Novartis Foundation Symposium
Edited by Derek J. Chadwick, Kate Widdows

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

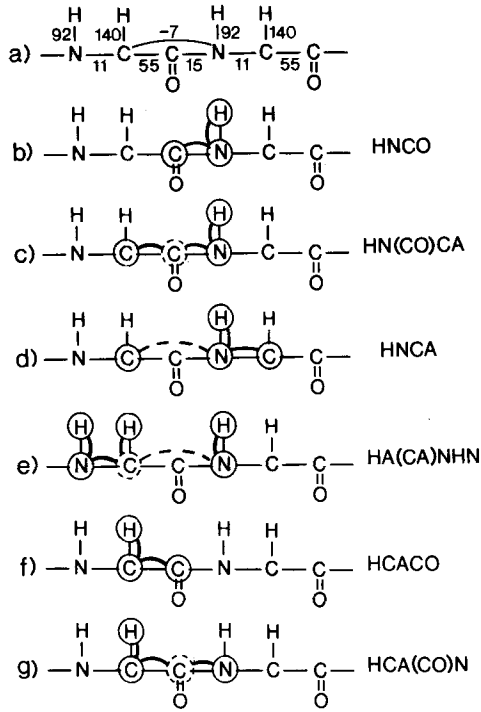
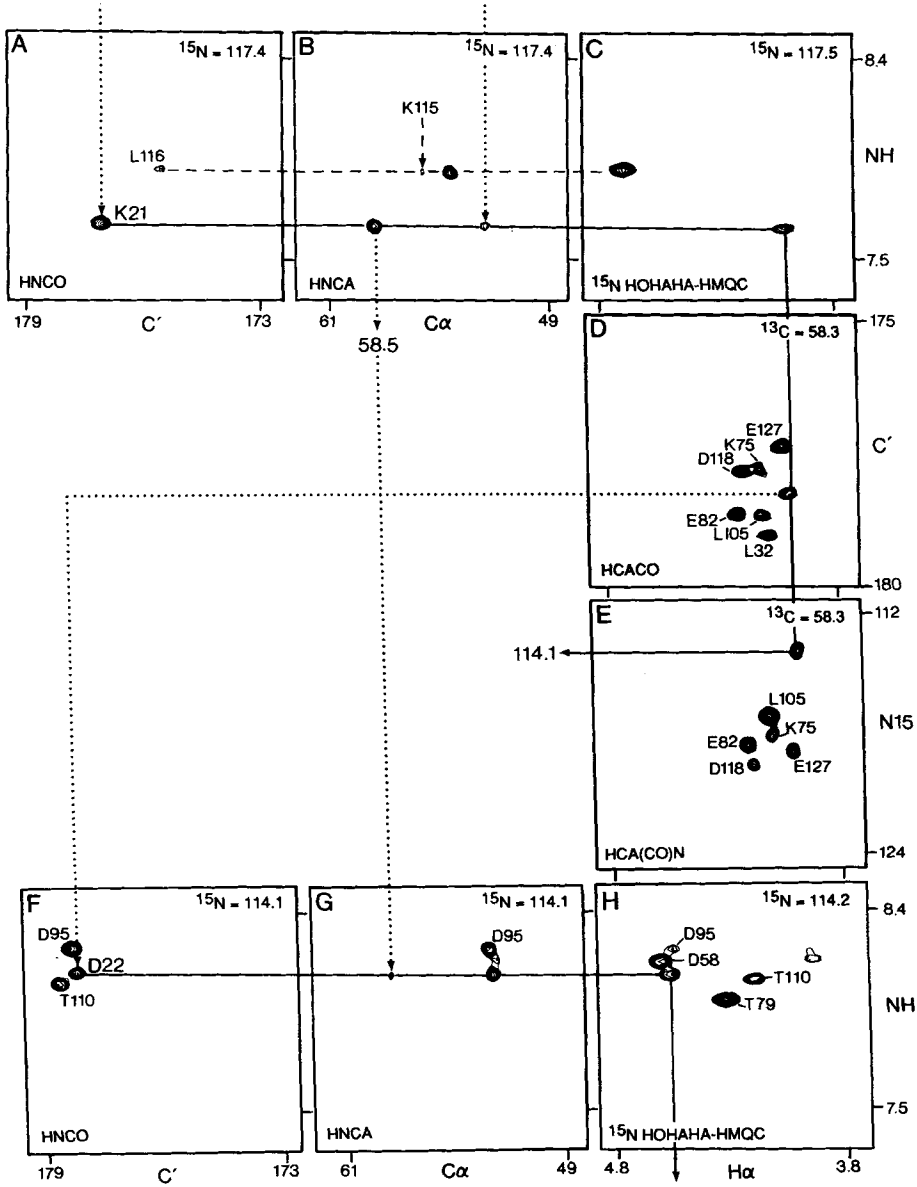


FIG. 1. Backbone atoms of two adjacent amino acids. (a) Size of relevant J couplings, in Hz; (b–g) connectivity diagrams indicating the correlations between different backbone atoms that can be observed with the pulse sequences shown on the right (CA, α-carbon). Circles mark the correlated resonances, 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 ^1H–^1H 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 ^1H 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_α 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 α-carbon of the preceding residue. The intensity of these two-bond correlations rapidly decreases when the molecular tumbling time, τ_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_α nuclei can be correlated with one another (HNCA spectrum; CA, α-carbon). In addition, the dashed line in this diagram indicates that the inter-residue connectivity between HN, N and the α-carbon of the preceding residue can also be observed (provided that τ_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_α 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_α 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_α and C_α 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_α, C_α and C′ chemical shifts. The slice of the HCACO spectrum taken at the C_α shift of Lys-21 shows a correlation at the Lys-21 H_α shift to a C′ resonance at 178.3 p.p.m. In addition, the HCA(CO)N spectrum shows that this H_α–C_α 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_α correlation that corresponds to this amide (Fig. 2G), and also shows a weak correlation to the α-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_α

and C_α 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

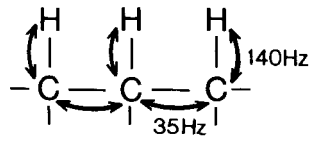


FIG. 3. Connectivity diagram for the so-called HCCH-type experiments (Kay et al 1990b, Fesik et al 1990, Bax et al 1990) which correlate side chain resonances utilizing one-bond J_{CH} and J_{CC} couplings. Arrows mark the magnetization transfer pathways.

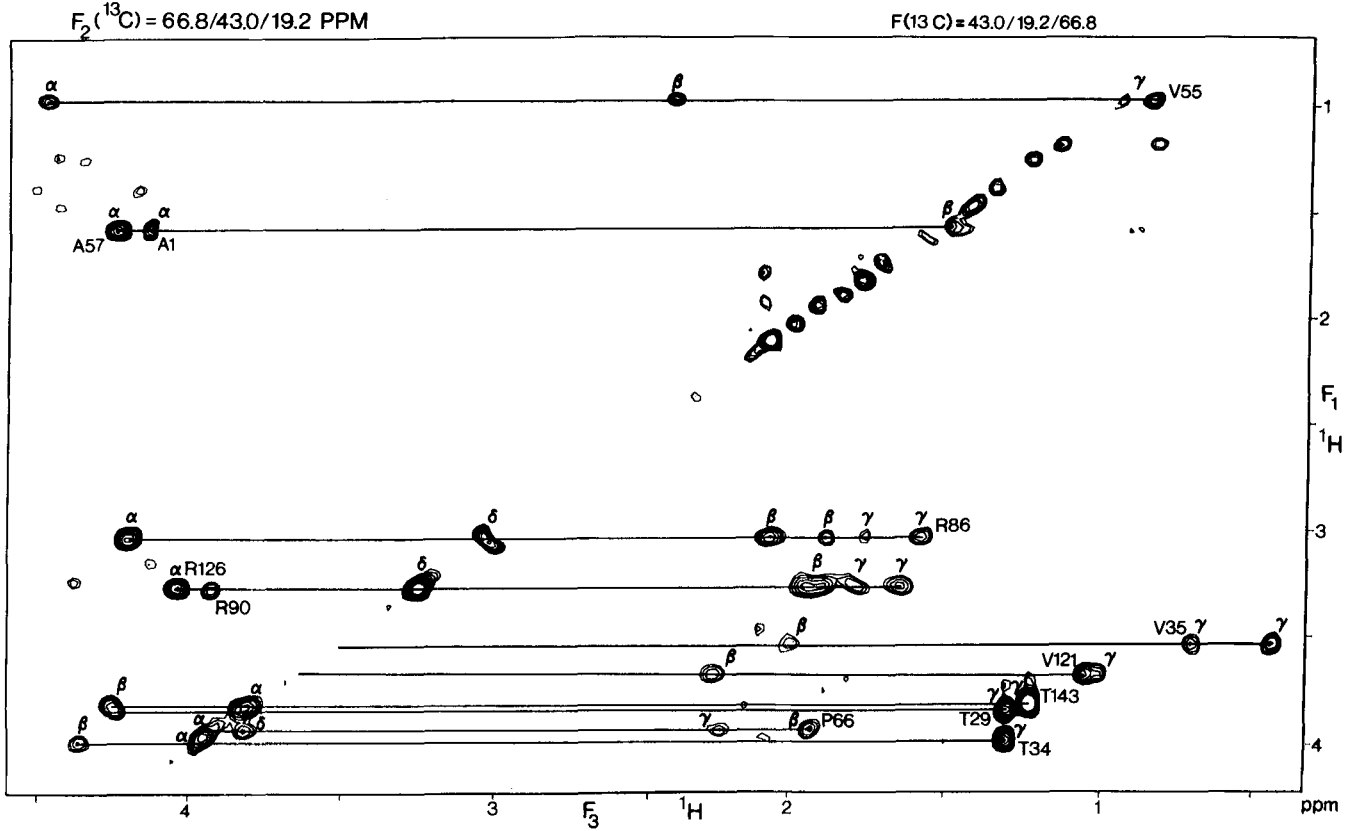


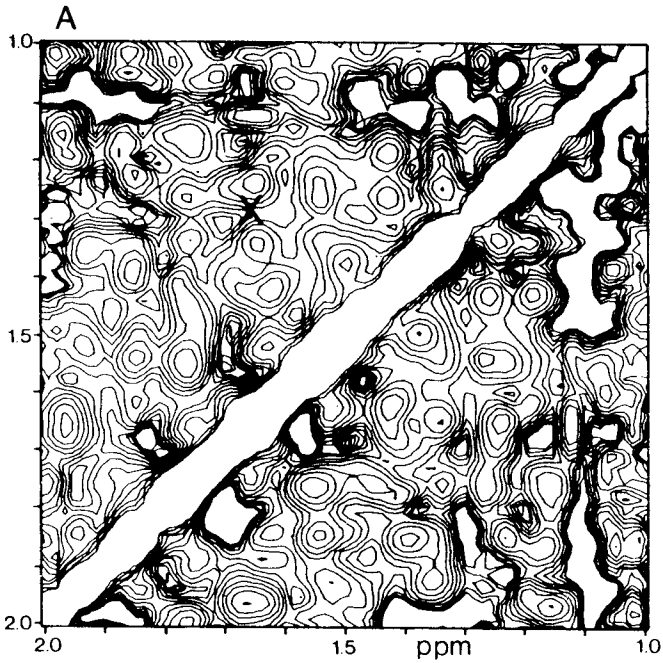
FIG. 4. Example of a slice taken from the three-dimensional HCCH-TOCSY spectrum (Bax et al 1990) of calmodulin. The diagonal resonances correspond to protons attached to carbons that resonate at 66.8, 43.0 or 19.2 p.p.m. A carbon isotropic mixing duration of 24 ms was used in this experiment, which is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

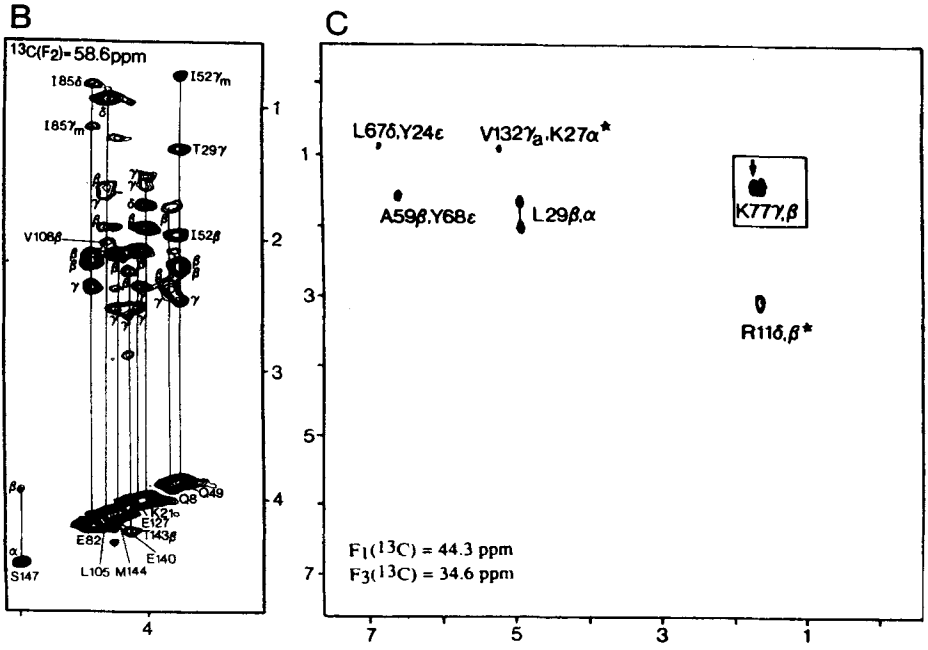
of H_α–C_α pairs this technique affords identification of the type of residue. Together with the amino acid linking method described above, this procedure provides a far overdetermined set of information for making complete and unambiguous spectral assignments of virtually all ^1H, ^{13}C and ^{15}N 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_α 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 ^1H 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 ^1H 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).

Acknowledgements

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

References

Bax A 1989 Two-dimensional NMR and protein structure. Annu Rev Biochem 58:223–256

Bax 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 spectra of isotopically ^{13}C-enriched proteins. J Magn Reson 88:425–431

Clore GM, Gronenborn AM 1989 Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. CRC Crit Rev Biochem Mol Biol 24:479–564

Clore GM, Kay LE, Bax A, Gronenborn AM 1991 Four-dimensional $^{13}C/^{13}C$-edited nuclear Overhauser enhancement spectroscopy of a protein in solution: application to interleukin 1β. Biochemistry 30:12–18

Fesik SW, Zuiderweg ERP 1988 Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplication of homonuclear two-dimensional NMR spectra. J Magn Reson 78:588–593

Fesik SW, Eaton HL, Olejniczak ET, Zuiderweg ERP, Dahlquist FW 1990 2D and 3D NMR spectroscopy employing ^{13}C–^{13}C magnetization transfer by isotropic mixing. Spin system identification in large proteins. J Am Chem Soc 112:886–888

Ikura M, Kay LE, Bax A 1990a A novel approach for sequential assignment of ^{1}H, ^{13}C, and ^{15}N spectra of larger proteins: heteronuclear triple-resonance NMR spectroscopy. Application to calmodulin. Biochemistry 29:4659–4667

Ikura M, Kay LE, Tschudin R, Bax A 1990b Three-dimensional NOESY-HMQC spectroscopy of a ^{13}C labeled protein. J Magn Reson 96:204–209

Kaptein R, Boelens R, Scheek RM, van Gunsteren WF 1988 Protein structures from NMR. Biochemistry 27:5389–5395

Kay LE, Ikura M, Tschudin R, Bax A 1990a Three-dimensional triple resonance spectroscopy of isotopically enriched proteins. J Magn Reson 89:496–514

Kay LE, Ikura M, Bax A 1990b Proton–proton correlation via carbon–carbon coupling: a three-dimensional NMR approach for the assignment of aliphatic resonances in proteins labeled with carbon-13. J Am Chem Soc 112:888–889

Kay LE, Clore GM, Bax A, Gronenborn AM 1990c Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1β in solution. Science (Wash DC) 249:411–414

Kay LE, Ikura M, Bax A 1991 The design and optimization of complex NMR experiments: application to a triple resonance pulse scheme correlating Hα, NH and ^{15}N chemical shifts in ^{15}N–^{13}C labeled proteins. J Magn Reson 91:84–92

LeMaster DM, Richards FM 1988 NMR sequential assignment of *Escherichia coli* thioredoxin utilizing random fractional deuteration. Biochemistry 27:142–150

Marion D, Kay LE, Sparks SW, Torchia DA, Bax A 1989a Three-dimensional heteronuclear NMR of ^{15}N labeled proteins. J Am Chem Soc 111:1515–1517

Marion D, Driscoll PC, Kay LE et al 1989b Overcoming the overlap problem in the 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

do it in that time provided you have a 1–1.5 mM concentration of the protein; a higher concentration wouldn't help because this is the minimum amount of time it would take to go through the procedure.

Richards: So, full time, on one machine, you could manage about six to eight proteins of that size per year.

Bax: If one wanted just to determine structure and stop there, eight proteins per year would probably be the theoretical limit. In practice, one usually wants to know more than only the structure, and a substantial amount of time can be spent on measurement of hydrogen exchange kinetics and on relaxation experiments to determine protein dynamics and conformational changes in the presence of ligands.

Richards: So if we wanted to know only protein structure, we could divide the cost of running a high field spectrometer for one year by eight to work out the cost per protein.

Bax: Yes; it's expensive, and the number you would get would be a lower limit that no one has yet reached.

Kollman: The crystal structure of calmodulin shows a 25-residue-long helix connecting the two globular domains (Babu et al 1988). Do you yet know whether this central helix is one long helix in solution?

Bax: We don't yet have the final answer to that question; there is some contradictory evidence. The NMR structure shows helical behaviour for residues 68–77 and 82–92, and rapid amide exchange for residues 76–81. The amides of these residues are protected from hydrogen exchange (compared with a free random coil peptide) by factors ranging from five to not protected at all (Spera et al 1991). This rapid hydrogen exchange is indicative of a quite flexible structure. However, when we use all the NOE and J coupling constraints measured for the central helix as the input for a structure calculation program we get a fairly narrow distribution of conformers. All calculated structures show a distinct kink near residue 80, bending the central helix with an angle of roughly 135° between the N-terminal and C-terminal half of this helix (M. Ikura, unpublished paper, Biophysics meeting, San Francisco, February 1991). Dr Ikura also found clear evidence for a tightly bound water molecule in this central helix, near the kink. This water molecule is not seen in the crystal structure. On the basis of this type of relatively rigid model of calmodulin, with a kinked helix, one would expect a substantial anisotropy in the molecular tumbling. Relaxation experiments conducted by my associates Gaetano Barbato and Lewis Kay do not show evidence for a large degree of anisotropic tumbling. Therefore, the degree of flexibility in the middle of the central helix must be considerably greater than the structure calculation program indicates.

Richards: How do you know the level of anisotropy of the motion?

Bax: We measure ^{15}N T_2 and T_1 relaxation times and we know the shapes of the individual domains, the two globular domains of calmodulin. We

know that amides with an NH bond vector parallel to the long axis should have a different relaxation from amides that are perpendicular to the long axis.

Richards: Do you do experiments separately on the isolated domains?

Bax: No. We can measure the relaxation properties of every individual nitrogen atom and we know which nitrogen is which and the size of the N–H dipolar interaction which determines the relaxation of the nitrogen; so if the N–H bond vector is parallel to the long axis of the molecule, the ^{15}N should relax differently from one where the bond vector is perpendicular. The difference in the measured correlation time is not nearly as great as we would predict on the basis of the model with the central kinked helix. The relaxation data are still in the final stages of analysis, so this might not be definite, but I am quite confident that this is the case.

Wüthrich: Do you see the amide protons in this helix?

Bax: Yes we do, but at pH 7 they are severely broadened because of exchange.

Schulz: Do you measure the dihedral angles from residue to residue?

Bax: No, within the residue itself you can measure only the ϕ angle—we cannot measure the ψ angle.

Schulz: So how do you establish the conformation of this long helix?

Bax: We use the proton–proton NOEs that determine the distance constraints. I can measure the proton–proton distances, but not the angles.

Fontana: The long helix connecting the two lobes of calmodulin should be quite special, because limited proteolysis with trypsin cleaves the calmodulin molecule at this helix, at Lys-77, producing fragments 1–77 and 78–146 (Draibikowski et al 1982). We have looked at some 30 different proteins of known crystal structure for which the sites of limited proteolysis, using a variety of proteolytic enzymes, have been reported. The general observation that emerged from this study was that exposed and flexible loops are the exclusive sites of proteolysis (Fontana et al 1986, Fontana 1989). Thus, the tryptic cleavage of a helix in calmodulin appears to be quite unusual and contradicts the generalization that only loops, and never elements of regular secondary structure (helices), are the favoured sites of limited proteolysis. However, in solution the long helix of calmodulin is 'bent' and thus has a turn/loop in its middle, which brings the two calcium-binding domains into closer contact (Heidorn & Trewhella 1988, Persechini & Kretsinger 1988). Moreover, in the original description of the crystal structure of calmodulin it was reported that the chain region 75–80 showed a weak electron density and poor visibility in the density map (Babu et al 1985). This indicates that the site of cleavage in calmodulin is characterized by some static–dynamic disorder, that is, by some chain flexibility. The long central helix of calmodulin is not as nice and straight as is usually depicted in the three-dimensional model of the molecule. Because of this, it seems that there are no exceptions to the rule

that limited proteolysis of globular proteins occurs at exposed and flexible loops (Fontana 1989).

Bax: Babu et al (1988) pointed out that there was something odd near the middle of the central helix and that there might be an artifact of crystal packing. Our data indicate that there is obviously flexibility there–it's the *degree* of flexibility that we are still trying to determine.

Richards: Are your anisotropy measurements good enough to provide an explicit axial ratio for an ellipsoid, for example?

Bax: Yes, I think they should be.

Richards: So you could model this?

Bax: That's what we are trying to do.

Hilbers: Do you think that the helix is transiently present, or not?

Bax: The chemical shift data, for example, are different from what would be expected for a random coil, which indicates that there is a preferred conformation; even in the crystal structure some central helix amides are not hydrogen-bonded. Our data indicate that the flexible region of the central helix does not behave like a random coil, but it's not a fully organized stable globular structure either.

Wüthrich: For a helix there are straightforward criteria—you either get the i to (i+3) connectivities, or you don't. If you do, then you have a helix.

Bax: But if the helix is present only 50% of the time you get weak i to (i+3) connectivities.

Wüthrich: You should still see it.

Bax: We do see it. We observe both i to (i+3) connectivities and the characteristic i to (i+1) dNN connectivities for amides that have a protection factor only 10-fold greater than a fully random coil-type peptide. In a regular α-helix you would expect a much higher degree of protection, of course.

Wüthrich: Do you mean slowed NH exchange?

Bax: Yes; if you interpret this result according to the Englander model (Englander & Kallenbach 1984) you would say that the helix is present 90% of the time, but I wouldn't dare to say that 90%, or 80% or 60% of the time a particular hydrogen bond is there—I don't think we can quantitate to that level of accuracy. There are three residues for which no dNN connectivities are observed, so we are confident that three of the residues in the central helix do not exist in a helical conformation.

Wüthrich: Do you see the medium-range NOEs or do you see only the sequential ones?

Bax: For most parts of the central helix we see i to (i+3) as well.

Wüthrich: Is that a continuous series of such NOEs?

Bax: No; we actually see an i, (i+5) somewhere, near the middle, where there is tightly bound water.

Wüthrich: That would indicate that it's not a helix.

Bax: No; there is a kink in the helix. There are three residues that are absolutely not helical, otherwise we wouldn't see i, (i + 5).

Richards: One is brought up in biochemistry to believe that there are canonical helices which differ by a specific jump in where the hydrogen bond is made (that is, to the third, or the fourth or the fifth residue). It seems to me that there is much more flexibility in this system, and that it may be much harder to distinguish between α-helices and 3_{10} helices and perhaps even π-helices than is normally thought, because of bifurcated bonds. One may be able to twist a helix relatively smoothly with relatively small energy barriers. I am guessing—I don't know this to be a fact.

Blundell: In 1983 we analysed α-helices for their ϕ–ψ parameters (Blundell et al 1983). The interesting feature that we noted was that in a plot of ϕ and ψ for amino acids there is no particular trend. However, for the ψ of residue i and the ϕ of residue (i + 1), there is a trend. This plot reflects the angle of the peptide plane with respect to the helix axis. Those on the outside rotate out so that the CO groups bind water. Furthermore, there is a relationship between the angle of the peptide plane and the curvature of the helix, and the curvature of the helix itself is related to its amphipathicity. The calmodulin helix is interesting because it is hydrophilic on all sides; it is not characteristic of those found in other proteins. It is expected to be more linear. For amphipathic helices we can predict the position of the centre of curvature. A helix with the Pauling parameters does not occur in real proteins. For the usual amphipathic helices the average ϕ and ψ angles are quite similar in many proteins, but they are an average of two classes that are characteristic of the hydrophobic and the hydrophilic residues.

Dobson: Is anything known about the integrity of the calmodulin helix under different conditions? Is, for example, the protein non-cooperative in its unfolding properties? Is the helix stable in the absence of the remainder of the protein's structure? Have any experiments been done to test ideas of this sort?

Bax: People have tried to make the helix without the domains. R. J. P. Williams (personal communication) told me that his group had synthesized a 17-amino acid residue section of the central helix that behaves like a random coil. We have made a 10-amino acid peptide, which also behaves like a random coil. There is no pH dependence for this conformation. We haven't looked at temperature dependence in as much detail as we would like to. On a decrease in temperature a number of the resonances broaden substantially, which probably means that the conformational exchange occurs on a time scale that is difficult to follow.

Dobson: I was thinking more about the cooperativity of the structure and whether, for example, the helical part of the structure could be lost before the globular regions unfold.

Bax: It is, I am sure. The globular domains are stable up to about 70 °C.

Schulz: If you deuterate a protein, the chemical properties change slightly. With complete ^{15}N and ^{13}C labelling, are you sure that the chemistry of the structure does not alter?

Bax: The effect of going from ^{12}C to ^{13}C is much smaller, of course, than that of going from ^{1}H to ^{2}H; it's a factor of 2% in comparison to 8%.

Wüthrich: You suggested that because there is so much empty space in the slices from the higher-dimensional spectra one should be able to deal with much bigger proteins. However, even if you go to 18 dimensions, there will be a limiting size at which there won't be any coherence left when you are through with the pulse sequence, and this will be independent of the amount of empty space left in the spectrum. If you label with ^{13}C and ^{15}N you actually make the situation worse, by introducing additional dipolar couplings. Do you really believe you will be able to handle proteins with a relative molecular mass of 40 000 using this approach?

Bax: This is a good point—going up to 18 dimensions, or even to five, which we've tried, is not very practical. With four dimensions you don't lose much more sensitivity than you do with two dimensions. You said that the incorporation of ^{13}C makes things worse, because there is another relaxation loss pathway—actually, that helps because we have to measure the dipolar interaction between two protons and we can measure this for only a short time otherwise we get indirect effects, so we have to use short mixing times. The initial build-up rate of the NOE between two protons is independent of whether you have ^{13}C or ^{15}N or not. Because we have to work in this relatively early region using short mixing times the relative NOE cross-peak intensity will be unaffected. What helps us is the fact that in ^{13}C-labelled proteins the protons relax faster, by as much as 30%, so we can repeat our experiments more quickly than we can for a ^{12}C protein. The small loss of cross-peak intensity that comes from not quite being in the linear region of NOE cross-peak build-up is offset by the advantage of being able to repeat the experiment more quickly than normal.

Wüthrich: At some point you will nonetheless run out of coherence after the experiment.

Bax: The NOE experiment is not where we will have a problem—measuring the distances is not going to be the problem—the problem will be the assignment. There we run into difficulties because the nitrogen–carbon coupling, for example, is only 11 Hz and above about 30 kDa you won't be able to use that interaction. I still think that 30 kDa is probably the upper limit of assignments with this kind of procedure in the absence of specific labels.

Wüthrich: You suggested earlier that 'conventional methods' will not work for higher molecular weight proteins. What you have just said, however, leads me to predict that the conventional methods using NOEs for the sequential assignments *will* work at higher molecular weights.

Bax: Of course, by combining the strategy that we are using now with conventional strategies one may be able to push things higher.

Wüthrich: You have simply added techniques that enable you to assign the side chains, but you will have to follow the old, conventional strategy using sequential NOEs.

Bax: Yes; so far we have been trying to show that we can do it without using sequential NOEs. When trying to raise the limits you use every available parameter. It's a combination of techniques that will allow us to study proteins about 20% larger, but nevertheless we will reach a limit close to about 35 kDa. I shouldn't define the limit in terms of the size of the protein. It really should be defined in terms of the tumbling correlation time. I think the limit would be around 20 nanoseconds. If you have a friendly 50 kDa protein that you can heat up to 50–60 °C it may still tumble rapidly enough to do those techniques: with a membrane-type protein that has to be kept in micelles, even if it's only 10 kDa, the technique may not work. The correlation time is the limiting factor.

Wüthrich: I would like to make one more point for the benefit of those who are not working with these techniques, that chemical shift degeneracy is common in work with proteins, even with small proteins of 60 amino acid residues. To overcome the ensuing difficulties, we simply start the structure determination without these ambiguous peaks. This means that of a total of perhaps 800 NOE cross-peaks in the spectrum we will initially assign 250–300 and use those to calculate a first structure. By reference to this structure many ambiguities in the NOESY cross-peak assignments can be resolved, which leads to the correct assignments. That's the way we work, going through three, five or six cycles of resonance assignments, refining the structure, and checking the structure for assignments of additional distance constraints to end up with a fully assigned spectrum and the refined structure. Going to isotope-labelling and higher-dimensional experiments is more expensive than going through several cycles of resonance assignment and structure determination, so that a careful evaluation of the different possible strategies is called for.

Bax: Of course, even when analysing our four-dimensional spectra we have to go through such a cyclic procedure. For our four-dimensional spectra we actually had to go through two cycles to obtain complete assignments of all NOE cross-peaks. With a four-dimensional experiment about half the cross-peaks, half the interactions, can be immediately and unambiguously identified as individual pairs. For the second half, there is still ambiguity so we have to go through a second cycle, but it's much faster to do it this way, plus you wouldn't be able to observe individual peaks in very crowded regions of the two-dimensional spectrum.

Wüthrich: Yes, that is true for bigger proteins.

Schulz: Where is the extra cost of higher dimension NMR; is it in the electronics or in the computing time?

Bax: For a 60-amino acid protein a week and a half of two-dimensional NMR may sometimes be sufficient to give you all the data you need for resonance assignment and structure determination. To get the same assignment and distance information for larger proteins from three-dimensional NMR the minimum time needed is of the order of a month and a half. Data collection is therefore much more costly. Modern spectrometers are suitable for these higher dimensional experiments almost without modification, so there is no extra cost there. You do have to pay for the isotopic labels, of course, and that can add up to a fair amount of money.

Wüthrich: It is important to remember that one can use higher-dimension techniques only for recombinant proteins. If the protein has not been cloned and expressed one has to work with the natural isotope distribution.

Hendrickson: Apart from the benefits that you have described about making the assignments and uniquely identifying the peaks, are there other fringe benefits—extra information about the structure—from the resonances associated with isotopic labelling?

Bax: Yes there are. The ^{13}C and ^{15}N relaxation rates can be interpreted in a straightforward fashion to give dynamic information about the protein. We have shown this with ^{15}N for the protein backbone of staphylococcal nuclease (Kay et al 1989). One can get local information about dynamics.

Richards: What sort of information can you get?

Bax: We can get the order parameter and information about the time scale on which the motion is taking place. The order parameter tells you how confined a particular carbon–proton bond vector is. There are rapid fluctuations of the NH and CH bond vectors occurring on a time scale of 0.1 ns or so. One can characterize these motions with carbon and nitrogen relaxation rates.

Hendrickson: Am I right in presuming that such information could be used directly in calculations of NOE intensities from relaxation matrix methods?

Bax: In principle, you could do that, but you have to be careful. The order parameter for a carbon–proton bond vector is quite different from the order parameter of a proton–proton bond vector. I would be scared to quantitate those NOEs more rigorously. I don't think that is the way to get better structures. To get more highly refined NMR structures one has to measure more NOE interactions rather than trying to say exactly what a distance is. We cannot define motional parameters accurately enough for that purpose.

I actually have a question for the crystallographers. With the relaxation data that we have been looking at, we see motions for side chain atoms that are significantly larger than what one would expect from crystallographic B factors. We heard earlier that with a 2 Å crystal structure one should be able to define B factors and say something about motion. I am concerned that the side chain motions are not characterized properly with 2 Å crystal structures. Are 'fudge factors' used that we don't know about?

Schulz: There are no fudge factors, but it is quite usual to observe higher mobility in the side chains. In X-ray analysis, of course, we can only observe where the side chain sits most of the time. Short-term positional changes go unrecognized.

Bax: B factors for interior leucine, valine or isoleucine side chains are typically comparable to those of the backbone. With NMR, we see a significant increase in mobility. ^{13}C relaxation data recorded by Richarz et al (1980) are also indicative of substantial mobility for several of the methyl groups in the core of bovine pancreatic trypsin inhibitor.

Schulz: BPTI doesn't really have a core.

Bax: It has some core.

Hendrickson: When we refine structures, of course we impose all kinds of stereochemical features in one way or another. If we were to use free atom refinements, there would be a helter-skelter splattering of the distances, for example. Likewise, the same thing would happen if we allowed the B values that are associated with each individual atom to vary freely. To cope with this ill-determination we have introduced techniques for linking B values together, with the idea that bonded atoms should have correlated motions. A sort of compromise has been made in many of our analyses which impose restraining linkages in an isotropic kind of model. This isotropy is really inappropriate and confining, but is forced on us for economy in parameterization. It means that the long side chains don't always reach the B values that they would if you allowed them to vary freely. When we use these restrained temperature factors we are typically seeing a dampening of the extreme values of the motions of the side chains that are exposed to the surface.

Bax: So when I look at crystal structures and find that the B factor of the methyl groups of leucine is the same as that of the α-carbon, does that really mean that there is no gradient in motion, or could this be an artifact caused by the constraints on B factors in the crystallographic refinement?

Hendrickson: Not in that case. If there's no gradient at all in B values, reasonable restraints could not have dampened a systematic progression.

Richards: While recognizing its necessity, I am nervous about the use of isotropic B factors and the effects this has on the refined structures. Some years ago we took the data from Martin Karplus's first molecular dynamics run on BPTI and plotted and contoured the position of the atomic centres to make the pseudo-electron density map. Some of the atoms were adequately represented by a three-dimensional Gaussian function, but others were wild in shape and appearance. Not even anisotropic ellipsoids would come close. This kind of behaviour is well recognized in small molecule structure, and undoubtedly is real in proteins. Assignment of multiple conformations for side chains and occasional main chain parts will help, but there will always be a residual anisotropy. X-ray analysis must take this into account.

van Gunsteren: In my paper I shall show how you can incorporate time-dependent restraints in both NMR refinement and crystallographic refinement. The anisotropy is included in the treatment. Basically, you fit an ensemble to your experimental data, which means that all possible anisotropy is allowed for. For example, for phospholipase the R value goes down from about 18% for one conformer to under 10% with an ensemble of conformers, and that's just because you have more structures; if you look at the structures you see that they form a very anisotropic distribution, which fits the experimental observations better.

Schulz: One should be careful to argue with the decrease of an R factor when one adds so many new parameters.

Richards: The argument here is about the difference in apparent mobility between that derived by NMR and that derived by crystallography.

van Gunsteren: Time-dependent restraints can be used in both cases.

Phillips: The mystery to me is that Ad Bax is saying that the differences are in internal core residues, whereas most of us would accept that it could be easily explained for surface residues.

van Gunsteren: If we allow for time-dependent NOE restraints, the mobility measured as root mean square fluctuations of the residues in the inner core doubles. There is more mobility, because you have satisfied the NOE data not with one model, but with an ensemble.

Bax: But my data on motions are based on ^{13}C relaxation data, not on proton–proton NOEs.

van Gunsteren: That doesn't matter. I am talking about whether you can have more mobility inside a protein than expected. The constraints that you impose damp down the motion.

Bax: It is the crystallographic constraints that concern me, whether the B factors correctly reflect motion or disorder.

Hendrickson: These parameters certainly reflect the feeblest piece of information in a diffraction experiment.

Dobson: Martin Karplus and I have looked in some detail at the results of simulations of BPTI protein dynamics, and have analysed the effects the motions observed in the simulation would have on experimental NMR parameters. The results are summarized in Karplus & Dobson 1986. One of the sets of parameters that we looked at particularly were the spin–spin coupling constants, which reflect the behaviour of bond torsion angles. Recently we have obtained experimental values of coupling constants between C_α and C_β protons for lysozyme, which depend on the χ_1 torsion angles (Smith et al 1991). For internal residues the extent of averaging about χ_1 is rather limited and actually fits in rather well with the predictions of the dynamics simulations. For the surface residues there is very extensive conformational averaging about χ_1 for a significant number of residues. Again, this is more or less what one would expect from the simulations. This suggests a rather

different picture of the surface from that expected on the basis of individual crystal structures. We therefore looked at lysozyme in different crystal forms and found that many of the residues that we see experiencing conformational averaging in solution are in different conformational states in different crystal forms. Whether this means that different refinement procedures emphasize different conformational states, or whether different conformers are really populated in different crystals, I don't know. There are, however, clearly more extensive dynamic events occurring for many surface residues in solution than one might expect from at least a superficial examination of crystallographic data.

Phillips: I think most crystallographers would agree that one sets out to find or to choose a dominant conformation for a surface side chain and one usually doesn't go beyond that. Sometimes there is other electron density in the vicinity which is interpreted as water. These problems of interpretation are functions of the resolution at which the analysis is done, and the method that is used for the refinement. None of us should be surprised by what Chris Dobson is saying, or by Wilfred van Gunsteren's findings. The details of surface side chains in normal crystallographic structure analyses are not to be taken terribly seriously.

Roberts: There is a fundamental difference between crystallography and NMR in the time scale. For example, with a side chain which exists as a mixture of two rotamers about χ_1, you should see the electron density for those separately in the crystal structure. In other words, you see a *superposition* of the structures that are present. In NMR you see an *average* of the structures present. It's not clear to me that you would necessarily expect a B factor and the ^{13}C relaxation time data to match exactly, because the time scales of the two experiments are completely different.

Bax: That's true, but we find that the NMR order parameter is smaller than the crystallographic data indicate, whereas if the problem were one of time scales you would expect the NMR order parameter to be too high. I should stress that this is preliminary data, so there might still be problems with interpretation of our relaxation data, but it does seem that some of the internal side chains have much greater mobility than crystallographic B factors suggest.

Roberts: There is also the problem of separating order parameter and correlation time in the analysis of the relaxation data, which is not entirely unambiguous.

Bax: It is unambiguous if you have three parameters available per carbon—that is, the T_1 and T_2 relaxation times and the NOE.

Roberts: That is correct, provided you assume that there is a single correlation time for that carbon–proton vector.

Bax: No, one can characterize motions on several different time scales because there is overdetermination (Clore et al 1990).

Hendrickson: What proteins are you referring to?

Bax: I am referring to data recorded by Lewis Kay, in collaboration with Dennis Torchia and his group, on staphylococcal nuclease, looking at leucine residues and comparing them with the 1.65 Å crystal structure with a 17% R factor determined by Loll & Lattman (1989).

Holmes: Micheal Levitt did molecular dynamics on BPTI, then analysed the motion into normal modes by spectral decomposition (Levitt et al 1985). On a nanosecond time scale only the lowest frequency modes are seen, so the structure appears to be vibrating coherently; internal motions seem to be moving along lines and look very anisotropic, whereas the total motion is in fact stochastic. The X-ray crystallographer sees this stochastic motion, with all possible time ranges, but what you are looking at is really just the lowest end of the spectrum.

Bax: Vibrational motions wouldn't affect the order parameter much because they don't change the orientation of a particular vector with respect to the magnetic field.

Roberts: It seems to me that it's the middle range of frequencies that affect ^{13}C relaxation. The very fast vibrations don't have any effect, as you say, and anything that is substantially slower than the overall correlation time of the protein similarly has no effect. There's a sort of 'window' of rates of internal motion of 10^{11}–$10^{8}\ s^{-1}$ that affect the relaxation rates.

Phillips: I'm intrigued by two technical issues that have arisen. In 1965 or 1966 I heard Oleg Jardetsky talk about NMR studies of interactions between lysozyme and sugars of various kinds. Soon after that he was using different isotopes in his NMR experiments. After dropping out of fashion, there seems to be a resurgence in the use of isotopes in NMR, at least in some schools. How important is the supply or incorporation of isotopes going to be in NMR studies? Also, Ad Bax said that although his experiments could have been done with a 400 MHz spectrometer, he did them with a 500 MHz. The question is, then, how important is the current work on the development of 750 MHz or 1000 MHz machines?

Bax: If you are pushing the limits, you want to go to higher field, even with a labelled protein. Calmodulin complexed with its target peptide has a molecular mass of 20 kDa; with more accumulation time we could have done the work at 400 MHz. However, for a 30 kDa protein we really would need a 600 MHz spectrometer.

Phillips: I suppose a 750 MHz machine would be even better.

Bax: 750 MHz would be better, although there are other problems with going to the higher field. The main advantage of going to 750 MHz would be where you want to look at proteins that cannot be labelled. If you wanted to study a 20 kDa protein that you cannot label, you could not possibly do it with a 500 or 600 MHz spectrometer, whereas you might be able to do it with the 750 MHz machine.

Richards: I had the pleasure of providing laboratory space for several years to David LeMaster. His attitude to NMR was that you should design the sample so that it will give you a spectrum that you can interpret and that will provide the answer to your questions. It's possible today, instead of doing uniform labelling, to label specifically where you want to see something. There may be a future in such an approach.

Bax: David LeMaster worked on thioredoxin from *E. coli* for about 10 years. With a 600 MHz spectrometer Julie Forman solved the structure of human thioredoxin in about a year (Forman-Kay et al 1991).

Richards: When David LeMaster started the 500 MHz spectrometer hadn't been invented. The entire field of NMR, in both megahertz and dimensions, developed at the time he was doing this work. Also, he did all the molecular biology with his own hands, he prepared the labelled amino acids and developed ways of separating them. It seems to me that for many purposes you don't want uniformly labelled proteins, you want highly labelled carbons, for example, at specific points. That is more difficult to do than uniform labelling and requires expertise in molecular biology or chemical synthesis. If there is a future in such an approach, we should plan for it, because few NMR spectroscopists will do organic synthesis, or develop expression systems, the way David LeMaster did. They will need to be provided with the materials, probably the labelled proteins.

Roberts: I think the nature of the labelling used depends on the kind of information you want. If you are looking for an NMR determination of the solution structure, then selective isotope labelling is an enormously lengthy business because you have to put in not only each individual kind of residue, but also all pair-wise combinations. For structure determination that's probably not a profitable way forward. What one might do is to combine David LeMaster's fractional deuteration with the general ^{13}C and ^{15}N labelling described by Ad Bax, because the deuteration would give sharper lines and longer T_2 relaxation times, to allow you to push a little bit further forward towards bigger proteins.

Richards: I have yet to be convinced that NMR is the way to solve structures for those proteins that will crystallize. You might as well start with the crystal structure and then use NMR for the things for which it is uniquely suited.

Bax: You may be able to solve the structure faster by NMR.

Blundell: We all have examples of structures that have taken a long time to solve by X-ray crystallography. For example, we have worked on the crystal structure of nerve growth factor for fourteen years. When there are no suitable groups for heavy atoms nor homologues for molecular replacement, solving the structure can take some time.

Phillips: I agree with Fred Richards on this issue. There is a disappointingly small number of NMR experiments that have addressed such things

as enzyme activity, where one would expect there to be some scope for NMR.

Wüthrich: NMR instruments have improved and there have been a lot of technical developments. At the moment, it is attractive to do heteronuclear experiments.

Bax: I agree. In 1988 two of the spectrometers in our laboratory were not capable of doing triple resonance experiments. The market is now reacting faster and the new spectrometers can do those experiments almost without modification, which is why isotopes are becoming so popular.

Wüthrich: It is clear that the next step will be fractional labelling. We have already got stuck with complete labelling.

Bax: I haven't got stuck yet.

Wüthrich: We have systems where we did, and we then had to resort to labelling individual residues selectively, which is even more expensive. What is your opinion about doing these experiments at natural isotope abundance?

Bax: It would probably be a waste of time. If you rely on natural abundance you lose a factor of 100 in the signal–noise ratio for carbon and of 300 for nitrogen. We need both of them so we would lose a factor of 30 000. Even a factor of 100 would be too great a loss at a 1 mM sample concentration. With a protein that is soluble at 10 mM you would be able to do the carbon experiments without labelling, but not the nitrogen, apart from recording regular shift correlation two-dimensional spectra.

Perham: This brings us back to a point made earlier, that 20 years ago the recombinant DNA revolution hadn't occurred—one couldn't express isolated genes in bacteria or do the isotopic labelling that is now relatively cheap and straightforward. Moreover, domains of complex proteins are now being expressed from sub-genes and these should be instantly accessible to attack by NMR. You might be lucky in crystallizing a protein in the first three months, but, equally, it might never crystallize.

Richards: One problem we have to face is in determining the structures of membrane proteins, which crystallographically are extraordinarily difficult to deal with. People do what they can. They use an enzyme to cut off the outer piece and they crystallize it, but the protein's function is intimately related to the fact that it is a membrane protein, frequently with domains on either side of the membrane. You learn nothing about signal transmission by looking at one of the domains by itself.

Perham: I am reminded of the joke about the man who was seen under a lamp-post looking for his wallet and his friend says 'Why are you looking for your wallet over here, you lost it over there?' and he replies 'I know, but there's no light over there, so I'd never be able to find it'!

Richards: With the examples of multidomain proteins we heard about

earlier, the actins and hexokinases etc., if you had cut one of those in the middle and looked at the two domains separately, some information would have been revealed but the interest level would have been orders of magnitude lower.

Dobson: Ad Bax made a very important point when he said that it is the motional correlation time, not the size of the protein, that determines what can be done by NMR spectroscopy. One of the interesting things about multidomain proteins is that the correlation times for different parts of the protein may differ significantly, because of flexibility in linker regions between domains. Where such flexibility exists, it should be possible to get detailed information from NMR about much larger proteins than one might otherwise have expected (Oswald et al 1989). Such proteins may well turn out to be the sorts of proteins that won't readily crystallize, because of their flexibility, making NMR studies particularly important. Despite doubts being expressed about the limitations apparently inherent in NMR methods, the size of proteins amenable to study has steadily increased over the last 20 years. I am very optimistic about our ability to study at least certain types of larger proteins, including those which, because of their dynamic behaviour, might be particularly interesting to study.

Richards: With a large domain protein how much motion does there have to be to make it easy to 'see' by NMR?

Dobson: That depends on what you are hoping to see in the NMR spectrum. To obtain a high resolution spectrum for a particular domain a lot of flexibility would be needed.

Richards: What would 'a lot' be?

Dobson: It seems likely that any linker between globular domains would need to be more flexible than that of calmodulin appears to be.

We have looked, for example, at several proteins involved in fibrinolysis. Even for plasminogen, which has a relative molecular mass of 100 000 kDa, we can see resonances from some of the domains, and we can get information about the way these interact with other domains (Teuten et al 1991).

Richards: Conceptually, it makes a lot of difference whether you rotate the molecule through 20° or 30°, which is possible in many cases, or whether you have to rotate through 180° to get the averaging effects you need.

Dobson: One needs a significant degree of spatial averaging. The types of motions that might be involved in hinge-bending, for example, are probably not sufficient to cause a very large change in resonance line widths. A greater degree of flexibility is likely to be needed.

Dodson: The analysis of isolated domains from membrane proteins or from large assemblies actually *is* useful. In a system that won't crystallize and is not tractable to NMR, the three-dimensional structures of the individual elements can be determined by X-ray analysis. We should pool techniques. NMR studies will help and perhaps molecular dynamics will be more tractable in the future. I must say that I don't agree that the study of subunits is uninformative.

In our laboratory we have just determined the crystal structure of a DNA gyrase B chain fragment (Wigley et al 1991). This crystallizes as a dimer. Completely unexpectedly, they found the ATP site, which involves structures from both components in the B chain, thus cross-linking the dimer. That important information will lead to useful experiments on the ATP–enzyme interactions. It cannot be obtained from the complete A_2B_2 structure, which is very large and has not been crystallized.

References

Babu YS, Sack JS, Greenhough TJ, Bugg CE, Means AR, Cook WJ 1985 Three-dimensional structure of calmodulin. Nature (Lond) 315:37–40

Babu YS, Bugg CE, Cook WJ 1988 Structure of calmodulin refined at 2.2 Å resolution. J Mol Biol 204:191–204

Blundell TL, Sibanda BL, Pearl L 1983 Three-dimensional structure, specificity and catalytic mechanism of renin. Nature (Lond) 304:273–275

Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, Gronenborn AM 1990 Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. J Am Chem Soc 112:4989–4991

Draibikowski W, Brzeska H, Venyaminov SY 1982 Tryptic fragments of calmodulin. J Biol Chem 257:11584–11590

Englander SW, Kallenbach NR 1984 Hydrogen exchange and structural dynamics of proteins and nucleic acids. Q Rev Biophys 16:521–655

Fontana A 1989 Limited proteolysis of globular proteins occurs at exposed and flexible loops. In: Kotyk A, Skoda J, Paces C, Kostka V (eds) Highlights of modern biochemistry (Proc 14th Int Congr Biochem). VSP, Utrecht, vol 2:1711–1726

Fontana A, Fassina G, Vita C, Dalzoppo D, Zamai M, Zambonin M 1986 Correlation between sites of limited proteolysis and segmental mobility in thermolysin. Biochemistry 25:1847–1851

Forman-Kay JD, Clore GM, Wingfield PT, Gronenborn AM 1991 High-resolution three-dimensional structure of reduced recombinant human thioredoxin in solution. Biochemistry 30:2685–2698

Heidorn DB, Trewhella J 1988 Comparison of the crystal and solution structures of calmodulin and troponin C. Biochemistry 27:909–915

Karplus M, Dobson CM 1986 Internal motion of proteins: nuclear magnetic resonance measurements and dynamic simulations. Methods Enzymol 131:362–389

Kay LE, Torchia DA, Bax A 1989 Backbone dynamics of proteins as studied by ^{15}N inverse detected heteronuclear NMR spectroscopy: application to Staphylococcal nuclease. Biochemistry 28:8972–8979

Levitt M, Sander C, Stern PS 1985 Protein normal mode dynamics—trypsin inhibitor, cranbin, ribonuclease, and lysozyme. J Mol Biol 181:423–447

Loll PG, Lattman EE 1989 The crystal structure of the ternary complex of staphylococcal nuclease, Ca^{2+}, and the inhibitor pdTp, refined at 1.65 Å. Proteins Struct Funct Genet 5:183–201

Oswald RE, Bogusky MJ, Bamberger M, Smith RAG, Dobson CM 1989 Dynamics of the multidomain fibrinolytic protein urokinase from two-dimensional NMR. Nature (Lond) 337:579–581

Persechini A, Kretsinger RH 1988 The central helix of calmodulin functions as a flexible tether. J Biol Chem 263:1275–1278

Richarz R, Nagayama K, Wüthrich K 1980 Carbon-13 nuclear magnetic resonance relaxation studies of internal mobility of the polypeptide chain in basic pancreatic trypsin inhibitor and a selectively reduced analogue. Biochemistry 19:5189–5196

Smith LJ, Sutcliffe MJ, Redfield C, Dobson CM 1991 Analysis of ϕ and χ_1 torsion angles for hen lysozyme in solution from ^{1}H NMR spin–spin coupling constants. Biochemistry 30:986–996

Spera S, Ikura M, Bax A 1991 Measurement of the exchange rates of rapidly exchanging amide protons: application to the study of calmodulin and its complex with a myosin light chain kinase fragment. J Biomol NMR 1:155–165

Teuten AJ, Smith RAG, Dobson CM 1991 Domain interactions in human plasminogen studied by proton NMR. FEBS (Fed Eur Biochem Soc) Lett 278:17–22

Wigley DB, Davies GD, Dodson EJ, Maxwell A, Dodson GG 1991 Crystal structure of an N-terminal fragment of the DNA gyrase B protein. Nature (Lond) 351:624–629