Triple-resonance multi-dimensional high-resolution NMR Spectroscopy

Dr. Geerten W. Vuister
Department of Biophysical Chemistry
NSRIM Center, University of Nijmegen
The Netherlands
http://www.nmr.kun.nl
Preface

This reading material is the result of ‘cut and paste’ of many pieces of text and figures that I collected over the various years, while giving different lectures and teaching courses related to high-resolution NMR, in particular that of proteins.

Comments to the text are appreciated and can be mailed to me at vuister@nmr.kun.nl.

Geerten Vuister
August 2003

Disclaimer

This material was assembled for educational purposes only. It is not to be distributed without prior consent of the author.
1. Introduction

In general, the structure determination process can be thought to consist of a number of stages:
1. Biochemical work such as molecular characterization, cloning and expression, purification, and sample preparation. Knowledge about the primary structure is also obtained during this stage.
2. Assignment stage: recording and processing of a set of NMR spectra, followed by analysis and interpretation.
3. Extraction of experimental restraints that can be used for the structure calculation process, e.g. NOE derived distances from NOE build-up curves, dihedral angle constraints derived from J-couplings.
4. Structure generation by algorithms such as Simulated annealing.
5. Structure refinement by Molecular Dynamics, Relaxation Matrix calculations, or energy minimization.
6. Relaxation measurements to study mobility.

1.1 The assignment problem.

Assignment is the process identifying the exact resonance frequency of each $^1$H, ($^{13}$C, and $^{15}$N) nucleus. The assignment of resonances constitutes an essential step in the structure determination process by high resolution NMR spectroscopy that precedes the actual calculation of structures. The $^1$H (and $^{13}$C, $^{15}$N) spectra of proteins will consist of many resonances. An example $^1$H spectrum of a protein is shown in Fig. 1.1. Indicated are the different regions where the different proton-resonances are found.

![Figure 1.1. $^1$H NMR spectrum of a protein](image)
Although the theory of chemical shifts is well known, in practice it is quite complicated to accurately predict the chemical shifts in proteins. Partially this is the result of the inaccuracy of the protein structures and the internal mobility. On the other hand the range of proton chemical shifts is fairly limited (ca. 12 ppm) and the exact geometry is relatively important.

The range of $^{13}$C chemical shifts is much larger (ca. 200 ppm) and the effects of the exact geometry are less important. $^{13}$C chemical shifts are therefore easier to predict and can be used in a more straightforward fashion for interpretation of spectra.

With increasing size of the biomolecule two problems become more and more pronounced:
1. An increase in the total number of resonances that need to be identified.
2. Broadening of the resonances.
As a result of both effects, the overlap of resonances becomes increasingly problematic.

A strategy based upon homonuclear two-dimensional (2D) experiments (COSY, TOCSY, and NOESY) was developed in the 80's and is usually referred to as the ‘Wüthrich’ approach (Wüthrich, 1986, Nobel price in Chemistry, 2002), because of the pioneering work from his laboratory (ETH, Zurich, Switzerland).

A more recent approach employed uniformly $^{15}$N and $^{15}$N, $^{13}$C labeled proteins and was to a large extent developed by the group of Bax at the National Institutes of Health (Bethesda, MD, USA) and the group of Fesik at Abbott Laboratories (Abbott Park, IL, USA). The strategy employs so-called triple-resonance$^1$ experiments to transfer magnetization through the polypeptide chain employing the large one-bond homo- and heteronuclear J-couplings. The applicability of this approach is extended even more by employing $^2$H-labeling and TROSY-type experiments. For an overview, see Table 1.1.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Labeling</th>
<th>Approximate $\tau_c$ (ns)</th>
<th>Approximate size (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homonuclear</td>
<td>-</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Heteronuclear</td>
<td>$^{15}$N</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Triple-Resonance</td>
<td>$^{15}$N/$^{13}$C</td>
<td>10-12</td>
<td>18-20</td>
</tr>
<tr>
<td>Triple-Resonance</td>
<td>$^{15}$N/$^{13}$C/$^2$H</td>
<td>18</td>
<td>30</td>
</tr>
<tr>
<td>Triple-Resonance +</td>
<td>$^{15}$N/$^{13}$C/$^2$H</td>
<td>25-30</td>
<td>80</td>
</tr>
<tr>
<td>TROSY</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^1$ Heteronuclear NMR experiments involving 3 (or more) nuclei.
1.2 Multidimensional NMR

All the experiments are based upon the concept of multi-dimensional (i.e. two-dimensional, 2D; three-dimensional, 3D; four-dimensional, 4D) NMR spectroscopy. In multi-dimensional NMR spectroscopy the interactions between spins are employed for obtaining correlations between the resonance frequencies of those spins in separate frequency domains.

1.2.1 Two-Dimensional NMR

The schematic outline of 2D NMR spectroscopy is shown in Fig. 1.2.1. During the preparation sequence the initial state of the spins is established. This can be as simple as a relaxation delay (RD) to re-established the Boltzmann equilibrium and a single 90 pulse, creating xy magnetization along the x- or y-axis in the rotating frame. After this, a time period $t_1$ follows, which is called evolution. During the evolution period each spin will precess with its own frequency $\omega_A$ in the rotating frame.

With a pulse, or a series of pulses, (a fraction of the) magnetization is transferred between the spins, e.g from spin $A$ to spin $B$ (denote as $T_{AB}$) and vice versa. This latter period is called the mixing period. A detection period ($t_2$) completes the experiment. During the detection period, the magnetization of spin $B$, will precess with the lamor frequency of spin $B$, $\omega_B$. The measured signal, $S$, becomes dependent upon two time variables: the time during which the signal is measured ($t_2$) and the so-called indirect dimension variable $t_1$.

The evolution time $t_1$ is systematically incremented; e.g a large number of experiments are recorded, resulting in a two-dimensional data matrix $S(t_1,t_2)$. For example: if we recorded FIDs of 2048 points, and we incremented $t_1$ over 512 values, we would obtain a data matrix of 512 x 2048 points (which could look as Fig. 1.2.2A).

![Figure 1.2.2](image)
The 2D spectrum is obtained from a two-dimensional Fourier-transformation. First all the 512 FIDs are transformed in the \( t_2 \) domain and a hybrid time-/frequency domain data matrix is obtained, \( S(t_1,F_2) \) where \( F_2 \) denotes that the data along the second axis are represented in the frequency domain (see Fig. 1.2.2B). This data matrix contains 512 spectra, each consisting of 1024 real and 1024 imaginary points

\[
S(t_1,t_2) \quad \rightarrow \quad \text{FT}(t_2) \rightarrow \quad S(t_1,F_2)
\]

After discarding the imaginary points, the data matrix becomes 512x1024 points. The signal is now modulated in \( t_1 \). The final 2D frequency spectrum is obtained after FT along the \( t_1 \) axis (see Fig. 1.2.2C)

\[
S(t_1,F_2) \quad \rightarrow \quad \text{FT}(t_1) \rightarrow \quad S(F_1,F_2)
\]

If no zero-filling would be used, a \( S(F_1,F_2) \) matrix of 512x1024 would be obtained, that contains 256 real points and 256 imaginary points along the \( t_1 \) axis for every column. In practice however, the \( S(t_1,F_2) \) data matrix of 512x1024 points is supplemented by zeros along the \( t_1 \) axis to a size of 2048x1024 points, Fourier transformed, and the imaginaries discarded, resulting in a 1024x1024 data matrix, \( S(F_1,F_2) \).

The spectra are usually plotted as maps of orthogonal axes \( F_1,F_2 \) whereas the intensity is represented as contour (compare with cartographic map). We will observe a so-called ‘cross peak’ with coordinates \( (F_1,F_2)=(\omega_A,\omega_B) \) in the 2D frequency map and an intensity proportional to \( T_{AB} M_{Aeq} \), i.e. proportional to the equilibrium magnetization of spin A and the transfer efficiency from A to B. Many 2D experiments are symmetric: during the mixing period magnetization is also transferred from spin B to spin A with transfer efficiency \( T_{BA} = T_{AB} \), resulting in a cross peak at \( (F_1,F_2)=(\omega_B,\omega_A) \).

The fraction of magnetization that was not transferred during the mixing period, \((1-T_{AB})\), will precess during \( t_2 \) with the frequency of spin A, and hence will result in a so-called diagonal peak in the 2D frequency domain spectrum, \((F_1,F_2)=(\omega_A,\omega_A)\). Analogously, a diagonal peak for spin B exists at \((F_1,F_2)=(\omega_B,\omega_B)\).
Several mechanisms for the transfer of magnetization exist:

1. Cross relaxation: used in the 2D NOE or NOESY experiment.
2. Chemical exchange: used in the EXSY (Exchange Spectroscopy).
3. J-couplings: homonuclear used in COSY (Correlation Spectroscopy) or TOCSY (Total Correlation Spectroscopy, also called the HOHAHA experiment: Homonuclear Hartman Hahn). Heteronuclear couplings used in HSQC (Heteronuclear Single Quantum Spectroscopy) and HMQC (Heteronuclear Multiple Quantum Spectroscopy) and triple-resonance experiments.

It will be clear the 2D NMR has two very important advantages over 1D NMR:

1. The chances of overlap in a 2D spectrum are much smaller because peaks are dispersed in a 2D frequency map rather than along a 1D frequency axis.
2. A cross peak directly indicates an interaction between two spins.
1.2.2 Multi-Dimensional NMR

Several protein structures have been solved using 2D NMR techniques. However, with increasing size of the molecules the overlap of resonances becomes more and more a problem, strongly interfering with the structure-determination process. In three- and four-dimensional NMR spectroscopy (3D and 4D), three or four spins are correlated with each other, respectively. The addition of the third or fourth dimension increases the resolution and reduces the overlap. The example of Fig. 1.2.4 may illustrate this point: Two or more peaks which (almost) overlap in the 2D frequency spectrum, can be separated in the 3D frequency spectrum by correlation with an unique third frequency.

![Image of 2D and 3D spectra]

*Figure 1.2.4*

The extension from 2D to 3D or 4D NMR spectroscopy is straightforward (cf. Fig. 1.2.5): in 2D we had in addition to the detection period one additional evolution period, $t_1$, separated by a mixing period. Analogously, in 3D NMR spectroscopy we have two additional evolution periods, $t_1$ and $t_2$, separated by two mixing periods, in 4D we have three additional evolution periods, $t_1$, $t_2$, and $t_3$, and three mixing periods. Analogous to the 2D experiment, the evolution time variables $t_1$, $t_2$, (3D) or $t_1$, $t_2$, $t_3$ (4D) are incremented independently in a systematic fashion, resulting in $S(t_1, t_2, t_3)$ or $S(t_1, t_2, t_3, t_4)$ data matrices, respectively.

The 3D frequency domain spectrum is obtained after independent Fourier transformation along the three orthogonal axes. The spectrum can be represented in a cube with axes $F_1$, $F_2$, and $F_3$. The 4D frequency domain spectrum is obtained after Fourier transformation of the four independent

---

2 For historic reasons the evolution time domains are numbered from left to right as they appear in the pulse sequence.
time-domain axes. The spectrum has four orthogonal frequency axes (a hypercube) that cannot easily be visualized.

<table>
<thead>
<tr>
<th></th>
<th>Preparation</th>
<th>t₁</th>
<th>Mix</th>
<th>t₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D</td>
<td></td>
<td></td>
<td>Mix1</td>
<td>t₂</td>
</tr>
<tr>
<td>4D</td>
<td></td>
<td></td>
<td>Mix1</td>
<td>t₂</td>
</tr>
</tbody>
</table>

Figure 1.2.5

Cross sections perpendicular to one of the frequency axes (2D plots at a certain coordinate), or strips parallel to one of the axes (pseudo 2D plots at a certain coordinate) are the method of choice for analysis of the resulting data sets (cf. Fig 1.2.6). In the case of 3D data set it is usually still possible to plot all cross sections perpendicular to a given frequency axis (typically 64-256 planes). In the case of a 4D spectrum this becomes problematic, since the number of planes increases rapidly (typically 1024-4096 for 32x32 to 64x64 points along two of the axes) and a large number of the planes are empty; i.e. they contain no cross peaks and hence no information. Also, a pure manual analysis of this large a number of planes becomes quite tedious.

Figure 1.2.6. Cross sections (A) and strips (B).
2. Assignment

2.1 Assignment using the Wüthrich approach.

The spectra of so-called COSY or TOCSY experiments yield information about proton that are scalar coupled, i.e. are connected by three or less chemical bonds. The different amino-acids give rise to different patterns in these spectra, one (or sometimes two) pattern(s) for each amino-acid of the primary sequence. However, for larger proteins, when more than one pattern corresponding to residue of a certain type is present, e.g. several valines giving rise to several valine patterns, the problem becomes more complicated. Together with the increasing linewidth, the overlap makes it harder to identify such patterns. Also, when more than one valine is present in the protein, it is also not a priori clear which valine in the primary sequence corresponds to the valine pattern in the 2D COSY or TOCSY spectrum.

In order to make a so-called ‘sequential assignment’, i.e. connecting the COSY or TOCSY patterns to individual amino acids in the primary sequence, we have to connect the COSY/TOCSY pattern of the valine to the COSY/TOCSY pattern of its sequential neighbor. Unfortunately, when using only proton NMR, no $^1$H-$^1$H J-couplings of appreciable size exist over the peptide bond and hence there exists no direct way of connecting the two sequential patterns. As a result we are forced to employ the fact that protons of sequential residues typically are in close spatial proximity. Hence, an experiment designed to yield correlation between protons in close proximity (i.e. a NOESY experiment), yield the cross-peaks connecting the sequential patterns. A statistical analysis (and the fact that it has worked), showed that in particular the amide-proton ($H^N$) is the key candidate to connects the patterns, since it most often is in close proximity to either intra-residual or sequential protons.

After linking two patterns, we can try to link more patterns together and try to make tripeptide, tetrapeptide, etc fragments. Statistical analysis also shows that the uniqueness of such di-, tri-, and tetra-peptide fragments in proteins increases with length. In practice, for proteins up to ~200 amino-acids, a tetra-peptide fragment is usually sufficiently unique to allow for identification in the polypeptide chain.
2.2 Assignments using triple-resonance experiments.

As mentioned earlier, the line width of the resonance lines increases with increasing molecular weight. As a result cross-peaks in COSY and TOCSY spectra are strongly attenuated. This is a fundamental problem that cannot easily be overcome without resorting to a different methodology.

A solution to this problem would be to use magnetization-transfer pathways that employ couplings that are large relative to the line width. In normal proteins such pathways are not available. Modern biochemical techniques allow uniform substitution of the natural occurring isotope $^{14}$N by $^{15}$N (spin 1/2) and of $^{12}$C by $^{13}$C (also spin 1/2). By doing this, a protein is obtained in which all nuclei have spin 1/2. Moreover, these nuclei are now coupled by one-, two-, and three-bond J-couplings. The one-bond couplings have a useful characteristic: they are (in a relative sense) independent of local conformation, and as such allow for the design of sequences that transfer magnetization efficiently. An overview for the homo- and heteronuclear J-couplings in the protein backbone is given in Fig. 2.2.1. Fig. 2.2.2 gives as estimate for the $T_2$ of the different nuclei in the interferon homo-dimer, with a total molecular weight of 31 kDa. It can be seen that the reciprocal of these $T_2$ values in general is of the same order of magnitude or smaller than the couplings that would be employed to transfer the magnetization. For example, the amide proton $T_2 = 13$ ms, $1/(\pi T_2) = 24.5$ Hz, whereas the $^1$H-$^15$N one-bond J-coupling is 92 Hz. It is therefore still possible to transfer magnetization from the $^1$H to the $^{15}$N (rule of thumb).
2.2.1 Sequential assignment using the HNCA and HN(CO)CA experiments.

The $^1$H-$^{15}$N correlations are a good starting point for the design of pulse sequences that correlate the spins of the protein backbone for two reasons: first, the $^{15}$N T$_2$ is relatively long (together with the carbonyl (C') T$_2$), and the dispersion in the $^1$H-$^{15}$N correlation spectrum is in general also relatively high, i.e. least amount of overlap.

![Fig. 2.2.3](image)

The HNCA$^3$ experiment (Fig. 2.2.3) correlates the amide proton (H), the $^{15}$N spin (N) and the $^{13}$C$^\alpha$ of residues $i$ and $i-1$ (CA) with each other. The nuclei whose frequencies are measured are indicated by the gray circles, whereas the transfer of magnetization is indicated by the curved lines. Similar to the HNCA experiment is the HNCO experiment, correlating the amide proton and the $^{15}$N spins of residue $i$ with the $^{13}$C' spin of residue $i-1$. The HN(CO)CA experiment correlates the $^1$H and $^{15}$N spin of residue $i$ with the $^{13}$C$^\alpha$ spin of residue $i-1$ and is one step more complicated. Magnetization is first transferred the C' of residue $i-1$, as in the HNCO experiment, after which it is followed by a transfer to the $^{13}$C$^\alpha i-1$. The C$'^{i-1}$ frequency is not measured (as is indicated by the open circle), although the magnetization is transferred via this nucleus.

Both the HNCA and HN(CO)CA yield spectra that have a $^1$H frequency axis, a $^{15}$N-frequency axis and a $^{13}$C$^\alpha$ frequency axis. The HN(CO)CA spectrum shows cross-peaks with coordinates ($^1$H$^{N,i}$, $^{15}$N$^i$, $^{13}$C$^\alpha,i-1$). The HNCA shows these same peaks, but in addition also yields the intra-residual ($^1$H$^{N,i}$, $^{15}$N$^i$, $^{13}$C$^{\alpha,i}$) peaks. A possible assignment procedure directly follows from these correlations (cf. Fig. 2.2.4): The HNCA spectrum shows at one $^1$H,$^{15}$N position (of residue $i$)

---

$^3$ The nomenclature of the experiments was introduced by Ikura et al. (1990) and indicates the magnetization transfer pathway and spin that are frequency labeled. Identifiers within brackets are involved in the magnetization transfer pathway, but are not frequency labeled.
correlations to two $^{13}\text{C}_\alpha$ spins (those of residue $i$ and residue $i-1$). The HN(CO)CA at this same $^{1}\text{H},^{15}\text{N}$ position identifies one of these two as the sequential connection $i-1$ (usually the weaker of the two), since the HN(CO)CA only gives correlations to the $\text{C}_\alpha$ of residue $i-1$. Hence, the other peak corresponds to the intra-residual $\text{C}_\alpha$ (residue $i$). Let's assume that this intra-residual peak in the HNCA spectrum is located at 53 ppm $\text{C}_\alpha$ frequency. To find the next residue, we now have to search all the cross peaks in the HN(CO)CA spectrum and identify those peaks with a $\text{C}_\alpha$ coordinate of 53 ppm. When found, this then identifies the $^{1}\text{H},^{15}\text{N}$ position for residue $i+1$, and we can repeat the above procedure.

\[53\text{ ppm}\]

- HN(CO)CA
- HNCA

\[53\text{ ppm}\]

\[\text{i-1}\]
\[\text{i}\]
\[\text{i+1}\]

**Figure 2.2.4. A) Schematic outline of the sequential assignment procedure using HNCA and HN(CO)CA data. B) Strips taken from actual data.**

**2.2.2 More triple resonance experiments.**

There is one unfortunate drawback to the procedure outlined above: the dispersion of the $^{13}\text{C}_\alpha$ spectrum is rather poor, and the chance that there will be more than one spin resonating at one particular frequency is very great, particularly in the most crowded region of the spectrum around 55 ppm. Naturally, the assignment procedure then becomes ambiguous. Possible solutions\(^4\) are to exploit the H$\alpha$ spin by transferring the magnetization one additional step. The resulting experiments are called HN(CA)HA and HN(COCA)HA (Fig. 2.2.6). The HN(CA)HA will show correlations between the intra-residual H$\alpha$ and the $^{15}\text{N}-\text{N}^{\text{H}}$, as well as a sequential correlation to the preceding H$\alpha$. This is analogous to the HNCA experiment that also showed intra-residual- and sequential $\text{C}_\alpha$ correlations. The HN(COCA)HA only will show correlations to the preceding H$\alpha$. This is analogous to the HN(CO)CA experiment. Clearly, an assignment procedure now resembles that

\(^4\) Experiments involving the C$'$ spin, e.g. HCACO, also offer alternative pathways for correlating the different residues.
explained earlier for the HNCA and HN(CO)CA experiments, only that correlation now can be established using the H$^{\alpha}$ chemical shift.

Also the C$^{\beta}$ nucleus is an obvious candidate. The resulting experiments are called HNCACB (first published as CBCANH by Gresziek et al.) and the CBCA(CO)NH (which can also be done as HN(CO)CACB). A schematic outline is shown in Fig. 2.2.6. The HNCACB experiment will show correlations to both intra-residual C$^{\alpha}$ and C$^{\beta}$, as well as sequential correlations to the C$^{\alpha}$ and C$^{\beta}$ of the previous residue. The CBCA(CO)NH (or HN(CO)CACB) experiment, on the other hand, only shows the sequential C$^{\alpha}$ and C$^{\beta}$ correlation to residue i-1. The sequential connection can thus also be made using both the C$^{\alpha}$ and C$^{\beta}$ chemical shifts, as outlined in Fig. 2.2.7.
2.3 Chemical shifts.

Although the relation between local structure and chemical shifts has been evident for a long time, it has found only limited direct use in the process of structure determination. This was mainly caused by the fact that the exact nature of the relationship was not known, models were computational expensive, and that parametrization of models was difficult due to the lack of sufficient data. However, with the increase in computational speed and the rapidly increasing number of assigned proteins for which also a structure is known, theoretical calculations and database analyses have great promise in translating chemical shifts into structural constraints. For example, direct calculation of the $\text{C}^\alpha$ and $\text{C}^\beta$ chemical shifts on the basis of quantum mechanical calculations for valine and alanine residues showed a very good correlation with the experimental shifts.

The non-uniform distribution of proton chemical shifts was already noted in the early stage of structure determination by NMR, and aided the Wuthrich assignment procedure.
2.3.1. $^{13}$C Chemical shifts for identifying secondary structure.

$^{13}$C chemical shifts are even more characteristic. Rather than looking only at the absolute value of the chemical shifts (giving information about the spin type), it was quickly realized that the deviation of a resonance from its random coil value, the so-called secondary shift, gave valuable information as well. Bax and Spera (1991) (cf. Fig. 2.3.1) investigated the $C_\alpha$ and $C_\beta$ secondary shift for a number of proteins and found a clear correlation with secondary structure. The average secondary $C_\alpha$ chemical shifts was $3.09 \pm 1.00$ ppm, and $-1.48 \pm 1.23$ ppm for $\alpha$-helix and $\beta$-sheet, respectively. The average secondary $C_\beta$ chemical shift was $-0.38 \pm 0.85$ ppm and $2.16 \pm 1.91$ ppm for $\alpha$-helix and $\beta$-sheet, respectively. The large rmsd spread for $\beta$ sheet residues was partially attributed to the larger spread in the backbone angles $\phi$ and $\psi$.

The secondary $H_\alpha$ chemical shift is known to give similar information about secondary structure, resonating downfield for $\beta$-sheets and upfield for $\alpha$-helices.

Fig. 2.3.1. Distribution of $C_\alpha$ and $C_\beta$ secondary shift.

---

5 The chemical shift for a specific spin when measured in a short peptide which has full conformational freedom.
2.3.2. $^{13}$C Chemical shifts in sequential assignment.

The $^{13}$C protein chemical shifts are highly indicative for the specific residue types. Knowledge about the $^{13}$C chemical shifts of a certain number of sequentially connected residues can be used for sequential assignment. Based upon the statistical analysis of 600 residues of different proteins, Bax & Grzesiek (1993) showed how C$\alpha$ and C$\beta$ chemical shifts could be used for the identification of short peptide fragments.

Fig. 2.3.2. Plot of the random coil chemical shifts for $^{13}$C$\alpha$ and $^{13}$C$\beta$ (solid dots) of the 19 natural amino acids which contain $\beta$-carbons. Superimposed onto the serine random coil value is a probability distribution of the deviation from the random coil values of 600 amino acids. The contour lines are shown for probabilities of 0.95, 0.90 and 0.60 of finding the secondary shift within the three contours.
Assignment exercises EMBO 2003

For these exercises you may want to look up sections 2.2 and 2.3 of the handout “Triple-resonance multi-dimensional high-resolution NMR spectroscopy”.

Exercise 1:
The following two planes of the HN(CO)CACB spectrum (marked I and II) define two sequential residues in Calmodulin. Which residues are these?

Sequence and secondary structure of Calmodulin. Thick lines denote α-helices.

\[ ^{13}C^\alpha,^{13}C^\beta \text{ axis (ppm)} \]

\[ ^1H^N \text{ axis (ppm)} \]
Exercise 2.
Eight strips along the $^{13}\text{C}$ axis of the HN(CO)CACB spectrum (denote as A-H, S1) and HN(CO)CACB spectrum (denoted as A-H, S2) taken at the $^1\text{H}-^{15}\text{N}$ frequencies of eight sequential residues of Calmodulin. Sort the strips in appropriate order and determine where the eight residues belong in the primary sequence.
Answers to assignment exercises EMBO 2003

Exercise 1:
Plane I has peaks at 67 ppm and 68 ppm. One of these is the $\mathrm{C}^\alpha$, the other the $\mathrm{C}^\beta$, but we do not know if the 67 ppm peaks is the $\mathrm{C}^\alpha$ or the other way around. However, Fig. 2.3.2 of the handout “Triple-resonance multi-dimensional high-resolution NMR spectroscopy”, shows us that this is most likely a Thr residue. Plane II has peaks at 60 ppm and 73 ppm. The assignment $\mathrm{C}^\alpha = 73$ ppm, $\mathrm{C}^\beta = 60$ ppm, seems unlikely, whereas $\mathrm{C}^\alpha = 60$ ppm, $\mathrm{C}^\beta = 70$ ppm most likely is also a Thr residue, potentially a Ser. Thus we are looking for a TT or TS fragment. Analysis of the primary sequence shows the presence of only a TT fragment (in binding loop 1).

Exercise 2.
The order of the eight strips is C-G-D-F-A-E-B-H (see below). Strips D and A are glycines (no $\mathrm{C}^\beta$ and characteristic ~40 ppm shifts), but the Gly-X-Gly fragment is not unique; it appears in all four Calmodulins binding loops. To assign the eight residues in the primary sequence we have to analyze which residue they all potentially could be: We can draw up the following table (using Fig. 2.3.2):

<table>
<thead>
<tr>
<th>Strip</th>
<th>Potential residue</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>Ile [, Val, Pro]</td>
</tr>
<tr>
<td>G</td>
<td>Leu, Asn, Asp</td>
</tr>
<tr>
<td>D</td>
<td>Gly</td>
</tr>
<tr>
<td>F</td>
<td>Leu, Asp Asn</td>
</tr>
<tr>
<td>A</td>
<td>Gly</td>
</tr>
<tr>
<td>E</td>
<td>Glu, Gln, Arg, Trp, .....</td>
</tr>
<tr>
<td>B</td>
<td>Pro, Val, Ile</td>
</tr>
<tr>
<td>H</td>
<td>Asp, Asn [, Leu]</td>
</tr>
</tbody>
</table>

The fragment Asp-Gly-Asx-Gly also appears in all four binding loops. It becomes unique for binding loop 4 when extending at the N-terminus:

Ile- Asp-Gly-Asp-Gly

or at the C-terminus:

Asp-Gly-Asp-Gly-Gln-Val.