

## HETERONUCLEAR 3D NMR AND ISOTOPIC LABELING OF CALMODULIN

### TOWARDS THE COMPLETE ASSIGNMENT OF THE $^1\text{H}$ NMR SPECTRUM

MITSUHIKO IKURA,\*† DOMINIQUE MARION,\*‡ LEWIS E. KAY,\* HENRY SHIH,\* MARIE KRINKS,§ CLAUDE B. KLEE§ and AD BAX\*†

\* Laboratory of Chemical Physics, NIDDK, and § Laboratory of Biochemistry, NCI, National Institutes of Health, Bethesda, MD 20892, U.S.A.

**Abstract**—New methods are described that permit detailed analysis of the NMR spectra of calmodulin, an  $\alpha$ -helical protein with a molecular weight of 16.7 kD. Two complementary approaches have been used: uniform labeling with  $^{15}\text{N}$  and labeling of specific amino acids with either  $^{15}\text{N}$  or  $^{13}\text{C}$ . It is demonstrated that uniform  $^{15}\text{N}$  labeling permits the recording of sensitive three-dimensional (3D) NMR spectra that show far better resolution than their conventional two-dimensional analogs. Selective  $^{15}\text{N}$  labeling of amino acids can be used for identifying the type of amino acid, providing information that is essential for the analysis of the 3D spectra. Simultaneous selective labeling with both  $^{15}\text{N}$  and  $^{13}\text{C}$  can provide a number of unique backbone assignments from which sequential assignment can be continued.

It has been well-established that solution structures of small proteins (<100 residues) can be determined by NMR. The first and most important step of the NMR structure determination process is to assign the  $^1\text{H}$  NMR spectrum of the protein. For small proteins, systematic analysis of the two-dimensional nuclear Overhauser effect spectroscopy (2D NOESY) and  $J$  correlation spectra often provides sufficient information for complete structure determination [1, 2]. If necessary, remaining ambiguities frequently can be resolved by recording the spectra at several different temperatures or pH values. For larger proteins the analysis becomes increasingly difficult because the density of resonances in the 2D spectra increases at least linearly with size, and also because the line width of individual resonances increases with molecular weight. Thus, complete analysis of the NMR spectrum of proteins larger than 100 residues is often impossible when following the standard procedure mentioned above. The assignment problem is particularly difficult for proteins rich in  $\alpha$ -helical structure, which typically show very narrow chemical shift dispersion for both the amide and the  $\text{C}\alpha$  protons. On the other hand, for a given molecular weight, proteins rich in  $\beta$ -sheet lend themselves more readily to complete NMR analysis because of the typically rather good shift dispersion [3–5]. Here we discuss new NMR methods, all relying on isotopic  $^{13}\text{C}$  or  $^{15}\text{N}$  enrichment, that can extend the molecular weight range of proteins for which a detailed solution structure can be obtained.

Two different approaches have been developed for simplifying the spectra of large proteins. The first approach requires isotopic labeling of selected amino acids. So-called editing experiments permit the regular 2D  $^1\text{H}$  NMR spectra of these selectively labeled

proteins to be simplified by exclusively monitoring interactions with protons attached to the  $^{13}\text{C}$  or  $^{15}\text{N}$ -labeled site [6–12]. Alternatively, heteronuclear chemical shift correlation of the labeled atoms with their directly or indirectly attached protons can provide useful information. These procedures have been tested exhaustively for three different proteins, staphylococcal nuclease (17.6 kD) [12, 13], T4 lysozyme (19 kD) [14, 15] and *Streptomyces* subtilisin inhibitor (23 kD) [16, 17].

The second approach utilizes uniform labeling of the protein with either  $^{15}\text{N}$ ,  $^{13}\text{C}$ , or both. These uniformly labeled protein samples permit the recording of sensitive three-dimensional heteronuclear NMR spectra, in which the overlap present in the corresponding 2D spectra is largely removed [18–21].

In this study, we report our strategy for assigning the  $^1\text{H}$  NMR spectrum of calmodulin (16.7 kD, 148 amino acid residues). The physiological function of calmodulin is to transmit a  $\text{Ca}^{2+}$  signal to proteins inside the cell [22]. The protein is rich in  $\alpha$ -helix and therefore it presents a challenge to modern NMR techniques, despite its relatively modest molecular weight. In the past, different research groups have applied a variety of 2D NMR techniques to calmodulin and its proteolytic fragments [23–36]. However, complete analysis of the  $^1\text{H}$  NMR spectrum of the intact protein has not been possible. We will demonstrate the combined use of heteronuclear 3D methods and selective labeling with  $^{15}\text{N}$  and/or  $^{13}\text{C}$  for assigning the  $^1\text{H}$  NMR spectrum, with special emphasis on the power of 3D NMR techniques for the analysis of  $\alpha$ -helical proteins.

† Address reprint requests to either M. Ikura or A. Bax, NIH, Building 2, Room 109, Bethesda, MD 20892.

‡ Present address: Centre de Biophysique Moleculaire, Centre National de la Recherche Scientifique, 45071 Orleans Cedex 2, France.

#### MATERIALS AND METHODS

The *Drosophila* calmodulin gene [31] is over-expressed in *Escherichia coli* strain AR58 using the

pas expression system which contains the temperature-sensitive phage  $\lambda$  P<sub>L</sub> promoter [32].\* The cells were grown at 32°, and prior to the stationary phase (O.D.<sub>650</sub> = 1.0–1.2) the temperature was shifted to 42° for thermal induction of the protein. Uniformly (~95%) <sup>15</sup>N-labeled protein was prepared using M9 minimal medium with <sup>15</sup>NH<sub>4</sub>Cl as the sole nitrogen source. Specific labeling with particular amino acids was carried out according to the procedure reported by Hibler *et al.* [33]; 50–70 mg of labeled L-amino acid(s) was added to 1.8 L of cell culture just 15 min prior to adding another 0.6 L of the M9 medium prewarmed at 65°. The protein induction period was 3.5 hr for uniform <sup>15</sup>N labeling and 1.5 hr for specific labeling. Isolation and purification of the protein from the cells were performed by a combination of procedures described previously [34, 35]. We typically obtained 15–25 mg of the homogeneous protein.

All NMR experiments reported here were performed using a 1.5 mM sample dissolved in 90% H<sub>2</sub>O/10% D<sub>2</sub>O, 0.1 M KCl, 6.2 mM Ca<sup>2+</sup>, pH 6.2, 45°. Spectra were recorded on modified Bruker AM-500 and AM-600 spectrometers. The 3D NMR spectra result from data matrices with 128 complex points in the F<sub>1</sub> dimension (<sup>1</sup>H), 32 complex points in the F<sub>2</sub> dimension (<sup>15</sup>N), and 512 real points in the F<sub>3</sub> dimension, corresponding to acquisition times of 25, 25 and 80 msec in the t<sub>1</sub>, to t<sub>2</sub> and t<sub>3</sub> dimension respectively. After zero filling, the digital resolution in the F<sub>1</sub>, F<sub>2</sub> and F<sub>3</sub> dimension was 20, 20 and 6.7 Hz respectively. The total recording time for each 3D spectrum was 2.5 days. The 3D data sets were processed on a SUN-4 computer using a simple home-written program for the F<sub>2</sub> Fourier transformation [36], combined with commercial 2D software (NMR2 from New Methods Research Inc., Syracuse, NY) for the F<sub>1</sub> and F<sub>3</sub> 2D Fourier transformation. 2D <sup>1</sup>H-<sup>15</sup>N shift correlation spectra were recorded using the Overboderhausen pulse sequence [37, 38]. These correlation spectra were recorded with substantial time-averaging (16 hr per spectrum) in order to also observe weak correlations from transaminated (cross-labeled) residues and to obtain accurate resonance intensities.

## RESULTS

### 3D NOESY-HMQC and 3D HOHAHA-HMQC of <sup>15</sup>N-labeled calmodulin

We have used two 3D experiments that combine the heteronuclear multiple quantum correlation (HMQC) method [39, 40] either with the homonuclear NOESY experiment [41] or with the homonuclear Hartmann-Hahn (HOHAHA) [42, 43] method. The pulse schemes used in this study and the technical details regarding the recording of the 3D NMR experiments and the data processing have been described elsewhere [19, 36]. Figure 1 gives a schematic view of the heteronuclear 3D NOESY-HMQC spectrum. This 3D spectrum can be considered as a set of 2D NOESY spectra that are

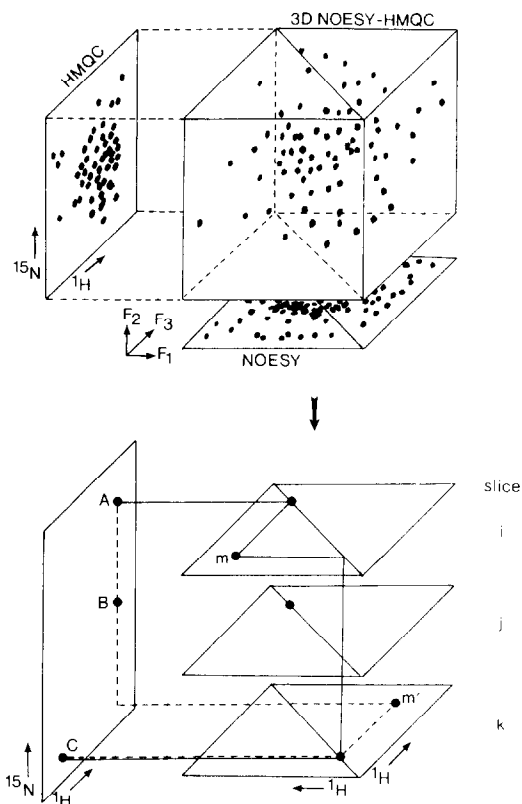


Fig. 1. Schematic representation of the NH-NH region of the heteronuclear 3D NOESY-HMQC spectrum. The projections of the 3D spectrum onto the F<sub>2</sub> = 0 plane (corresponding to 2D NOESY) and onto the F<sub>1</sub> plane (corresponding to 2D <sup>1</sup>H-<sup>15</sup>N HMQC) are shown. The bottom half of the figure illustrates the use of "mirror image" cross peaks to overcome degeneracy of the <sup>1</sup>H chemical shift of two amide protons (e.g. peaks A and B). The cross peak labeled m' alone would not distinguish between CB or CA connectivity. However, the "mirror image" cross peak (labeled m) shows that the correlation is between C and A.

separated along the F<sub>2</sub> axis according to the <sup>15</sup>N shift of the nitrogen attached to the amide proton that appears on the F<sub>1</sub> = F<sub>3</sub> diagonal. The projection of the 3D spectrum onto the F<sub>2</sub> = 0 plane corresponds to the regular 2D NOESY spectrum; the projection onto the F<sub>1</sub> = 0 plane corresponds to the 2D <sup>1</sup>H-<sup>15</sup>N HMQC correlation spectrum. The number of resonances present in each (F<sub>1</sub>, F<sub>3</sub>) NOESY slice of the 3D spectrum is greatly reduced relative to the regular 2D NOESY spectrum, solving the overlap problem.

In the regular NOESY spectrum, cross peaks appear as pairs, at mirror image positions with respect to the diagonal. In slices through the 3D spectrum, taken perpendicular to the F<sub>2</sub> axis (see, for example, Fig. 2, b and c), cross peaks only occur at the F<sub>3</sub> chemical shift position of the <sup>1</sup>H detected during the acquisition time, t<sub>3</sub>, of the 3D experiment. This results from the HMQC step prior to data acquisition, which ensures that only amide proton frequencies appear in the F<sub>3</sub> dimension. Therefore, for NH-CH NOE interactions, the mirror image

\* Maune JF, Klee CB and Beckingham K, manuscript in preparation.

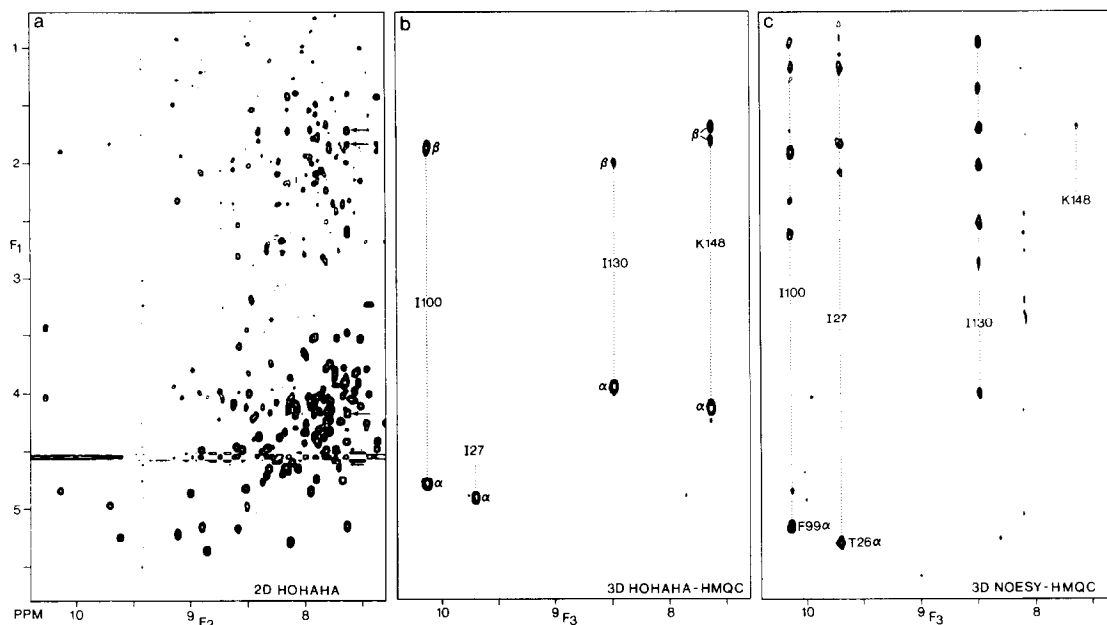


Fig. 2. (a) Aliphatic region of the homonuclear 2D HOHAHA spectrum recorded with  $^{15}\text{N}$  decoupling of a uniformly  $^{15}\text{N}$ -labeled sample of 1.5 mM calmodulin. Corresponding regions are shown of ( $F_1, F_3$ ) slices taken at  $F_2 = 127.2$  ppm from (b) the heteronuclear 3D HOHAHA-HMQC spectrum and (c) the 3D NOESY-HMQC spectrum. 2D HOHAHA and 3D HOHAHA-HMQC spectra were recorded at 500 MHz  $^1\text{H}$  frequency; the 3D NOESY-HMQC spectrum was recorded at 600 MHz. All spectra were recorded at  $45^\circ$  and pH 6.2, in the presence of 6.2 mM  $\text{Ca}^{2+}$  and 100 mM KCl. For the HOHAHA experiments 39-msec mixing with a WALTZ17 sequence [45] was used. For the NOESY experiment a 120-msec mixing period was used.

cross peak present in 2D NOESY is absent in 3D NOESY-HMQC spectra. However, for two interacting amide protons (Fig. 1), the "mirror images" with respect to the  $F_1 = F_3$  diagonal [44] are found in the two planes that correspond to the  $^{15}\text{N}$   $F_2$  frequencies of the two amides. The presence of this mirror image is extremely useful because it yields unambiguous NOE correlations between resonances present in the 2D  $^1\text{H}$ - $^{15}\text{N}$  shift correlation spectrum, i.e. for every NH-NH NOE interaction four frequency coordinates determine the position of the two "mirror image" cross peaks, the two  $^1\text{H}$  chemical shifts and the two  $^{15}\text{N}$  chemical shifts.

Figure 2 compares a small region of the regular 2D HOHAHA spectrum of calmodulin with slices taken from the 3D HOHAHA-HMQC and the 3D NOESY-HMQC spectra. Comparison of panels (a) and (b) shows the dramatic spectral simplification that is obtained with the 3D method. Figure 2 also indicates that the sensitivity of the 2D and 3D methods is comparable. Note that the cross peaks between the Lys-148 NH resonance and the  $C\alpha$  and  $C\beta$  protons (marked by arrows) could not be identified uniquely in the 2D HOHAHA spectrum because the Lys-21 NH proton shift overlaps with several other protons. This is the case for the vast majority of the NH proton resonances, most of them present in the 7.5–8.5 ppm region, as expected for a protein rich in  $\alpha$ -helix. To make a detailed analysis of the NMR spectrum of such a protein, it is essential to improve resonance separation in this region. Since

most of the NH resonances between 7.5 and 8.5 ppm are reasonably well-resolved in the 2D  $^1\text{H}$ - $^{15}\text{N}$  shift correlation spectrum (Fig. 4b), spreading the 2D spectrum in a third ( $^{15}\text{N}$ ) dimension provides an ideal avenue for removing resonance overlap. Thus, with the  $^{15}\text{N}$  3D approach, overlap of amide proton resonances in the fingerprint region can be removed almost completely.

In the 3D NOESY-HMQC slice shown in Fig. 2c, relatively strong  $\text{CaH}_i\text{-NH}_{i+1}$  ( $d_{\alpha\text{N}}$ ) NOEs are observed for Phe-99/Ile-100 and for Thr-26/Ile-27, whereas intraresidue  $\text{CaH}_i\text{-NH}_i$  NOEs are very weak. This NOE pattern is characteristic for extended  $\beta$  strand conformations, and indeed these residues are parts of two short antiparallel  $\beta$ -sheets, present between  $\text{Ca}^{2+}$  binding loops I and II and between loops III and IV [23, 28, 46].

On the other hand,  $\alpha$ -helices exhibit relatively strong NOE interactions between sequential amide protons ( $d_{\text{NN}}$  connectivity) and weak  $d_{\alpha\text{N}}$  connectivity. Analysis of the  $\text{NH}_i\text{-NH}_{i+1}$  cross peaks in the 7.5–8.5 ppm region is therefore necessary to obtain sequential assignments of NH protons in  $\alpha$ -helices. Figure 3 illustrates this assignment procedure for the residues Phe-68 to Met-72. As mentioned before, unambiguous identification of  $\text{NH}_i\text{-NH}_{i+1}$  cross peaks can be made based on symmetric pairs of cross peaks (labeled a and a' for Phe-68/Leu-69, b and b' for Leu-69/Thr-70, and c and c' for Thr-70/Met-71). Despite the removal of overlap in the 3D spectra, analysis of NOE connectivities is still very laborious

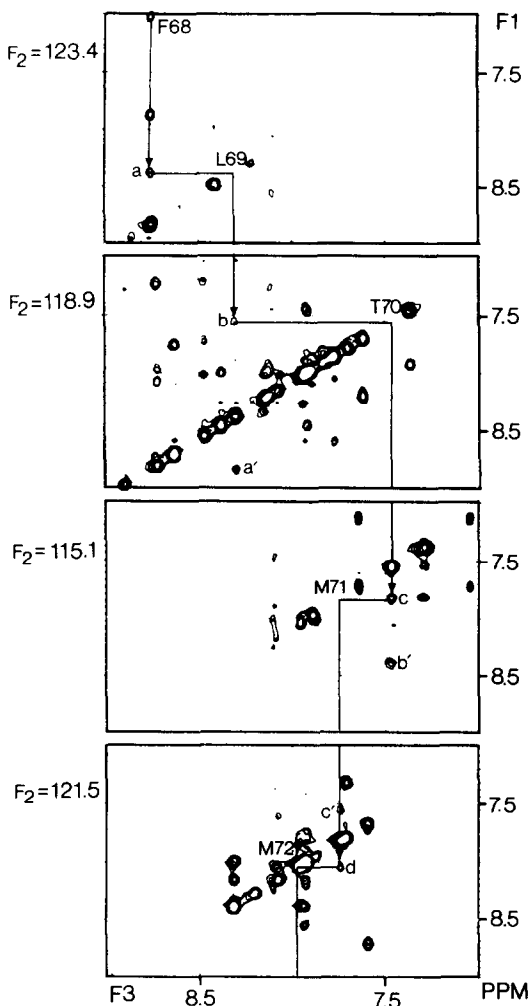


Fig. 3. Example of sequential assignment via  $\text{NH}_i\text{-NH}_{i+1}$  connectivities in the heteronuclear 3D NOESY-HMOC spectrum. Four different slices were taken at  $^{15}\text{N}$  chemical shifts of 123.4, 118.9, 115.1 and 121.5 ppm, corresponding to the  $^{15}\text{N}$  chemical shifts of Phe-68, Leu-69, Thr-70 and Met-71 respectively. These slices illustrate the sequential connectivities for residues Phe-68 to Met-72. The "mirror image" cross peaks (mentioned in the text) are labeled a and a' for Phe-68/Leu-69, b and b' for Leu-69/Thr-70, and c and c' for Thr-70/Met-71.

because of the large number of resolved cross peaks in the 3D NOESY-HMOC spectrum (*ca.* 1200 in the case of calmodulin). To expedite this task, we have developed relatively simple computer programs to aid the analysis of 3D NMR spectra. So far, routines have been written for "pseudo 3D peak picking," a program that utilizes the information of the high-resolution 2D  $^1\text{H}\text{-}^{15}\text{N}$  correlation spectrum to distinguish real peaks from noise and artifacts, and various "pattern search" programs for identifying  $\text{NH}_i\text{-NH}_{i+1}$  and  $\text{C}\alpha\text{H}_i\text{-NH}_{i+1}$  connectivities and routines that search for particular amino acid spin systems.

#### Specific labeling

As demonstrated above, 3D NOESY-HMOC and

HOHAHA-HMOC methods using uniformly  $^{15}\text{N}$ -labeled calmodulin are very useful for finding stretches of sequential  $\text{C}\alpha\text{H}_i\text{-NH}_{i+1}$  and  $\text{NH}_i\text{-NH}_{i+1}$  connectivities. However, for reliable sequential assignment it is essential to have unambiguous residue type assignments for a substantial number of backbone  $^1\text{H}$  resonances. For smaller proteins this information usually can be extracted from the  $J$  connectivity experiments such as COSY and HOHAHA. However, for calmodulin, the 3D HOHAHA- $(^{15}\text{N})$ HMOC experiment provides only very weak relayed connectivity between NH and  $\text{C}\beta\text{H}$  protons, making it difficult to establish the type of amino acid. These indirect connectivities are for the most part unobservable in calmodulin because for this size protein the  $^1\text{H}$  line width is too large for efficient transmittal of  $J$  coupling information. The COSY and HOHAHA spectra of the protein dissolved in  $\text{D}_2\text{O}$  (without the amide protons present) show very severe overlap and relatively weak cross peaks because of the high molecular weight.

To solve the problem of obtaining residue type assignments, two different approaches are possible. The first one uses incorporation in the protein of a small number of  $^{13}\text{C}$ -labeled amino acids. The NMR spectrum then can be edited to show only resonances of these labeled amino acids. This technique has been used very extensively for the proteins staphylococcal nuclease [11, 13] and T4 lysozyme [14, 15]. More recently, a second approach has been developed, relying on heteronuclear 3D NMR of uniformly  $^{13}\text{C}$ -labeled proteins [47, 48]. This latter approach provides efficient  $J$  correlation between vicinal protons by utilizing a pathway that consists of the well resolved  $^1J_{\text{CH}}$  and  $^1J_{\text{CC}}$  couplings for transferring magnetization from one spin to another. In addition, this method removes overlap by spreading the regular COSY-type spectrum into a third dimension, the  $^{13}\text{C}$  chemical shift. Although this approach appears to be extremely promising, its general utility has not yet been proven, and here we limit our discussion to the selective labeling approach.

Figure 4a shows the most crowded region of the  $^1\text{H}\text{-}^{15}\text{N}$  correlation spectrum of a calmodulin sample labeled with  $[^{15}\text{N}]\text{Met}$ ,  $[^{15}\text{N}]\text{Glu}$ , and  $[1\text{-}^{13}\text{C}]\text{Lys}$ . The corresponding region of the spectrum recorded for uniformly  $^{15}\text{N}$ -labeled calmodulin is shown in Fig. 4b. As expected, a comparison of these two spectra shows that the number of correlations in the selectively labeled sample is much lower. However, the total number of correlations observed in the spectrum of Fig. 4a (about 90 when the entire spectrum is plotted at a low contour level) is much larger than the sum of the number of methionine (9) and glutamic acid (21) residues. This larger than expected number of resonances results from transamination and other metabolic processes that cause  $^{15}\text{N}$  labeling (at a lower level) of many other residues [49, 50]. The level of cross labeling depends strongly on the types of amino acids involved and also on the bacterial strain used and on the duration of the induction period. As will be discussed below, this cross labeling can be used to great advantage.

The intensities in the uniformly  $^{15}\text{N}$ -labeled spectrum show some variation, caused by differences in (a)  $^1\text{H}$  and  $^{15}\text{N}$  line widths, (b) hydrogen exchange

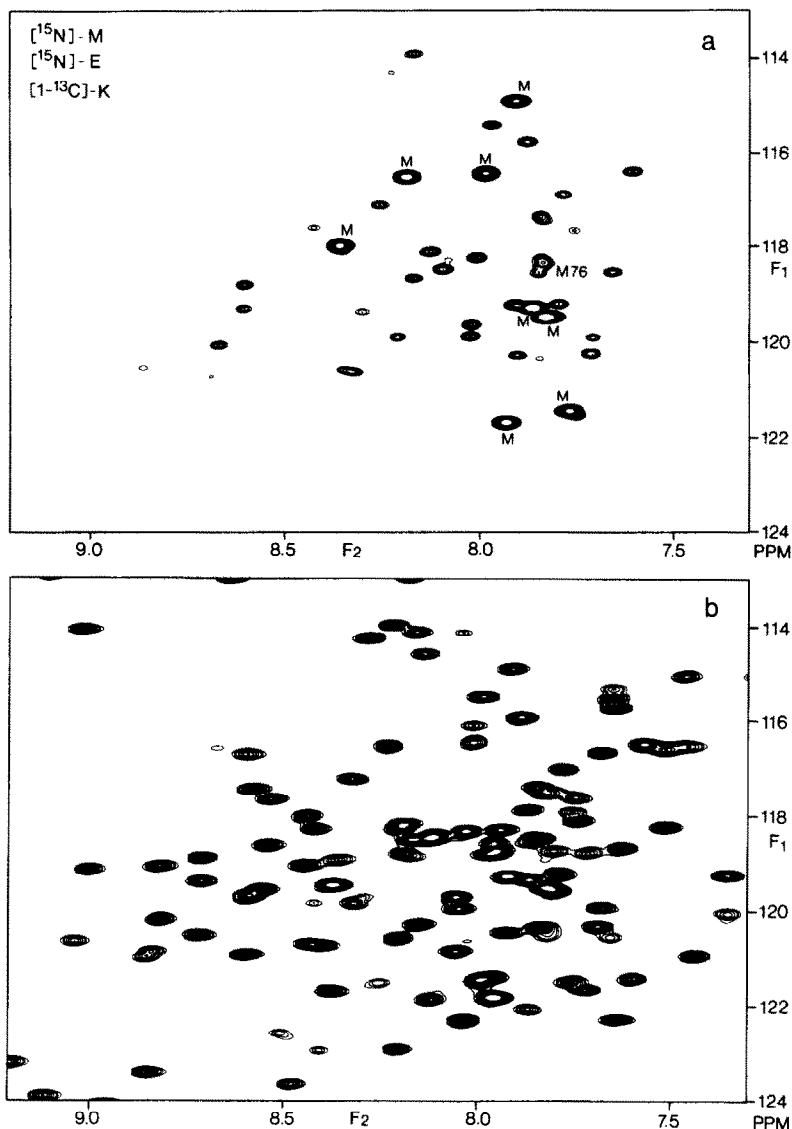


Fig. 4. The most crowded region of the  $^1\text{H}$ - $^{15}\text{N}$  correlation spectrum of 1.5 mM calmodulin (a) labeled with  $[^{15}\text{N}]\text{Met}$ ,  $[^{15}\text{N}]\text{Glu}$ , and  $[1-^{13}\text{C}]\text{Lys}$  and (b) uniformly labeled with  $^{15}\text{N}$ . Both spectra were recorded with the Overbundenhuasen pulse sequence [37, 38], using acquisition times of 128 msec ( $t_2$ ) and 150 msec ( $t_1$ ). Identical Lorentzian-Gaussian  $t_2$  filtering and  $45^\circ$  phase-shifted sinebell  $t_1$  filtering were used for the two spectra.

rates with solvent, and (c) NH-CaHJ couplings. The same factors, combined with the level of  $^{15}\text{N}$  labeling, determine the resonance intensity in the spectrum of Fig. 4a. Therefore, if the ratios of the peak intensities are calculated for corresponding resonances in the two spectra, this ratio is expected to be constant for a given type of amino acid. As can be clearly seen in Fig. 4a, there are eight very intense correlations (each with a peak intensity ratio Fig. 4a/Fig. 4b near unity), all corresponding to methionine residues. A further analysis shows that the intensity ratios for Glu, Gln, Asp and Asn residues are nearly indistinguishable (all about 20%). Much lower levels of  $^{15}\text{N}$  labeling are found for Phe (6%), Ser (5%), Ala (4%), Tyr (4%), Gly (2%), Val (1–2%), Ile (1–2%),

Leu (1–2%), and Lys (1%). No correlation for any of the other residues could be observed.

As mentioned above, the protein preparation also included  $[1-^{13}\text{C}]\text{Lys}$ . Because the  $^{15}\text{N}$  of Met-76 is directly bound to the carbonyl of Lys-75, the  $^{15}\text{N}$  resonance is expected to show the  $^1J_{\text{NC}}$  coupling ( $\approx 5$  Hz) [16, 51]. Indeed this resonance is observed in Fig. 4a at  $(F_1, F_2) = 118.6$  ppm, 7.85 ppm). However, because of overlap of one of the two doublet components with another correlation, the doublet appears asymmetric in intensity. This double-labeling procedure [11, 16] is extremely useful because it provides an unambiguous sequence-specific assignment for the backbone amide proton that can serve as an anchor point for further assignments.

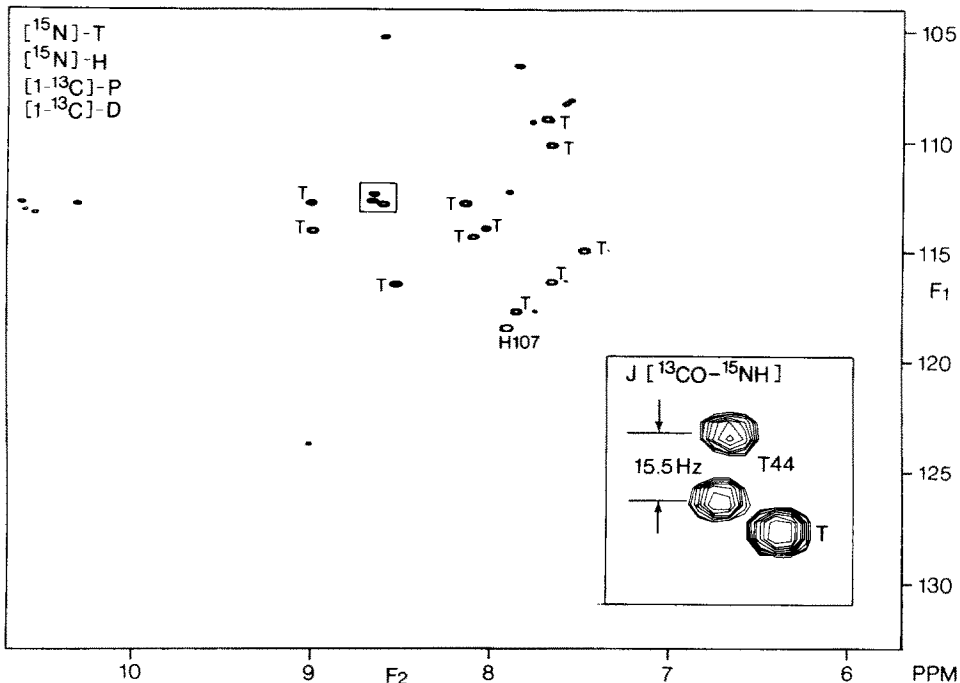


Fig. 5. The  $^1\text{H}$ - $^{15}\text{N}$  shift correlation spectrum of 1.5 mM calmodulin labeled with  $^{15}\text{N}$ Thr,  $^{15}\text{N}$ His,  $^{13}\text{C}$ Pro, and  $^{13}\text{C}$ Asp. The correlations of Thr residues and of His-107 are labeled. The acquisition and processing parameters were the same as for the spectra of Fig. 4.

Figure 5 shows another example of a  $^1\text{H}$ - $^{15}\text{N}$  shift correlation spectrum, obtained for a sample labeled with  $^{15}\text{N}$ Thr,  $^{15}\text{N}$ His,  $^{13}\text{C}$ Pro, and  $^{13}\text{C}$ Asp. In this case, comparison with the spectrum of uniformly  $^{15}\text{N}$ -labeled calmodulin shows that only a single peak (His-107) has a high level (>90%) of  $^{15}\text{N}$  labeling, whereas the Thr residues are easily recognized by their intermediate level ( $\approx 40\%$ ) of enrichment. The NH correlation (enlarged in the inset) of residue Thr-44, adjacent to Pro-43, is clearly recognized by the  $^1J_{\text{NC}}$  doublet splitting. There are several weaker cross peaks visible in the spectrum, presumably originating from cross labeling of  $^{15}\text{N}$ Thr. These much weaker correlations all correspond to Gly and Ser residues, consistent with our results from the 3D NMR analysis. It should be emphasized that these independent residue-type confirmations are important for obtaining reliable assignments in the early phase of the spectrum analysis. This particular protein preparation was intended to also incorporate  $^{13}\text{C}$ Asp. However, the spectrum suggests that the incorporation was unsuccessful since Thr-79 (adjacent to Asp-78, and identified at a later stage) shows up as a singlet, even at very low contour levels, indicating that the level of  $^{13}\text{C}$ Asp incorporation must have been less than about 20%.

#### DISCUSSION

Below, we briefly summarize our assignment strategy. The good spectral separation in  $^1\text{H}$ - $^{15}\text{N}$  correlation spectrum is crucial to our approach. For calmodulin, the  $^1\text{H}$ - $^{15}\text{N}$  correlation spectrum, when

recorded with the kind of technique first developed by Bodenhausen and Ruben [37] or variations thereof [38], shows that the vast majority of NH correlations are well resolved, despite a high  $\alpha$ -helical content (nearly 65% in the X-ray crystal structure).

The first step of our analysis gives identification (ID) numbers to each of the correlations in the  $^1\text{H}$ - $^{15}\text{N}$  shift correlation spectrum. These correlations directly correspond to the peptide backbone amide groups in the protein. In our case, 144 such correlations could be identified. This is a very large number considering that the maximum number of such correlations that might be expected for calmodulin (148 residues) is 145; it contains two prolines and the N-terminal  $\text{NH}_2$  group is invisible because its protons are in rapid exchange with solvent. For each ID number, the  $^1\text{H}$  and  $^{15}\text{N}$  chemical shifts are directly available from this NH correlation spectrum. For nearly all ID numbers the corresponding  $\text{CaH}$  shift could then be determined from the 3D HOHAHA-HMOC spectrum, and for a much smaller fraction the 3D spectrum also yielded  $\text{C}\beta$  proton shifts. The 3D NOESY-HMOC spectrum then identifies the NH or  $\text{CaH}$  shift (or both) of the preceding amino acid residue, providing the start of the sequential assignment procedure. At this stage, side chain spin system assignments also could be made for the single Tyr residue and for several of the Gly, Ala, Thr, Val, Ile, and Phe residues by using conventional 2D homonuclear purged COSY [52], HOHAHA, and NOESY spectra recorded in  $\text{D}_2\text{O}$ . Many of these sometimes tentative residue type assignments were confirmed by specific labeling experiments of the type described above. Specific labeling with the long

side chain amino acids such as Arg, Leu, Lys, Met, and Pro which are difficult to identify (even from the 3D spectra) greatly facilitates the sequential assignment process. Double-labeling with [<sup>15</sup>N]-enriched amino acids and [1-<sup>13</sup>C]-enriched amino acids also is important, mainly for establishing unambiguous anchor points in the sequential assignment process. In fortunate cases, sequence-specific assignment could be extended by up to nine amino acid residues starting from a single marker, before coincidental overlap or the lack of NOE connectivity interrupted the chain. As demonstrated elsewhere,\* by judiciously choosing the types of amino acids that are labeled, and by deliberately using different levels of enrichment for different amino acids, we were able to make nineteen such sequence-specific NH assignments plus a much larger number of type-specific assignments from a single protein preparation.

As mentioned before, the search for symmetric pairs of cross peaks in the NH region works extremely well for establishing d<sub>NN</sub> connectivity in  $\alpha$ -helical proteins. Indeed, this type of search provided the majority of the sequential connectivities found in calmodulin.

So far, we have obtained sequence-specific assignments for the backbone protons of 123 residues as well as a substantial number of side chain proton assignments. The main remaining problem concerns the assignment of several amino acid residues in three homologous sequences at Glu6-Glu7-Gln8, Glu82-Glu83-Glu84, and Asp118-Glu119-Glu120. New 3D techniques, relying on triple resonance spectroscopy of calmodulin labeled uniformly with both <sup>15</sup>N and <sup>13</sup>C, are under development to complete the remainder of the assignment task.

**Acknowledgements**—We thank Kathy Beckingham and John F. Maune for providing us with the *Drosophila* calmodulin coding construct, and Paul Driscoll, Steven W. Sparks, and Dennis A. Torchia for many valuable 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 financial support from the Medical Research Council of Canada and the Alberta Heritage Trust Foundation.

#### REFERENCES

1. Wüthrich K, *NMR of Proteins and Nucleic Acids*. John Wiley, New York, 1986.
2. Clore GM and Gronenborn AM, Three-dimensional structures of proteins in solution by nuclear magnetic resonance spectroscopy. *Protein Seq Data Anal* 2: 1–8, 1989.
3. Clayden NG and Williams RJP, Peptide group shifts. *J Magn Reson* 49: 383–396, 1982.
4. Dalgarno DC, Levine BA and Williams RJP, Structural information from NMR secondary chemical shifts of peptide  $\alpha$  C-H protons in proteins. *Biosci Rep* 3: 443–452, 1983.
5. Szilagy L and Jardetzky O,  $\alpha$ -Proton chemical shifts and secondary structure in proteins. *J Magn Reson* 83: 441–449, 1989.
6. Wilde JA, Bolton P, Stolowich NJ and Gerlt JA, A method for the observation of selected proton NMR resonances on proteins. *J Magn Reson* 68: 168–171, 1986.
7. McIntosh LP, Dahlquist FW and Redfield AG, Proton NMR and NOE structural and dynamic studies of larger proteins and nucleic acids aided by isotope labels: T4 lysozyme. *J Biomol Struct Dyn* 5: 21–34, 1987.
8. Senn H, Otting G and Wüthrich K, Protein structure and interactions by combined use of sequential NMR assignments and isotope labeling. *J Am Chem Soc* 109: 1090–1092, 1987.
9. Bax A and Weiss MA, Simplification of two-dimensional NOE spectra of proteins by <sup>13</sup>C labeling. *J Magn Reson* 71: 571–575, 1987.
10. Fesik SW, Luly JR, Erickson JW and Abad-Zapatero C, Isotope-edited proton NMR study on the structure of a pepsin/inhibitor complex. *Biochemistry* 27: 8297–8301, 1988.
11. Torchia DA, Sparks SW and Bax A, Staphylococcal Nuclease: Sequential assignments and solution structure. *Biochemistry* 28: 5509–5524, 1989.
12. Torchia DA, Sparks SW and Bax A, Proline assignments and identification of the *cis* K116/P117 peptide bond in liganded *S. nuclease* using isotope edited 2D NMR spectroscopy. *J Am Chem Soc* 111: 8315–8317, 1989.
13. Wang J, Hinck AP, Loh SN and Markley JL, Two-dimensional NMR studies of staphylococcal nuclease: 2. Sequence-specific assignments of carbon-13 and nitrogen-15 signals from the (nuclease H124L) deoxythymidine-3',5'-bisphosphate Ca<sup>2+</sup> ternary complex. *Biochemistry*, in press.
14. Griffey RH, Redfield AG, Loomis RE and Dahlquist FW, Nuclear magnetic resonance observation and dynamics of specific amide protons in T4 lysozyme. *Biochemistry* 24: 817–822, 1985.
15. Griffey RH and Redfield AG, Proton-detected heteronuclear edited and correlated nuclear magnetic resonance and nuclear Overhauser effect in solution. *Q Rev Biophys* 19: 51–82, 1987.
16. Kainosho M and Tjuji T, Assignment of the three methionyl carbonyl carbon resonances in *Streptomyces* subtilisin inhibitor by a carbon-13 and nitrogen-15 double-labeling technique. A new strategy for structural studies of proteins in solution. *Biochemistry* 21: 6273–6279, 1982.
17. Westler WM, Kainosho N, Nagao H, Tomonaga N and Markley JL, Two-dimensional NMR strategies for carbon-carbon correlation and sequence-specific assignments in carbon-13 labeled proteins. *J Am Chem Soc* 110: 4093–4095, 1988.
18. Fesik SW and Zuiderweg EPR, Heteronuclear three-dimensional NMR spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra. *J Magn Reson* 78: 588–593, 1988.
19. Marion D, Kay LE, Spark SW, Torchia DA and Bax A, Three-dimensional heteronuclear NMR of <sup>15</sup>N-labeled proteins. *J Am Chem Soc* 111: 1515–1517, 1989.
20. Zuiderweg EPR and Fesik SW, Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* 28: 2387–2391, 1989.
21. Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM and Clore GM, Overcoming the overlap problem in the assignment of <sup>1</sup>H-NMR spectra of larger proteins by use of three-dimensional heteronuclear <sup>1</sup>H-<sup>15</sup>N Hartmann-Hahn multiple quantum coherence and nuclear Overhauser multiple quantum coherence spectroscopy: Application to interleukin 1 $\beta$ . *Biochemistry* 28: 6150–6156, 1989.
22. Cohen P and Klee CB (Eds.), *Molecular Aspects of Cellular Regulation*, Vol. 5. Elsevier, Amsterdam, 1988.
23. Dalgarno DC, Klevit RE, Levine BA, Williams RJP,

\* Ikura M, Klee CB, Krinks M, Torchia DA and Bax A, manuscript in preparation.

- Dobrowolski Z and Drabikowski W, <sup>1</sup>H-NMR studies of calmodulin: Resonance assignments by use of tryptic fragments. *Eur J Biochem* **138**: 281–289, 1984.
24. Evans JS, Levine BA, Williams RJP and Wormald HR, NMR studies of calmodulin in solution. Structure and dynamics in relation to function. In: *Molecular Aspects of Cellular Regulation* (Eds. Cohen P and Klee CB), Vol. 5, p. 57–82. Elsevier, Amsterdam, 1988.
  25. Kleivit RE, Dalgarno DC, Levine BA and Williams RJP, <sup>1</sup>H-NMR studies of calmodulin. The nature of the Ca<sup>2+</sup>-dependent conformational change. *Eur J Biochem* **139**: 109–114, 1984.
  26. Aulabaugh A, Niemczura WP, Blundell TL and Gibbons WA, A study of the interactions between residues in the C-terminal half of calmodulin by one and two-dimensional NMR methods and computer modelling. *Eur J Biochem* **143**: 409–418, 1984.
  27. Thulin E, Andersson A, Drakenberg T, Forsén S and Vogel HJ, Metal ion and drug binding to proteolytic fragments of calmodulin: Proteolytic, cadmium-113, and proton nuclear magnetic resonance studies. *Biochemistry* **23**: 1862–1870, 1984.
  28. Ikura M, Minowa O and Hikichi K, Hydrogen bonding in the carboxyl-terminal half fragment 78–148 of calmodulin as studied by two-dimensional nuclear magnetic resonance. *Biochemistry* **24**: 4264–4269, 1985.
  29. Ikura M, Minowa O, Yazawa M, Yagi K and Hikichi K, Sequence-specific assignments of downfield-shifted amide proton resonances of calmodulin. Use of two-dimensional NMR analysis of its tryptic fragments. *FEBS Lett* **219**: 17–21, 1987.
  30. Seeholzer S and Wand J, Structural characterization of the interactions between calmodulin and skeletal muscle myosin light chain kinase: Effect of peptide (576–594) G binding on the Ca<sup>2+</sup> binding domains. *Biochemistry* **28**: 4011–4020, 1989.
  31. Smith VL, Doyle KE, Maune JF, Munjaal RP and Beckingham K, Structure and sequence of the *Drosophila melanogaster* calmodulin gene. *J Mol Biol* **196**: 471–485, 1987.
  32. Shatzman AR and Rosenberg M, Efficient expression of heterologous genes in *Escherichia coli*. *Ann NY Acad Sci* **478**: 233–248, 1985.
  33. Hibler DW, Stolowich NJ, Reynolds MA, Gerlt JA, Wilde JA and Bolton PH, Site-directed mutants of staphylococcal nuclease. Detection and localization by <sup>1</sup>H NMR spectroscopy of conformational changes accompanying substitutions for glutamic acid-43. *Biochemistry* **26**: 6278–6286, 1987.
  34. Yazawa M, Sakuma M and Yagi K, Calmodulins from muscles of marine invertebrates, scallop and sea anemone. *J Biochem (Tokyo)* **87**: 1313–1320, 1980.
  35. Klee GB, Conformational transition accompanying the binding of Ca<sup>2+</sup> to the protein activator of 3',5'-cyclic adenosine monophosphate phosphodiesterase. *Biochemistry* **16**: 1017–1024, 1977.
  36. Kay LE, Marion D and Bax A, Practical aspects of 3D heteronuclear NMR of proteins. *J Magn Reson* **84**: 72–84, 1989.
  37. Bodenhausen G and Ruben DJ, Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem Phys Lett* **69**: 185–189, 1980.
  38. Bax A, Ikura M, Kay LE, Torchia DA and Tschudin R, Comparison of different modes of two-dimensional reverse correlation NMR for the study of proteins. *J Magn Reson* **86**: 304–318, 1990.
  39. Bendall MR, Pegg DT and Doddrell DM, Pulse sequences utilizing the correlated motion of coupled heteronuclei in the transverse plane of the doubly rotating frame. *J Magn Reson* **52**: 81–117, 1983.
  40. Bax A, Griffey RH and Hawkins BL, Correlation of proton and nitrogen-15 chemical shifts by multiple quantum NMR. *J Magn Reson* **55**: 301–315, 1983.
  41. Jeener J, Meier BH, Bachmann P and Ernst RR, Investigation of exchange processes by two-dimensional NMR spectroscopy. *J Chem Phys* **71**: 4546–4553, 1979.
  42. Braunschweiler L and Ernst RR, Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy. *J Magn Reson* **53**: 521–528, 1983.
  43. Bax A and Davis DG, MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J Magn Reson* **65**: 355–360, 1985.
  44. Ikura M, Kay LE, Tschudin R and Bax A, Three-dimensional NOESY-HMQC spectroscopy of a <sup>13</sup>C-labeled protein. *J Magn Reson* **86**: 204–209, 1990.
  45. Bax A, Homonuclear magnetization transfer experiments using isotropic and nonisotropic mixing schemes. *Isr J Chem* **28**: 309–317, 1988.
  46. Babu YS, Bugg CE and Cook WJ, Structure of calmodulin refined at 2.2 Å resolution. *J Mol Biol* **204**: 191–204, 1988.
  47. Kay LE, Ikura M and Bax A, Proton-proton correlation via carbon-carbon couplings: A three-dimensional NMR approach for the assignment of aliphatic resonances in proteins labeled with carbon-13. *J Am Chem Soc* **112**: 888–889, 1990.
  48. Fesik SW, Eaton HL, Olejniczak ET, Zuiderweg ERP, McIntosh LP and Dahlquist FW, 2D and 3D NMR spectroscopy employing <sup>13</sup>C-<sup>13</sup>C magnetization transfer by isotropic mixing. Spin system identification in large proteins. *J Am Chem Soc* **112**: 886–888, 1990.
  49. Muchmore DC, McIntosh LP, Russell CB, Anderson DE and Dahlquist FW, Expression and <sup>15</sup>N labeling of proteins for proton and nitrogen-15 NMR. *Methods Enzymol*, in press.
  50. Gottschalk G, *Bacterial Metabolism*, 2nd Edn. Springer, New York, 1986.
  51. Kainosho M, Nagao H and Tsuji T, Local structural features around the C-terminal segment of *Streptomyces subtilisin* inhibitor studied by carbonyl carbon nuclear magnetic resonances of three phenylalanyl residues. *Biochemistry* **26**: 1068–1075, 1987.
  52. Marion D and Bax A, P-COSY, a sensitive alternative for double-quantum filtered COSY. *J Magn Reson* **80**: 528–533, 1988.