{15}N$–$^{13}C$ labeled proteins. J Magn Reson 91:84–92
faster DM, Richards FM 1988 NMR sequential assignment of *Escherichia coli* ioredoxin utilizing random fractional deuteration. Biochemistry 27:142–150
ion D, Kay LE, Sparks SW, Torchia DA, Bax A 1989a Three-dimensional eteronuclear NMR of $^{15}N$ labeled proteins. J Am Chem Soc 111:1515–1517
ion D, Driscoll PC, Kay LE et al 1989b Overcoming the overlap problem in ie assignment of $^{1}H$ NMR spectra of larger proteins using three-dimensional



homonuclear Hartmann-Hahn and nuclear Overhauser $^{1}H$–$^{15}N$ heteronuclear multiple quantum coherence spectroscopy. Biochemistry 28:6150–6456
McIntosh LP, Griffey RH, Muchmore DC, Nielson CP, Redfield AG, Dahlquist FW 1987 Proton NMR measurements of bacteriophage T4 lysozyme aided by $^{15}N$ isotopic labeling: structural and dynamic studies of larger proteins. Proc Natl Acad Sci USA 84:1244–1248
Messerle BA, Wider G, Otting G, Weber C, Wüthrich K 1989 Solvent suppression using a spin lock in 2D and 3D NMR spectroscopy with $H_2O$ solutions. J Magn Reson 85:608–613
Niemczura WP, Helms GL, Chesnick AS, Moore RE, Bornemann V 1989 Carbon-detected correlation of carbon-13-nitrogen-15 chemical shifts. J Magn Reson 81:635–640
Oh BH, Westler WM, Derba P, Markley JL 1988 Protein carbon-13 spin systems by a single two-dimensional nuclear magnetic resonance experiment. Science (Wash DC) 240:908–911
Senn H, Otting G, Wüthrich K 1987 Protein structure and interactions by combined use of sequential NMR assignments and isotopic labeling. J Am Chem Soc 109:1090–1092
Torchia DA, Sparks SW, Bax A 1989 Staphylococcal nuclease: sequential assignment and solution structure. Biochemistry 28:5509–5524
Wüthrich K 1986 NMR of proteins and nucleic acids. Wiley, New York
Zuiderweg ERP, McIntosh LP, Dahlquist FW, Fesik SW 1990 Three-dimensional $^{13}C$-resolved proton NOE spectroscopy of uniformly $^{13}C$-labeled proteins for the NMR assignment and structure determination of larger molecules. J Magn Reson 86:210–216

## DISCUSSION

*Richards:* When will the automation for assignments be available?

*Bax:* The software is available to anybody who wants it, but it hasn't been written by a professional programmer—it was written by Dr Ikura because he wanted to simplify the analysis of the spectrum of calmodulin.

*Richards:* So it should be generally available in a year or two.

*Bax:* Yes, I think so. It's really quite straightforward software; a good programmer could write it in two months.

*Richards:* There's an enormous amount of data collection here; how does this convert into dollars? What machine do you use?

*Bax:* The higher dimension experiments could easily be done at 400 MHz. The resolution is limited not by the magnetic field strength but by the digitization. We used 500 and 600 MHz spectrometers, because those are the ones we have.

*Richards:* How long does the data collection take for a 15 kDa protein?

*Bax:* To get the complete assignments for side chain and backbone residues including some time for technical failures and incorrect setting up of parameters would probably take 3–4 weeks full-time measuring. It would take another two weeks to get all the NOE data. It depends a little on concentration— you could