

NMR INVESTIGATIONS OF PROTEIN STRUCTURE

GERHARD WAGNER

Biophysics Research Division, University of Michigan, Ann Arbor, MI 48109, U.S.A.

(Received 15 December 1989)

CONTENTS

1. Introduction	102
2. ^1H Assignments	102
2.1. First sequential assignments	102
2.2. Homonuclear sequential assignment strategies	105
2.3. Assignments using heteronuclear editing techniques with ^{15}N or ^{13}C labeled proteins	107
2.4. Sequential assignments using heteronuclear correlation	109
2.4.1. Techniques based on $^{15}\text{N}/^{13}\text{C}$ double labeling	109
2.4.2. $^1\text{H}-^{13}\text{C}-^{15}\text{N}-^1\text{H}$ triple resonance techniques	110
2.4.3. Sequential assignments via $^{15}\text{N}-^1\text{H}^*$ long-range correlation	112
2.4.4. Homonuclear 3D spectroscopy	112
2.4.5. Heteronuclear 3D spectroscopy	112
2.5. Summary of assigned proteins	113
3. ^{13}C and ^{15}N Assignments	113
3.1. Assignment methods	114
3.2. Summary of assignments	114
3.3. Impact of ^{13}C and ^{15}N assignments	115
4. NOE: Virtues and Weaknesses for Structure Determinations	115
4.1. Identification of a large number of NOEs	115
4.2. Relations between NOE cross peaks, cross relaxation rates, distances, spectral density functions, distance ranges	115
4.3. Quantitative distances from the initial slopes	118
4.4. The practical problem of integrating all cross peaks in NOESY spectra	120
4.5. Back calculation of NOESY spectra with the complete relaxation matrix approach	121
4.6. NOEs and mobility: averaging effects vs. correlation time effects	122
5. Coupling Constants	123
5.1. Accurate measurements of heteronuclear coupling constants in ^{15}N labeled proteins	124
5.2. Accurate measurements of homonuclear coupling constants	125
5.3. Stereospecific assignments and determination of the angle χ^1	127
6. Hydrogen Bonds	128
7. Identification of Secondary Structures	129
7.1. Procedures for characterizing secondary structures	129
7.2. Impact of the identification of secondary structures on structure calculations	130
8. Tertiary Structures	131
8.1. Typical data sets for structure calculations	132
8.2. Algorithms for model building	133
8.3. Refinements	133
9. Criteria for Judging the Quality of Structures	133
9.1. Residual violations of constraints and average root mean square distances (rmsd)	133
9.2. Back calculation and <i>R</i> -factors	134
10. Limitations for Structural Studies of Proteins by NMR	134
10.1. Size of the protein	134
10.2. Bad spectra—large line width—small chemical shift dispersion	134
10.3. Paramagnetic centers	135
10.4. Protein degradation	135
10.5. Mobility	135
11. Conclusions	136
Acknowledgements	136
References	136

1. INTRODUCTION

NMR is the only technique that can determine protein structures in solution. The procedures and strategies that are common today have been developed in the course of the last ten years.⁽¹⁻⁸⁾ Earlier, detailed structural data on proteins had been available only from X-ray or neutron diffraction studies of single crystals. This development of protein NMR was due to three achievements; (i) the development of strategies for sequential assignments,⁽³⁻⁸⁾ (ii) the advent of 2D NMR techniques,⁽⁹⁻¹⁶⁾ and (iii) the development of software to build protein models from distance data.⁽¹⁷⁻²⁵⁾ The challenging task to analyze complex protein spectra has also stimulated the recent progress in NMR technology, and techniques developed with proteins are now used in other applications of NMR concerned with structural aspects at the molecular level. Today more than 80 proteins have been nearly completely assigned,⁽²⁶⁾ and for more than 30 proteins solution structures have been derived by NMR.⁽²⁶⁾ For most proteins where an NMR structure and a crystal structure have been determined, the results of both techniques are similar. A few of the NMR studies were sufficiently detailed to show minor differences between solution and crystal structures, mainly residues on the protein surfaces. Most of these differences were found to be due to protein-protein contacts in the crystals.⁽²⁷⁻²⁹⁾ There are a few cases showing major differences between solution and crystal structures.⁽³⁰⁻³²⁾ The success of the structural studies has overshadowed the power of NMR to study dynamic effects although NMR is probably the most powerful technique to study this very aspect of proteins.

This review first describes techniques and strategies for resonance assignments. Secondly, techniques to obtain structural parameters are reviewed, together with methods to derive structural models. Protein mobility will be discussed only in connection with structural aspects.

2. ¹H ASSIGNMENTS

Complete ¹H assignments are the basis for all the recent achievements of protein NMR in the areas of structure analysis and studies of protein mobility. There are two levels of so called 'complete' proton assignments. (i) The assignments can be complete in a sense that at least one or two resonances are identified per residue. Such assignments are valuable for studies of protein dynamics where a dense set of probes is required in order to monitor protein mobility with spatial resolution. Such assignments are also sufficient to characterize regular secondary structures. (ii) Truly complete assignments require identification of all non-labile protons and all labile protons that do not exchange rapidly with the water. This is a requirement for a good structure determination. If there are some resonances not assigned, every NOE involving signals in a spectral region containing unassigned resonances is ambiguous.

Prior to 1978, individual resonance assignments in proteins were very difficult to obtain. They were based mainly on comparison of homologous proteins, selective chemical modification, selective isotope labeling, and use of shift or broadening reagents together with reference to crystal structures. These techniques became obsolete with the development of the strategies of sequential assignments and the arrival of 2D NMR techniques, in particular 2D NOE (NOESY) and 2D correlated spectroscopy (COSY).

2.1. First Sequential Assignments

When the first sequential assignment strategies were developed they were based on two analytical techniques namely, spin decoupling and the measurement of nuclear Overhauser effects. It was realized that NOE values provide selective distance information only when short irradiation times are used.⁽³³⁾ Applying such truncated NOE, we realized from inspection of a model of a protein that, in a β -sheet, the amide proton of residue i is always close to the α -proton of residue $i - 1$ (see Fig. 1) and shows an intense NOE when the H^N of residue i is irradiated (Fig. 2).⁽³⁾ NOE measurements together with results from spin decoupling, allowed longer segments of the basic pancreatic trypsin inhibitor to be assigned. In these studies, spin decoupling provided intraresidue connectivities, while NOE values yielded the sequential connectivities.⁽³⁾

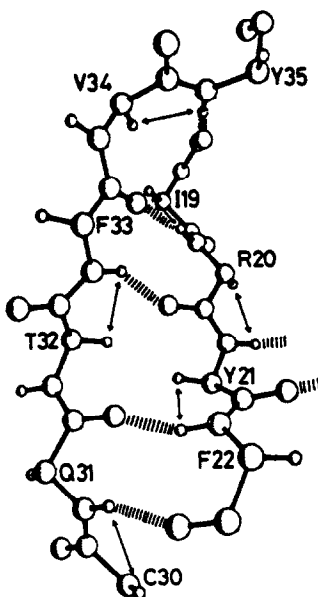


FIG. 1. The antiparallel β -sheet of the basic pancreatic trypsin inhibitor (BPTI). The short $d_{\alpha N}(i, i+1)$ distances for which first sequential assignments were obtained are indicated with arrows (adapted from Ref. 3).

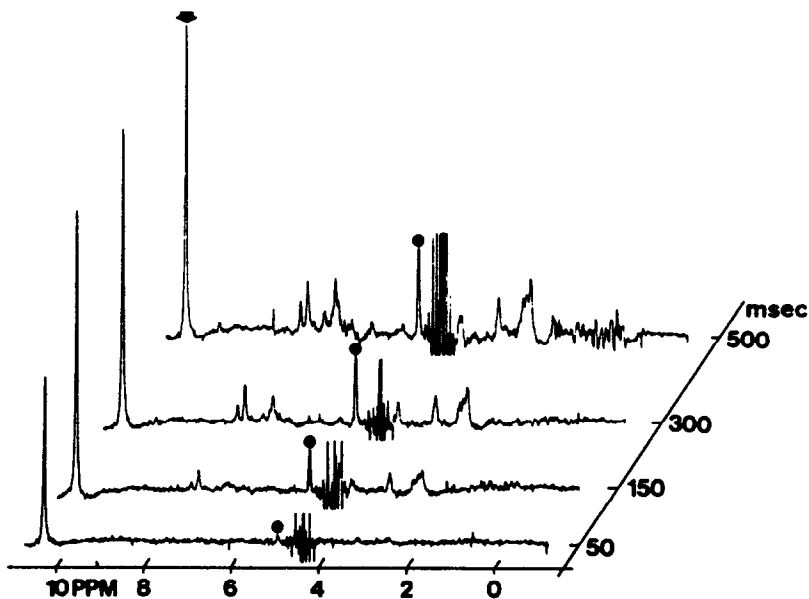


FIG. 2. 1D truncated driven NOE spectra of BPTI. The irradiation times are indicated on the right. The build up of the strong NOE at the H^{α} of Phe 22 after irradiating the H^N of Tyr 23 is indicated with a dot.⁽³⁾

In particular, these 1D experiments on BPTI resulted in the assignment of five consecutive residues of the central strand of a triple stranded antiparallel β -sheet (Arg 20 to Asn 24). All of these residues have slowly exchanging amide protons. A number of other residues were assigned in the same study, making reference to the crystal structure. The NOE values across the β -sheet which are typical for

antiparallel β -sheets were then first characterized and used for extending the assignments. The extent of these assignments is shown in Fig. 3A. The sequential assignments could not be extended to those residues without slowly exchanging amide protons since the standard spectrometers available at that time were not equipped for simultaneous presaturation of the solvent line and presaturation of a

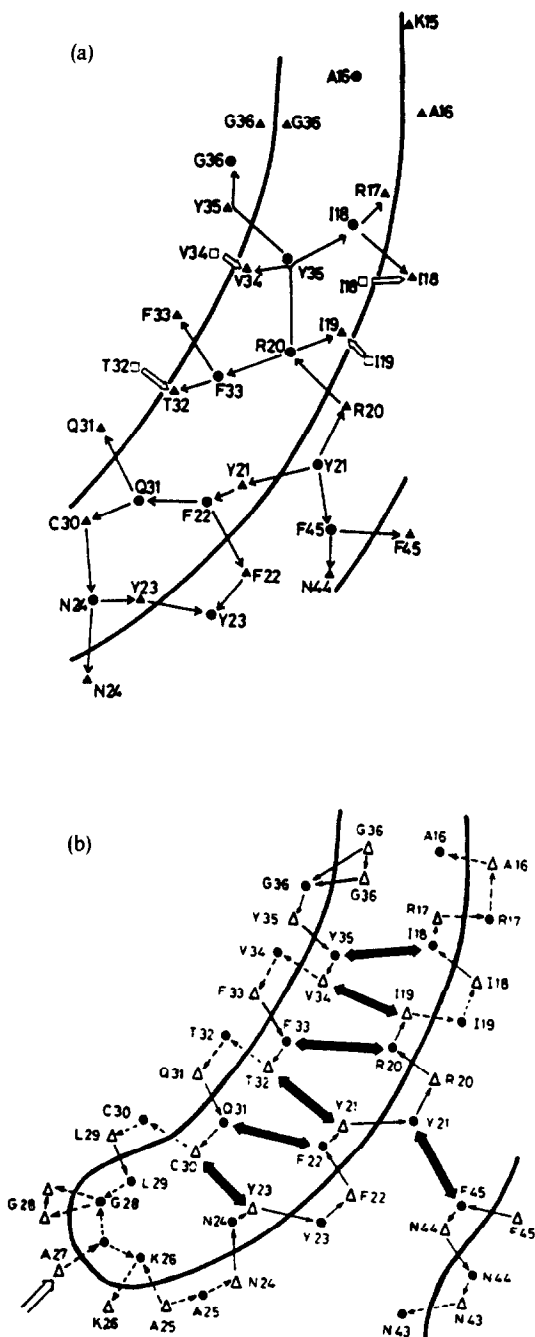


FIG. 3. (a) Assignment pathways in the β -sheet of BPTI by 1D experiments in D_2O ⁽¹³⁾ and (b) by 2D experiments in D_2O and H_2O ⁽¹⁴⁾. Dots and triangles represent H^N and H^2 , respectively.

protein resonance to cause a NOE. This problem was solved with the introduction of 2D experiments where the selective inversion of protein resonances is achieved by the t_1 evolution period, and the decoupler channel is free for solvent irradiation.⁽³⁴⁾ This together with the resolution in the second dimension obtained with the 2D experiments allowed an almost complete assignment of the not yet identified ^1H resonances of BPTI.^(4,7) Figure 4A shows the 2D version⁽⁴⁾ of the assignments achieved previously by 1D techniques in D_2O ⁽³⁾ (see Figs 2 and 3A). This is a combined plot of a COSY (lower triangle) providing intraresidue $\text{H}_i^{\text{N}}-\text{H}_i^{\alpha}$ connectivities, and a NOESY (upper triangle) providing sequential $\text{H}_{i-1}^{\alpha}-\text{H}_i^{\text{N}}$ connectivities. The assignment pathway was called a β -snail.⁽⁴⁾ The possibility of now running these experiments also in H_2O allowed one to extend these assignments as shown in Figs 4B and 3B. In the course of this work, it had also been realized that sequential assignments can be obtained in helices and turns via NOE values between the amide protons of sequentially adjacent residues (see for example the β -hairpin in Fig. 3B), and it was argued that sequential assignments can be obtained, generally, independent of the secondary structure. This hypothesis was later verified in a theoretical study.⁽⁶⁾

2.2. Homonuclear Sequential Assignment Strategies

The assignment technique that evolved in the work with BPTI was soon developed to a systematic assignment strategy. This strategy is illustrated in Fig. 5. The procedure consists of the following steps.

(i) *Spin system assignments in D_2O .* The spin systems of the non-labile protons are usually elucidated in D_2O solution.⁽³³⁾ In order to simplify the spectra, one tries to exchange all labile protons with deuterons. Heating up the sample in D_2O solution or variation of the pH may be required to achieve this. Methods used in this part of the spin system assignment include SECSY,⁽³⁶⁾ COSY,^(12,13,16) DQF-COSY,⁽³⁷⁻³⁹⁾ Multiple Quantum Filtered COSY,^(40,41) Double Quantum Spectroscopy,⁽⁴²⁻⁴⁶⁾ RELAY⁽⁴⁷⁻⁵¹⁾ spectra, Double RELAY spectra⁽⁵⁰⁾ and TOCSY⁽⁵²⁾ (HOHAHA⁽⁵³⁾). Often only TOCSY and COSY may be required to identify the type of spin system. However, for establishing unambiguous assignments of all resonances of longer side chains RELAY experiments may be necessary. Sometimes, only one cross peak between an α -proton and a β -proton is found for residues that have a β -methylene group, and it is not clear whether the two β -protons are degenerate or the second cross peak is just too weak to be observed. This question can be clarified by recording Double Quantum Spectra. These spectra contain remote cross peaks at ($\omega_2 = \delta(\text{H}^{\alpha})$, $\omega_1 = \delta(\text{H}^{\beta 1}) + \delta(\text{H}^{\beta 2})$). If the resonance position of one β -proton is known the other can be determined from this remote peak.^(45,46) In this step of the analysis, spin systems can readily be classified to certain types of residues. Some spin systems are unique (Gly, Ala, Thr, Ile, Val, Leu), others have to be assigned to groups of residues, such as AMX spin systems (Tyr, Phe, His, Trp, Cys, Asp, Asn, Ser), 5-spin systems (Glu, Gln, Met), and long spin systems (Arg, Lys, Pro). Arginine and proline spin systems can often be assigned completely and distinguished from the longer lysines. Arginines can be identified later since they have a connectivity to a peptide H^{N} , in contrast to prolines, and the δ -protons are coupled to a labile H^{ϵ} . The spin systems of the aromatic side chains are independent and have to be connected to the β -methylene groups via NOESY experiments. This is also necessary for connecting the NH_2 groups of Asn and Gln side chains.⁽⁵⁻⁸⁾

(ii) *Intraresidue $\text{H}^{\text{N}}-\text{H}^{\alpha}$ connectivities.* Connectivities between the spin systems of the non-labile protons and the labile amide protons are obtained from COSY-type experiments in H_2O , combined with proper water-elimination techniques. Usually, DQF-COSY, TOCSY and RELAY experiments are applied at this step. Particularly helpful are DQ-spectra. On the one hand, they do not suffer from the bleach-out effect due to water presaturation,⁽⁵⁴⁾ on the other hand they help to identify glycine residues since these are the only residues with remote cross peaks at ($\omega_2 = \delta(\text{H}^{\text{N}})$, $\omega_1 = \delta(\text{H}^{\alpha 1}) + \delta(\text{H}^{\alpha 2})$).^(45,46) To overcome the bleach-out effect of water presaturation, techniques such as Pre-TOCSY COSY⁽⁵⁵⁾ and SCUBA⁽⁵⁶⁾ have been proposed and have been found useful. The step of establishing connectivities to the amide protons is a crucial one since the amide protons are most important for establishing sequential assignments and for identification of long-range NOE.

(iii) *Sequential assignments.* Sequential connectivities between the spin systems of the amino acid residues are identified in NOESY spectra in H_2O . This relies on identifying NOE values from the H^{N} of

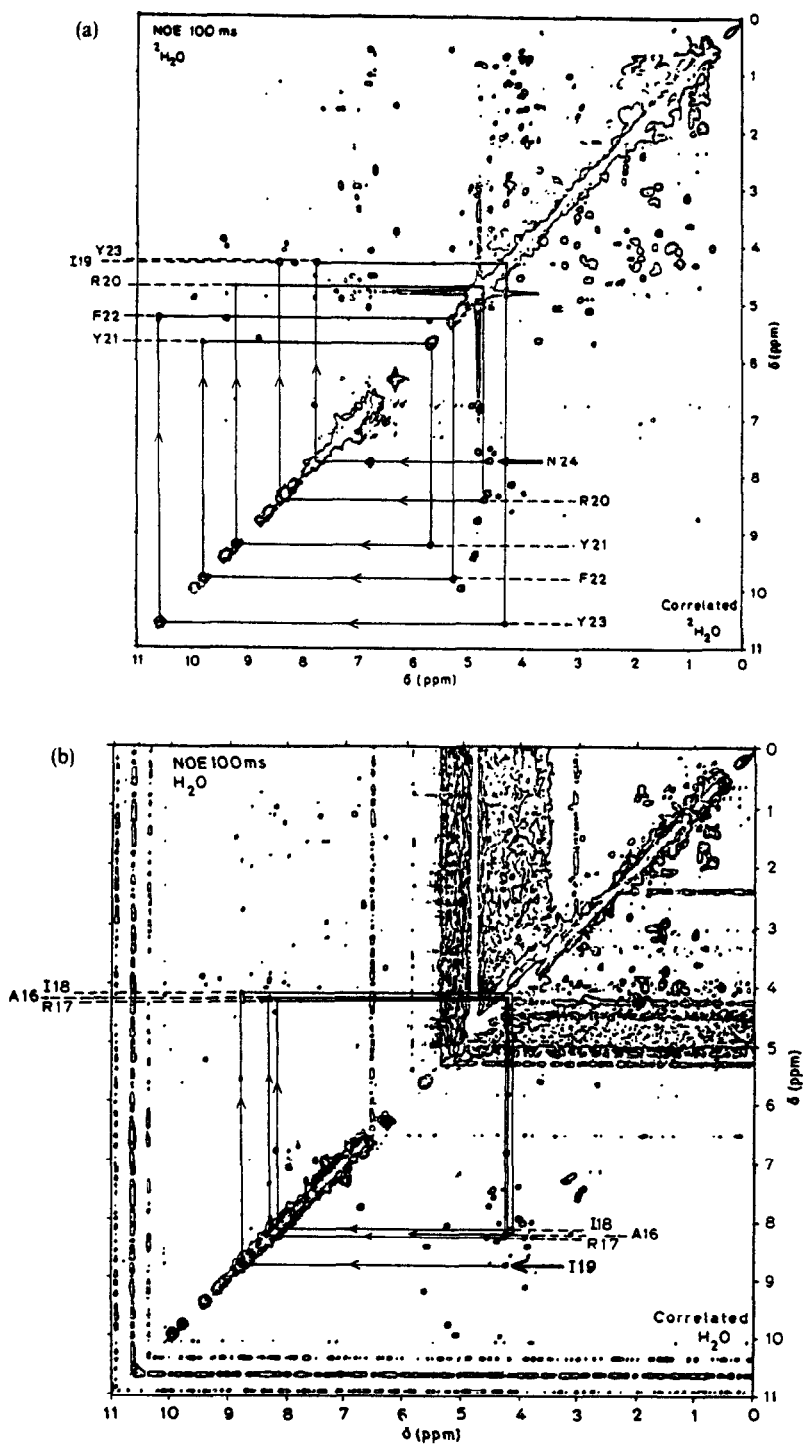


FIG. 4. First sequential assignments via $d_{2N}(i, i+1)$ connectivities in BPTI in a combined COSY-NOESY diagram (a) in D_2O , (b) in H_2O .⁽⁴⁴⁾

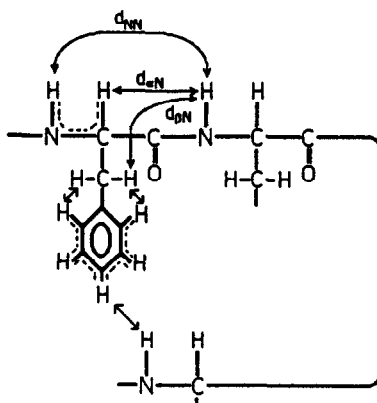


FIG. 5. Schematic of a polypeptide structure with the connectivities $d_{\alpha N}$, d_{NN} , $d_{\beta N}$.

residue i to the H^N , H^α and H^β of residue $i-1$. To increase the probability of finding the sequentially neighboring residue one tries to identify more than one NOE from the H^N to different protons of the sequentially preceding residue. This can often be achieved as is shown in Fig. 6 for the protein eglin c.⁽¹⁶³⁾ This is very important since a H^N usually shows many NOEs, not only to sequentially neighboring residues. However, if more than one NOE is observed from the H^N of one residue to another spin system the probability is very high that this is the sequentially preceding residue. The statistics of identifying the sequential neighbor have been worked out in detail.⁽⁶⁾ Obviously, for prolines and glycines, slightly different rules apply.^(6, 7) For prolines, the δ -protons play the role of the H^N , and for glycines, the second H^α has a similar function as the β -protons. Pre-TOCSY NOESY⁽⁵⁵⁾ and SCUBA⁽⁵⁶⁾ techniques are also useful at this step to overcome bleach-out effects due to preirradiation of the water.

Steps (i) and (ii) can be combined depending on the complexity of the spectrum. On the other hand, if the spectra are very crowded, steps (ii) and (iii) may require recording spectra of proteins with partially deuterated peptide groups. In particular, spectra may be recorded in D_2O of protein preparations where only the slowly exchanging amide protons are present, or alternatively, spectra may be recorded in H_2O where the slowly exchanging amide protons were exchanged with deuterons prior to the experiment so that the spectrum of the rapidly exchanging amide protons is seen. This strategy was applied for several proteins,^(4, 7, 57) most systematically in a recent study of lysozyme.⁽⁵⁸⁾

Recently, a variant of the above described assignment strategy has been propagated. It is called main-chain-directed (MCD) assignment strategy.⁽⁵⁹⁾ It could also be classified as a secondary structure directed assignment technique whereas the original strategy is primary-structure directed. This technique uses essentially the same principles as the technique described above. However, there are some minor differences. (i) Less emphasis is given to identifying all spin systems of the amino acid side chains at the very beginning. (ii) Sequential NOE and NOE typical for regular secondary structures are analyzed simultaneously to obtain sequential assignments. At this step there is a search for closed assignment loops. Obviously such loops were analyzed in previous work (see for example, the loop formed by the residues Phe 22, Tyr 23, Asn 24, Cys 30 and Gln 31 of BPTI in Fig. 3A). However, they were not made targets in the assignment procedure. The MCD assignment strategy may be advantageous in cases where there is already information about the secondary structures available. It remains to be seen whether this strategy can be applied if no information is available about the secondary structure, or if the protein has no regular secondary structure. It has been pointed out that very complete assignments are desirable if a structure determination is intended.⁽⁶⁰⁾

2.3. Assignments Using Heteronuclear Editing Techniques with ^{15}N or ^{13}C Labeled Proteins

The homonuclear assignment strategies described above have been applied successfully for proteins with up to 129 residues. The largest proteins assigned so far by these techniques are lysozyme⁽⁵⁵⁾ and

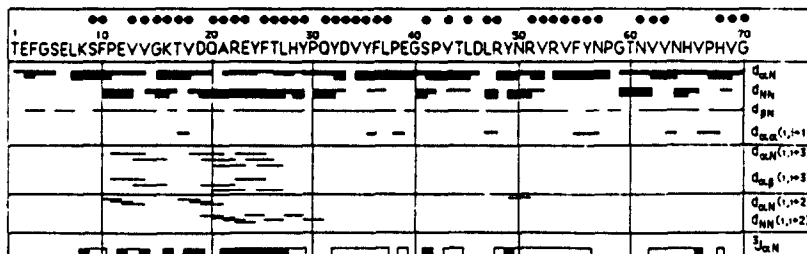


FIG. 6. Summary sequential assignments for the elastase inhibitor eglin c. Medium-range NOEs are also indicated. The dots (top) indicate slow hydrogen exchange. Filled and open bars (bottom) indicate small and large coupling constants $^3J(\text{H}^\alpha\text{-H}^N)$.⁽¹⁶³⁾

ribonuclease.⁽⁶¹⁾ For larger proteins and for proteins with heavy overlap of resonances these techniques may not lead to complete assignments. A major improvement has come with the availability of isotope labeling. Since for many proteins efficient expression systems are available proteins can be produced uniformly labeled with ^{15}N or ^{13}C , or more importantly, proteins can be produced with only a single or a few amino acid types labeled (semiselective labeling). Several excellent descriptions are available in the literature about the labeling procedures.⁽⁶²⁻⁶⁷⁾

So far, semiselective ^{15}N labeling of single amino acid types combined with heteronuclear editing have had the highest impact. The NMR strategies were developed in several laboratories, starting with the work of Redfield and coworkers.⁽⁶⁸⁾ It was realized that 1D and 2D spectra can be edited for signals of protons directly coupled to ^{15}N . Originally, this was achieved in 1D proton spectra recorded with and without ^{15}N decoupling.⁽⁶⁹⁾ The two spectra were then subtracted and the difference spectrum contains only signals from protons coupled to ^{15}N .⁽⁶⁹⁾ This method has the disadvantage that the signals left are a difference between a doublet and a singlet. Experiments were then developed where the proton coherence was allowed to evolve to antiphase coherence with respect to ^{15}N .⁽⁷⁰⁾ A 180° (^{15}N) pulse applied in every other scan, together with inversion of the receiver phase selects for coherence of ^{15}N bound protons. In a 2D experiment, this editing can be applied before or after the mixing period. In the former case the technique has been called an ω_1 -half filter.⁽⁷⁰⁾ It selects for cross peaks where the protons along ω_1 are coupled to ^{15}N . Figure 7A shows such a pulse sequence for an ω_1 -half filtered COSY. The spectral region selected is shown in Fig. 7a'. If the filter is applied after the mixing period we have an ω_2 -half filter as shown in Fig. 7b and 7b'. For selection of heteronuclear antiphase coherence the delay τ has to be tuned to $1/4J_{\text{HN}}$. This method has the disadvantage that the heteronuclear doublets are antiphase and ^{15}N cannot be decoupled during t_2 . This problem was eliminated by tuning τ to $1/2J_{\text{HN}}$.⁽⁷¹⁾ In this case the filter is a bilinear rotation decoupling (BIRD)⁽⁷²⁻⁷⁴⁾ sequence. The multiplet components are in-phase and decoupling can be applied.

Semiselective ^{15}N labeling simplifies the assignment process at the level of spin system assignments, and it reduces overlap problems in the H^N region of 2D spectra. If, for example, the protein is labeled with ^{15}N at all leucine residues, ^{15}N edited COSY/TOCSY type experiments will usually lead to complete assignments of the leucine spin systems, and ^{15}N edited NOESY experiments will provide assignments of the sequentially preceding residues. With a sufficiently large number of residues labeled, complete assignments can usually be obtained. However, if a complete set of NOEs needs to be acquired for a structure determination, semiselective ^{15}N labeling does not solve some ambiguities in the assignments of cross peaks to aromatic and aliphatic protons. Usually, the aliphatic and aromatic spectrum is more crowded than the amide proton spectrum. Therefore, semiselective carbon labeling would be much more valuable (but more expensive).

Heteronuclear editing techniques have been used extensively for the assignments of T4 lysozyme,^(63, 65, 66, 69) staphylococcal nuclease,⁽⁶⁷⁾ P22 c2 repressor,^(70, 75) and thioredoxin.⁽⁶²⁾ Staphylococcal nuclease with 141 residues and T4 lysozyme with 164 residues are so far the largest proteins for which nearly complete assignments are available. For 132 of 141 residues in staphylococ-

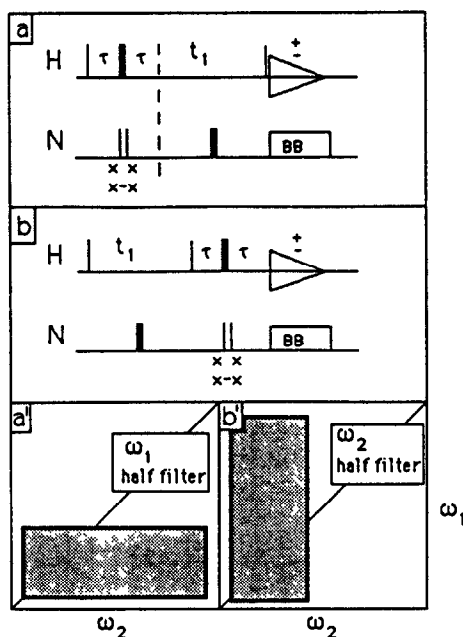


FIG. 7. Pulse sequences for ^{15}N edited COSY experiments. (a) ω_1 half filter, (b) ω_1 half filter. (a') and (b') regions selected with the sequences (a) and (b), respectively. Thin and thick vertical bars indicate 90° and 180° pulses, respectively.⁽⁷⁰⁾

cal nuclease, and for 161 of the 164 residues in T4 lysozyme at least one of the backbone resonances were assigned.⁽⁶⁶⁾

2.4. Sequential Assignments Using Heteronuclear Correlation

2.4.1. Techniques Based on $^{15}\text{N}/^{13}\text{C}$ Double Labeling. Availability of uniformly ^{15}N and/or ^{13}C labeled proteins has led to the development of assignment techniques relying on heteronuclear correlation. Markley and coworkers were the first to produce proteins uniformly labeled with both ^{13}C and ^{15}N , with the purpose of achieving sequential assignments.^(76,77) Their strategy was first to achieve carbon-carbon connectivities using carbon-carbon double-quantum experiments. For this step, they found isotope enrichments of 25% most advantageous avoiding coupling to a third carbon nucleus. The carbon-carbon connectivity patterns are unique for all amino acid types (except that Gln/Glu and Asn/Asp cannot be distinguished). Thus, unique spin system identifications can be achieved at the very beginning of the analysis. Next sequential connectivities are established via carbon-nitrogen double-quantum spectroscopy. After they have achieved complete sequence-specific assignments of the nitrogen and carbon resonances they establish ^1H assignments by heteronuclear proton-carbon and proton-nitrogen correlation experiments. In contrast to the homonuclear proton assignment strategies, this technique is independent of the protein conformation, relying only on one-bond couplings. On the other hand, each sequential connectivity depends on a single correlation. Degeneracy of carbon and nitrogen resonances may be more hampering than degeneracy of proton resonances in the case of the NOE-based proton assignments where observation of more than one of the $d_{\alpha\text{N}}$, d_{NN} and $d_{\beta\text{N}}$ connectivities can resolve ambiguities. Therefore, these techniques may be combined with homonuclear NOESY experiments. This method requires double labeling with ^{13}C and ^{15}N . The preparation is expensive and has been done only for very few proteins at the present time. However, the technique may become important for large proteins where homonuclear COSY/TOCSY experiments are difficult.

2.4.2. ^1H - ^{13}C - ^{15}N - ^1H Triple Resonance Techniques. An alternative assignment technique relying on heteronuclear correlation has been proposed recently which also relies on isotope enrichment of ^{15}N or ^{13}C .⁽⁷⁸⁾ The basic experiment is a ^1H - ^{13}C - ^{15}N - ^1H triple resonance experiment. The basic pulse sequence is outlined in Fig. 8a and the coherence transfer pathways are shown in Fig. 8c. First polarization is transferred from the α -protons to the α -carbons via a refocused INEPT⁽⁷⁹⁻⁸²⁾ transfer. The carbon coherence is frequency labeled during t_1 while protons and nitrogens are decoupled. Next an INEPT-type sequence transfers coherence from the α -carbons of residue i and $i-1$ to the nitrogen of residue i . The intraresidue one-bond coupling is 11 Hz, and the interresidue two-bond coupling is 7 Hz so that the delay b can be tuned sufficiently well for both coupling constants. Finally, the nitrogen coherence is transferred to the amide proton via an inverse INEPT sequence. A schematic drawing of the resulting spectrum is shown in Fig. 8d. It contains intraresidue and sequential cross peaks between the α -carbons and the amide protons. This experiment evolved from attempts to obtain accurate homonuclear coupling $\text{H}^{\text{N}}\text{-H}^{\alpha}$ constants.⁽⁸³⁾ In this case, the last 90° (^1H) is replaced with a TANGO⁽⁸⁴⁾ pulse (Fig. 8b). In this application this is a 90° pulse selective for the amide proton coupled to ^{15}N . In addition, there is no proton decoupling during t_1 . This results in different multiplet

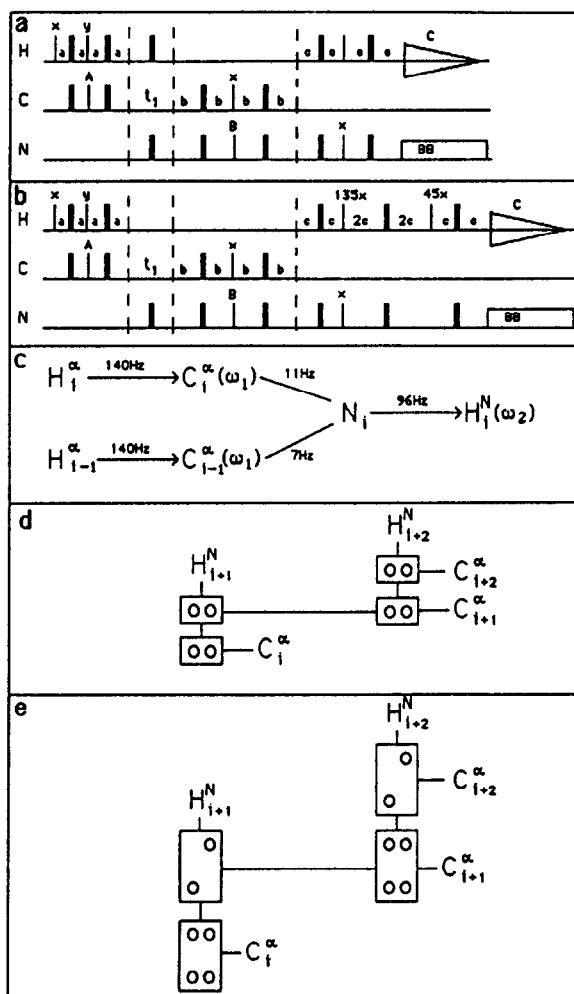


FIG. 8a-e

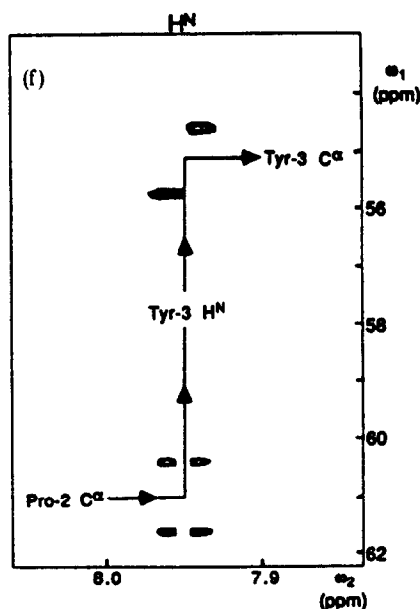


FIG. 8. Triple resonance experiments for sequential assignments in ^{15}N or ^{13}C enriched polypeptides.⁽⁷⁸⁾ (a) With proton decoupling during t_1 , (b) without proton decoupling during t_1 . The last proton 90° pulse of (a) is replaced with a TANGO⁽⁸⁰⁾ pulse. (c) Intraresidue and sequential coherence transfer pathways obtained with the sequences (a) and (b). (d) and (e) show schematics of spectra obtained with the pulse sequences (a) and (b), respectively. Note that the sequence (b) yields different multiplet patterns for intraresidue and sequential cross peaks. (f) Experimental cross peaks recorded with the pulse sequence of Fig. 8b for the tripeptide Ac-Asn-Pro- (^{15}N) Tyr-NHMe.⁽⁷⁸⁾

structures being observed for intraresidue and sequential cross peaks as outlined in Fig. 8e. This feature may be attractive for automated sequential assignments. So far this technique has only been applied to a ^{15}N enriched peptide,⁽⁸³⁾ at natural abundance ^{13}C . The pulse sequence suffers from the long delays b which need to be tuned for the small $\text{C}_i^\alpha\text{-N}_{i+1}$ coupling constant of 7 Hz whereas T_2 s of α -carbons in proteins of ca. 6 kDa are around 50 ms, and ^{15}N T_2 s are ca. 100 ms.

Obviously, the pulse sequence of Fig. 8a can be modified in several ways.⁽⁸³⁾ (i) The carbon and nitrogen channels can be interchanged. In this case we have the coherence transfer pathway:

$$\text{H}_{i+1}^{\text{N}} \rightarrow \text{N}_{i+1}(\omega_1) \rightarrow \text{C}_{i+1}^\alpha, \text{C}_i^\alpha \rightarrow \text{H}_{i+1}^\alpha, \text{H}_i^\alpha(\omega_2).$$

This experiment yields cross peaks between the nitrogen of the residue $i+1$ (along ω_1) and the α -protons of residues i and $i+1$ (along ω_2). (ii) The t_1 period can be placed immediately after the first proton pulse of Fig. 8a. In this case we have the coherence transfer pathway:

$$\text{H}_{i+1}^\alpha, \text{H}_i^\alpha(\omega_1) \rightarrow \text{C}_{i+1}^\alpha, \text{C}_i^\alpha \rightarrow \text{N}_i \rightarrow \text{H}_i^{\text{N}}(\omega_2).$$

This experiment yields cross peaks between the H^α of the residues i and $i+1$ (along ω_1) and the H^{N} of residue i (along ω_2). These experiments have been applied only to a ^{15}N labeled peptide. Apparently, these experiments can be expanded to 3D techniques in order to resolve overlap problems. Obvious choices would be coherence transfer pathways:

$$\text{H}_{i+1}^\alpha, \text{H}_i^\alpha(\omega_1) \rightarrow \text{C}_{i+1}^\alpha, \text{C}_i^\alpha(\omega_2) \rightarrow \text{N}_i \rightarrow \text{H}_i^{\text{N}}(\omega_3)$$

or

$$\text{H}_{i+1}^\alpha, \text{H}_i^\alpha(\omega_1) \rightarrow \text{C}_{i+1}^\alpha, \text{C}_i^\alpha \rightarrow \text{N}_i(\omega_2) \rightarrow \text{H}_i^{\text{N}}(\omega_3).$$

2.4.3. *Sequential Assignments via ^{15}N - ^1H Long-range Correlation.* The three-bond coupling between the peptide nitrogen and the H^α of the preceding residue depends on the dihedral angle ψ . It varies between +1 and -6.5 Hz.⁽⁸⁵⁾ In many cases sequential cross peaks can be observed in 2D heteronuclear long-range correlation experiments when ^{15}N enriched protein is available. The most successful experiment used for this purpose so far is a proton detected heteronuclear multiple bond correlation experiment (Fig. 9) with a heteronuclear multiple quantum evolution period containing a purge pulse to eliminate the one-bond correlations.^(86,87) This technique has been used for part of the sequential assignments of the Ner protein.⁽⁸⁸⁾ The technique requires only ^{15}N but no ^{13}C enrichment. The pulse sequence and a schematic spectrum are shown in Fig. 9.

2.4.4. *Homonuclear 3D Spectroscopy.* The development of modern computers has made the extension to 3D NMR spectroscopy possible. 3D spectroscopy can be used to resolve ambiguities present in 2D spectra or to resolve a 2D spectrum in a third dimension. Homonuclear 3D spectra were the first to be developed.⁽⁸⁹⁻⁹⁴⁾ The first experiment combined a COSY with a J -resolved experiment so that the multiplet splitting was perpendicular to the COSY plane.⁽⁸⁹⁾ Then NOESY and TOCSY spectra were combined so that NOE and J connectivities are in perpendicular planes.^(93,94) The experiments are, in some way, 3D extensions of 2D relayed NOESY experiments⁽⁹⁵⁾ where a tuned delay is replaced by an incremented delay. The experiments increase the number of cross peaks compared to homonuclear experiments manifold. Due to the increased complexity, homonuclear 3D experiments seem to be less important than heteronuclear experiments which can simplify the spectra by dispersing the cross peaks in a third dimension, without increasing the total number of cross peaks. To reduce the amount of data 3D experiments with selective pulses were used.^(96,97) The combination of 2D spectra with a third dimension will lead to a large number of different 3D experiments.⁽⁸⁹⁻⁹⁸⁾ All 2D experiments containing tuned delays can obviously be extended to 3D experiments. An initial overview of possible experiments and a classification has been given by Griesinger *et al.*⁽⁹⁶⁾ An attractive aspect of homonuclear 3D experiments not yet generally recognized is that projections of homonuclear 3D experiments in a 2D plane may be used to obtain homonuclear ω_2 -decoupled homonuclear 2D correlated spectra, such as COSY, NOESY, TOCSY, and 2D projections of homonuclear 4D experiments can lead to truly homonuclear ω_2 and ω_1 decoupled homonuclear correlated 2D spectra.⁽⁹⁹⁾

2.4.5. *Heteronuclear 3D Spectroscopy.* The NOE-based homonuclear assignment technique often suffers from overlap problems in the homonuclear 2D experiments when larger proteins are studied. If uniformly ^{15}N labeled proteins are available the overlap problem can be resolved by inserting a ^{15}N evolution period before or after the ^1H evolution period.⁽¹⁰⁰⁻¹⁰³⁾ Two possible sequences for ^{15}N

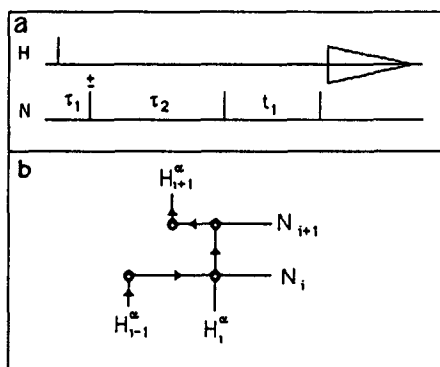


FIG. 9. (a) Pulse sequence for heteronuclear multiple bond correlation via multiple quantum coherence (HMBC).^(82,83) (b) Schematic spectrum with intrasidue and sequential connectivities between amide nitrogens and α -protons.⁽⁸⁴⁾

dispersed NOESY spectra are shown in Figs 10a and 10b. They use a heteronuclear multiple quantum evolution period. The sequence of Fig. 9a has the ^{15}N evolution period preceding the ^1H evolution period. Analogous to the ω_1 half-filtered 2D spectra (Fig. 7), the H^{N} subspectrum is selected along ω_2 . The selected spectral region is shaded in the front ω_2 - ω_3 plane of the schematic 3D spectrum of Fig. 10a'. The pulse sequence of Fig. 10b has the ^{15}N evolution period after the ^1H evolution period. The H^{N} subspectrum is selected along ω_3 (Fig. 10b'). The sequence of Fig. 9b has the advantage that the t_1 -noise of the water does not interfere with the selected spectrum whereas the sequence of Fig. 10a has the better resolution in the spectrum of protons that show NOEs to the amide protons. The shaded plane inclined by 45° relative to the ω_2 - ω_1 plane contains the 2D nitrogen-proton correlated spectrum.

2.5. Summary of Assigned Proteins

The first protein that was completely assigned is the basic pancreatic trypsin inhibitor (BPTI),⁽⁷⁾ An overview of assigned proteins has been published recently.⁽²⁶⁾ A database of the assigned proteins including the assignments is being developed by Markley and coworkers.⁽¹⁰⁴⁾

3. ^{13}C AND ^{15}N ASSIGNMENTS

Assignments of ^{13}C and ^{15}N have generally been considered of lower priority. Obviously, ^1H assignments are more important since they can yield NOE distance constraints for structural studies. Nowadays, assignments of heteronuclei are becoming more abundant since they are by-products of advanced assignment techniques, in particular from 3D resolved heteronuclear experiments. They can help in obtaining additional conformational parameters via heteronuclear coupling constants and are also useful for studies of protein mobility.

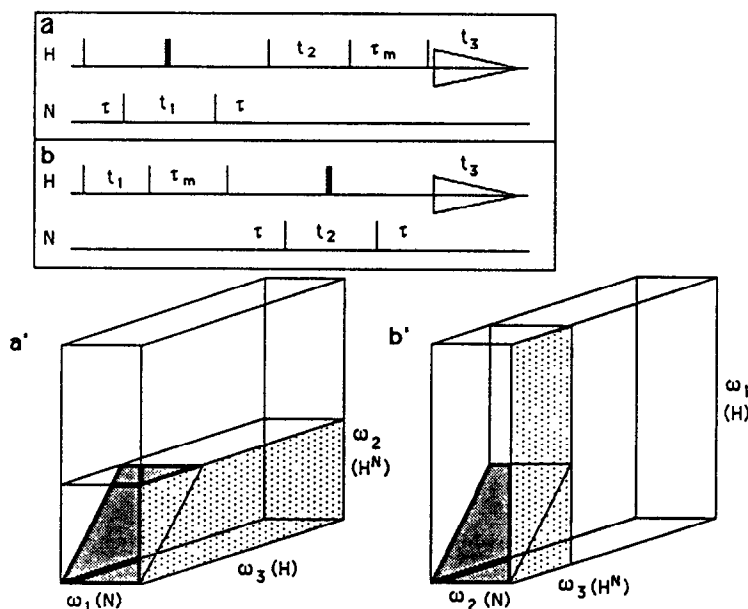


FIG. 10. Heteronuclear 3D ^{15}N dispersed NOESY experiments. (a) and (b) are analogous to the ω_1 and ω_2 half filters⁽⁷⁰⁾ of Fig. 7. Correspondingly, different regions of the 3D space are selected. Note that (a, a') avoid interference with the t_1 -noise of the water while (b, b') have the better resolution along ω_3 .

3.1. Assignment Methods

Assignments of protonated ^{13}C and ^{15}N nuclei can be made in a straightforward manner by using proton detected heteronuclear correlation experiments,^(105–108) after the ^1H assignments have been obtained. The most commonly used experiments are outlined in Fig. 11. The double INEPT experiment of Fig. 11a⁽¹⁰⁵⁾ is very sensitive and the spectra are easy to phase. Figure 11b shows the popular heteronuclear multiple quantum experiment^(109,110) with refocusing of the proton chemical shift. This is the most sensitive experiment. A similar experiment without a refocusing period was described earlier by Müller.⁽¹¹¹⁾ The heteronuclear multiple quantum experiment is usually more difficult to phase along ω_1 , and ^1H – ^1H couplings are active during the t_1 period which results in a splitting of the carbon or nitrogen signals along ω_1 , and reduces the resolution along ω_1 . Refocused double INEPT^(105,110,112) (Fig. 11c) and double DEPT⁽¹⁰⁵⁾ (Fig. 11d) are less sensitive but easy to phase and show no ^1H – ^1H couplings along ω_1 . In addition, editing of the spectra for CH , CH_2 and CH_3 is possible with double DEPT⁽¹⁰⁵⁾ and refocused double INEPT.⁽¹⁰⁵⁾ The refocused INEPT (Fig. 11c) was actually the first proton detected heteronuclear experiment applied to a protein.^(112,113) Heteronuclear 3D experiments analogous to those of Fig. 9 provide nitrogen or carbon assignments directly.

Assignments of non-protonated ^{13}C and ^{15}N nuclei have so far only been obtained when the proteins were isotope enriched. Heteronuclear multiple-bond correlation or carbon–carbon and carbon–nitrogen double-quantum experiments have been used for these assignments.^(76,77) Only for the protein Tendamistat have extensive assignments of non-protonated carbons been obtained without isotope enrichment.⁽¹⁰⁸⁾

3.2. Summary of Assignments

So far extensive ^{13}C assignments have been obtained only for a few proteins. Except for the assignments obtained by the Markley group^(26,76,77) they were all obtained at natural abundance of ^{13}C . Extensive ^{15}N assignments have been obtained at natural abundance only for BPTI.⁽¹⁰⁷⁾ For a

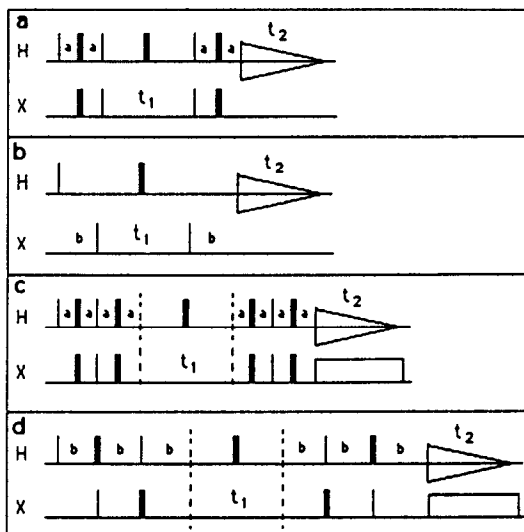


FIG. 11. Pulse sequences for heteronuclear one-bond correlation experiments: (a) double INEPT transfer with antiphase evolution period, (b) heteronuclear multiple quantum evolution period, (c) double refocused INEPT with inphase evolution period of the X-nucleus, (d) double DEPT transfer with inphase evolution of the X-nucleus. The lengths of the delays are: $a = 1/(4J)$, $b = 1/(2J)$.

number of proteins with ^{15}N enrichment the ^{15}N assignments were obtained together with the proton assignments.^(62–67, 76, 77)

3.3. Impact of ^{13}C and ^{15}N Assignments

^{13}C and ^{15}N assignments will probably soon become an integral part of advanced assignment procedures. They will resolve ambiguities of proton assignments, in particular in combination with heteronuclear 3D experiments. ^{13}C and ^{15}N assignments will aid measurements of homonuclear and heteronuclear vicinal coupling constants and open the resource of a new class of conformational constraints for structure calculations. With carbon and nitrogen assignments we get access to extensive measurements of relaxation parameters, such as T_1 , T_2 and the cross relaxation rates σ_{HC} and σ_{HN} .

4. NOE: VIRTUES AND WEAKNESSES FOR STRUCTURE DETERMINATIONS

Resonance assignments provide the essential first stage for structure calculations. The quality of the protein models derived depends on the extent and the quality of conformational parameters derived from the NMR data. So far, NOE analysis has dominated conformational studies of proteins. For the best NMR structures of proteins more than 10 NOEs per residue were identified on average. These NOEs are, however, not usually uniformly distributed over the protein. Interior residues may be characterized by up to 40 interresidue NOEs, whereas surface residues may show as few as two or three interresidue NOEs and no long-range NOEs. Therefore, while NOEs may provide a good means of determining the overall folds of proteins in solution, and be good for characterizing the protein interior and the regular secondary structures of β -sheets and helices, the NOE analysis is not usually sufficient for characterizing conformations on the protein surfaces. In this case it is necessary to measure additional conformational parameters, such as coupling constants in order to obtain a more complete characterization of the protein conformation.

4.1. Identification of a Large Number of NOEs

The strength of NMR for the determination of protein conformations is based on the ability to identify many distance constraints which define the overall topology of the polypeptide. It is thus more important to increase the number of assigned NOEs than to measure fewer NOEs very precisely. Therefore, it is also advisable to record NOESY spectra with long mixing times, although the NOEs identified in these spectra can be associated only with loose distance limits. It is desirable that every single cross peak in the NOESY spectra be assigned, even if this can be achieved only during the process of model building when ambiguities in peak assignments can sometimes be resolved by backcalculating NOESY spectra from model structures. Ambiguities in peak assignments due to resonance overlap can be resolved at the primary step of resonance assignments by variation of the temperature (chemical shift of amide proton resonances often have different temperature coefficients) by partial exchange of labile protons, or by recording more advanced NOESY-type experiments. These are, for example, a NOESY followed or preceded by relayed coherence transfer,⁽⁹⁵⁾ or the extension of this experiment in a 3D NOESY–COSY or COSY–NOESY.⁽⁹⁶⁾ NOESY spectra preceded by a double-quantum evolution period have also been used.⁽¹¹⁴⁾ Most promising appear to be heteronuclear 3D experiments where the NOESY experiment is combined with a ^{15}N evolution period. Obviously, this requires ^{15}N enrichment of the protein.

4.2. Relations Between NOE Cross Peaks, Cross Relaxation Rates, Distances, Spectral Density Functions and Distance Ranges

The functional dependence of NOE cross peak intensities, a_{kl} , on cross relaxation rates, R_{kl} , proton–proton distances, r_{kl} , and spectral density functions $J_{kl}(\omega)$ of rotational reorientation of the inter-proton vectors, \mathbf{r}_{kl} , is clearly understood.^(2, 115–119) However, accurate measurements of distances from NOEs is not trivial, and it is not clear how accurately internuclear distances can be

measured from NOEs. This is a crucial aspect because observation of a certain NOE is routinely associated with a certain upper and lower distance limit compatible with this NOE. It is this very distance range (uncertainty of the measurement) that cannot be measured but only guessed. On the other hand, violations of these guessed upper and lower distance limits (residual distance violations) are used to judge the quality of the protein models obtained from computer algorithms.

In a NOESY experiment, peak intensities, $a_{kl}(\tau_m)$, can be measured as a function of the mixing time.⁽¹¹⁹⁾ These intensities can be calculated according to:

$$\dot{\mathbf{A}} = -\mathbf{R}\mathbf{A} \quad (4.1)$$

where \mathbf{A} is the matrix of the elements $a_{kl}(\tau_m)$. The solution of eqn. (4.1) is:

$$\mathbf{A}(\tau_m) = \exp(-\mathbf{R}\tau_m)\mathbf{A}(0) \quad (4.2)$$

or

$$a_{kl}(\tau_m) = [\exp(-\mathbf{R}\tau_m)]_{kl} a_{ll}(0) \quad (4.3)$$

where $a_{kl}(\tau_m)$ represents the intensity of the peak correlating resonance l (at ω_1) and k (at ω_2). It is assumed that the system is in equilibrium prior to each scan so that $\mathbf{A}(0)$ is a diagonal matrix, and the diagonal elements are Boltzmann populations. \mathbf{R} is the relaxation matrix. Equation (4.2) can be written in a power series:

$$a_{kl}(\tau_m) \approx (\delta_{kl} - R_{kl}\tau_m + 1/2 \sum_j R_{kj}R_{jl}\tau_m^2 + \dots) a_{ll}(0). \quad (4.4)$$

R_{kk} is the relaxation rate and is always positive:

$$R_{kk} = 2(n_k - 1)(w_1^{kk} + w_2^{kk}) + \sum_{k \neq j} (w_0^{kj} + 2w_1^{kj} + w_2^{kj}). \quad (4.5)$$

R_{kj} ($j \neq k$) is the cross relaxation rate. It is positive for small molecules ($w_2^{kj} > w_0^{kj}$) and negative for large molecules ($w_2^{kj} < w_0^{kj}$).

$$R_{kj} = n_k(w_2^{kj} - w_0^{kj}); \quad (4.6)$$

n_i is the number of equivalent spins in a group and w_0^{kj} , w_1^{kj} and w_2^{kj} are the zero-, single- and double-quantum transition probabilities for the two-spin systems formed by the spins k and j . The terms $R_{kj}R_{jl}\tau_m^2$ ($j \neq k, l$) in the sum of eqn. (4.4) represent spin diffusion and they are all positive. The terms $R_{kk}R_{kl}\tau_m^2$ and $R_{kl}R_{ll}\tau_m^2$ are all negative for macromolecules. They represent the curvature of the NOE build-up rates due to relaxation of the diagonal peaks ($R_{kl}R_{ll}\tau_m^2$) and the cross peaks ($R_{kk}R_{kl}\tau_m^2$).

Assumption of a rigid molecule. The main purpose of recording NOESY spectra is to measure intramolecular distances. This implicitly makes the assumption that there are defined distances, that the distances are fixed and time independent, and that there is no internal or segmental motion. This assumption has essentially been made in all the NMR structure determinations carried out to date, in the absence of better working models. Under this assumption, the cross relaxation rate, R_{kj} , is proportional to r_{kj}^{-6} , and can be obtained from the initial slope of the build-up curve of the cross peak intensity, $a_{kl}(\tau_m)$. Measuring the initial slopes is not trivial, and different approaches will be discussed below.

Transition probabilities per unit time of eqns (4.5) and (4.6) are obtained with time-dependent first-order perturbation theory:^(115, 116)

$$\overline{w_{mn}^{kj}} = \frac{4\pi^2}{\hbar^2} \left| \int_0^t \overline{\langle n | \mathcal{H}'(t') | m \rangle e^{-i\omega_{mn} t'}} dt' \right|^2 \quad (4.7a)$$

$$\overline{w_{mn}^{kj}} = \frac{4\pi^2}{\hbar^2} \int_0^t \int_0^{t'} \overline{\langle n | \mathcal{H}'(t') | m \rangle \langle n | \mathcal{H}'^*(t'') | m \rangle e^{-i\omega_{mn}(t'-t'')}} dt' dt'' \quad (4.7b)$$

m and n refer to the levels in Fig. 12. The bar indicates the ensemble average. It is obvious from eqns (4.7a) and (4.7b) that the transition probability is only non-zero if the dipolar Hamiltonian, $\mathcal{H}'(t')$, or rather the autocorrelation function of the dipolar Hamiltonian oscillates with the frequency ω_{mn} . This means the vector connecting the nuclei k and j has to oscillate in length, or orientation relative to the magnetic field with the zero-, single- or double-quantum frequency (energy differences in Fig. 12) to induce the corresponding transitions. The probability for a rotational motion of the internuclear vector with a frequency ω is usually expressed in the form of a spectral density function, $J_{kj}(\omega)$, which can be shown to be proportional to the Fourier transform of the autocorrelation function of the dipole-dipole Hamiltonian.^(115, 116) For fixed distances r_{kj} we obtain:^(115, 116)

$$\omega_0^{kj} = \frac{2\gamma^4 \hbar^2}{20r_{kj}^6} J_{kj}(\omega_k - \omega_j); \quad (4.8)$$

$$\omega_1^{kj} = \frac{3\gamma^4 \hbar^2}{20r_{kj}^6} J_{kj}(\omega_j); \quad \omega_1^k = \frac{3\gamma^4 \hbar^2}{20r_{kj}^6} J_{kj}(\omega_k);$$

$$\omega_2^{kj} = \frac{12\gamma^4 \hbar^2}{20r_{kj}^6} J_{kj}(\omega_k + \omega_j).$$

This assumes that all relaxation and cross relaxation is caused by the modulation of the dipole-dipole interaction between the two nuclei with rotational motions of the internuclear vector, relative to the external field. For isotropic rotational diffusion of a spherical molecule, the spectral density function has the following form:⁽¹¹⁵⁾

$$J_{kj}(\omega) = \frac{\tau_c^{kj}}{1 + (\omega\tau_c^{kj})^2}. \quad (4.9)$$

If there is internal motion the spectral density function can be expected to be different. This means that the relative contributions of protein motions to the cross relaxation rates at the frequencies $\omega_k - \omega_j$, ω_k , ω_j and $\omega_k + \omega_j$ are different from what would be expected from eqn. (4.9).

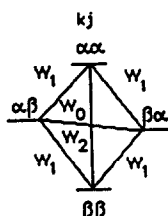


FIG. 12. Energy diagram for a two-spin system.

4.3. Quantitative Distances from Initial Slopes

It is important to obtain good initial guesses for distances from the initial slopes of NOE build-up curves, even if, after the first model building, distance measurements may be improved by back calculation of the NOESY spectra with the complete relaxation matrix approach.^(118,120,121) There are some major obstacles for measuring initial slopes of NOE build-up curves in NOESY spectra:

(i) *The linear range of the τ_m build-up curves is short.* This is because at longer mixing times the slope of the NOE build-up curves decreases and becomes negative due to the relaxation of the diagonal peaks which are the magnetization source for the cross peaks. This effect could be eliminated by dividing the cross peak intensities by the intensities of the diagonal peaks at every mixing time. This approach, in a slightly modified form, has been proposed by Macura *et al.*⁽¹²²⁾ They propose to divide the cross peak intensities by the mean of the two diagonal peaks. The effect of this can be visualized starting from eqn. (4.4):

$$\begin{aligned} \frac{a_{kl}(\tau_m)}{0.5(a_{kk}(\tau_m) + a_{ll}(\tau_m))} &\approx \frac{(-R_{kl}\tau_m + 1/2\sum_j R_{kj}R_{jl}\tau_m^2 + \dots)a_{ll}(0)}{(1 - 0.5(R_{kk}\tau_m - R_{ll}\tau_m) + 1/4\sum_j R_{kj}R_{jk}\tau_m^2 + 1/4\sum_j R_{lj}R_{jl}\tau_m^2 + \dots)a_{ll}(0)} \\ &\approx -R_{kl}\tau_m + 1/2\sum_j R_{kj}R_{jl}\tau_m^2 - 0.5R_{kl}R_{kk}\tau_m^2 - 0.5R_{kl}R_{ll}\tau_m^2 + \dots \quad (4.10) \\ &\approx -R_{kl}\tau_m + 1/2\sum_{j \neq k,l} R_{kj}R_{jl}\tau_m^2 + \dots \end{aligned}$$

This shows that there are no relaxation contributions but only cross relaxation and spin diffusion contributions in the terms up to second order in τ_m . A more elaborate proof is given by Macura *et al.*⁽¹²²⁾ The cross peak intensities normalized with the diagonal peak intensities show a significantly longer linear growth regime. Finally they flatten and approach a constant steady state value.⁽¹²²⁾ Of course, the signal-to-noise ratio decreases with longer τ_m values. In principle this approach should be better for obtaining initial slopes but, in practical applications with proteins it is impossible to integrate diagonal peaks because of considerable resonance overlap. A partial solution to this problem can be found in a ^{15}N relayed NOESY. This is a NOESY preceded by a ^{15}N evolution period; it requires ^{15}N labeled protein and can only be applied to analyze NOEs with amide protons. Figure 13a shows a pulse sequence we have used⁽¹²³⁾ for such experiments. It is similar to an experiment developed independently by Gronenborn *et al.*⁽¹²⁴⁾ (Fig. 13b). The former experiment has some advantages over the latter. In the former we transfer polarization from the amide protons to ^{15}N via an INEPT step and let the ^{15}N coherence evolve during t_1 in antiphase to the amide proton. After a reverse INEPT the inphase coherence of the amide proton is converted into the z-axis so that NOEs can develop during τ_m . This experiment contains direct $\text{H}^{\text{N}}\text{-}^{15}\text{N}$ cross peaks and cross peaks due to the NOEs between the amide proton and other protons, at the ω_1 frequencies of the ^{15}N nuclei (see the schematics of Fig. 13c). In this experiment, the intensities corresponding to the diagonal peaks in a regular NOESY can be obtained from the direct cross peaks which are now no longer overlapped as found on the diagonal of a regular NOESY, and the quantity $a_{ll}(0)$ of eqn. (4.4) can readily be measured. However, the mean of $a_{ll}(\tau_m)$ and $a_{kk}(\tau_m)$ cannot be obtained since, while proton l may be coupled to a ^{15}N nucleus, proton k will usually not be coupled to ^{15}N .

(ii) *Calibration of cross peak intensities between different mixing times.* Because it is important to conduct NOESY experiments with different mixing times, the performance of the spectrometer may change with time, for example, because of environmental effects. Thus, a comparison of peak intensities recorded with different mixing times is problematical, unless the peak intensities can be calibrated within each experiment. This can be achieved by dividing the cross peak intensities by the intensities of the diagonal peaks, or better with the intensities of the direct peaks in the ^{15}N relayed NOESY of Fig. 13.⁽¹²³⁾

(iii) *Calibration of cross relaxation rates with distances.* If the cross relaxation rates R_{kj} could be measured precisely, measurements of the distances r_{kj} would also require knowledge of the spectral density functions $J_{kj}(\omega)$ for the rotational motions of the vector r_{kj} . This seems to be an unsolvable problem. At present, it can only be circumvented by the crude assumption that the spectral density

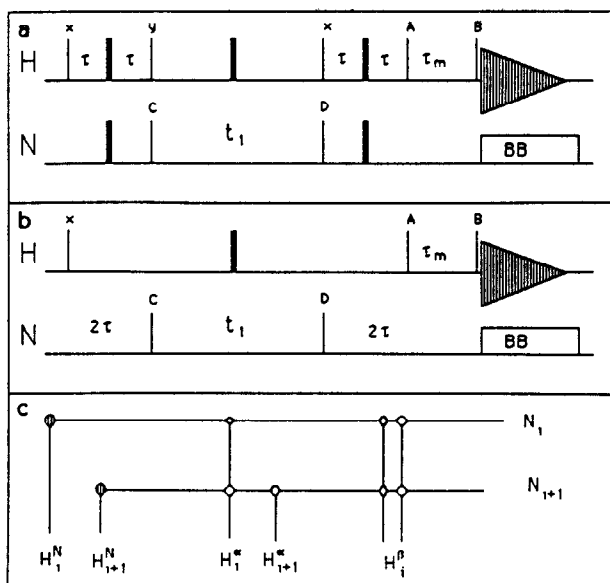


FIG. 13. Pulse sequences for ^{15}N relayed NOESY. (a) Heteronuclear antiphase evolution period,⁽¹²³⁾ (b) heteronuclear multiple quantum evolution period.⁽¹²⁴⁾

function is the same for all interproton vectors of the protein, this means:

$$J_{kj}(\omega) = J(\omega). \quad (4.11)$$

This essentially assumes that the protein is rigid. If, in addition, it is assumed that the protein is spherical, the form of the spectral density function of eqn. (4.9) is appropriate, and the correlation time τ_c can be obtained from ^{13}C relaxation time measurements of protonated carbons. The distance r_{kj} can be calculated from:

$$r_{kj} = \left(\frac{\gamma^4 \hbar^2}{10R_{kj}} \left(\frac{6\tau_c}{1 + (\omega_k + \omega_j)^2 \tau_c^2} - \frac{\tau_c}{1 + (\omega_k - \omega_j)^2 \tau_c^2} \right) \right)^{1/6}. \quad (4.12)$$

However, even this has not been done systematically for a protein.

Another approach, which makes the same assumption about the rigidity of the protein, is to use NOEs between protons separated by a fixed distance, r_0 , to calibrate the NOEs. The distance to be measured is then given by:

$$r_{kj} = r_0 (R_0/R_{kj})^{1/6}. \quad (4.13)$$

R_0 is the cross relaxation rate measured for the known distance r_0 . A possible standard distance in proteins is that between methylene protons which are separated by 1.8 Å. Distances between protons of aromatic side chains are also of fixed length. However, in practice, the NOEs between analogous protons of different aromatic side chains vary too much to be usable for calibrations. This indicates that the rigid-molecule approximation is not appropriate for longer side chains. Generally in proteins, protons with fixed distances have similar chemical shifts, and the cross peaks are close to the diagonal so that the peak integration is difficult. In addition, they are always J -coupled. This requires that zero-quantum contributions to NOESY cross peaks be eliminated carefully (see below) which is often a difficult task. For these reasons this approach to calibrating NOEs has not been used extensively.

A more qualitative calibration is more widely used. NOESY spectra are recorded with short mixing times (30 to 80 ms), so that spin diffusion is not significant. The spectra plotted with contours above a certain cut-off intensity are then searched for certain NOEs that must fall within certain distance ranges. For example, the distance between H_{i+1}^N and H_i^H must fall between 2.2 Å and 3.6 Å.^(1,6) The lowest contour level can be selected in a way that ca. 80% of the $d_{\alpha N}(i, i+1)$ NOEs are seen. This tells us that the cut-off distance must be lower than 3.6 Å, and an upper distance limit of 3.6 Å, for the NOEs observed in this plot seems to be a safe estimate. Another NOE with a narrow distance range that can be used for this kind of calibration is the intraresidue $d_{\alpha N}(i, i)$ which must fall between 2.2 and 2.9 Å.^(1,6) Other distances have also been discussed.^(1,6,125)

(iv) *Zero-quantum contributions in NOESY cross peaks of J-coupled nuclei.* The phase cycles of normal NOESY experiments and homospoil pulses in the mixing period cannot eliminate zero quantum contributions to NOESY peaks of *J*-coupled nuclei.^(119,126,127) *J*-peaks are antiphase in both dimensions, and have zero net intensity. Therefore, it might seem that their elimination is not necessary. However, the antiphase patterns of *J*-peaks are wider than the inphase NOE peaks and have perturbing sidelobes. For this reason, they interfere unfavorably with neighboring NOESY peaks and may obscure parts of the NOESY spectra. Some early attempts to eliminate *J*-cross peaks include random variation of the mixing time^(126,127) within each t_1 -value, or incrementation of the mixing time, proportional to t_1 .^(126,127) More recently, it has been proposed to co-add NOESY spectra in which the mixing times are increased in equal increments in order to eliminate zero-quantum frequencies $\Delta\omega$ in a certain range $\Delta\omega_{\min} < \Delta\omega < \Delta\omega_{\max}$. The increments $\Delta\tau$ have to be set to $1/(\Delta\omega_{\max} + \Delta\omega_{\min})$. The larger the number of increments, the better the suppression.⁽¹²⁸⁾ If only cross peaks to amide protons are concerned, a complete elimination of *J*-peaks can be achieved with the ^{15}N relayed NOESY of Fig. 13a.⁽¹²³⁾ Since the t_1 -period is a ^{15}N evolution, no zero-quantum coherence can be created by the 90° proton pulse prior to the mixing time. In contrast to the sequence of Fig. 13a, zero-quantum peaks are present in the sequence of Fig. 13b.

(v) *Spin diffusion.* For long mixing times NOEs between two spins can be relayed by a third spin.^(33,117-120,129,130) As a consequence, the NOEs are more intense than expected for a direct dipole-dipole interaction. The NOEs increase with τ_m^2 .⁽¹¹⁷⁻¹¹⁹⁾ The problem is most serious for NOEs involving methylene protons. Probably, the best method of taking account of spin diffusion in interpretation of NOESY peaks is by back calculation of NOESY spectra from molecular models and making iterative readjustments of distances.^(120,121)

4.4. The Practical Problem of Integrating All Cross Peaks in NOESY Spectra

The most limiting factor for quantitative analysis of NOESY spectra is the complexity of the spectra. A typical NOESY spectrum of a protein with a molecular weight of 8000 Da contains several thousand cross peaks. Integrating all these cross peaks in spectra with different mixing times and determining the initial slopes is a formidable task, and no one has yet tried this approach. To reduce the amount of work, it has been proposed that the spectra of a NOESY τ_m series be fitted point by point to a polynomial, prior to peak integration.⁽¹³¹⁾ The series of NOESY spectra recorded with different mixing times is then transformed into a series of spectra of Taylor coefficients. According to the Taylor expansion of eqn. (4.4), the spectrum of the zero-order coefficients should contain only diagonal peaks and base plane artefacts, the spectrum of the first-order coefficients should contain only cross relaxation rates, the spectrum of the second-order coefficients should contain spin diffusion and relaxation contributions, etc. To obtain the cross relaxation rates, it is only necessary to integrate the cross peaks of the first-order spectrum. This approach has been tried with the protein eglin c, and a series of NOESY spectra with 14 different mixing times has been used to perform this fit. It appears that this approach provides reliable values for the cross relaxation rates. The limiting factor is that the Taylor series has to be truncated after the second- or third-order term to achieve a decent signal-to-noise ratio.⁽¹³¹⁾ Nevertheless, a second order polynomial seems to be an acceptable approximation of the initial part of a NOESY build-up curve.

4.5. Back Calculation of NOESY Spectra with the Complete Relaxation Matrix Approach

The most promising avenue to quantitative analysis of NOESY spectra is the back calculation of NOESY spectra with the complete relaxation matrix approach.^(120, 121) The back calculation can only be done after an initial protein model is available. The spectra computed from the protein model can then be compared with the experimental spectra and the distances readjusted iteratively until a match between the calculated and the experimental spectrum is achieved. Calculation of NOESY spectra with eqn. (4.2) would be cumbersome. Bodenhausen and Ernst,⁽¹¹⁸⁾ and later on Keepers and James⁽¹²⁰⁾ have pointed out that eqn. (4.2) can be simplified by transformation of the relaxation matrix into a diagonal matrix:

$$R = T\lambda T^{-1}. \quad (4.13)$$

Obviously, T is the eigenvector matrix and λ is the diagonal matrix of eigenvalues. The series expansion of eqn. (4.2) shows that the solution simplifies to:^(118, 120)

$$A(\tau_m) = T e^{-\lambda\tau_m} T^{-1}. \quad (4.14)$$

With this approach NOESY spectra can be calculated by diagonalizing the relaxation matrix R and evaluating eqn. (4.14). This has been done, for example in the software packages CORMA⁽¹²⁰⁾ and COMATOSE.⁽¹²¹⁾ NOESY spectra for proteins of 9 kDa can be calculated within a few hours on a modern computer workstation⁽¹³²⁾ (see Fig. 14). With this tool in hand, distances can be readjusted iteratively until a good fit between the experimental and computed spectra is achieved. This may in future replace the tedious task of integrating many NOESY cross peaks. Obviously, this approach can also be used to check on models of internal motions.

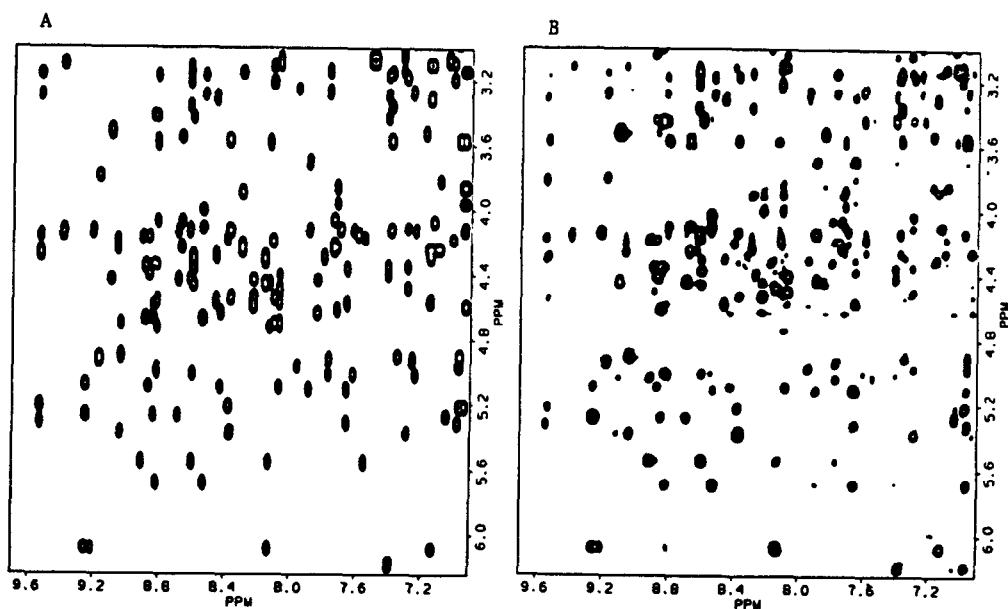


FIG. 14. Comparison of a calculated and an experimental NOESY spectrum of the protein eglin c.⁽¹³²⁾ Only a small part of the spectrum is shown.

4.6. NOEs and Mobility: Averaging Effects vs. Correlation Time Effects

The effect of internal motions on NOEs has been discussed extensively in the literature.^{1, 2, 133, 134} An excellent comprehensive treatment has been presented by Neuhaus and Williamson.⁽¹³⁵⁾ Internal motions can act on the effective correlation time. If the motions reorient the internuclear vector relative to the external field with a rate faster than the overall tumbling of the protein then they will decrease the cross relaxation rates in this region of the protein. Variation of the length of the internuclear vector r_{kj} causes an averaging of the cross relaxation rates. If the correlation time for the internal motion, τ_{int} , causing the length of the vector to fluctuate is much shorter than the overall correlation time ($\tau_{int} \ll \tau_c$), the averaged distance appears as $\langle r^3 \rangle^2$ in the cross relaxation rate. If the internal motion is much slower than the overall correlation time ($\tau_{int} \gg \tau_c$), the averaged distance appears as $\langle r^6 \rangle$ in the cross relaxation rate.⁽¹³⁵⁾ In any case, the apparent distance derived when neglecting internal motions is shorter than the average distance $\langle r \rangle$. Thus, the effects of internal motions on the correlation time and averaging have opposite tendencies.

Internal motions will also have the consequence of producing NOE patterns that are inconsistent with a single conformation. A proton may hop between different environments due to internal motions, experiencing NOEs from all these environments. It appears to be at different environments simultaneously so that the structure calculations searching for a single conformation will always show violations of distance constraints.⁽²⁷⁾

To separate motional from static aspects of protein structure it will be necessary to measure relaxation parameters. Let us consider the relation between the cross relaxation rate and the spectral density function. From eqns (4.6) and (4.7) we see that measuring cross relaxation rates samples the frequency spectrum of rotational motions of the internuclear vector r_{kj} ($J_{kj}(\omega)$) at the frequencies $(\omega_k + \omega_j)$ and $(\omega_k - \omega_j)$, with the relative weights of +12 and -2, respectively. This is illustrated in Fig. 15. The sampling of the spectral density function with measurements of other relaxation parameters is also indicated. For molecules of the size of proteins, there are few rotational motions with frequencies as high as $(\omega_k + \omega_j)$. Therefore, the motions at frequencies $(\omega_k - \omega_j)$ dominate the cross relaxation and make it negative (eqn. (4.6)). The sampling of the spectral density function with measurements of different relaxation parameters is summarized below, and is visualized in Figs 15a and b. The best known relaxation parameters are selective and non-selective longitudinal relaxation times, transverse relaxation time and cross relaxation rates. They sample the spectral density functions in a complex way at various frequencies with different weights. For proton relaxation parameters we have:

$$1/T_1^{k, sel} = \sum_j c_{kj} [2J(\omega_k - \omega_j) + 6J(\omega_k) + 12J(\omega_k + \omega_j)] \quad (4.15)$$

$$c_{kj} = \frac{\hbar^2 \gamma_k^2 \gamma_j^2}{20r_{kj}^6} \quad (4.16)$$

$$1/T_2^k = \sum_j c_{kj} [4J(0) + J(\omega_k - \omega_j) + 6J(\omega_j) + 3J(\omega_k) + 6J(\omega_k + \omega_j)] \quad (4.17)$$

$$1/T_1^{k, nonselect} = \sum_j c_{kj} [2J(\omega_k - \omega_j) + 6J(\omega_k) + 12J(\omega_k + \omega_j)] \quad (4.18)$$

$$R^{kj} = c_{kj} [-2J(\omega_k - \omega_j) + 12J(\omega_k + \omega_j)]. \quad (4.19)$$

Analogous parameters can be measured for the relaxation of carbon nuclei. Note that the longitudinal and transverse relaxation times sample an ensemble of spectral density functions of the various internuclear vectors. Only the cross relaxation rates sample a single spectral density function (Fig. 15a, b). At the high magnetic fields used for protein NMR chemical shift anisotropy contributions to the relaxation have to be considered for ^{13}C and ^{15}N . CSA relaxation contributes to T_1 and T_2 relaxation but not to cross relaxation. Therefore, measurements of σ_{HX} is important for separation of dipole-dipole and CSA relaxation.^(11,5)

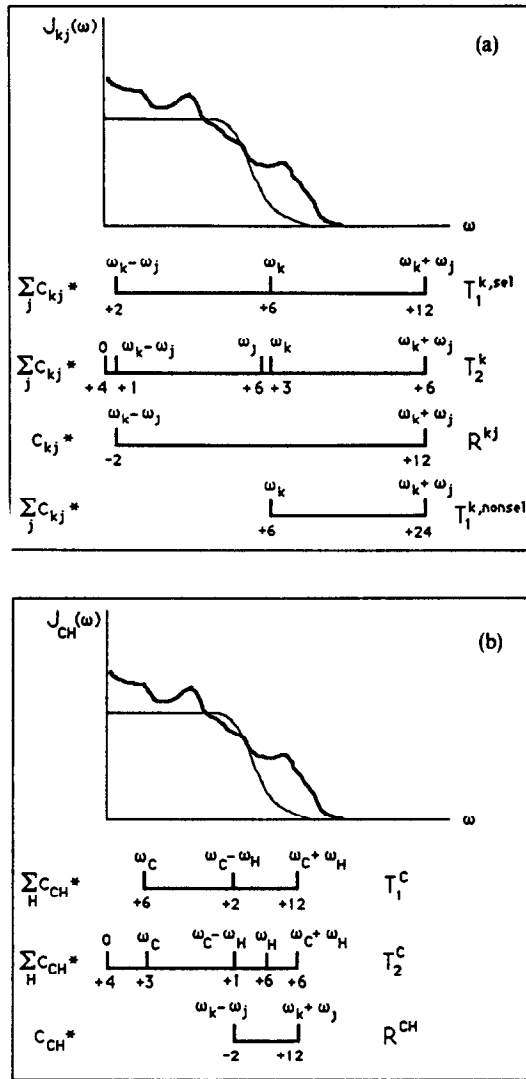


FIG. 15. Sampling of the spectral density function $J(\omega)$ with measurements of different relaxation parameters. A hypothetical spectral density function for the motion of the internuclear vector r_{kj} is shown on top. The frequencies at which different relaxation rates sample the spectral density function are indicated with the forks in the lower half of the figure. The relative weights with which the respective frequencies contribute to the relaxation rates are written underneath each fork. In addition the contributions are weighted with the factors c_{kj} which contain r_{kj}^{-6} . This is indicated on the left hand side of the figure. Note that $1/T_1^{k,sel}$, $1/T_2^k$, $1/T_1^{k,nonsel}$ sample an ensemble of spectral density functions while the cross relaxation rate R^{kj} samples a single spectral density function. (a) Proton relaxation parameters, (b) carbon relaxation parameters. In the latter case, a single spectral density function is sampled for methine carbons.

5. COUPLING CONSTANTS

Coupling constants have been investigated intensively for conformational analysis of peptides.⁽⁸⁵⁾ However, they have proved to be less useful for studies of proteins, firstly, because of the difficulty in making accurate measurements when the line widths are comparable or even larger than the coupling

constants and secondly because there are insufficient accessible coupling constants to characterize polypeptide dihedral angles uniquely. The development of successful determinations of protein conformations based on NOEs also made it appear that coupling constants were not really necessary for structural analysis of proteins. However, more recently, coupling constants have been combined with the NOE data to provide valuable information for refining protein structures, particularly at the local level on the protein surface, where fewer NOEs are available, and also in non-regular secondary structures. It is, however, necessary to extend the number of measurable coupling constants to heteronuclear coupling constants, and to find ways of making accurate measurements of coupling constants, even if they are smaller than the line widths.

5.1. Accurate Measurements of Heteronuclear Coupling Constants in ^{15}N Labeled Proteins

It has been realized for some time that knowledge of heteronuclear vicinal coupling constants between ^1H and ^{15}N would be valuable for characterizing the dihedral angles ψ and χ^1 .⁽⁸⁵⁾ Methods used for measuring these in the past have included one-dimensional and two-dimensional techniques where the coupling constants would be measured from inphase or antiphase splittings.^(136,137) However, if the line widths are comparable to or larger than the coupling constants, the apparent splittings are smaller or larger than the true coupling constants for antiphase and inphase multiplets, respectively.⁽¹³⁷⁾

With the availability of ^{15}N labeled proteins, a simple way has been found of measuring vicinal ^{15}N - ^1H coupling constants accurately.⁽¹³⁸⁾ A straight forward 2D experiment that results in intrasidial cross peaks between the amide proton and the β -protons (NOESY, TOCSY, RELAY) allows one to measure the intrasidial coupling constants $^3J(\text{N}_i\text{-H}_i^\beta)$, and 2D experiments that provide sequential $d_{\alpha\text{N}}$ cross peaks allow measurement of the coupling constant $^3J(\text{N}_i\text{-H}_{i-1}^\alpha)$.⁽¹³⁸⁾ No ^{15}N decoupling is applied during evolution and detection, and no pulses on ^{15}N other than 180° pulses are applied between evolution and detection. If the ^{15}N enrichment is significantly less than 100% a BIRD⁽⁷²⁻⁷⁴⁾ sequence can be inserted prior to each scan to detect couplings in the ^{15}N labeled protein. The pulse sequence and a schematic spectrum are shown in Fig. 16a and c, respectively. If no ^{15}N mixing pulse is applied between evolution and detection, two subspectra are observed for each set of signals associated with one NH group, corresponding to the up or down orientation of the ^{15}N spin,

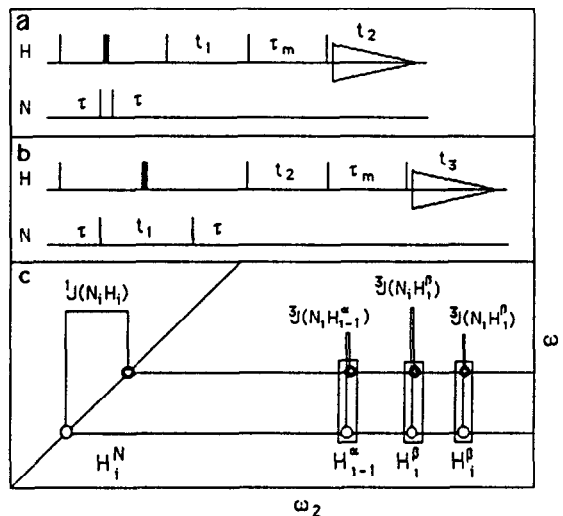


FIG. 16. (a) NOESY pulse sequence for measurements of vicinal coupling constants $^3J(\text{N-H})$. The BIRD sequence preceding the NOESY is to select for ^{15}N labeled proteins.^(138,140) (b) 3D version of the experiment.⁽¹³⁹⁾ (c) Schematic of the lower part of a NOESY of ^{15}N labeled protein.

respectively. The cross peaks are split along ω_1 by the large one-bond heteronuclear coupling $^1J(\text{N-H})$, and the vicinal heteronuclear couplings of interest can be measured accurately from the separation of the two non-overlapping peaks along ω_2 . Figure 16c is a schematic drawing of the cross peaks associated with a single amide NH group. Figure 17 shows an experimental NOESY spectrum of uniformly ^{15}N labeled transforming growth factor α (TGF α).⁽¹³⁸⁾ The region of cross peaks between amide protons and aliphatic protons is shown. If there is a problem of resolving all cross peaks of interest, the experiment of Fig. 16a can be expanded to a three-dimensional experiment by replacing the delay between the pair of ^{15}N pulses of Fig. 16a with a ^{15}N evolution period, as shown in Fig. 16b.⁽¹³⁹⁾ The spectrum drawn in Fig. 16c would correspond to a single ω_2 - ω_3 cross plane at the ω_1 frequency of a single nitrogen.

The principle used for measurements of vicinal ^{15}N - ^1H coupling constants can obviously be extended to measurements of vicinal coupling constants between protons and protonated carbons. An attractive coupling constant is that between the amide proton and the β -carbon. This $^3J(\text{H}^{\text{N}}-\text{C}^{\beta})$ coupling, combined with the homonuclear $^3J(\text{H}^{\text{N}}-\text{H}^{\alpha})$ coupling constant allows a more reliable determination of the dihedral angle ϕ . We have performed such an experiment on a 100 mM solution of a cyclic tetrapeptide consisting of a leucine, a modified glutamic acid and two non-standard amino acids. The experiments were performed at natural abundance of ^{13}C .⁽¹⁴⁰⁾ Figure 18 shows the H^{β} - H^{N} cross peak of a leucine residue. The heteronuclear coupling constant $^3J(\text{H}^{\text{N}}-\text{C}^{\beta})$ is 1.5 Hz and the homonuclear coupling constant $^3J(\text{H}^{\text{N}}-\text{H}^{\alpha})$, measured independently is 6.9 Hz. The homonuclear coupling constant is compatible with ϕ angles around -160° , -80° , $+30^\circ$ and $+90^\circ$. The $^3J(\text{H}^{\text{N}}-\text{C}^{\beta})$ coupling of 1.5 Hz excludes all but the range around $+30^\circ$.⁽⁸⁵⁾

5.2. Accurate Measurements of Homonuclear Coupling Constants

The best homonuclear techniques for accurate measurements of homonuclear coupling constants are ECOSY,⁽¹⁴¹⁾ COSY45⁽¹⁴⁶⁾ and P. E. COSY.⁽¹⁴²⁾ These techniques produce cross peaks that are 1:1 combinations of antiphase cross peaks and double-antiphase cross peaks (Fig. 19). This results in two square antiphase cross peak patterns displaced by the passive coupling in both dimensions. This passive coupling can then be measured from the outer components of these squares without cancellation effects of antiphase components as in regular COSY cross peaks. This spectral feature can be obtained by combination of different multiple quantum filtered COSY experiments⁽¹⁴¹⁾ or by COSY with flip angles of the mixing pulses different from 90° . However, these techniques require

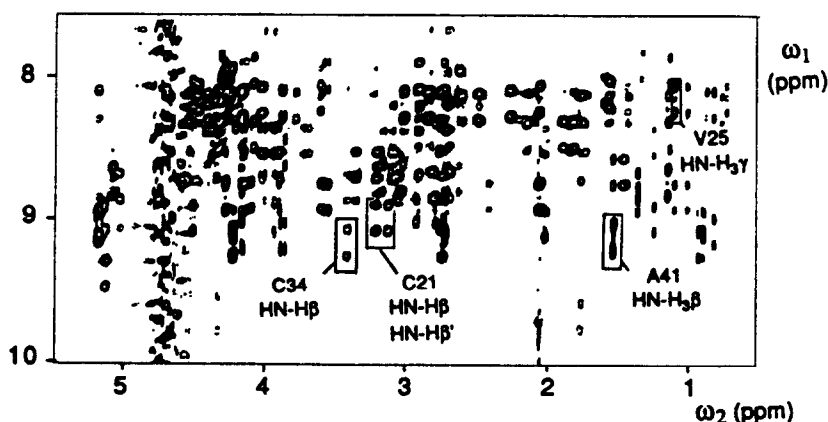


FIG. 17. Region of a NOESY spectrum in H_2O of a 5 mM sample of uniformly 99% ^{15}N enriched human transforming growth factor α (TGF α) containing cross peaks between amide protons and α -protons. Some of the intraresidue $\text{H}^{\text{N}}-\text{H}^{\beta}$ cross peaks with the multiplet structure indicated in Fig. 16 are identified with rectangular frames. The mixing time was 350 ms (pH = 3.5, 30°C) (adapted from Montelione *et al.*⁽¹³⁸⁾).

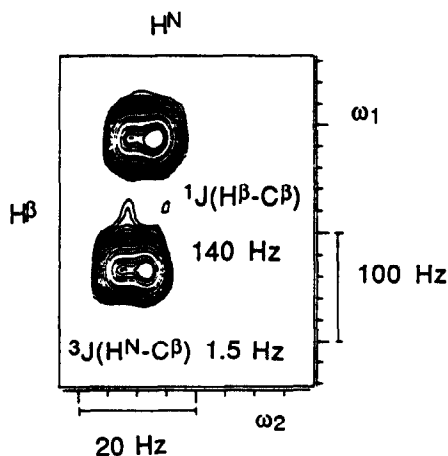


FIG. 18. Measurement of ${}^3J(\text{H}^{\text{N}}-\text{C}^{\beta})$ coupling constants in a 100 mM tetrapeptide, using the sequence of Fig. 16a. The cross peak is between H^{β} and H^{N} . The coupling constant measured is 1.5 Hz. The experiment was recorded in approximately 12 h.⁽¹⁴⁰⁾

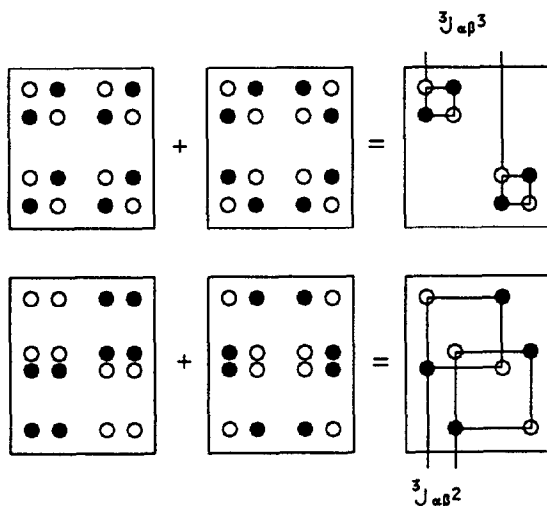
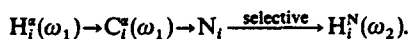


FIG. 19. ECOSY cross peak as combination of an antiphase and a double antiphase cross peak.

systems of at least three mutually coupled spins, such as H^{α} , $\text{H}^{\beta 2}$ and $\text{H}^{\beta 3}$ of an amino acid, and then the coupling constant ${}^3J(\text{H}^{\alpha}, \text{H}^{\beta 2})$ can be measured from the $(\text{H}^{\alpha}, \text{H}^{\beta 3})$ cross peak, and vice-versa. The ${}^3J(\text{H}^{\alpha}, \text{H}^{\text{N}})$ coupling constants, however, cannot be measured by using these techniques since the H^{N} is coupled only to one H^{α} and to no other proton. Glycines are an exception to this, of course.

The simple concept used for measurements of heteronuclear vicinal coupling constants, as described above, has been extended to measurements of homonuclear coupling constants, for example ${}^3J(\text{H}^{\text{N}}-\text{H}^{\alpha})$.⁽⁸³⁾ It requires an experiment where the two multiplet components are separated by a large one-bond coupling constant along ω_1 , and the coupling constant of interest is measured as the displacement of two multiplet components along ω_2 . This has been achieved with the pulse sequence of Fig. 8b. A short notation of the pulse sequence is:



The experiment yields intraresidue cross peaks between α -carbon and amide proton resonances. The TANGO⁽⁸⁴⁾ pulse (Fig. 8b) is a selective 90° pulse for protons coupled to ^{15}N , and does not mix the two orientations of the α -proton. Figure 20 shows the intraresidue cross peak for the tripeptide Ac-Asn-Pro (^{15}N)Tyr-NHMe. The experiment is proportional to the concentration of ^{15}N and ^{13}C . In order to achieve a good signal-to-noise ratio, ^{15}N and/or ^{13}C enrichment is required.

5.3. Stereospecific Assignments and Determination of the Angle χ^1

Stereospecific assignments of β -methylene protons, γ -methyl protons of valines and δ -methyl protons of leucines have been shown to be important for the quality of structure determinations of proteins.⁽¹⁴³⁻¹⁴⁹⁾ Stereospecific assignments of β -methylene protons and determination of the dihedral angle χ^1 are closely related. For β -methylene protons, stereospecific assignments have been achieved by measuring homonuclear $^3J(\text{H}^\alpha\text{-H}^\beta)$ coupling constants by direct analysis of multiplet patterns of COSY cross peaks or from spectral simulations.⁽¹⁴⁴⁾ In addition, intraresidue and sequential NOEs have been measured quantitatively to distinguish between different side chain conformations and between alternative stereospecific assignments. If we assume that the side chain adopts only one of the three staggered rotamers with $\chi^1 = 60^\circ$, 180° or -60° , the procedure for stereospecific assignments can be described following Fig. 20. This presents a view along the $\text{C}^\beta\text{-C}^\alpha$ axis of an amino acid residue containing a β -methylene group. Small (~ 3 Hz) or large (~ 12 Hz) $^3J(\text{H}^\alpha\text{-H}^\beta)$ coupling constants are expected if the H^β is *gauche* or *trans* to the H^α , respectively. Thus observation of two small $^3J(\text{H}^\alpha\text{-H}^\beta)$ coupling constants identifies the *gauche-gauche* conformation (a), one large and one small coupling constant identify either of the two other side chain conformations (b) and (c). However, homonuclear coupling constants alone cannot distinguish between $\chi^1 = 180^\circ$ and $\chi^1 = -60^\circ$, and they are not sufficient for stereospecific assignments. Figure 21 suggests that this might be possible from quantitative measurements of NOEs between the H^N and the β -protons. For the *gauche-gauche* conformation, for example, one expects a strong NOE between H^N and $\text{H}^{\beta 3}$, etc. Indeed, this kind of

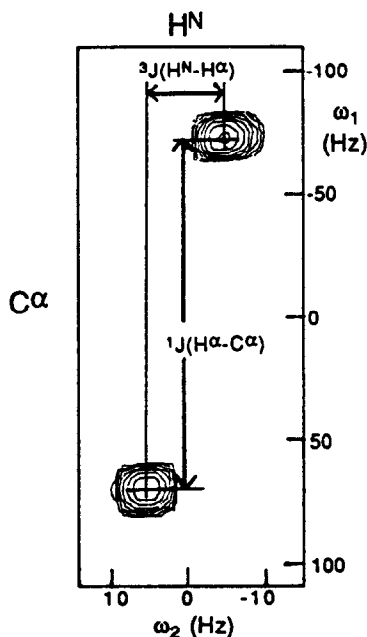


FIG. 20. Measurement of the $^3J(\text{H}^\text{N}\text{-H}^\alpha)$ coupling constant using the pulse sequence of Fig. 8b. The intraresidue $\text{H}^\text{N}\text{-H}^\alpha$ cross peak of Tyr 3 of the tripeptide Ac-Asn-Pro- (^{15}N) Tyr-NHMe is shown. The coupling constant is 9.7 ± 0.8 Hz.⁽⁸³⁾

analysis has led to stereospecific assignments of ca. 50% of the β -methylene protons.⁽¹⁴⁴⁻¹⁴⁹⁾ However, the intraresidue H^N-H^β distances depend on the dihedral angle ϕ , and this analysis is not fully satisfactory.

The availability of accurate measurements of heteronuclear $^{15}N-H^\beta$ coupling constants, as described above, makes the stereospecific assignments of β -methylene protons trivial. As for homonuclear coupling constants, small (1 Hz) and large (6 Hz) values are expected for the heteronuclear coupling constants, if the nitrogen is *gauche* and *trans* to the β -proton, respectively. For example, for the situation of Fig. 21a, the β -proton with the small heteronuclear coupling constant is identified as $H^{\beta 3}$ and the one with the large heteronuclear coupling constant is $H^{\beta 2}$. Figure 22 shows an experimental example. The figure is an expansion of the spectrum of Fig. 17. It shows the two H^N-H^β cross peaks of Cys 21. The left hand cross peak shows two large homonuclear coupling constants ($H^\beta-H^\beta$ and $H^\alpha-H^\beta$) and the right hand cross peak shows one large ($H^\beta-H^\beta$) and one small homonuclear coupling constant ($H^\alpha-H^\beta$, not resolved). Both $^{15}N-H^\beta$ coupling constants, as measured from the displacements of the upper from the lower multiplet components, are small (~ 1.5 Hz).⁽¹³⁸⁾ This identifies this residue as case b of Fig. 21, with $\chi^1=180^\circ$, and since the downfield β -proton is *trans* to H^α , it must be $H^{\beta 3}$.

If the observed homonuclear and heteronuclear coupling constants are different from the standard values for the three staggered rotamers this is an indication of rotational averaging. The simplest model to take care of this is a hopping between the three rotamers. Each of the two homonuclear and the two heteronuclear coupling constants can then be written as a weighed average of the coupling constants of the three staggered rotamers, and the relative populations of the three rotamers, p_a , p_b and p_c (with $p_a + p_b + p_c = 1$) can be determined.⁽⁶⁵⁾ Two fitting parameters can be determined from four experimentally measured coupling constants.

6. HYDROGEN BONDS

NMR is not a good method for identifying hydrogen bonds. Although slow exchange of an amide proton usually indicates that it is involved in an intramolecular hydrogen bond, there is no easy way of identifying the hydrogen bond acceptor. Only if there is additional information from NOEs that the residue of the slowly-exchanging amide proton is part of a regular secondary structure, and if all typical NOEs for this regular secondary structure are observed, is it safe to assume that the hydrogen bond acceptor is the one expected in this regular secondary structure.⁽¹⁾

Hydrogen bonds of amide protons with titratable side chains can sometimes be identified when the protein can be titrated around the pK of the titratable group in the side chain. If the pK has a unique value (for example determined by pH titration of carboxyl groups in ^{13}C spectra,⁽¹⁴⁵⁾) and the H^N shows a large titration shift of more than ~ 0.2 ppm, the hydrogen bond partners can be identified. This has been achieved for some surface groups of BPTI,⁽¹⁴⁵⁾ and the results showed some significant differences of the surface structure of this protein from the crystal structure.⁽¹⁴⁵⁾ Thus, these 'old-fashioned' titration experiments provide valuable information on surface conformations, complementary to X-ray diffraction data.

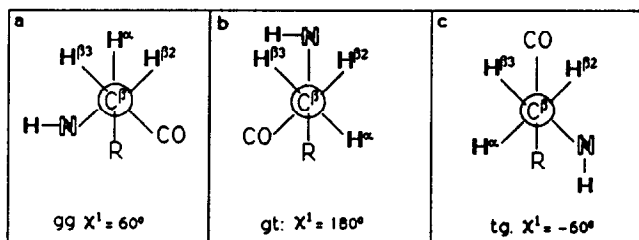


FIG. 21. The three staggered rotamers of residues with β -methylene protons.

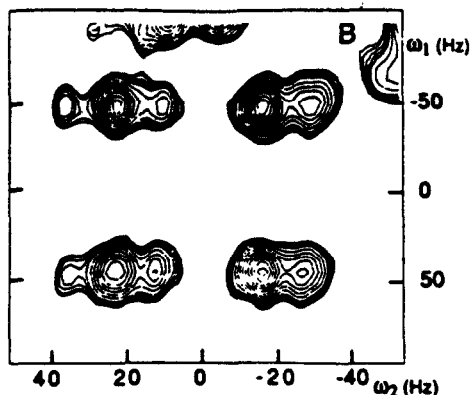


FIG. 22. H^N - H^β NOESY cross peaks of Cys 21 of uniformly ^{15}N labeled TGF α .⁽¹³⁸⁾

7. IDENTIFICATION OF SECONDARY STRUCTURES

7.1. Procedures for Characterizing Secondary Structures

It has been known for several years that regular secondary structures have typical NOE patterns.^(3,4) Figure 23 shows a NOESY spectrum of BPTI where the NOEs most typical of antiparallel β -sheets were first observed.⁽⁴⁾ These NOEs correspond to those indicated in the schematic drawing of the antiparallel β -sheet of Fig. 24. Subsequently, the short interproton distances characteristic of regular secondary structures have been analyzed systematically.^(1,2,5) Table 1 is a list of all these distances shorter than 4.5 Å. Indications of regular secondary structures from the NOE patterns can be corroborated by analysis of H^N - H^α coupling constants and by analysis of amide proton exchange. For regular helices and β -sheets, the amide protons involved in hydrogen bonds are expected to exchange slowly. The criteria for the different regular secondary structures can be summarized as follows.

α -helix and 3_{10} -helix. Both helices have strong $d_{NN}(i, i+1)$ NOEs, which are usually not observed in β -sheets. The $d_{\alpha N}(i, i+2)$, $d_{\alpha N}(i, i+3)$ and $d_{NN}(i, i+2)$ NOEs are often observable, in part probably due to efficient spin diffusion pathways. The $d_{\alpha N}(i, i+4)$ NOE is rarely observed in the α -helix. However, the $d_{\alpha\beta}(i, i+3)$ NOE is often present, although it may be overlapped in the crowded region of α - β cross peaks. There is no reliable way of distinguishing the α -helix from the 3_{10} -helix. The $d_{\alpha N}(i, i+2)$ and the $d_{\alpha N}(i, i+4)$ NOEs which may seem to be suitable for this distinction are so weak that they can be observed only at longer mixing times where the distance selectivity is lost due to spin diffusion. The distinction of the two helices can only be made at the level of the structure calculations. It is therefore dangerous to use H-bond constraints for helices at the first round of distance geometry calculation. However, from experience, longer helices are always α -helices, while short ones may be of either type.^(1,50) The coupling constants $^3J(H^\alpha H^N)$ usually are small in helices. However, some residues in helices may have coupling constants as large as in residues in β -sheets. The heteronuclear coupling constants $^3J(H^\alpha N_{i+1})$ are large in helices (-6Hz) while they are small in β -sheets (-2Hz) (see Table 1).^(1,2,5)

Tight turns are more difficult to identify by means of NMR than are helices and β -sheets, except when they are the hairpin bend of an antiparallel β -sheet. There are four types of tight turns as drawn in Fig. 25.^(1,50,151) The NMR parameters expected are summarized in Fig. 26. The classical type I and type II turns are rarely observed, while type I' and type II' turns occur frequently. The parameter most distinctive of tight turns, the H-bond between the carbonyl oxygen of residue 1 and the amide proton of residue 4, cannot reliably be identified by NMR. An isolated $d_{\alpha N}(i, i+2)$ or $d_{NN}(i, i+2)$ connectivity may be indicative of a tight turn with residue i being number 2 of the turn. Type I and I' turns have two consecutive $d_{NN}(i, i+1)$ connectivities (Fig. 26), whereas the type II and type II' turns have only one. Whether the homonuclear and heteronuclear coupling constants quoted in Fig. 26 will have impact on the identification of turns remains to be seen.

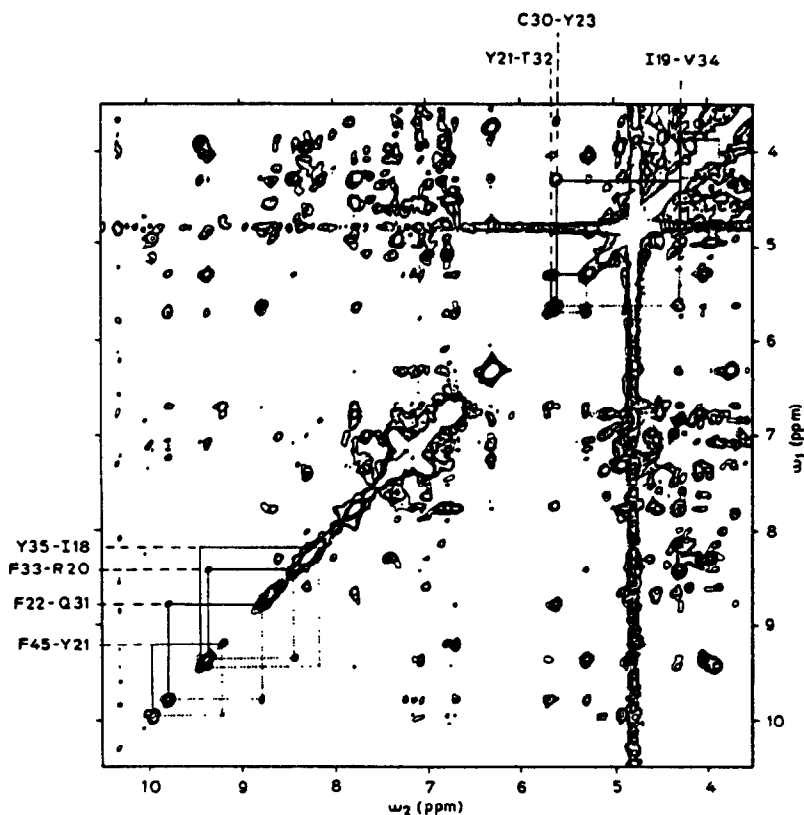


FIG. 23. NOESY spectrum of the basic pancreatic trypsin inhibitor. Some d_{NN} and d_{zz} NOEs typical for antiparallel β -sheets are identified (from Ref. 4).

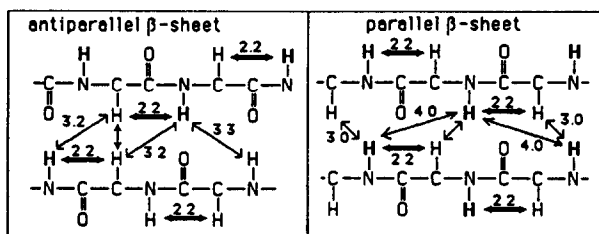


FIG. 24. Characteristic distances (in Å) for regular antiparallel and parallel β -sheets.

β -sheets can be identified very reliably from NOEs because of the short distances listed in Fig. 24. In addition, residues in β -sheets have, in most cases large coupling constants $^3J(\text{H}^2\text{H}^N)$. Slow NH exchange is observed for residues involved in intramolecular hydrogen bonds.

7.2. Impact of the Identification of Secondary Structures on Structure Calculations

The identification of secondary structures prior to distance geometry calculations seems, at a first glance, to be unnecessary, because the computer algorithms would provide this information anyway. However, knowledge of the secondary structures allows one to add H-bond constraints in regions

TABLE 1. Useful sequential and medium range ^1H - ^1H distances (in Å) and vicinal coupling constants for identification of secondary structures in polypeptide chains

	α -helix $\phi = -57^\circ$ $\psi = -47^\circ$	3_{10} -helix $\phi = -60^\circ$ $\psi = -30^\circ$	β -antiparallel $\phi = -139^\circ$ $\psi = +135^\circ$	β -parallel $\phi = -119^\circ$ $\psi = +113^\circ$
$d_{\alpha\text{N}}(i, i)$	2.7	2.7	2.8	2.8
$d_{\alpha\text{N}}(i, i+1)$	3.5	3.4	2.2	2.2
$d_{\alpha\text{N}}(i, i+2)$	4.4	3.8		
$d_{\alpha\text{N}}(i, i+3)$	3.4	3.3		
$d_{\alpha\text{N}}(i, i+4)$	4.2	(>4.5)		
$d_{\text{NN}}(i, i+1)$	2.8	2.6	4.3	4.2
$d_{\text{NN}}(i, i+2)$	4.2	4.1		
$d_{\text{NN}}(i, i+3)$	(4.8)	(5.2)		
$d_{\alpha\beta}(i, i+3)$	2.5-4.4	3.1-5.1		
$^3J(\text{H}^{\alpha}\text{H}^{\text{N}})$	4	4	10	10
$^3J(\text{H}^{\alpha}\text{N}_{i+1})$	-6	-6	-2	-2

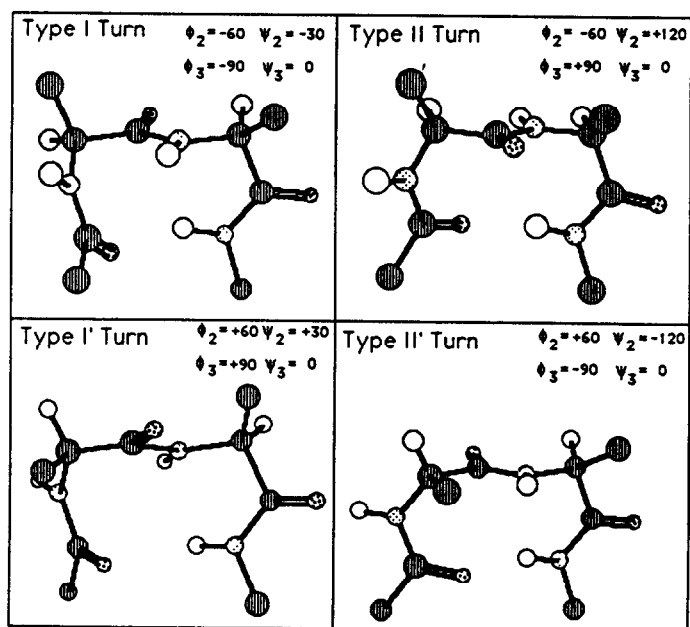


FIG. 25. Structures of the four most common types of tight turns.

assigned to regular secondary structures. In addition, convergence of distance geometry runs may be improved when starting with dihedral angles around those for the regular secondary structures previously identified. Furthermore, lower distance constraints may be used for residues on far ends of the secondary structures, corresponding to the length of a standard helix or a standard β -sheet.

8. TERTIARY STRUCTURES

After complete assignments and analysis of secondary structures extensive sets of data are collected as input for the distance geometry calculations. The quality of the structures depend on the amount of

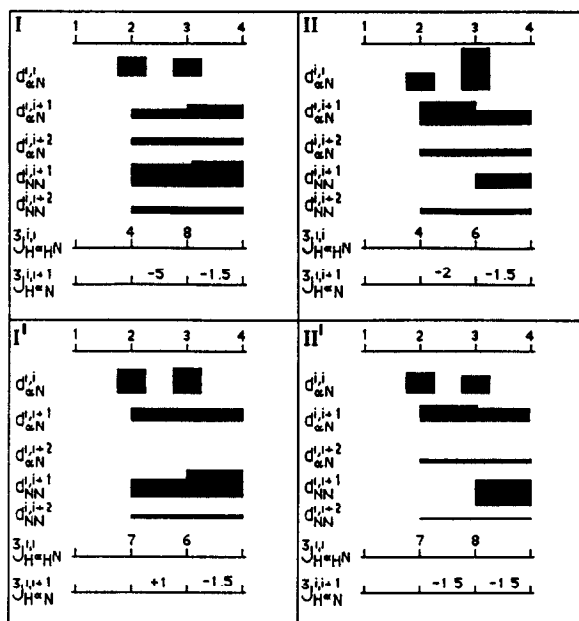


FIG. 26. Characteristic parameters to identify tight turns. The expected NOE intensities are indicated by the height of the vertical bars which are drawn as proportional to the logarithm of r_{ij}^{-6} . The coupling constants are estimated according to data in the literature.^(1,85)

conformational parameters, the precision of the parameters determined, the conversion of the distance geometry algorithms, and the refinement procedures.

8.1. Typical Data Sets for Structure Calculations

NOE values. The best determined NMR structures of proteins have used data sets of approximately ten interresidue NOEs per residue, with more than 30 interresidue NOEs for some of the internal residues. Surface residues often have fewer interresidue NOEs. In the best cases, NOEs have been classified in six groups corresponding to six upper distance limits, such as 2.4 Å, 2.6 Å, 2.9 Å, 3.2 Å, 3.8 Å and 4.6 Å.⁽¹⁵²⁾ For NOEs involving methyl groups, an additional correction of 0.5 Å is usually added.⁽¹⁴⁵⁾ In most cases the lower distance limits are set to the Van der Waals distances. This is based on the assumption that some NOEs may be quenched by mobility.

Stereospecific assignments have been obtained for up to 50% of the β -methylene protons for several proteins using the conventional assignment technique based on NOEs and homonuclear coupling constants (Fig. 21).⁽¹⁴⁴⁻¹⁴⁹⁾ There is no experience yet with the use of heteronuclear ^1H - ^{15}N coupling constants for stereospecific assignments in proteins. Valine γ -methyl groups can usually all be assigned stereospecifically by analysis of NOEs and the α - β coupling constant, except when the two methyl groups are degenerate.^(143, 144) Additional stereospecific assignments and stereospecific assignments of δ -methyl groups of leucines have been obtained during the final stage of structure calculations where the two alternative stereospecific assignments can be distinguished on the basis of the NOEs. Usually only one alternative is consistent with the observed NOEs.⁽¹⁵²⁾ The same approach has been used for stereospecific assignments of α -methylene protons of glycines, δ -methylene protons of prolines and some γ -methylene protons of glutamines.⁽¹⁵²⁾

Dihedral angle constraints. The conventional way of measuring coupling constants from the antiphase splittings of COSY cross peaks only allows one to measure coupling constants which are at least twice as large as the line width.⁽¹³⁷⁾ A detailed list of dihedral constraints was used by Kline *et*

al.⁽¹⁵²⁾ For ϕ -angle constraints they chose:

$$\begin{aligned} {}^3J(\text{H}^N\text{H}^\alpha) < 5.5 \text{ Hz} & \quad \phi = [-90^\circ, -40^\circ] \\ {}^3J(\text{H}^N\text{H}^\alpha) > 8.0 \text{ Hz} & \quad \phi = [-160^\circ, -80^\circ] \\ {}^3J(\text{H}^N\text{H}^\alpha) > 10.0 \text{ Hz} & \quad \phi = [-140^\circ, -100^\circ]. \end{aligned}$$

The more accurate methods for measuring ${}^3J(\text{H}^N\text{H}^\alpha)$ coupling constants described above have not yet been applied to proteins. Constraints for the angle χ^1 have been obtained extensively and, in the best determined cases, such constraints were available for ca. half the residues. It was always assumed that the side chain adopts one of the three staggered conformations (Fig. 21), and that the χ^1 angle was constrained within $\pm 60^\circ$ around the angle of the staggered conformation.

Hydrogen bond constraints. Depending on the amount of regular secondary structure, and the number of NOEs characterizing the secondary structure, a few hydrogen bond constraints can be identified. For proteins between 50 and 70 residues, up to 20 hydrogen bond constraints have been identified.^(145,153)

8.2. Algorithms for Model Building

The first algorithm developed that can handle the construction of a protein model from distance constraints was based on the EMBED method.⁽¹⁷⁻²⁰⁾ The program is called DISGEO and is written in Pascal.^(23,24) The algorithm creates a set of Cartesian coordinates from a set of distance constraints. The initial distances are chosen at random within the distance limits. The distance limits are corrected to obey triangle and tetrangle inequalities. From the distances a metric matrix is constructed, and from the eigenvalues and eigenvectors of the metric matrix the coordinates are obtained. The second algorithm developed was a Variable Target Function Method and this program is called DISMAN.⁽²⁵⁾ It acts in the dihedral angle space. First, a starting structure is created by choosing all dihedral angles at random. Then, a target function is constructed which consists essentially of the violations of the distance limits by the model structure. The target function initially only extends over two or three residues. The dihedral angles of the model structure are then varied to minimize the target function. The target function is moved along the polypeptide chain while optimizing the model over the limited length covered by the target function. The length of the segment covered by the target function is then increased and the model is again optimized until finally the target function covers the whole polypeptide chain. The DISGEO and the DISMAN algorithms have been shown to produce essentially equivalent results⁽¹⁴⁵⁾ although DISMAN seems to have more convergence problems.

8.3. Refinements

The models obtained with the distance geometry programs are often refined by including energy terms. This has been achieved with restrained molecular dynamics calculations (RMD) where the distance constraints were included as pseudo-energy potentials. This approach was first used by Kaptein and coworkers as an alternative to distance geometry calculations.⁽¹⁵⁴⁾ Nowadays, RMD methods are mainly used for refinement of models obtained from other calculations.^(155,156) RMD calculations are computationally expensive and to overcome this problem all non-bonding potentials may be removed with only pseudo potentials from distance constraints being retained. This is called simulated annealing.⁽¹⁵⁷⁾ Simulated annealing has been used for a number of structure calculations.^(158,159) Restrained refinements have also been performed with the AMBER⁽¹⁶⁰⁾ program package.⁽¹⁶¹⁾

9. CRITERIA FOR JUDGING THE QUALITY OF STRUCTURES

9.1. Residual Violations of Constraints and Average Root Mean Square Distances (rmsd)

The first criterion is the extent to which the experimental constraints on distances and dihedral angles are violated. This is essentially a criterion to check on the performance of the algorithm. Based

on the residual violations, structures are either accepted as converged, or they are rejected. Only the accepted structures are compared and an average root mean square distance is computed for the class of converged structures. For example, in the case of Tendamistat, only 9 out of 100 structures were accepted. For these nine structures the average rmsd was as low as 0.85 Å for backbone atoms.⁽¹⁵²⁾

Residual violations and rmsds are good criteria for judging the performance of the algorithms, but they are not ideal measures of the quality of the structure. The upper distance limits and the limits on the dihedral angles are chosen somewhat arbitrarily. Choosing (too) narrow limits of constraints may leave sizeable residual violations of constraints and may result in very precise structures with low rmsds. Thus the structures may appear to be well-defined even though they are not the correct structures. Choosing too wide ranges of constraints may result in no residual violations, but the structures may have a large average rmsd.

While it has been recognized that rmsds are not good measures for assessing the quality of NMR structures, there is no better alternative available as yet.

9.2. Back Calculation and R-Factors

It now seems possible to compute the NOESY spectra from the model structures, using the complete relaxation matrix approach.⁽¹²⁰⁾ The experimental and the computed spectra can be scaled and subtracted, and the residual can be used as an *R*-factor. Obviously, regions such as the diagonal and the region around the solvent line have to be excluded from this comparison.

10. LIMITATIONS FOR STRUCTURAL STUDIES OF PROTEINS BY NMR

The structural analysis of proteins by NMR is limited to a certain class of proteins. NMR can determine structures only of those proteins that have a unique 3D structure, are small, soluble to at least 1 mM, homogeneous, and have spectra with narrow lines. At least 1 μmol of protein (0.5 ml of a 1 mM solution) is required. Only certain proteins fulfil these requirements, mainly proteins with rather rigid structures. Proteins that act by changing their structure, or that have to be adaptable to interaction with a target molecule, may not be accessible to the analysis.

10.1. Size of the Protein

So far proteins up to 140 residues have been assigned, and structures of proteins of up to 100 residues have been determined. With increasing molecular weight the number of signals increases linearly, and so does approximately, the number of cross peaks. The spectra are overcrowded and signal overlap and signal degeneracy become more serious. Methods of isotope labeling and 3D techniques discussed above can overcome the problem of overcrowding to some extent. It may be that proteins up to 200 residues can be analyzed by NMR.

10.2. Bad Spectra—Large Line Width—Small Chemical Shift Dispersion

Narrow lines are crucial for detailed structural analysis of proteins. If the line widths are larger than homonuclear coupling constants, antiphase cross peaks are weak or undetectable, and all through-bond cross peaks which arise due to homonuclear scalar couplings are generally difficult to observe. Longer pulse sequences cannot be applied because of short T_2 s. If large line widths are caused only by dipole-dipole interactions and the slow tumbling rates of the proteins then random deuteration⁽¹⁶²⁾ may help to narrow the lines. This method has been applied to thioredoxin⁽¹⁶²⁾ where the protein was randomly deuterated to 75%. With this procedure, the multiplet structures of all proton resonances that give rise to COSY cross peaks are essentially doublets, and the multiplet structures giving rise to NOESY cross peaks are essentially singlets. In addition to the simplification of the multiplet structures, the line widths also decrease due to the dilution of the protons and the reduction of the dipole-dipole interaction. The approach of random deuteration is, however, very expensive and as yet, has not been generally applied.

A large molecular weight is only one of the factors causing large line widths. Other factors are equally important but are not well understood. These include aggregation and exchange between multiple conformations. The quality of poor protein spectra may be improved by variation of temperature, pH, ionic strength, buffer, concentration, or cofactors. Attempts to improve poor spectra are comparable with the attempts to make a protein crystallize for diffraction studies. The conditions that are best for NMR studies may not always be physiologically relevant. There are proteins with molecular weights well above 20 kDa that show well-resolved COSY and NOESY spectra, whereas small proteins below 10 kDa have poor spectra and cannot be analyzed in detail.

pH range. The best pH range for NMR is between pH 2 and pH 6. At these conditions the amide proton exchange is slowest (minimum at around pH 3). Above and below this range the amide exchange rates increase linearly with the concentration of OH^- (base catalysis) and H_3O^+ (acid catalysis). This pH range may not be available for many proteins because (i) they may not be soluble at this pH, such as proteins with a low isoelectric point, (ii) they may be unfolded at this pH, (iii) metal binding proteins may lose their metals at acidic pH, (iv) enzymes may be inactive at low pH. The factors determining the amide proton exchange on protein surfaces are not well understood. There are some proteins where all surface amide protons exchange so slowly that there is no saturation transfer from the saturated water line even a little above neutral pH (for example in the basic pancreatic trypsin inhibitor), while there are other proteins where significant saturation transfer occurs even at pH 5. This may be related to electrostatic effects which are not yet properly understood.

Chemical shift dispersion. A good dispersion of chemical shifts is important to resolve the resonances so that cross peaks can be assigned unambiguously. Proteins containing many aromatic residues are easier to analyze than others. From experience, proteins with a lot of β -sheet secondary structures have more chemical shift dispersion than helical proteins or even proteins with many turns or non-regular secondary structures. Proteins with a large variety of residues are obviously easier to handle than proteins with excess of a few types of residues.

10.3. Paramagnetic Centers

Structures of proteins containing paramagnetic centers may never be completely determined by NMR unless an isomorphous replacement of the paramagnetic group with a diamagnetic one can be achieved. The paramagnetic center usually broadens the proton resonances in its immediate environment by dipole-dipole interaction of the electron spin with the nuclear spins such that these cannot be detected.

10.4. Protein Degradation

Many proteins are unstable and this is a serious problem when NMR data acquisition may require keeping the protein for days or even weeks in solution. Protein degradation may be caused by self digestion (proteases) or by desamidation of glutamines and asparagines. The latter phenomenon is most serious at neutral pH but less of a problem at low pH. Some partial sequences are particularly prone to desamidation. Modification of such sequences by site-directed mutagenesis can help to eliminate desamidation sites.

10.5. Mobility

Mobility is a generic property of proteins. Some proteins need mobility to function. For the process of structure determination mobility is a limitation. Mobility may lead to NOEs inconsistent with a single conformation. This may be difficult to distinguish from poor convergence of the algorithm. On the other hand, mobility may lead to broad lines. This would also cancel antiphase cross peaks. A combined analysis that takes account of conformational and dynamic aspects of protein structure is a challenging task, and it will take some time to solve this aspect satisfactorily.

11. CONCLUSIONS

Several protein structures have been determined by NMR during the last five years. However, the technology is still at a very early stage. Many of the procedures used are not as yet fully satisfactory. The data analysis is mainly made by hand. Mobility is neglected. 3D NMR is at a very early stage, and isotope labeling has not yet been used extensively. Considering this, the success of the present technology is amazing. It can be anticipated that computer assisted data analysis or even automated assignments and automated parameter determination will be developed. Coupling constants will be included in the analysis to a much larger extent. It can be expected that structures of proteins up to 20 or even 30 kDa may be solved by NMR in a few years. It is also likely that surface structures of proteins will be characterized more reliably. Parameters characterizing the mobility will probably be determined in the future, similar to the B-factors in X-ray analysis.

Acknowledgements—The author thanks Dr. G. Montelione, Dr. V. Thanabal, Dr. N. R. Nirmala, S. Hyberts, and N. Beeson for fruitful discussions on the subject of this manuscript. The preparation of this manuscript was in part supported by grants from NSF (project DMB86-16059) and NIH (project GM38608).

REFERENCES

1. K. WÜTHRICH, *NMR of Proteins and Nucleic Acids*, J. Wiley, New York (1986).
2. R. R. ERNST, G. BODENHAUSEN and A. WOKAUN, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford Science Publications, Oxford (1986).
3. A. DUBS, G. WAGNER and K. WÜTHRICH, *Biochim. Biophys. Acta* **577**, 177 (1979).
4. G. WAGNER, ANIL KUMAR and K. WÜTHRICH, *Eur. J. Biochem.* **114**, 375 (1981).
5. K. WÜTHRICH, G. WIDER, G. WAGNER and W. BRAUN, *J. Mol. Biol.* **155**, 311 (1982).
6. M. BILLETER, W. BRAUN and K. WÜTHRICH, *J. Mol. Biol.* **155**, 321 (1982).
7. G. WAGNER and K. WÜTHRICH, *J. Mol. Biol.* **155**, 347 (1982).
8. G. WIDER, H. K. LEE and K. WÜTHRICH, *J. Mol. Biol.* **155**, 367 (1982).
9. J. JEENER, Ampère International Summer School, Basko Polje, Yugoslavia (1971).
10. R. R. ERNST, *Chimia* **29**, 179 (1975).
11. L. MÜLLER, ANIL KUMAR and R. R. ERNST, *J. Chem. Phys.* **63**, 5490 (1975).
12. W. P. AUE, E. BARTHOLDI and R. R. ERNST, *J. Chem. Phys.* **64**, 2229 (1976).
13. K. NAGAYAMA, ANIL KUMAR, K. WÜTHRICH and R. R. ERNST, *J. Magn. Reson.* **40**, 321 (1980).
14. J. JEENER, B. H. MEIER, P. BACHMANN and R. R. ERNST, *J. Chem. Phys.* **71**, 4546 (1979).
15. ANIL KUMAR and R. R. ERNST, *Biochem. Biophys. Res. Commun.* **95**, 1 (1980).
16. A. BAX and R. FREEMAN, *J. Magn. Reson.* **44**, 542 (1981).
17. I. D. KUNTZ, G. M. CRIPPEN, P. A. KOLLMAN and D. KIMMELMAN, *J. Mol. Biol.* **106**, 983 (1976).
18. I. D. KUNTZ, G. M. CRIPPEN and P. A. KOLLMAN, *Biopolymers* **18**, 939 (1979).
19. G. M. CRIPPEN, *J. Comp. Phys.* **24**, 96 (1977).
20. G. M. CRIPPEN and T. F. HAVEL, *Acta Crystallogr.* **A3A** 282 (1978).
21. H. WAKO and H. SCHERAGA, *J. Prot. Chem.* **1**, 85 (1981).
22. T. F. HAVEL, I. D. KUNTZ and G. M. CRIPPEN, *Bull. Math. Biol.* **45**, 665 (1983).
23. T. F. HAVEL and K. WÜTHRICH, *Bull. Math. Biol.* **45**, 673 (1984).
24. T. F. HAVEL and K. WÜTHRICH, *J. Mol. Biol.* **182**, 281 (1985).
25. W. BRAUN and N. GO, *J. Mol. Biol.* **186**, 611 (1985).
26. J. L. MARKLEY, *Methods Enzymol.* **176**, 12 (1989).
27. G. WAGNER, W. BRAUN, T. F. HAVEL, T. SCHAUMANN, N. GO and K. WÜTHRICH, *J. Mol. Biol.* **196**, 611 (1987).
28. W. J. CHAZIN, T. E. HUGLI and P. E. WRIGHT, *Biochemistry* **27**, 9139 (1988).
29. D. G. NETTESHEIM, R. P. EDALJI, K. W. MOLLISON, J. GREER and E. R. P. ZUIDERWEG, *Proc. Natl. Acad. Sci. USA* **85**, 5036 (1988).
30. M. H. FREY, G. WAGNER, M. VASAK, O. W. SØRENSEN, D. NEUHAUS, E. WÖRGÖTTER, JHR. KÄGI and K. WÜTHRICH, *J. Am. Chem. Soc.* **107**, 6847 (1985).
31. P. SCHULTZE, E. WÖRGÖTTER, W. BRAUN, G. WAGNER, M. VASAK, JHR. KÄGI and K. WÜTHRICH, *J. Mol. Biol.* **203**, 251 (1988).
32. R. E. KLEVIT and E. B. WAYGOOD, *Biochemistry* **25**, 7774 (1986).
33. G. WAGNER and K. WÜTHRICH, *J. Magn. Reson.* **33**, 675 (1979).
34. ANIL KUMAR, G. WAGNER, R. R. ERNST and K. WÜTHRICH, *Biochem. Biophys. Res. Commun.* **96**, 1156 (1980).
35. K. NAGAYAMA and K. WÜTHRICH, *Eur. J. Biochem.* **114**, 365 (1981).
36. K. NAGAYAMA, K. WÜTHRICH and R. R. ERNST, *Biochem. Biophys. Res. Commun.* **90**, 305 (1979).
37. U. PIANTINI, O. W. SØRENSEN and R. R. ERNST, *J. Am. Chem. Soc.* **104**, 6800 (1982).

38. A. J. SHAKA and R. FREEMAN, *J. Magn. Reson.* **51**, 169 (1983).
39. M. RANCE, O. W. SØRENSEN, G. BODENHAUSEN, G. WAGNER, R. R. ERNST and K. WÜTHRICH, *Biochem. Biophys. Res. Commun.* **117**, 479 (1983).
40. N. MÜLLER, R. R. ERNST and K. WÜTHRICH, *J. Am. Chem. Soc.* **108**, 6482 (1986).
41. M. RANCE and P. WRIGHT, *J. Magn. Reson.* **66**, 372 (1986).
42. A. BAX, R. FREEMAN and S. P. KEMPEL, *J. Am. Chem. Soc.* **102**, 4849 (1980).
43. A. BAX, R. FREEMAN, T. A. FRENKIEL and M. H. LEVITT, *J. Magn. Reson.* **43**, 478 (1981).
44. T. H. MARECI and R. FREEMAN, *J. Magn. Reson.* **51**, 531 (1983).
45. L. BRAUNSCHEILER, G. BODENHAUSEN and R. R. ERNST, *Mol. Phys.* **48**, 535 (1983).
46. G. WAGNER and E. R. P. ZUIDERWEG, *Biochem. Biophys. Res. Commun.* **113**, 854 (1983).
47. G. EICH, G. BODENHAUSEN and R. R. ERNST, *J. Am. Chem. Soc.* **104**, 3731 (1982).
48. P. H. BOLTON and G. BODENHAUSEN, *Chem. Phys. Lett.* **89**, 139 (1982).
49. G. WAGNER, *J. Magn. Reson.* **55**, 151 (1983).
50. G. KING and P. E. WRIGHT, *J. Magn. Reson.* **54**, 328 (1983).
51. A. BAX and G. DROBNY, *J. Magn. Reson.* **61**, 306 (1985).
52. L. BRAUNSCHEILER and R. R. ERNST, *J. Magn. Reson.* **53**, 521 (1983).
53. A. BAX and D. G. DAVIS, *J. Magn. Reson.* **65**, 355 (1986).
54. G. OTTING and K. WÜTHRICH, *J. Magn. Reson.* **66**, 359 (1986).
55. G. OTTING and K. WÜTHRICH, *J. Magn. Reson.* **75**, 546 (1987).
56. S. C. BROWN, P. L. WEBER and L. MÜLLER, *J. Magn. Reson.* **77**, 166 (1988).
57. R. M. KELLER, R. BAUMANN, E.-H. HUNZIKER-KWIK, F. J. JOUBERT and K. WÜTHRICH, *J. Mol. Biol.* **163**, 623 (1983).
58. C. REDFIELD and C. M. DOBSON, *Biochemistry* **27**, 122 (1987).
59. S. W. ENGLANDER and A. S. WAND, *Biochemistry* **26**, 5953 (1987).
60. W. J. CHAZIN, M. RANCE and P. E. WRIGHT, *J. Mol. Biol.* **202**, 603 (1988).
61. A. D. ROBERTSON, E. O. PURISMA, M. A. EASTMAN and H. A. SCHERAGA, *Biochemistry* **28**, 5930-5938 (1989).
62. D. M. LEMASTER and F. M. RICHARDS, *Biochemistry* **24**, 7263 (1985).
63. L. P. MCINTOSH, R. H. GRIFFEY, D. C. MUCHMORE, C. P. NIELSON, A. G. REDFIELD and F. W. DAHLQUIST, *Proc. Natl. Acad. Sci. USA* **84**, 1244 (1987).
64. B. J. STOCKMAN, W. M. WESTLER, E. S. MOOBERRY and J. L. MARKLEY, *Biochemistry* **27**, 136 (1988).
65. D. C. MUCHMORE, L. P. MCINTOSH, C. B. RUSSELL, D. E. ANDERSON and F. W. DAHLQUIST, *Methods Enzymol.*, in press.
66. L. P. MCINTOSH and F. W. DAHLQUIST, *Quat. Rev. Biophys.*, in press.
67. D. A. TORCHIA, S. W. SPARKS and A. BAX, *Biochemistry* **28**, 5509 (1989).
68. R. H. GRIFFEY and A. G. REDFIELD, *Quat. Rev. Biophys.* **19**, 51 (1987).
69. R. H. GRIFFEY, A. G. REDFIELD, R. E. LOOMIS and F. W. DAHLQUIST, *Biochemistry* **24**, 817 (1985).
70. G. OTTING, H. SENN, G. WAGNER and K. WÜTHRICH, *J. Magn. Reson.* **70**, 500 (1986).
71. A. BAX and M. A. WEISS, *J. Magn. Reson.* **71**, 571 (1987).
72. J. R. GARBOW, D. P. WEITEKAMP and A. PINES, *Chem. Phys. Lett.* **93**, 504 (1982).
73. A. BAX, *J. Magn. Reson.* **53**, 517 (1983).
74. V. RUTAR, *J. Magn. Reson.* **56**, 87 (1984).
75. H. SENN, G. OTTING and K. WÜTHRICH, *J. Am. Chem. Soc.* **109**, 1090 (1987).
76. B. J. STOCKMAN, M. D. REILY, W. M. WESTLER, E. L. ULRICH and J. L. MARKLEY, *Biochemistry* **28**, 230 (1989).
77. W. M. WESTLER, B. J. STOCKMAN, Y. HOSOYA, Y. MIYAKE, M. KAINOSHO and J. L. MARKLEY, *J. Am. Chem. Soc.* **110**, 6256 (1988).
78. G. T. MONTELIONE and G. WAGNER, *J. Magn. Reson.*, in press.
79. G. A. MORRIS and R. FREEMAN, *J. Am. Chem. Soc.* **101**, 760 (1979).
80. G. A. MORRIS, *J. Am. Chem. Soc.* **102**, 428 (1980).
81. D. P. BURUM and R. R. ERNST, *J. Magn. Reson.* **39**, 163 (1980).
82. G. A. MORRIS, *J. Magn. Reson.* **41**, 185 (1980).
83. G. T. MONTELIONE and G. WAGNER, *J. Am. Chem. Soc.* **111**, 5474 (1989).
84. S. WIMPERIS and R. FREEMAN, *J. Magn. Reson.* **58**, 348 (1984).
85. V. F. BYSTROV, *Prog. NMR Spectrosc.* **10**, 41 (1976).
86. A. BAX and M. F. SUMMERS, *J. Am. Chem. Soc.* **108**, 2093 (1986).
87. A. BAX and D. MARION, *J. Magn. Reson.* **78**, 186 (1988).
88. A. M. GRONENBORN, P. T. WINGFIELD and G. M. CLORE, *Biochemistry* **28**, 5081 (1989).
89. G. W. VUISTER and R. BOELEN, *J. Magn. Reson.* **73**, 328 (1987).
90. C. GRIESINGER, O. W. SØRENSEN and R. R. ERNST, *J. Am. Chem. Soc.* **109**, 7227 (1987).
91. C. GRIESINGER, O. W. SØRENSEN and R. R. ERNST, *J. Magn. Reson.* **73**, 574 (1987).
92. G. W. VUISTER, R. BOELEN and R. KAPTEIN, *J. Magn. Reson.* **80**, 176 (1988).
93. H. OSCHKINAT, C. GRIESINGER, P. J. KRAULIS, O. W. SØRENSEN, R. R. ERNST, A. M. GRONENBORN and G. M. CLORE, *Nature* **332**, 374 (1988).
94. H. OSCHKINAT, C. CIESLAR, A. M. GRONENBORN and G. M. CLORE, *J. Magn. Reson.* **81**, 212 (1989).
95. G. WAGNER, *J. Magn. Reson.* **57**, 497 (1984).
96. C. GRIESINGER, O. W. SØRENSEN and R. R. ERNST, *J. Magn. Reson.* **84**, 14 (1989).

97. H. OSCHKINAT, C. CIESLAR, T. A. HOLAK, G. M. CLORE and A. M. GRONENBORN, *J. Magn. Reson.* **83**, 450 (1989).
98. R. BOELEN, G. W. VUISTER, T. M. G. KOMIG and R. KAPTEIN, *J. Am. Chem. Soc.* **111**, 8525 (1989).
99. G. WAGNER, V. THANABAL and S. G. HYBERTS, manuscript in preparation.
100. S. W. FESIK and E. R. P. ZUIDERWEG, *J. Magn. Reson.* **78**, 588 (1988).
101. S. W. FESIK, R. T. GAMPE, E. R. P. ZUIDERWEG, W. E. KOHLBRENNER and D. WEIGL, *Biochem. Biophys. Res. Commun.* **159**, 842 (1989).
102. L. E. KAY, D. MARION and A. BAX, *J. Magn. Reson.* **84**, 72 (1989).
103. D. MARION, L. E. KAY, S. W. SPARKS, D. A. TORCHIA and A. BAX, *J. Am. Chem. Soc.* **111**, 1515 (1989).
104. E. L. ULRICH, J. L. MARKLEY and Y. KYOGOKU, *Protein Sequences Data Anal.* **2**, 23 (1989).
105. G. WAGNER, *Methods Enzymol.* **176**, 93 (1989).
106. G. WAGNER and D. BRÜHWILER, *Biochemistry* **25**, 5839 (1986).
107. J. GLUSHKA and D. COWBURN, *J. Am. Chem. Soc.* **109**, 7879 (1987).
108. H. KESSLER, P. SCHMIEDER and W. BERMEL, *Biopolymers*, in press.
109. A. BAX, R. H. GRIFFEY and B. L. HAWKINS, *J. Magn. Reson.* **55**, 301 (1983).
110. M. R. BENDALL, D. T. PEGG and D. M. DODDRELL, *J. Magn. Reson.* **52**, 81 (1983).
111. L. MÜLLER, *J. Am. Chem. Soc.* **101**, 4481 (1979).
112. D. A. VIDUSEK, M. F. ROBERTS and G. BODENHAUSEN, *J. Am. Chem. Soc.* **104**, 5452 (1982).
113. G. BODENHAUSEN and D. J. RUBEN, *Chem. Phys. Lett.* **69**, 185 (1980).
114. F. J. M. VAN DE VEN, C. A. G. HAASNOT and C. W. HILBERS, *J. Magn. Reson.* **61**, 181 (1985).
115. A. ABRAGAM, *Principles of Nuclear Magnetism*, Oxford University Press, London (1961).
116. I. SOLOMON, *Phys. Rev.* **99**, 559 (1955).
117. S. MACURA and R. R. ERNST, *Mol. Phys.* **41**, 95 (1980).
118. G. BODENHAUSEN and R. R. ERNST, *J. Am. Chem. Soc.* **104**, 1304 (1982).
119. ANIL KUMAR, G. WAGNER, R. R. ERNST and K. WÜTHRICH, *J. Am. Chem. Soc.* **103**, 3654 (1981).
120. J. W. KEEPERS and T. L. JAMES, *J. Magn. Reson.* **57**, 404 (1984).
121. B. A. BORGIAS and T. L. JAMES, *J. Magn. Reson.* **59**, 493 (1984).
122. S. MACURA, B. T. FARMER and L. R. BROWN, *J. Magn. Reson.* **70**, 493 (1986).
123. G. T. MONTELINE and G. WAGNER, to be published.
124. A. M. GRONENBORN, A. BAX, P. T. WINGFIELD and G. M. CLORE, *FEBS Lett.* **243**, 93 (1989).
125. K. WÜTHRICH, M. BILLETER and W. BRAUN, *J. Mol. Biol.* **180**, 715 (1984).
126. S. MACURA, K. WÜTHRICH and R. R. ERNST, *J. Magn. Reson.* **47**, 351 (1982).
127. S. MACURA, Y. HUANG, D. SUTER and R. R. ERNST, *J. Magn. Reson.* **43**, 259 (1981).
128. M. RANCE, G. BODENHAUSEN, G. WAGNER, K. WÜTHRICH and R. R. ERNST, *J. Magn. Reson.* **62**, 497 (1985).
129. A. KALK and H. J. C. BERENDSEN, *J. Magn. Reson.* **24**, 343 (1976).
130. E. T. OLEJNICZAK, F. M. POULSEN and C. M. DOBSON, *J. Magn. Reson.* **59**, 518 (1984).
131. S. G. HYBERTS and G. WAGNER, *J. Magn. Reson.* **81**, 418 (1989).
132. N. BEESON, S. G. HYBERTS, T. WEYMOUTH and G. WAGNER, to be submitted.
133. E. T. OLEJNICZAK, F. M. POULSEN and C. M. DOBSON, *J. Am. Chem. Soc.* **103**, 6576 (1981).
134. W. BRAUN, C. BÖSCH, L. R. BROWN, N. GO and K. WÜTHRICH, *Biochim. Biophys. Acta* **667** (1981).
135. D. NEUHAUS and M. WILLIAMSON, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, VCH Publishers, New York (1989).
136. A. PARDI, M. BILLETER and K. WÜTHRICH, *J. Mol. Biol.* **180**, 741 (1984).
137. D. NEUHAUS, G. WAGNER, M. VASAK, J. H. R. KÄGI and K. WÜTHRICH, *Eur. J. Biochem.* **151**, 257 (1985).
138. G. T. MONTELINE, M. E. WINKLER, P. RAUENBÜHLER and G. WAGNER, *J. Magn. Reson.* **82**, 198 (1989).
139. G. T. MONTELINE, D. HARE and G. WAGNER, to be submitted.
140. V. THANABAL and G. WAGNER, to be submitted.
141. C. GRIESINGER, O. W. SØRENSEN and R. R. ERNST, *J. Am. Chem. Soc.* **107**, 6394 (1985).
142. L. MÜLLER, *J. Magn. Reson.* **72**, 191 (1987).
143. E. R. P. ZUIDERWEG, R. BOELEN and R. KAPTEIN, *Biopolymers* **24**, 601 (1985).
144. S. G. HYBERTS, W. MÄRKI and G. WAGNER, *Eur. J. Biochem.* **164**, 625 (1987).
145. G. WAGNER, W. BRAUN, T. F. HAVEL, T. SCHAUMANN, N. GO and K. WÜTHRICH, *J. Mol. Biol.* **196**, 611 (1987).
146. P. C. DRISCOLL, A. M. GRONENBORN and G. M. CLORE, *FEBS Lett.* **243**, 223 (1989).
147. P. L. WEBER, R. MORRISON and D. HARE, *J. Mol. Biol.* **204**, 483 (1988).
148. A. ARSENIIEV, P. SCHULTZE, E. WÖRGÖTTER, W. BRAUN, G. WAGNER, M. VASAK, J. H. R. KÄGI and K. WÜTHRICH, *J. Mol. Biol.* **201**, 637 (1988).
149. P. GÜNTERT, W. BRAUN, M. BILLETER and K. WÜTHRICH, *J. Am. Chem. Soc.* **111**, 3997 (1989).
150. J. S. RICHARDSON, *Adv. Protein Chem.* **34**, 167 (1981).
151. B. L. SIBANDA and J. M. THORNTON, *Nature* **316**, 170 (1985).
152. A. D. KLINE, W. BRAUN, and K. WÜTHRICH, *J. Mol. Biol.* **204**, 675 (1988).
153. M. P. WILLIAMSON, T. F. HAVEL, and K. WÜTHRICH, *J. Mol. Biol.* **182**, 295 (1985).
154. R. KAPTEIN, E. R. P. ZUIDERWEG, R. M. SCHEEK, R. BOELEN and W. F. VAN GUNSTEREN, *J. Mol. Biol.* **182**, 179 (1985).
155. G. M. CLORE, A. M. GRONENBORN, A. T. BRÜNGER and M. KARPLUS, *J. Mol. Biol.* **186**, 435 (1985).

156. G. M. CLORE, M. NILGES, D. K. SUKUMARAN, A. T. BRÜNGER, M. KARPLUS and A. M. GRONENBORN, *EMBO J.* **5**, 2729 (1986).
157. S. KIRKPATRICK, C. D. GELATT and M. P. VECCHI, *Science* **220**, 671 (1983).
158. M. NILGES, G. M. CLORE and A. M. GRONENBORN, *FEBS Lett* **239**, 129 (1988).
159. M. NILGES, A. M. GRONENBORN, A. T. BRÜNGER and G. M. CLORE, *Protein Eng.* **2**, 27 (1988).
160. S. J. WEINER, P. A. KOLLMAN, D. A. CASE, C. SINGH, C. GHIO, G. ALAGONA, S. PROFETA and P. WEINER, *J. Am. Chem. Soc.* **106**, 765–784 (1984).
161. H. WIDMER, M. BILLETER and K. WÜTHRICH, *Proteins*, in press.
162. D. LEMASTER and F. M. RICHARDS, *Biochemistry* **27**, 142 (1988).
163. S. G. HYBERTS and G. WAGNER, *Biochemistry*, **29**, 1465 (1990).