

A brief introduction to NMR spectroscopy of proteins

**By Flemming M. Poulsen
2002**



Introduction

Nuclear magnetic resonance, NMR, and X-ray crystallography are the only two methods that can be applied to the study of three-dimensional molecular structures of proteins at atomic resolution. NMR spectroscopy is the only method that allows the determination of three-dimensional structures of proteins molecules in the solution phase.

In addition NMR spectroscopy is a very useful method for the study of kinetic reactions and properties of proteins at the atomic level.

In contrast to most other methods NMR spectroscopy studies chemical properties by studying individual nuclei. This is the power of the methods but sometimes also the weakness.

NMR spectroscopy can be applied to structure determination by routine NMR techniques for proteins in the size range between 5 and 25 kDa. For many proteins in this size range structure determination is relatively easy, however there are many examples of structure determinations of proteins, which have failed due to problems of aggregation and dynamics and reduced solubility.

It is the purpose of these notes to introduce the reader to descriptions and applications of the methods of NMR spectroscopy most commonly applied in scientific studies of biological macromolecules, in particular proteins.

The figures 1,2 and 11 are copied from “Multidimensional NMR in Liquids” by F.J.M de Ven (1995)Wiley-VCH

The Figure13 and Table 1 have been copied from “NMR of Proteins and Nucleic acids” K. Wüthrich (1986) Wiley Interscience

The Figures 20, 21 and 22 have been copied from J. H Prestegard, Nature Structural Biology 5, 517-522 1998

The principle of NMR spectroscopy

The atomic nuclei with **odd** mass numbers has the property spin, this means they **rotate** around a given axis. The nuclei with even numbers may or may not have this property. A **spin angular momentum vector** characterizes the spin. The nucleus with a spin is in other words a charged and spinning particle, which in essence is an electric current in a closed circuit, well known to **produce a magnetic field**. The magnetic field developed by the rotating nucleus is described by a nuclear magnetic moment vector, $\vec{\mu}$, which is proportional to the spin angular momentum vector.

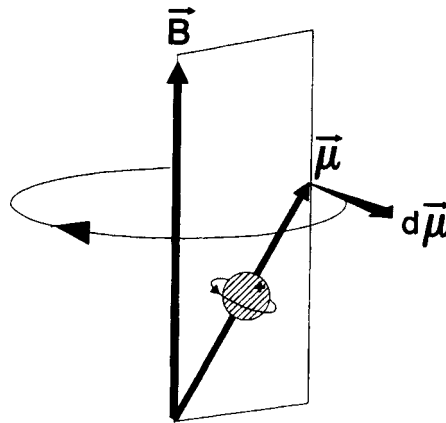


Figure 1, the spinning nucleus with a charge precessing in a magnetic field.

Larmor precession

When a nucleus with a nuclear magnetic moment is placed in an **external magnetic field B** , Figure 1, the magnetic field of the nuclei will not simply be oriented opposite to the orientation of the magnetic field. Because the nucleus is rotating, the **nuclear magnetic field** will instead **precess** around the axis of the external field vector. This is called **Larmor precession**. The frequency of this precession is a physical property of the nucleus with a spin, and it is proportional to the strength of the external magnetic field, the higher the external magnetic field the higher the frequency.

Making Larmor precession of a nuclear spin observable

The magnetic moment of a single nucleus cannot be observe. In a sample with many nuclei the magnetic moment of the individual nuclei will add up to one component. In a sample of nuclei in a magnetic field the component of the magnetic moment of the precessing nuclei around the external field axis will be a vector in the direction of the field axis. This nuclear magnetization is impossible to observe directly. In order to observe the nuclear magnetization we want to bring the nuclear magnetization perpendicular to the applied field. Applying a radio frequency pulse, which is perpendicular to the external magnetic field, can do this. If this pulse has the same frequency as the Larmor frequency of the nuclei to be observed and a well defined length the component of the nuclear magnetization can be directed from the direction of the magnetic field to a direction perpendicular to this.

After the pulse the nuclear magnetization vector will be rotating in the plane perpendicular to the magnetic field. The angular rate of this rotation is the Larmour frequency of the nuclei being observed. The rotating magnetic field will induce an electric current, which can be measured in a circuit, the receiver coil, which is placed around the sample in the magnet.

The free induction decay, FID

The nuclear magnetization perpendicular to magnetic field will decrease with time, partly because the nuclear magnetization gets out of phase and because the magnetization returns to the direction of the external field. This will be measured in the receiver coil as a fluctuating declining amplitude with time. This measures a frequency and a decay rate as a function of time. This is the direct result of the measurement and referred to as the free induction decay, FID, Figure 2.

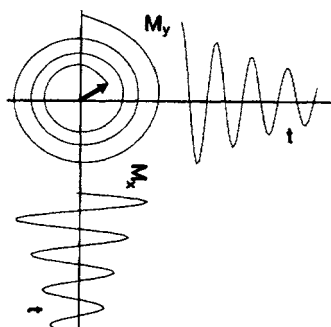


Figure 2. The free induction decay, FID, as measured as a function of time in the x - and y -directions perpendicular to magnetic field.

Fourier analysis

In order to get from the FID to an NMR spectrum the data in the FID are subjected to Fourier analysis. The FID is a function of time; the Fourier transformation converts this to a function of frequency. This is the way the NMR spectrum is normally displayed.

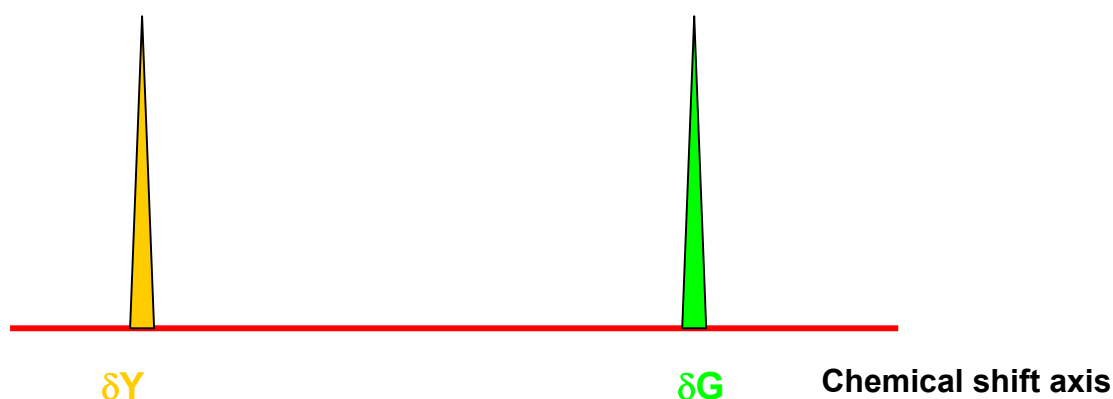


Figure 3. Schematic presentation of an NMR spectrum of a compound with two nuclei of different chemical shift

The origin of Chemical shift

When a molecule is placed in an external magnetic field, this will induce the molecular electrons to produce local currents. These currents will produce an alternative field, which opposes the external magnetic field. The total effective magnetic field that acts on the nuclear magnetic moment will therefore be reduced

depending on strength of the locally induced magnetic field. This is called a screening effect, a shielding effect, or more commonly known as the chemical shift. The shielding from the external magnetic field depends on the strength of the external magnetic field, on the chemical structure and the structural geometry of the molecule. The NMR active nuclei of the molecule, therefore, experience slightly different external magnetic fields, and for this reason the resonance conditions are slightly different. The precession of the nuclear magnetic moment vector of the individual nuclei will have different angular velocities. The FID of a molecule will therefore have several frequency components, and the Fourier transformation will produce an NMR spectrum with signals from each of the different types of nuclei in the molecule.

Measuring the relative chemical shift

The chemical shift for an NMR signal is normally measured in Hz shifted relative to a reference signal. In order to compare NMR spectra obtained at different external magnetic fields the difference in chemical shift is divided by the spectrometer frequency, which is the Larmour frequency of the observed nucleus type for the strength of the magnetic field of the spectrometer magnet. The chemical shift is normally a very small number in Hz, divided by a very large number the spectrometer frequency also in Hz. Therefore the resulting small number is multiplied by one million and given in the dimension less unit of parts per million.

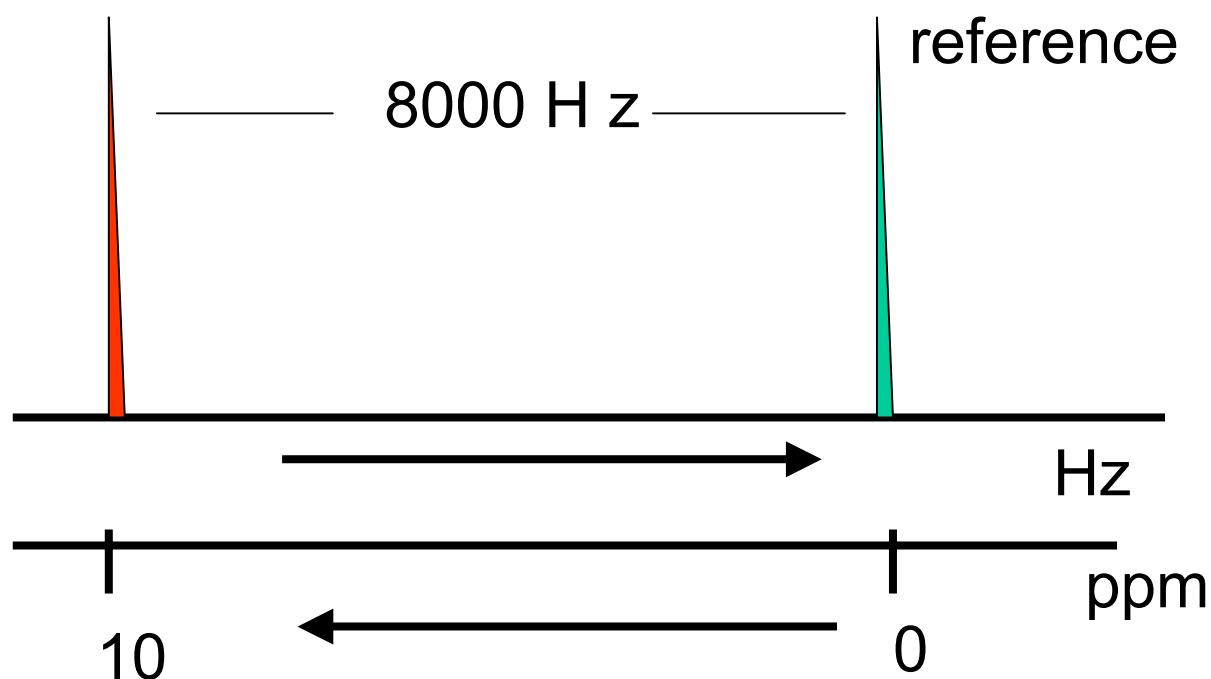


Figure 4. Schematic presentation of the chemical shift axis. The top shows an NMR spectrum with two signals. One signal (right) is from a chemical compound, which has been added to the sample for reference. The other signal is from a nucleus in the compound of interest. The two signals are 8000 Hz apart and the spectrometer frequency is 800 MHz. The two signals are $8000 \cdot 10^6 / 800 \cdot 10^6$ ppm = 10 ppm apart. If the reference is set to 0 ppm the signal of interest is at 10 ppm, bottom axis.

Increasing field and decreasing ppm values, figure 4

In contrast to the conventional axis, the ppm axis in NMR spectroscopy is a left oriented axis, counting increasing numbers towards left. Originally the axis was used to describe the effective magnetic field at the nuclei observed. In this respect the axis is a fully conventional right oriented axis, with increasing effective field towards right.

NMR of the 20 common amino acids

The ^1H (proton) NMR summary of the 20 common amino acids is shown in Table 1 as a listing of the chemical shifts for the hydrogen atoms of the residues in a random coil peptide.

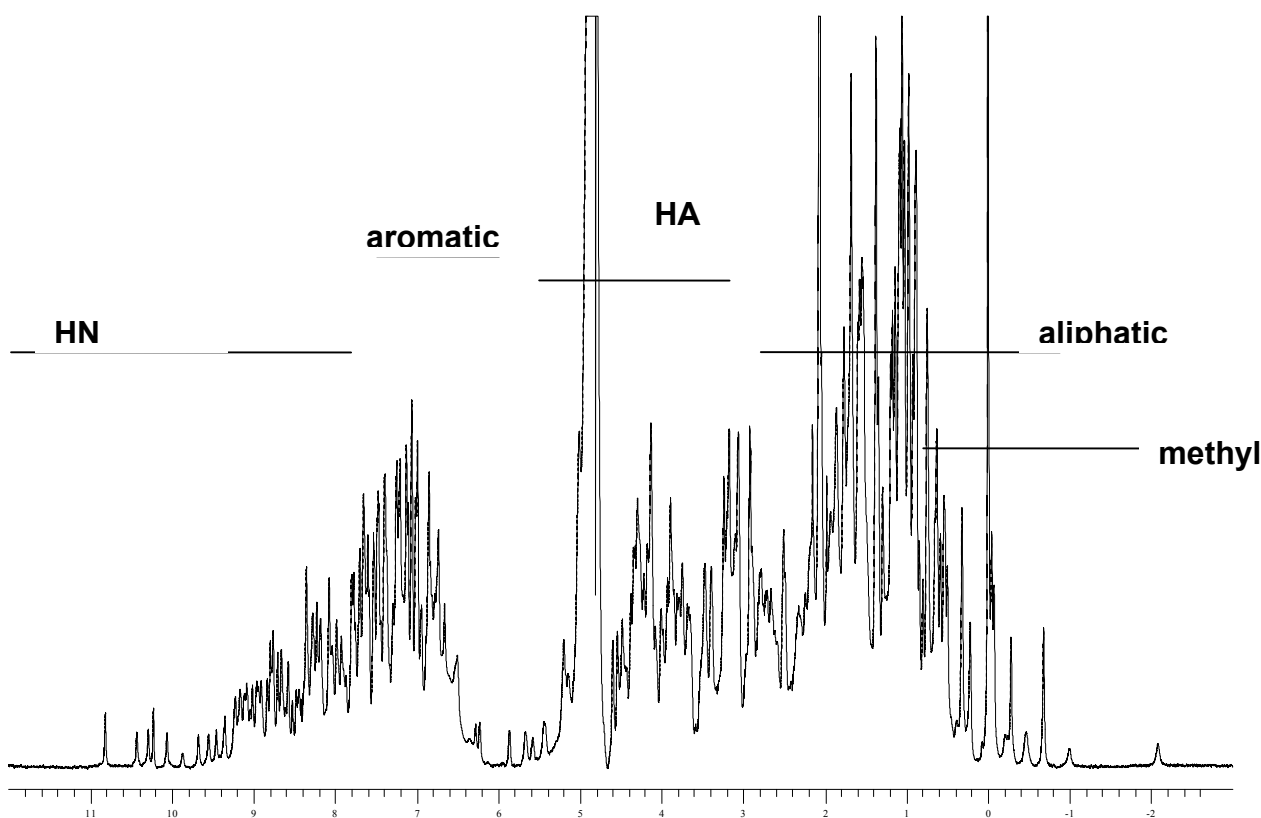


Figure 5. ^1H -NMR spectrum of hen egg white lysozyme.

The NMR spectrum of a folded protein

In a protein structure the individual residues are packed into chemical environments, which are very different from the random coil situation. This is seen in Figure 5 and 6 that shows the NMR spectrum of a folded protein and a comparison with the NMR spectrum of the same protein recorded at conditions where the protein is unfolded. The spectrum of the unfolded protein corresponds to a spectrum, which in essence is the sum of the random coil spectra of the amino acid residues in the proteins as given in Table 1. The dispersion of signals in the spectrum of the folded protein is far beyond the envelope of signals seen in the spectrum of the unfolded protein. This clearly reflects that nuclei in the folded form are subject to a many different types of microenvironments of chemical screens.

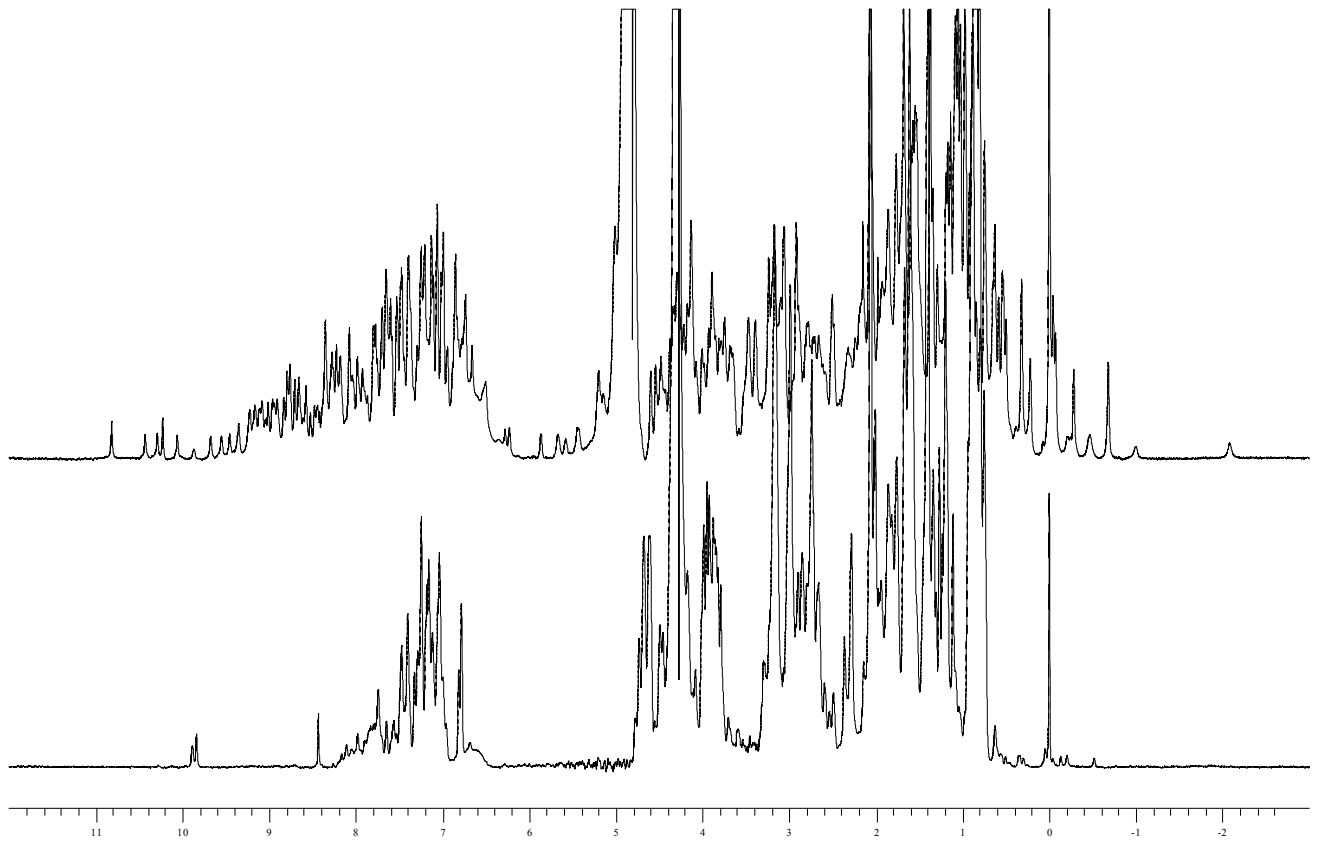


Figure 6. Comparison of NMR spectra of folded (top) and unfolded (bottom) protein

Small and large deviation from random coil shifts for the amino acid residues in proteins

For NMR spectra of proteins the NMR signals of the nuclei of the individual residues are in most cases seen in the vicinity of the random coil shift value. However, due to structural diversity in a protein structure, the signals are often shifted significantly away from the random coil value. In some cases, a proton may be in a chemical environment of a protein structure, where the chemical shift of the external magnetic field is so strong that the corresponding signals are observed several ppm away from the random coil values. The signal observed at -2 ppm in the spectrum of lysozyme is from one of the gamma methylene protons of isoleucine 98. This is shifted more than 3 ppm upfield from the random coil position at 1.48 ppm (Table 1) due to the proximity of a tryptophan residue.

Determination of secondary structure from chemical shift analysis

The chemical shift can be used in structure determination of proteins. In particular the analysis of the chemical shift of $^1\text{H}^\alpha$, and the ^{13}CO , $^{13}\text{C}^\alpha$, and $^{13}\text{C}^\beta$ of the peptide backbone can be used to determine the secondary structure type of a given peptide segment, figure 8.

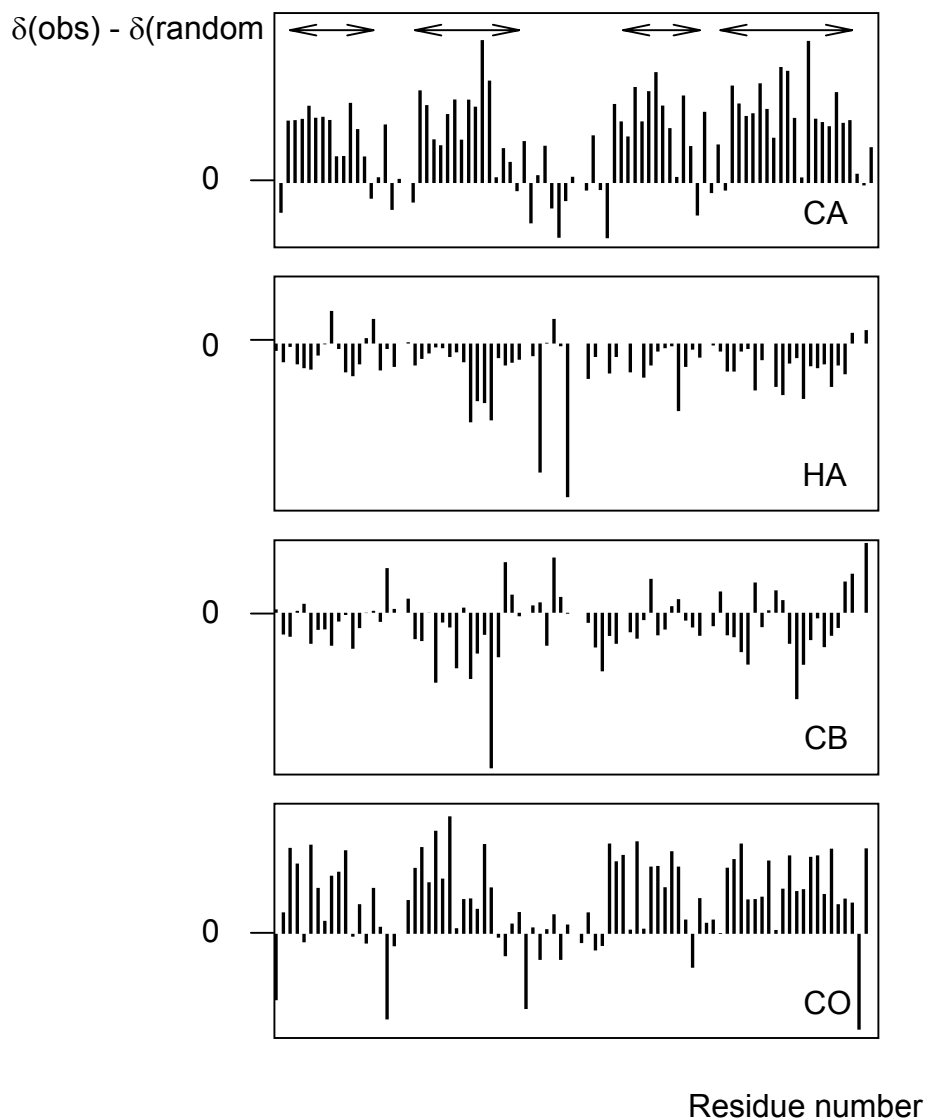


Figure 7. Chemical shift analysis of the peptide backbone NMR signals. The protein studied has four α -helices as marked by the arrows in the top panel. Each panel represent a chemical shift analysis of the individual residues comparing the observed shift with the shift observed for the same residue type in random coil model peptides.

Residue	NH	α H	β H	Others
Gly	8.39	3.97		
Ala	8.25	4.35	1.39	
Val	8.44	4.18	2.13	γ CH ₃ 0.97, 0.94
Ile	8.19	4.23	1.90	γ CH ₂ 1.48, 1.19 γ CH ₃ 0.95 δ CH ₃ 0.89
Leu	8.42	4.38	1.65, 1.65	γ H 1.64 δ CH ₃ 0.94, 0.90
Pro ^b		4.44	2.28, 2.02	γ CH ₂ 2.03, 2.03 δ CH ₂ 3.68, 3.65
Ser	8.38	4.50	3.88, 3.88	
Thr	8.24	4.35	4.22	γ CH ₃ 1.23
Asp	8.41	4.76	2.84, 2.75	
Glu	8.37	4.29	2.09, 1.97	γ CH ₂ 2.31, 2.28
Lys	8.41	4.36	1.85, 1.76	γ CH ₂ 1.45, 1.45 δ CH ₂ 1.70, 1.70 ϵ CH ₂ 3.02, 3.02 ϵ NH ₃ ⁺ 7.52
Arg	8.27	4.38	1.89, 1.79	γ CH ₂ 1.70, 1.70 δ CH ₂ 3.32, 3.32 NH 7.17, 6.62
Asn	8.75	4.75	2.83, 2.75	γ NH ₂ 7.59, 6.91
Gln	8.41	4.37	2.13, 2.01	γ CH ₂ 2.38, 2.38 δ NH ₂ 6.87, 7.59
Met	8.42	4.52	2.15, 2.01	γ CH ₂ 2.64, 2.64 ϵ CH ₃ 2.13
Cys	8.31	4.69	3.28, 2.96	
Trp	8.09	4.70	3.32, 3.19	2H 7.24 4H 7.65 5H 7.17 6H 7.24 7H 7.50 NH 10.22
Phe	8.23	4.66	3.22, 2.99	2,6H 7.30 3,5H 7.39 4H 7.34
Tyr	8.18	4.60	3.13, 2.92	2,6H 7.15 3,5H 6.86
His	8.41	4.63	3.26, 3.20	2H 8.12 4H 7.14

^a Data for the nonterminal residues X in tetrapeptides GGXA, pH 7.0, 35°C [from Bundi and Wüthrich (1979a), except that more precise data were obtained for Leu, Pro, Lys, Arg, Met, and Phe using new measurements at 500 MHz].

^b Data for *trans*-Pro.

Table 1, Random chemical shift of the 20 common amino acid residues in proteins

Multidimensional NMR spectroscopy

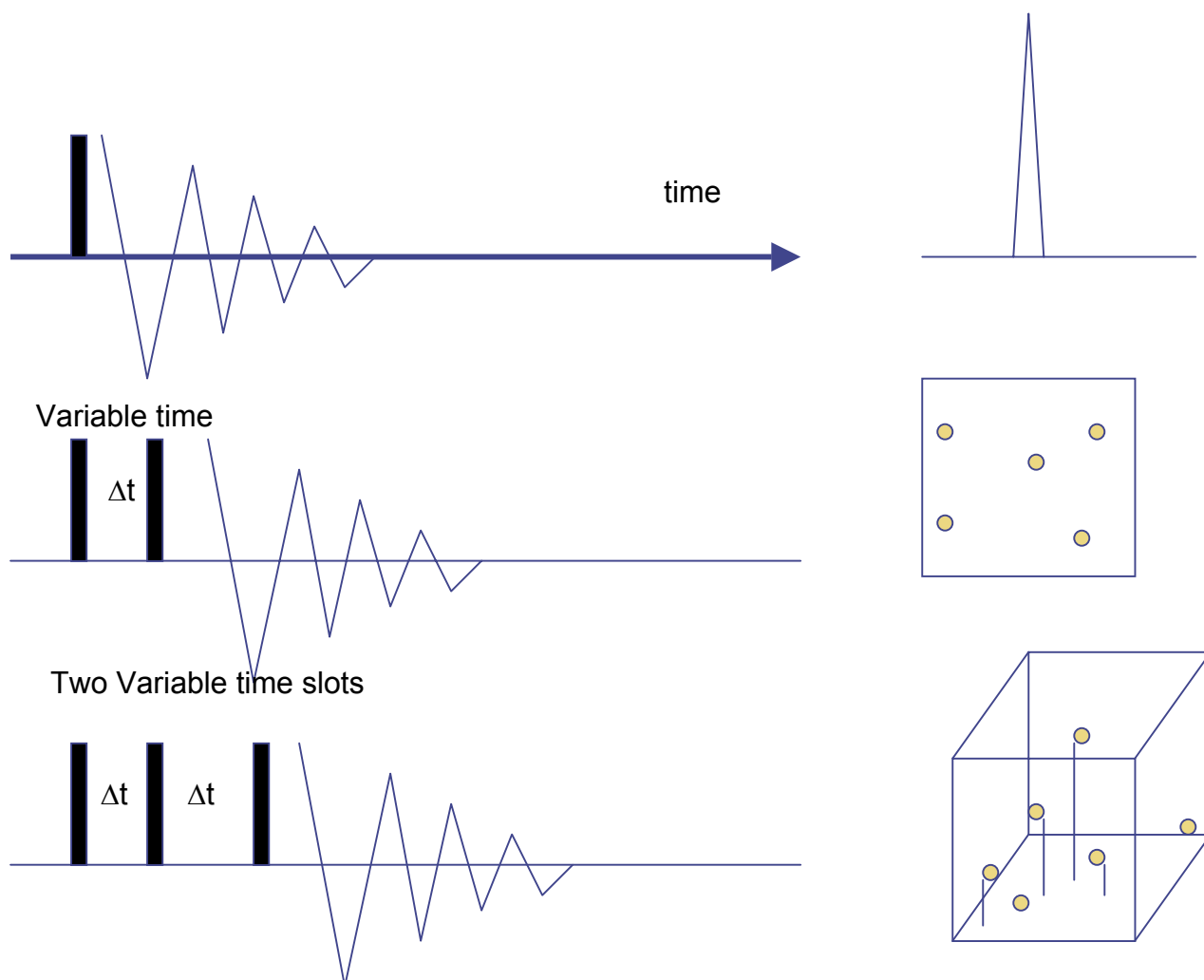


Figure 8. Schematic presentation of NMR spectra in one-dimension, top, two-dimensions, middle, and three-dimensions.

The fundamental NMR spectrum is a one-dimensional frequency spectrum, figure 8. The FID of one-dimensional NMR spectrum represents only one time dimension. In two-dimensional NMR spectroscopy the second dimension might be another time domain, which can be established by introducing an additional pulse and systematically increasing the time between the two pulses, And equivalently for the three-dimensional NMR spectrum. This can be combined with the observation of one or two more nuclei for instance ^{13}C and ^{15}N .

J coupling origin and application in structure determination.

J coupling is also referred to as spin-spin coupling and scalar coupling. Two nuclei in a molecule, which are connected by one, two, and three bonds, can be seen to be coupled in the NMR spectrum. The coupling is observed by a splitting of the NMR signal. The origin of this coupling is a process in which the two nuclei perturb the respective valence electrons of the molecule. Here the electron spin and the nuclear spin at one atom are being aligned anti-parallel to each other and as a response to this the electrons of the coupling nucleus orient either parallel or anti-parallel to that of A and vice versa. For two nuclei which can each exist in two states there will be four different types of interactions, which is reflected in the A and B signals are both splitting into two signals.

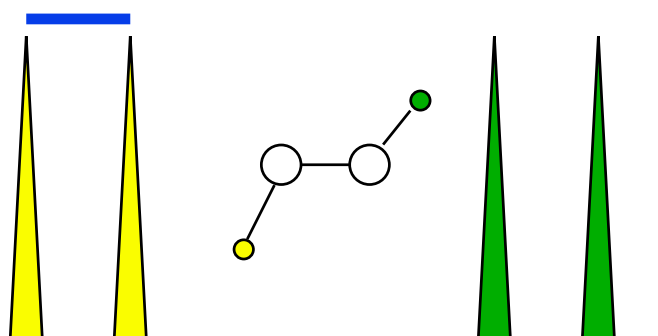


Figure 9. J-coupling of two nuclei separated by three bonds of compound of the type Y-C-C-G. The blue bar shows the size of the coupling.

The coupling constant. J

The separation between the two components of the split signal is the coupling, figure 9. This is normally measured in Hz and referred to as the coupling constant. The coupling constant is independent of the magnetic field. The size of the coupling depends on many structural properties. For structure determination we are mainly concerned with the relationship between the coupling constant and dihedral angles.

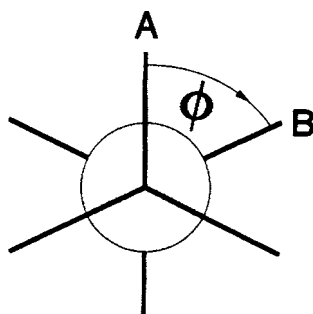


Figure 10. The definition of a dihedral angle ϕ as seen in Newman projection

J-couplings and determination of dihedral angles

The three-bond coupling constant depends on the dihedral angle defined by rotation around the middle bond in the coupling system. The definition of the dihedral angle is shown in the Newman projection. The J-coupling may also be used to distinguish between trans and gauche conformations. One particularly important application of

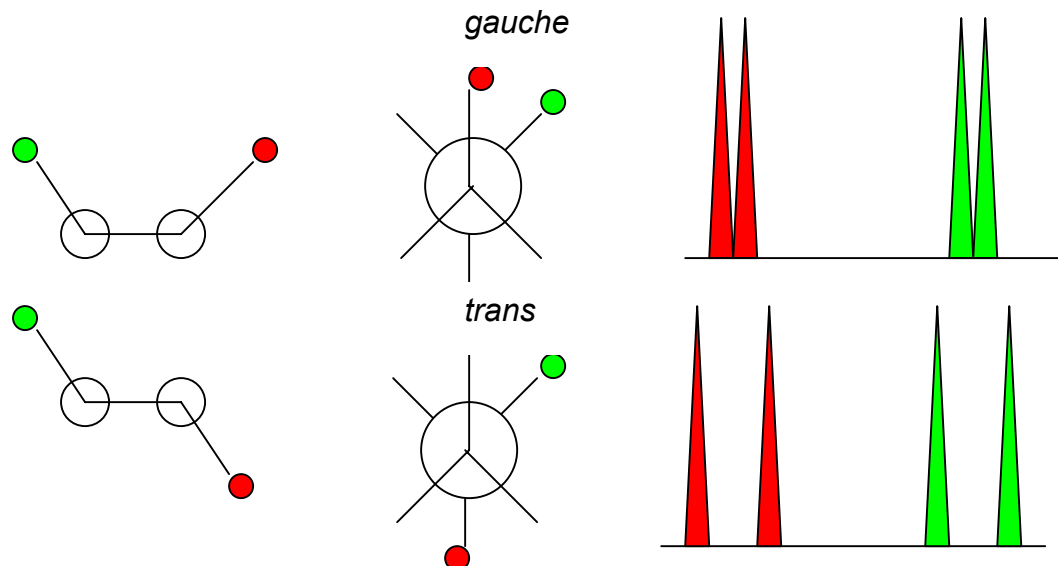


Figure 11. The coupling constant can distinguish between the gauche and trans conformation

the coupling constant is as a measure of the coupling between the H^α and the H^N in the peptide backbone. This coupling depends on the ϕ - angle in the peptide bond. The coupling may also be measured by the coupling constant between the H^N and C^β . The Karplus curve shown in figure 12 shows the correlation between the coupling between H^α and the H^N and the ϕ - angle. It is seen that coupling constants are around 4 Hz for peptide segments in α helices where the ϕ - angle is around -60° , and it is between 8 and 12 Hz for peptide segments in β -structures, where the ϕ - angle is in the -120° range.

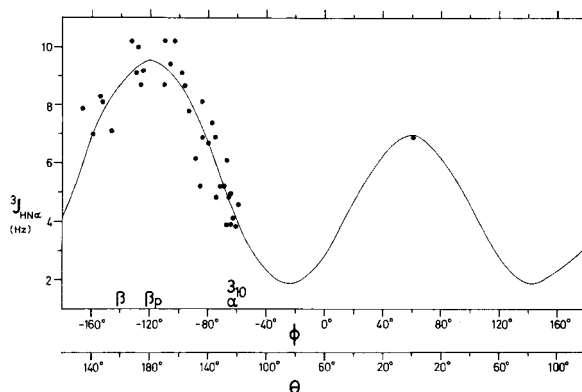


Figure 12. The Karplus function shows the correlation between the ϕ -angle in the peptide bond and coupling constant between H^α and H^N .

The use of J-coupling for identification of the amino acid residues in the NMR spectra of proteins

The one dimensional ^1H NMR spectrum of Valine

Most of the methods used in NMR spectroscopy are designed to measure the presence of coupling between nuclear spins. If we consider the amino acid residue of valine in a peptide this will have NMR signals from the H^{N} , the H^{α} , the H^{β} , and the two triple intensity signals of $\text{H}^{\gamma 1}$ and $\text{H}^{\gamma 2}$. The schematic ^1H NMR spectrum of a valine residue is shown in figure 13, where $\delta\text{H}^{\text{N}}$ is the chemical shift of the H^{N} signal etcetera. This one-dimensional spectrum shows the individual signals of the ^1H nuclei, however, it does not provide information about the spin-spin coupling partners of the five corresponding nuclei. This information can be obtained by two-dimensional correlation NMR spectroscopy, COSY.

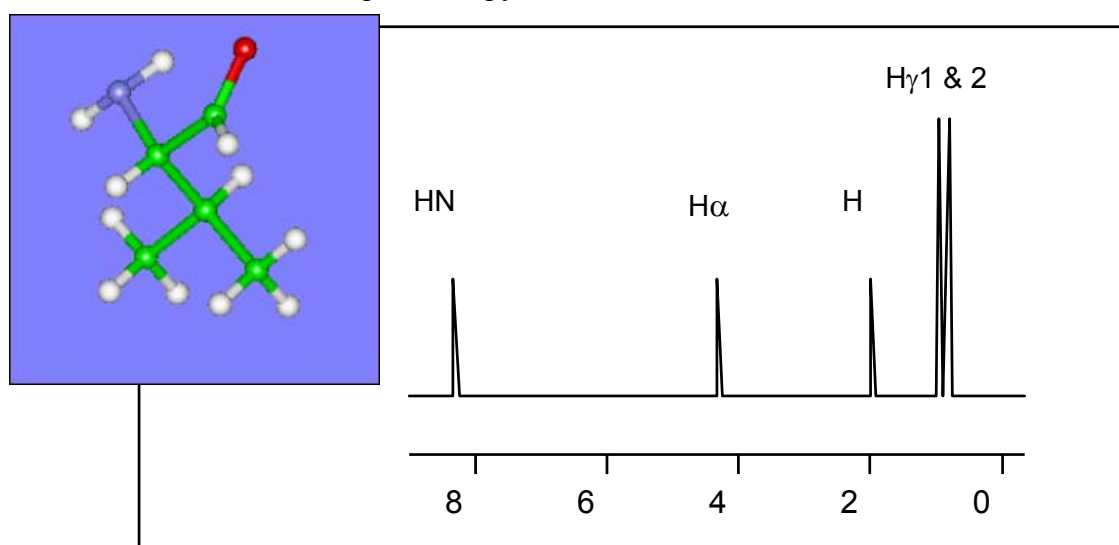


Figure 13. Schematic presentation of the NMR spectrum of Valine

The COSY spectrum of Valine

From the covalent structure of the valine residue it is seen that the H^{N} spin couples to the H^{α} spin, which couples to the H^{β} spin which couples to the two sets of H^{γ} spins. The two-dimensional COSY spectrum, figure 14, is recorded so that the spectrum contains two types of signals, diagonal peaks and off-diagonal peaks often called cross peaks. The diagonal peaks represent the signals from each of the ^1H types in valine (H^{N} , H^{α} , H^{β} , and the two H^{γ}) and their position in spectrum are at $(\delta\text{H}^{\text{N}}, \delta\text{H}^{\text{N}})$, $(\delta\text{H}^{\alpha}, \delta\text{H}^{\alpha})$, $(\delta\text{H}^{\beta}, \delta\text{H}^{\beta})$, $(\delta\text{H}^{\gamma 1}, \delta\text{H}^{\gamma 1})$ and $(\delta\text{H}^{\gamma 2}, \delta\text{H}^{\gamma 2})$. The cross peaks are, however, the most interesting, they report the couplings between pairs of nuclei. The coupling between H^{N} and H^{α} , show at the two symmetry related positions $(\delta\text{H}^{\text{N}}, \delta\text{H}^{\alpha})$ and $(\delta\text{H}^{\alpha}, \delta\text{H}^{\text{N}})$, figure 15. Similarly the coupling between H^{α} and H^{β} are at $(\delta\text{H}^{\alpha}, \delta\text{H}^{\beta})$ and $(\delta\text{H}^{\beta}, \delta\text{H}^{\alpha})$, and the couplings between H^{β} , and the $\text{H}^{\gamma 1}$ and $\text{H}^{\gamma 2}$, respectively, at $(\delta\text{H}^{\beta}, \delta\text{H}^{\gamma 1})$ and $(\delta\text{H}^{\gamma 1}, \delta\text{H}^{\beta})$, and $(\delta\text{H}^{\beta}, \delta\text{H}^{\gamma 2})$ and $(\delta\text{H}^{\gamma 2}, \delta\text{H}^{\beta})$.

The TOCSY spectrum of Valine

By recording the correlation NMR spectrum in a way so that all the spins in a spin system of an amino acid are all correlated results in a more complicated spectrum.

However this type of spectrum becomes very useful in particular in the heteronuclear NMR experiments a comparison of the TOCSY and COSY spectrum can be seen in figure 15.

COSY and TOCSY spectra of the amino acid residues in proteins

Many of the types of residues of the amino acids have characteristic patterns of couplings in combinations with characteristic chemical shifts, which make these unique for identification. Other residue types are quite similar and not so easy to identify just on the basis of their coupling patterns and chemical shifts. The COSY and TOCSY pattern for each of the individual amino acid residues can easily be drawn schematically from the information in table 1.

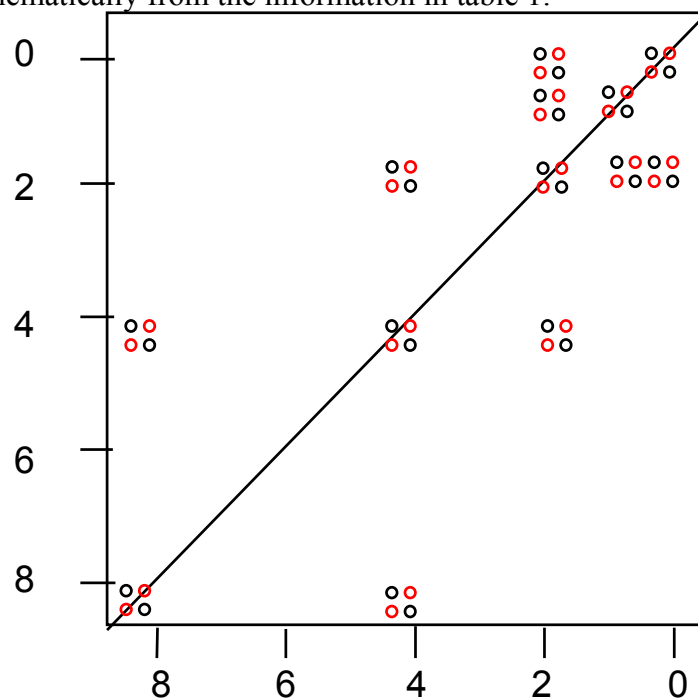
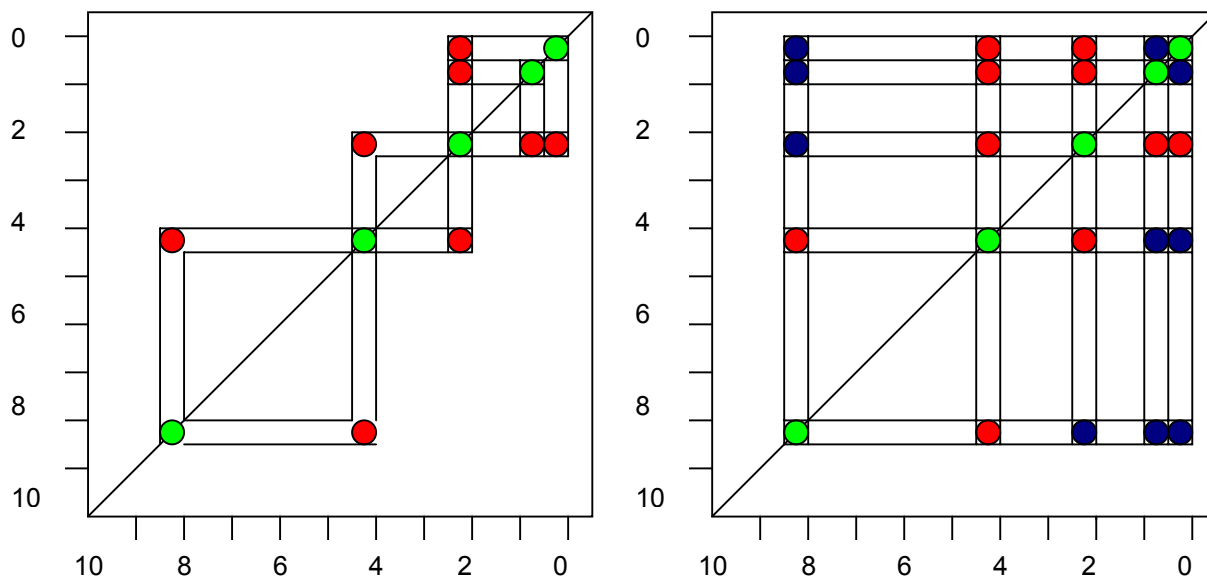


Figure 14. Schematic presentation of a COSY spectrum of the amino acid residue valine. The pattern of four circles represents on peak. The peaks on the diagonal are the diagonal peaks. The off-diagonal peaks are the cross peaks. Try to assign the spectrum.

Assigning the amino acid residue spin system by COSY and TOCSY

The assignment of the ^1H NMR spectrum of a protein will normally use a starting point in the coupling between H^{N} and H^{α} of a residue often appearing as well resolved signal in the $(\delta\text{H}^{\text{N}}, \delta\text{H}^{\alpha})$ and $(\delta\text{H}^{\alpha}, \delta\text{H}^{\text{N}})$ region of the ^1H NMR spectrum. The assignment may then continue to identify the remaining signals of the residue. On the basis of the number of coupling spins and their chemical shifts it is possible to assign the systems of spins to an amino acid type or to a group of similar amino acid types, which have similar number of spins with similar chemical shifts.

Figure 15. Schematic presentation of COSY (left) and TOCSY (right) spectra of a valine residue. Green circles are diagonal peaks, red circles are two. And three-bond couplings seen only in the COSY spectrum, blue circles are TOCSY peaks representing long range coupling in the spin system



Introducing NMR active stable isotopes ^{13}C and ^{15}N

It is often a great advantage for the analysis of a protein by NMR to introduce the NMR active stable isotopes ^{13}C and ^{15}N . With the introduction of these two NMR active nuclei the spins in a protein are almost all being connected by one-bond couplings, and this facilitates the study tremendously. The preparation of proteins enriched with the two nuclei are accomplished by heterologous expression of the protein in micro-organisms grown in minimal growth medium where the carbon source is fully ^{13}C labelled and the nitrogen source is fully ^{15}N labelled.

The ^1H - ^{15}N coupling in the peptide bond is the starting point for the heteronuclear NMR analysis of proteins

The one-bond coupling ^1H - ^{15}N is the most important starting point for the heteronuclear NMR analysis of proteins. This bond is present in every amino acid residue in a protein except the N-terminal and the proline residues. The correlation spectroscopy method used to record this coupling is called a ^1H - ^{15}N HSQC spectrum (heteronuclear single quantum correlation). An example of this spectrum is shown in figure 16. In this spectrum there is a ^1H -chemical shift axis and a ^{15}N chemical shift axis. The signals report the coupling between the H^{N} and N and the signal appears at $(\delta\text{H}^{\text{N}}, \delta\text{N})$.

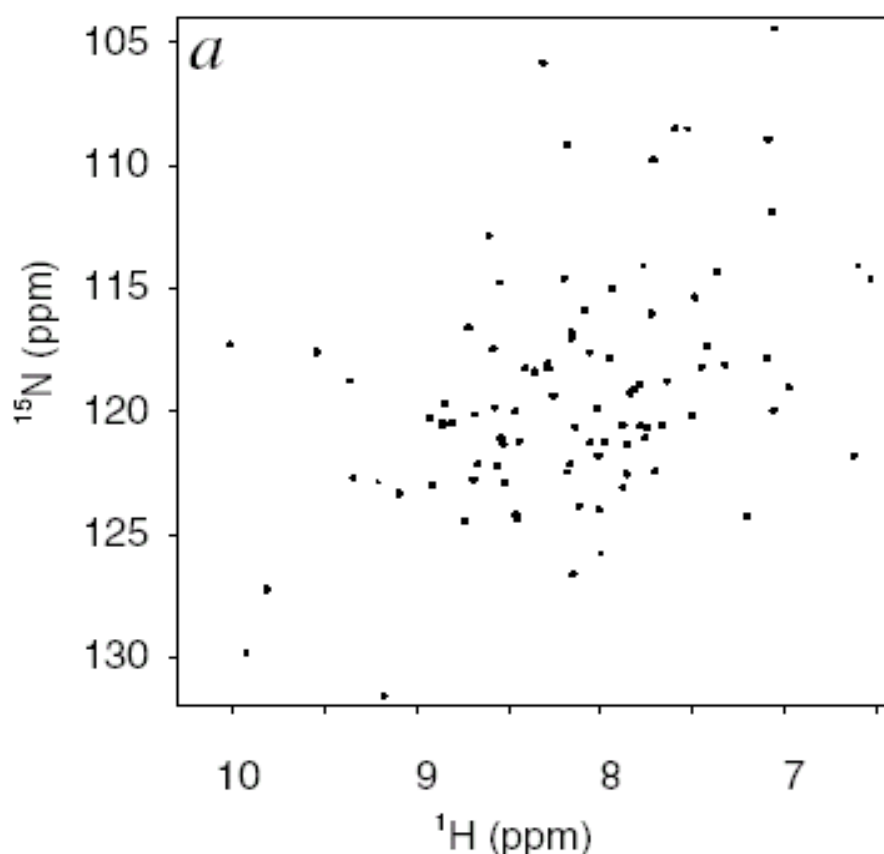


Figure 16. ^1H - ^{15}N HSQC spectrum of the ^{15}N labelled protein ACBP. Each “spot” is an NMR signal representing the ^1H - ^{15}N coupling from one of the amino acid residues in the protein.

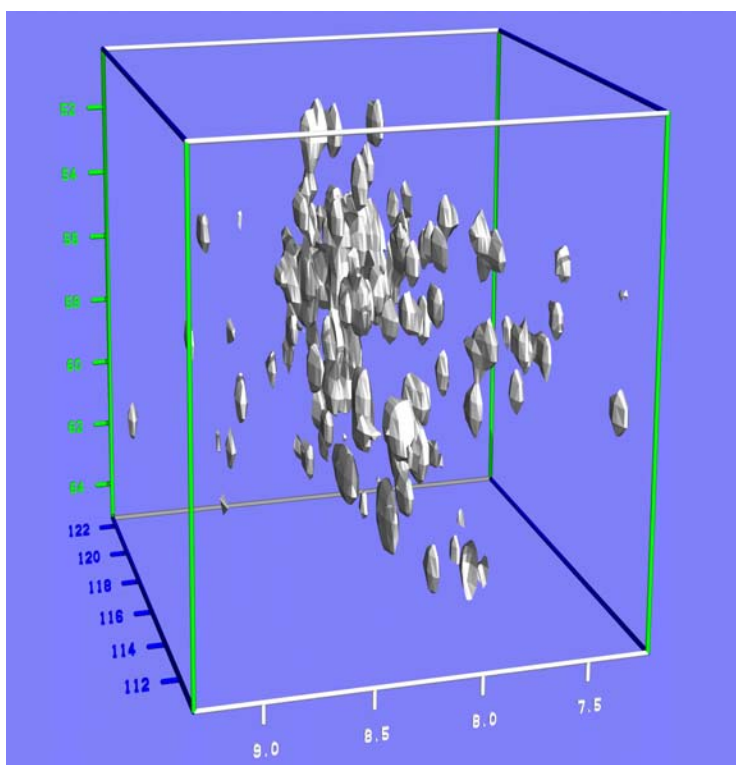


Figure 17. Triple-resonance heteronuclear NMR spectrum of a protein
Each signal represents the coupling between H^N and N with C^α in one residue. The projection of the signals into the $^1H^N$ and ^{15}N plane will show a spectrum as shown in figure 15.

With three NMR active nuclei fully incorporated in a protein experiments in three dimensions are possible, one for each type of isotope. This is a heteronuclear triple-resonance experiment. The list of experiments that can be used to identify couplings between the three nuclei in amino acid residues are many, and it is beyond the scope of this presentation to mention all these experiments here. Above in figure 17 is shown a triple resonance heteronuclear correlation spectrum, which correlates the coupling between H^N and N with C^α in one residue. The three-dimensional spectrum has three chemical shift axes a 1H -axis, a ^{13}C -axis and a ^{15}N -axis, and the signals appear at $(\delta H^N, \delta N, \delta C^\alpha)$.

Sequential assignment using heteronuclear scalar coupling.

Sequential assignment is a process by which a particular amino acid spin system identified in the spectrum is assigned to a particular residue in the amino acid sequence. In a protein, which is ^{13}C - and ^{15}N labelled almost the entire protein is one continuous spin system. In particular the peptide backbone is one long series of scalar coupled nuclei. Specific types of heteronuclear correlation spectroscopy can record these individual types of coupling. There are several principles for sequential assignments using heteronuclear scalar coupling. One principle is illustrated in figure 18. It is based on the recording of two different heteronuclear correlation spectra. One 1H -, ^{13}C -, and ^{15}N -heteronuclear three-dimensional NMR spectrum, which records the one bond coupling between $^1H^N$ and ^{15}N and the one and two bond coupling between ^{15}N and $^{13}C^\alpha$ and $^{13}C^\beta$ in one residue. The spectrum is called a HNCACB spectrum. This type of experiment also records the coupling across $^{13}C'$ to the $^{13}C^\alpha$ and $^{13}C^\beta$ in the preceding residue. Another similar experiment measures specifically the heteronuclear coupling between $^1H^N$ and ^{15}N in one residue and the coupling across $^{13}C'$ to the $^{13}C^\alpha$ and $^{13}C^\beta$ in the preceding residue. This spectrum is called a CBCA(CO)NH spectrum. In a combined analysis of these two types of spectra it is

possible from each individual $^1\text{H}^{\text{N}} - ^{15}\text{N}$ peak in the $^1\text{H}^{\text{N}} - ^{15}\text{N}$ correlation spectrum, Figure 16 to identify the $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ in the same residue and the preceding residue. The principle in the method is demonstrated in figure 18 and 19.

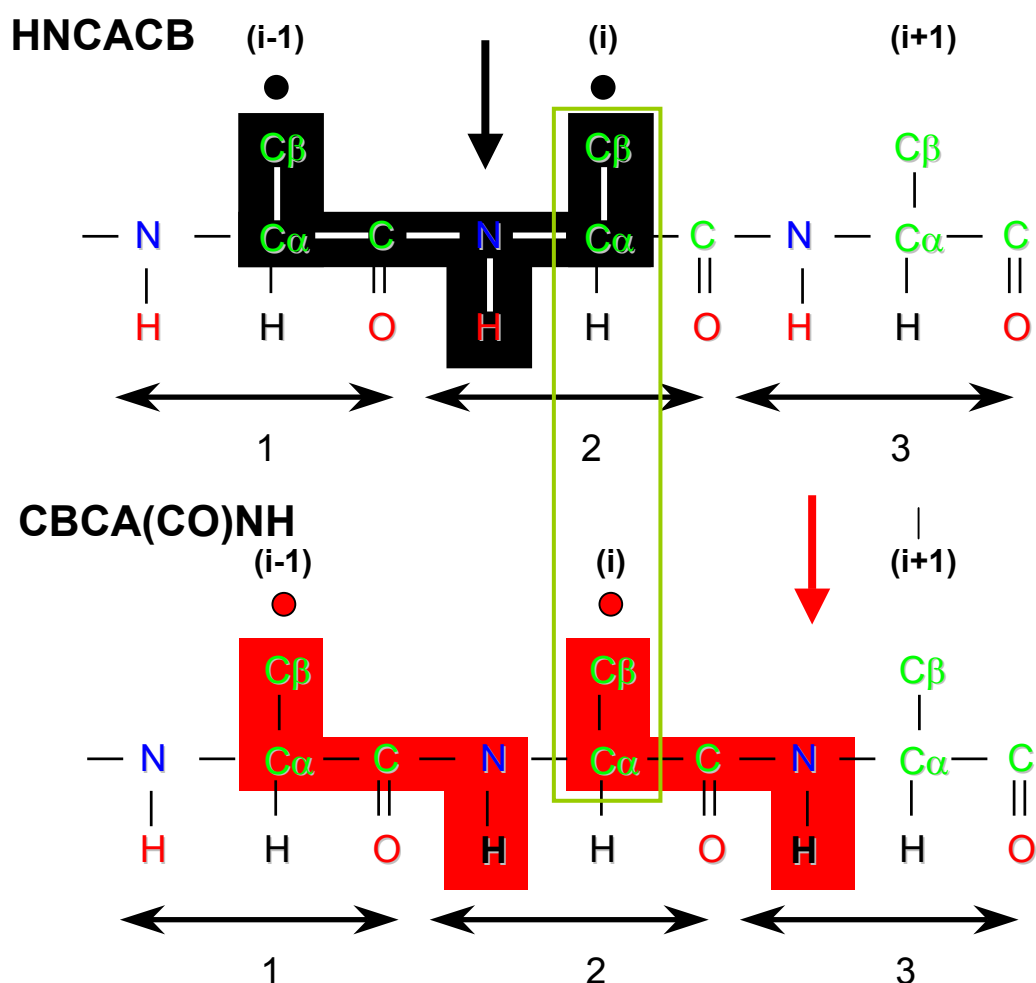


Figure 18. The two panels show the scalar coupling correlation, which is measured by the HNCACB (top) and by the CBCA(CO)NH (bottom).

In the HNCACB (top) the coupling is mediated through the chemical bonds shown on a black background. The $^1\text{H}^{\text{N}} - ^{15}\text{N}$ coupling pair of residue (i) is correlated to the $^{13}\text{C}^{\alpha} - ^{13}\text{C}^{\beta}$ pair of residue (i) and (i-1).

In the CBCA(CO)NH (bottom) the coupling is mediated through the bonds shown on a red background. Here the $^1\text{H}^{\text{N}} - ^{15}\text{N}$ coupling pair of residue (i) is correlated to the $^{13}\text{C}^{\alpha} - ^{13}\text{C}^{\beta}$ in residue (i-1).

If the same $^{13}\text{C}^{\alpha} - ^{13}\text{C}^{\beta}$ pair, as shown in the open green frame, are seen to couple to two different pairs of $^1\text{H}^{\text{N}} - ^{15}\text{N}$ couplings as indicated by the black and red arrows in the two panels, they may be assigned as signals from neighbouring residues.

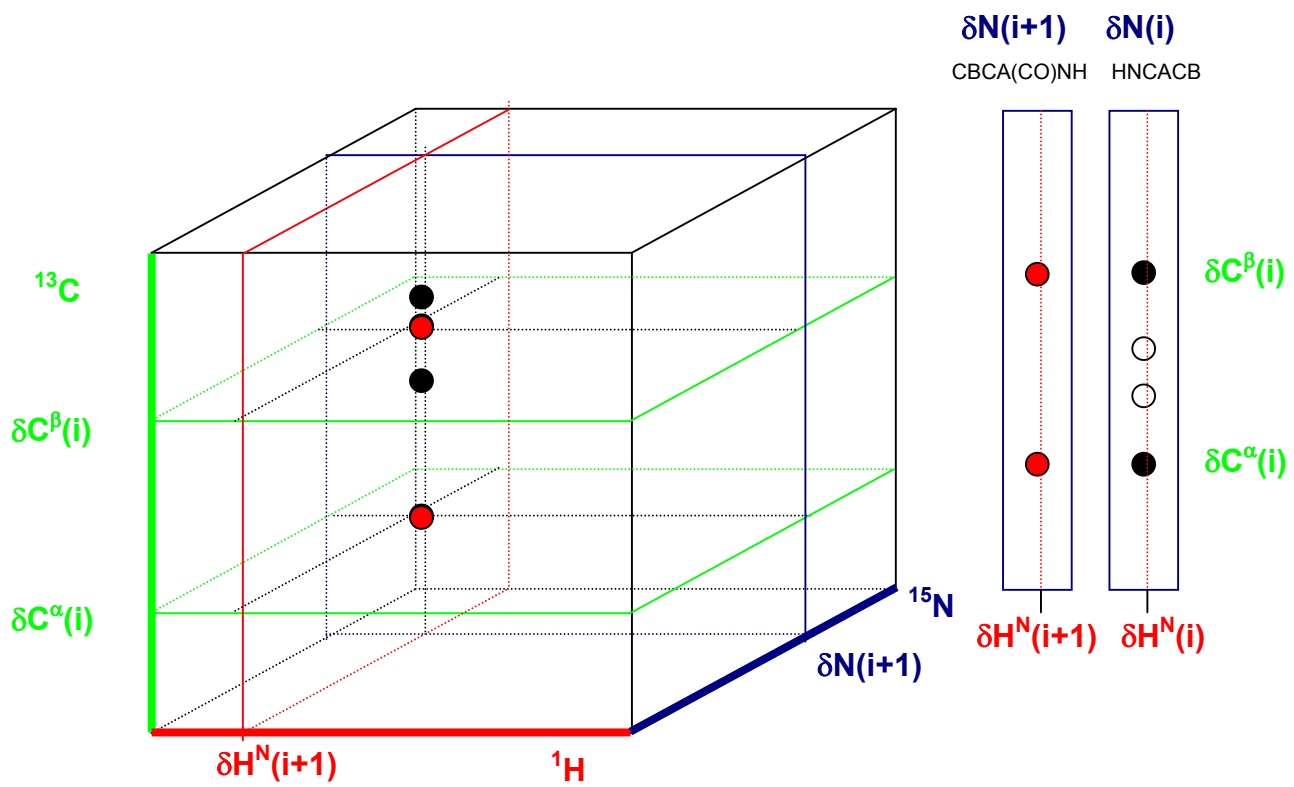
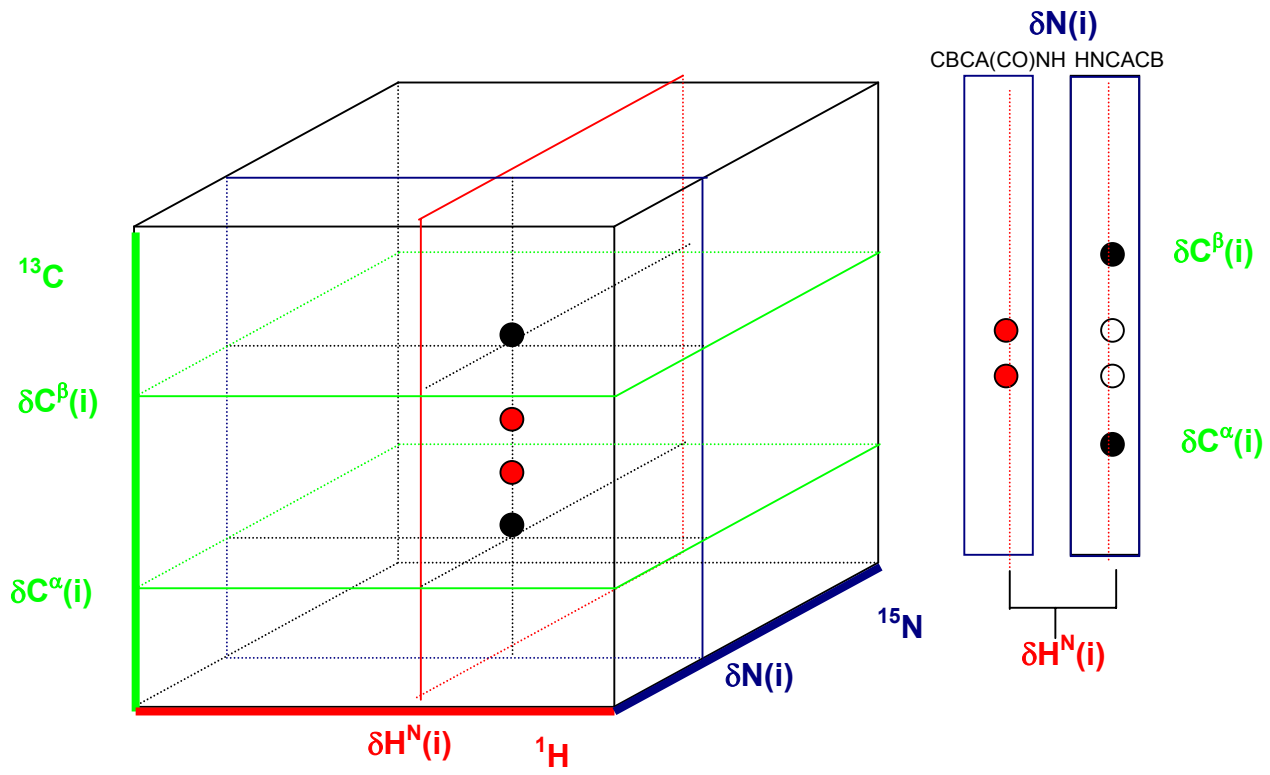


Figure 19. Schematic presentation of the combined spectrum analysis of the three-dimensional HNCACB and a CBCA(CO)NH spectra. The ^{15}N axis and the frames in the ^{15}N dimension are coloured blue. The ^1H axis and the frames in the ^1H dimension are coloured red. The ^{13}C axis and the frames in the ^{13}C dimension are green. Two planes $\delta\text{N}(i)$, top-left, and $\delta\text{N}(i+1)$, bottom-left, are highlighted.

The cross peaks in HNCACB are ● and in CBCA(CO)NH are ●

To the right are shown the segments of the planes from the two types of spectra. In the top-right the HNCACB and the CBCA(CO)NH from the same plane, $\delta\text{N}(i)$, of the two spectra are compared.

In the bottom-right $\delta\text{N}(i)$ plane from the HNCACB is compared to the $\delta\text{N}(i+1)$ planes of the CBCA(CO)NH.

The $\delta\text{N}(i)$ plane of the CBCA(CO)NH spectra has two signals at the $\delta\text{H}(i)$. These are the red cross peaks and origins from C^β and C^α of residue $(i-1)$.

The $\delta\text{N}(i)$ plane of the HNCACB has four cross peaks (black) at the $\delta\text{H}(i)$. Two of these superimpose with the red cross peaks of the CBCA(CO)NH are from the preceding residue $(i-1)$. The observation establishes a sequential assignment.

The $\delta\text{N}(i)$ plane of the HNCACB and the $\delta\text{N}(i+1)$ plane of the CBCA(CO)NH have one pair of C^β and C^α cross peaks in common. The observation establishes a sequential assignment.

As described in figure 18 and 19 the sequential assignment brings groups of spin systems together in a sequence. This sequence can be imbedded in the amino acid sequence of the protein by a chemical shift analysis of the spin systems. Several of the residue types have typical chemical shifts and this can be used for a residue assignment. For instance a series of sequentially assigned spin systems the chemical shift analysis identifies

GXX(T/S)XXXAXXGXX

which can be imbedded between residue 132 and 144 in the amino acid sequence

132 GNRSKDVAIVGLL 144

leading to the assignment of the intervening residues, which were not assigned directly by the chemical shift analysis.

Dipole-dipole interactions

A nuclear spin is a magnetic dipole with a magnetic field. In a hypothetical situation in the presence of an external field the total magnetic field at the spin will be the sum of the local field and the external field. If two identical nuclear spins are close to each other in an external magnetic field, the two nuclear spins will be influenced by the external field and by the magnetic field of the other nuclear spin. The local field may be either aligned or parallel to the external field, and this will give rise to two signals. The separation between the two signals depends on the distance between the two spins and the angle between directions of the vector connecting the two spins and the vector representing the external magnetic field, figure 20.

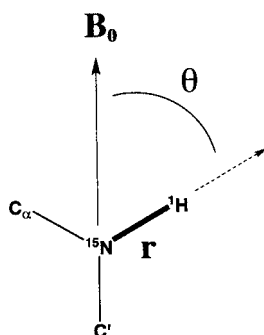


Figure 20. The dipole-dipole coupling between the ^{15}N and the ^1H of the NH one bond vector of the peptide backbone depends on the length of the bond and the angle θ between the direction of the bond vector and the direction of the external magnetic field vector.

In samples of solid material where nuclear spins are heterogeneously maintained in fixed positions relative to each other this gives rise to extremely broad lines. In solution where there is molecular motion the dipole-dipole interactions are averaged away and not observed.

Residual dipolar coupling

It is possible to re-establish the dipolar coupling in solutions of charged colloid particles. In very strong magnetic fields it is possible to orient homogeneously charged

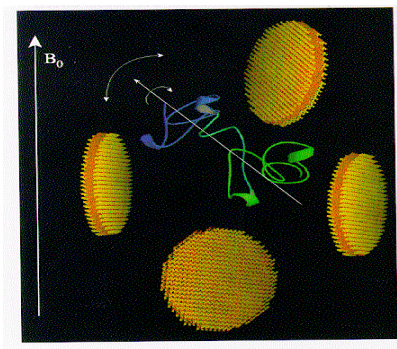


Figure 21. Oriented bicelles in a magnetic field. The protein molecule dissolved in the colloidal solution has anisotropic motion, which reintroduces dipolar coupling.

particles. Protein molecules placed in such environments will no longer have free molecular motion and the dipole-dipole interactions will be partly re-established. The degree to which the residual dipolar coupling is re-established can be controlled by the concentration of colloidal solution. The reason for being interested in re-establishing the dipolar coupling is the angular dependence to the magnetic field, figure 21. In the isotope enriched protein molecule there are several dipole-dipole interactions between spins of fixed distances maintained in single bonds, for instance the H^N -N bond vector of the peptide backbone, figure 22. By measuring residual dipolar coupling for spin-pairs with fixed distances it is possible to relate all the angles of these bonds to the direction vector of the magnetic field, and subsequently to each other. The use of residual dipolar coupling has in many cases been proven to have an enormous effect on the accuracy of the structure determination by NMR spectroscopy.

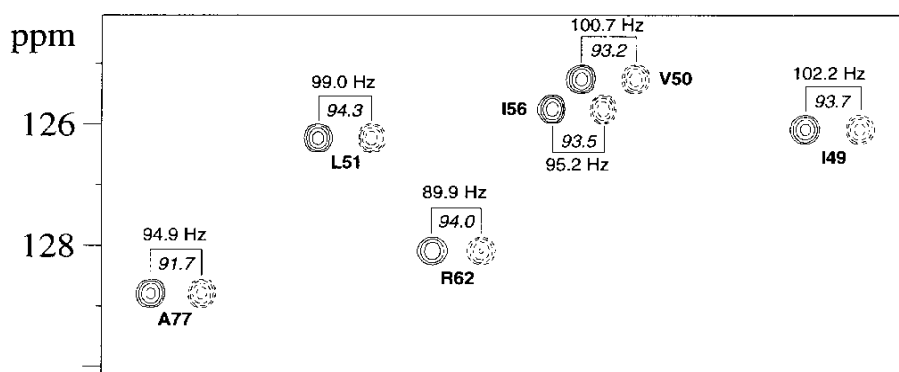


Figure 22. An example of measuring a residual dipolar coupling. The one bond coupling between 1H and ^{15}N in the peptide bond has been measured in a magnetically oriented colloidal solution and in the freely tumbling protein. The coupling is the sum of the J-coupling and the residual dipolar coupling.

Cross relaxation and the nuclear Overhauser effect

Cross relaxation is a result of dipole-dipole interaction between proximate nuclear spins. When two spins are very close in space, they experience each other's magnetic dipole moment. It is possible by NMR pulse techniques to either reverse the direction of the magnetic dipole of the nuclear spin or to "turn off the magnetic dipole" of one of the spins and to measure, how this affects the other spin. The rate by which this effect is transmitted to the other nuclear spin is the cross relaxation rate, and this is inversely proportional to the sixth power of the distance between the two nuclear spins. The cross relaxation rate and the related nuclear Overhauser effect can be used to estimate distances between two nuclei in a protein molecule. The effect is not only depending on the distance between the two nuclei it also depends on the overall molecular motion and on internal motions in the protein, which change the distance between the two nuclei with time.

The nuclear Overhauser effect is one of the most important tools in structure determination by NMR spectroscopy. The experiment used to determine the nuclear Overhauser effect is called NOESY derived from Nuclear Overhauser Effect Spectroscopy. The most common experiment is a homo-nuclear two-dimensional NMR experiment. In figure 23 is shown a schematic presentation of a two-dimensional NOESY spectrum between two nuclei R and G.

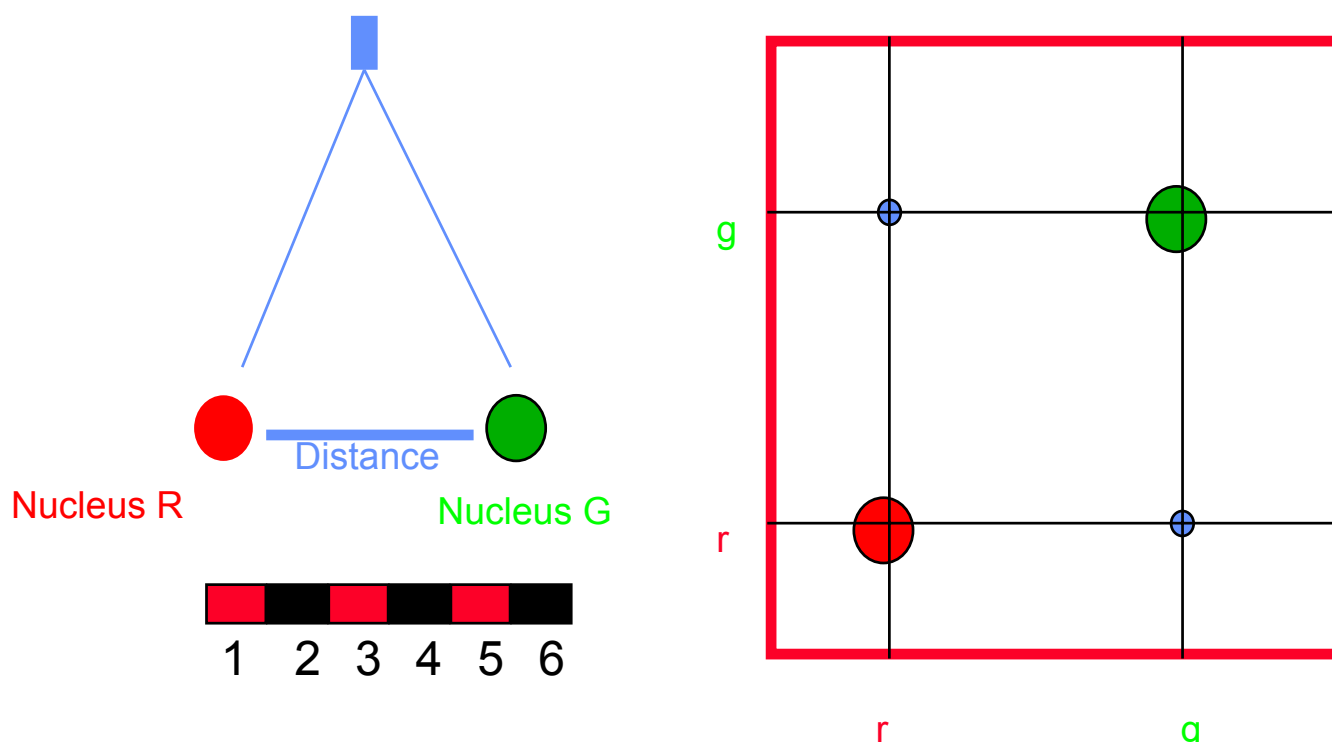


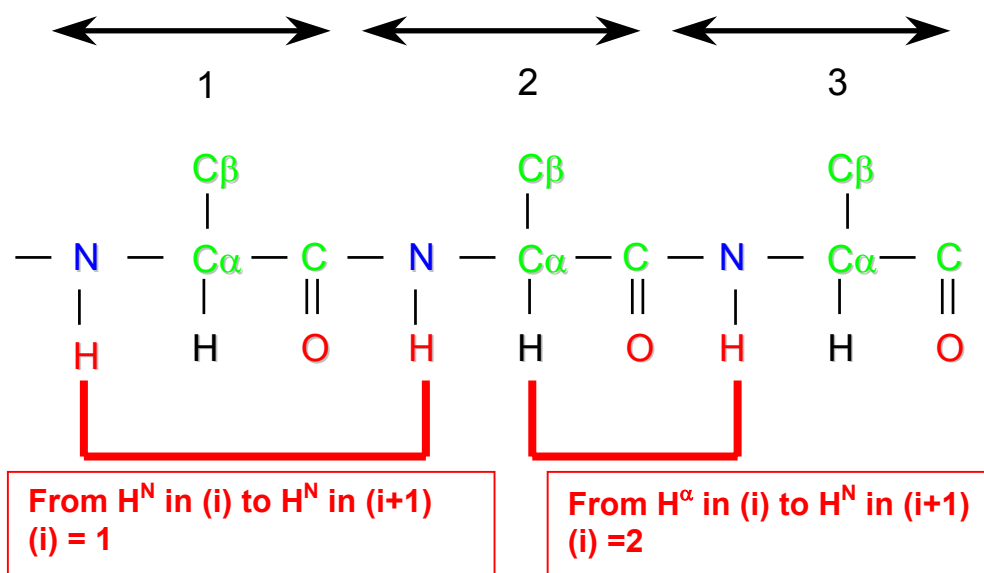
Figure 23. Schematic presentation of a two-dimensional NMR spectrum recording the Nuclear Overhauser effect between nucleus R and nucleus G. The effect is measured by the intensity of the blue cross peaks at (r,g) and (g,r) .

The nuclear Overhauser effect can typically be measured between nuclei, which are less than 0.5 nanometers (10^{-9} meter) apart. Many structural chemists prefer to use Ångström as the unit of length (1 Ångström = 0.1 nanometer). The NOE decline with the distance between the two nuclei being inversely proportional to the sixth power of

the distance between the two nuclear spins. This means that if a nuclear Overhauser effect is measured to be one for two nuclei, which are 5 Ångströms apart, the NOE between two nuclei, which are 2 Ångströms apart, will be $(5/2)^6 \approx 244$ times larger.

Sequential assignment using homonuclear ^1H NMR NOE spectroscopy.

It is not always feasible to produce a ^{13}C , ^{15}N double labelled protein sample. In this case samples may either be studied by ^1H -NMR or by a combination of ^{15}N and ^1H NMR spectroscopy. Here the nuclear Overhauser effect can be used to make sequential assignments. In the two most common types of secondary structure the peptide chain accommodates conformation which bring ^1H of the peptide backbone and the side chains of neighbouring residues so close together that they are observable by NOE spectroscopy. In the α -helix the neighbouring H^{N} are 2.8 Ångström apart, and in the β -strand the distance from H^{α} in residue (i) to H^{N} in residue (i+1) is only 2.2 Ångström.



Two important sequential assignment tools using sequential assignment by NOEs

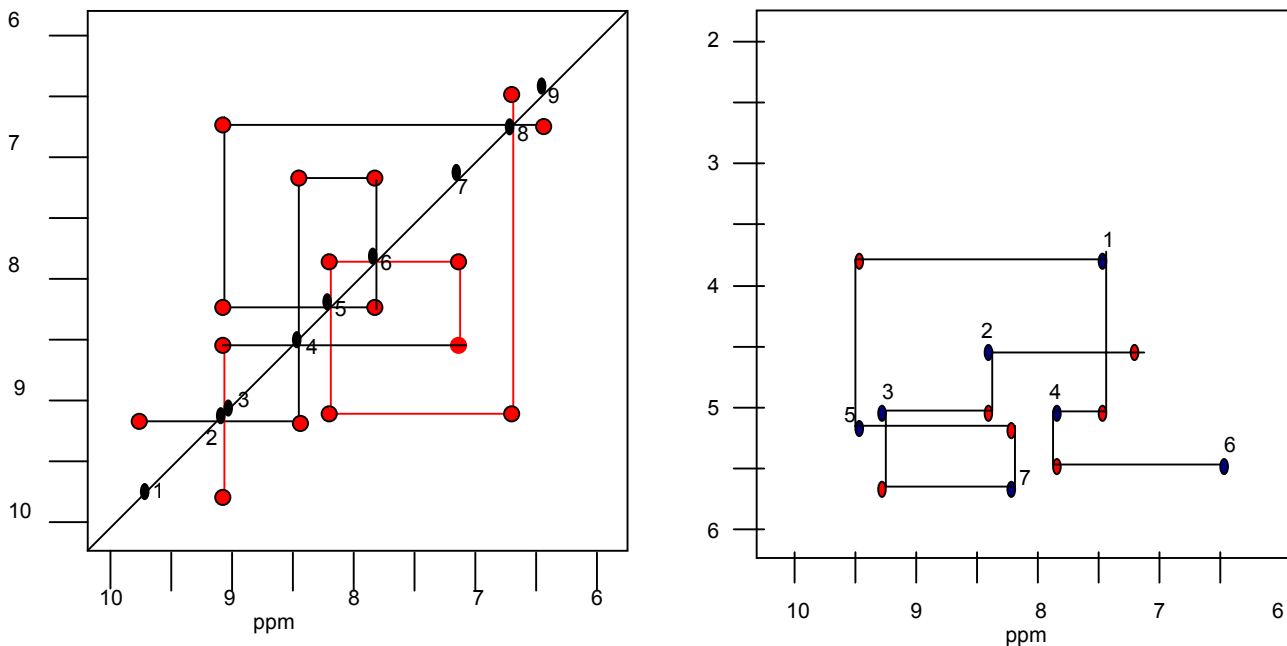


Figure 24. left. Schematic presentation of a NOESY spectrum in the region where nuclear Overhauser effects between H^N atoms are observed. The diagonal cross peaks have been marked as black signals and labelled 1 to 9. The red cross peaks are sequential NOEs.

Figure 24 right. Schematic presentation of sequential assignment of by the NOE between H^α in residue (i) to H^N in residue (i+1). The schematic presentation shows the superimposition of a COSY spectrum with blue annotated cross peaks and a NOESY spectrum with red cross peaks. The blue cross peaks are the scalar (through bond) coupling between H^α and H^N in one residue. The red cross peaks are the sequential NOE between H^α in residue (i) to H^N in residue (i+1).

The red and black lines in figure 25 left show the sequential assignment of nine H^N residues of a helix, where the order of sequence of the H^N atoms in the peptide chain is revealed by the NOE connectivity of the signals in the order 1, 2, 4, 7, 6, 5, 3, 8, 9 or the reverse.

If it is known that signal 2 is from a valine, 5 from a tryptophan and 9 from a glycine the sequential assignment may be performed by matching the sequences

XVXXXWXXG or the reverse **GXXWXXXVX** to the known amino acid sequence of the protein, so a fit would be the sequence

983567421 sequence of signals
GAKWSR**YV**P amino acid sequence

Using the signal annotation the sequential in figure 25 right the assignment goes 6, 4, 1, 5, 7, 3 2. Here the direction of the sequence is unambiguous because the direction is always from H^α in residue (i) to H^N in residue (i+1). As above the signal sequence can be imbedded in the amino acid sequence if two or three of the spin systems of the connected signals have been assigned.

Structure determination by NMR spectroscopy

The structure determination by NMR spectroscopy depends critically on the assignment of the NMR signals of the NMR active nuclei of the protein. NMR spectroscopy has information about the structural geometry around every NMR observable nuclei in the protein. This information may be read from the NMR signals of the nuclei. It is therefore important that the signals in the NMR spectrum have been assigned to the correct nuclei, those, which give rise to the signals. From the NMR signals the distance to the near-by nuclei in the structure can be read using NOE spectroscopy, the dihedral angles can be determined by coupling constants and chemical shifts, the relative angles of a number of bond vectors can be determined by residual dipolar coupling.

NMR spectroscopy provides several sets of information, which are all being used for the structure calculation. The computer programs, which handles the structure calculation is a molecular dynamics program. As a starting point the program has information about the covalent structure geometry of all the common amino acids regarding bond lengths and angles and potential functions to ensure that the bonds are maintained within realistic limits. For non-defined distances the program has a set of potentials that controls electrostatic interactions and van der Waals interactions ensuring that atoms do not get closer to each other than the Lennard-Jones potential permits.

The structural information from NMR is typically entered into the program either as a distance range or an angular range. The computer uses potential functions to ensure that every single NMR derived input structure data stay in the set range. The molecular dynamics program is using a process called simulated annealing where the protein atoms are heated and cooled successively, while the potential functions are turned on to form the structure. The repeated heating and cooling process is meant to help energetically unfavourable structures to overcome energy barriers and end up in energetically more favourable structures, which may resist the subsequent heating process.

Because the NMR derived structural input cannot provide a unique structure input for all the atoms in the protein, the NMR data set cannot define the structure unambiguously. Therefore, the structure calculation based on NMR derived structure constraints takes its origin in a randomly generated structure. The structure determination is repeated several times with new and different starting points. Typically one hundred structures are calculated, and those structures, which comply best to the NMR input data and are energetically most favourable, are selected as group of structures often referred to as an NMR bundle. The bundle is the result of a structure determination by NMR, figure 27. The superimposed structures reveal how reproducibly the structure calculation program calculates the structure. This is measured as the root mean square deviation, RMSD, of every single atom in the structure. The rmsd for the peptide backbone atoms is the sum of the RMSDs for all the atoms in the peptide backbone divide by the number of atoms.

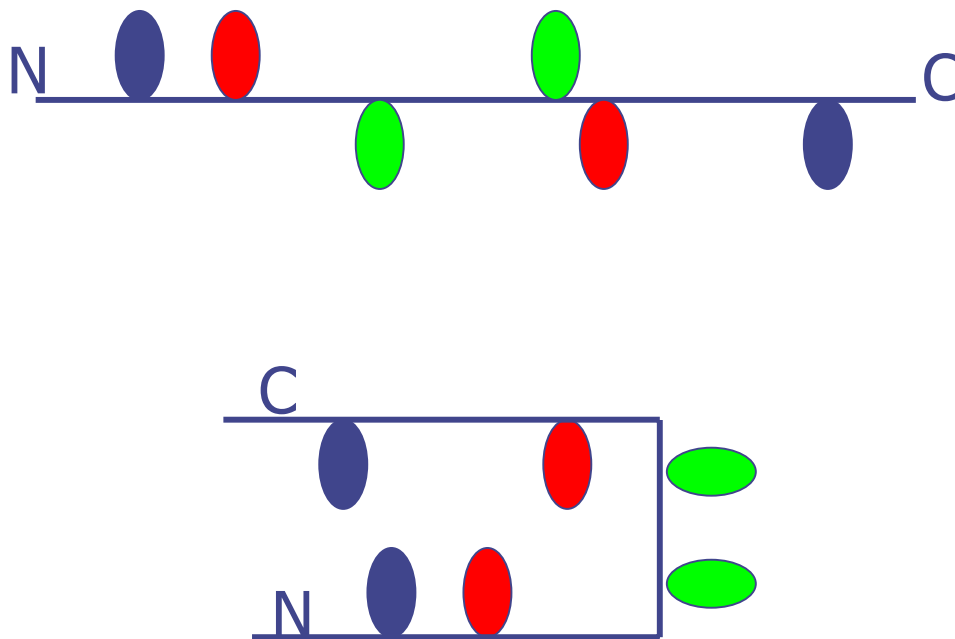
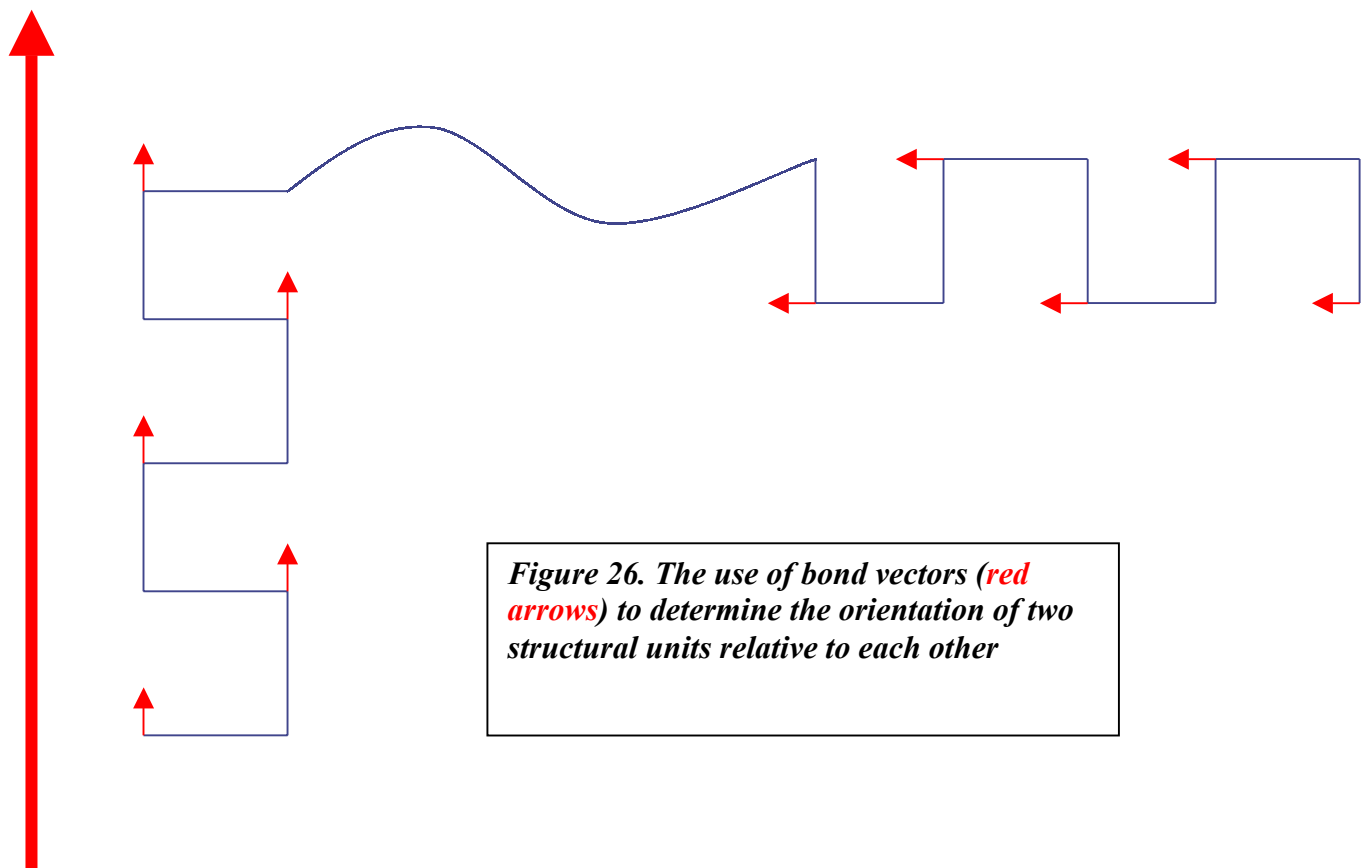


Figure 25. Schematic demonstration of using distance constraints to calculate a structure. In the upper figure is shown a linear strand with beads. The green, red and blue pairs of beads, respectively have been shown to be close to each other. The structure below represents one solution to determining the structure based on the three pieces of distance information.



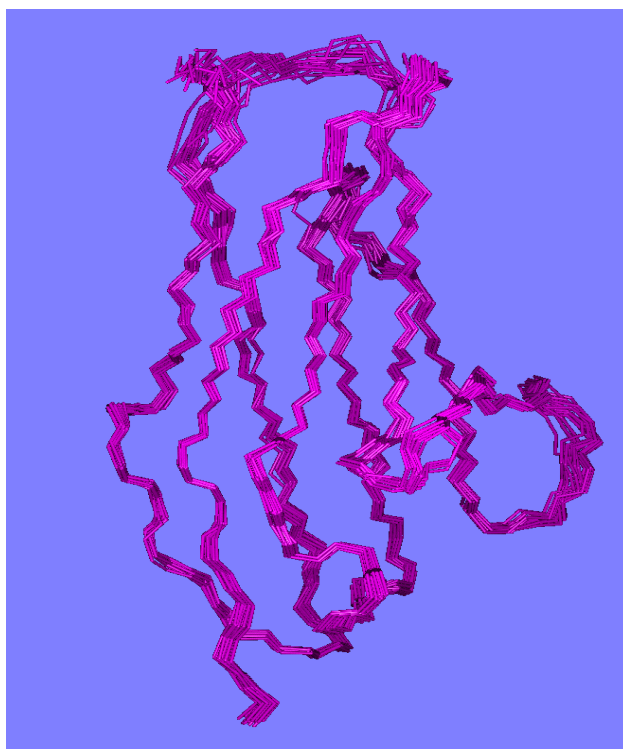


Figure 27. NMR bundle, 20 structures superimposed.

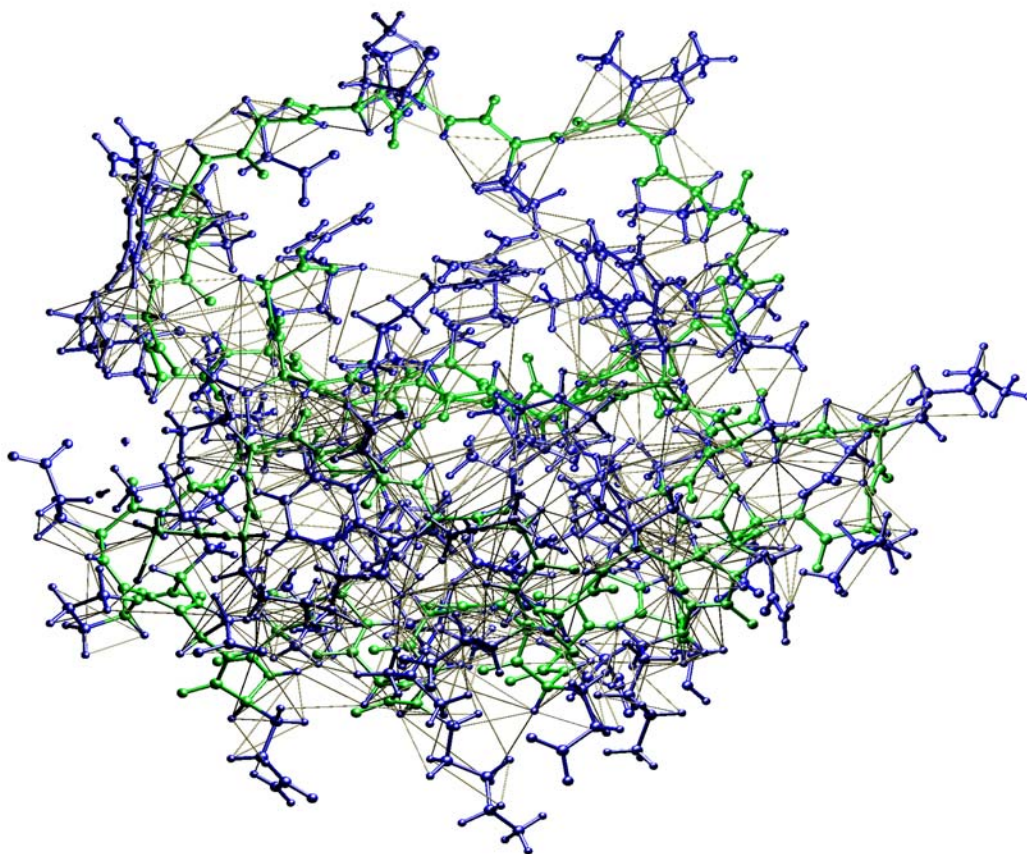


Figure 28. Illustration of the NOE based distance constraints used to determine the structure of the protein chymotrypsin inhibitor 2. The distance constraints are drawn as thin lines between the two atoms, whose distances are constrained. The backbone atoms and the bonds between them are green, the side chain atoms and the bonds between them are blue.

Application of NMR spectroscopy to study chemical and properties of proteins

When an NMR signal of a given nucleus has been assigned in the NMR spectrum, the signal represents a specific reporter for the environment of the nucleus in the protein structure. The chemical shift is a very sensitive parameter, which might report even subtle conformational changes near the nucleus. The scalar coupling, represented by the coupling constant may report on conformational changes in dihedral angles. The binding of a ligand or the pH dependent change of a charge may result in conformational changes and a change in the NMR parameters for the nuclei in the environment of the chemical modification. Since NMR is specific to the atomic level NMR is by far superior to any other spectroscopic technique to register the changes, which occur as a result of the protein binding to another molecule.

Chemical exchange by NMR

When a ligand binds to a protein, the equilibrium of the interaction is determined by the rate constants for the formation of the complex and rate constants for the dissociation of the complex. It is often of interest to characterize the properties of the protein-ligand complex, and knowing the on and off rates of the reaction is an important part of this characterization. NMR can measure reaction rates. This is illustrated in figure 29.

In a sample where a ligand is in slow exchange with the protein and the signal of a given nucleus has a chemical shift for the bound conformation and a chemical shift for the free conformation. NMR will record the sample as containing two different species the bound and the free form, bottom situation in figure 29.

If this sample is heated and the rate constants increase the two signals get broader. This is because the nucleus is rapidly transferred from one magnetization condition to another, leading to line broadening. As it is seen in the middle panel the line broadening effect is so strong that the NMR signal essentially disappears.

At even higher exchange rates the transfer of magnetization is faster than the difference in the chemical shift frequencies of the two different exchange sites.

The magnetization will not precess with either frequency but will be observed as an average. The position of the average chemical shift depends on the fractions of the time the nuclear spin spends in each of the sites.

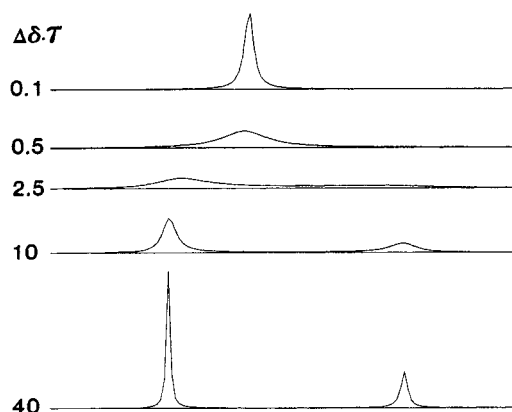


Figure 29. A model calculation of a two-site exchange system for the ratio between the chemical shift difference $\Delta\delta$ and the rate constant $1/\tau$ varying between 40 and 0.1

The two extremes of fast and slow exchange, respectively, are both encountered in studies of protein ligand interactions when studied by NMR. In the fast exchange situation the ligand titration will behave as seen in figure 30.

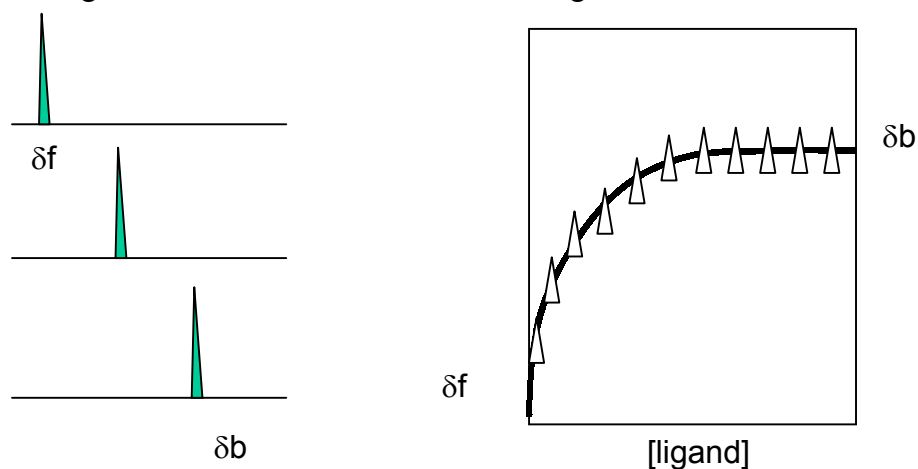


Figure 30. Schematic presentation of a “fast exchange” protein ligand titration by NMR. The top left: spectrum represents a start situation where no ligand has been added. The NMR signal is at δ_f . The middle line is the situation where enough ligand has been added to saturate half of the binding sites in the protein. The bottom line: ligand has been added to occupy all binding sites and the NMR signal is now at δ_b .

Figure 30 right. Binding curve of a titration. The binding constant can be determined from the binding curve.

In the fast exchange titration the signal will shift with increasing concentration from the chemical shift position of the free protein, δ_f , to the chemical shift position δ_b . It is possible to determine the binding constant for the protein ligand interaction by fitting the observed binding curve to the theoretical expression for the binding. It is also possible from a line shape analysis to determine on- and off-rate constants.

Protein ligand interactions in slow exchange gives a completely different titration pattern, figure 31.

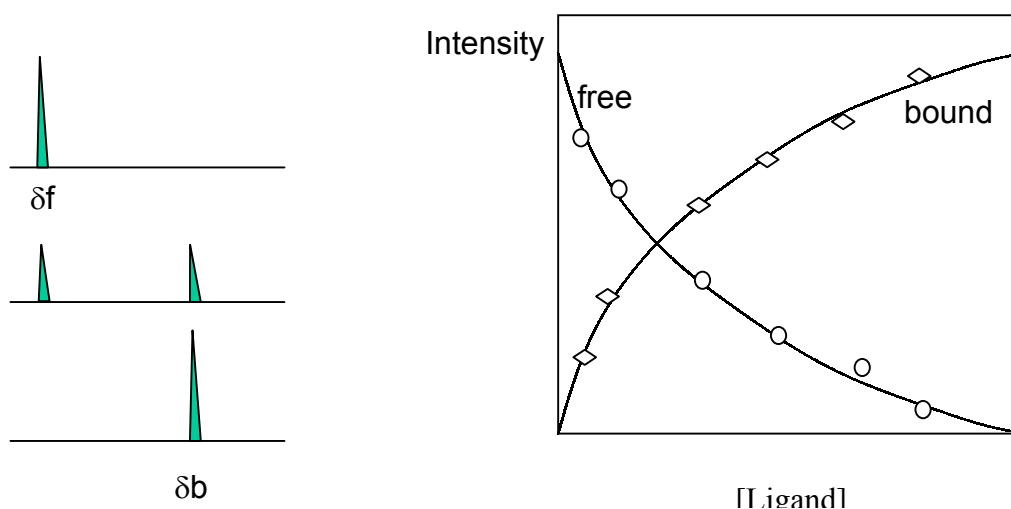


Figure 31. Schematic presentation of a “slow exchange” protein ligand titration. Left: change in spectrum with increasing ligand concentration. Right: change in signal intensities with ligand concentration

Here the signal of the protein ligand complex will increase with the addition of ligand from no intensity to full intensity when all sites are occupied. At the same time intensity of the free protein signal will decrease. The sum of the intensity of the two peaks will always be constant. In principle the binding constant can be obtained from binding curves.

Application of NMR spectroscopy for studies of chemical reactions

For chemical processes, where chemicals are being produced, NMR spectroscopy is an ideal method for studying the time courses of the formation of a product and of the disappearance of the starting material. An NMR spectrum can be recorded in a fraction of a second, and it is therefore possible to study processes with half-life times from a second and up.

Enzymatic processes are often easily studied by NMR. One enormous advantage of NMR spectroscopy is its ability to study biochemical processes directly in living organisms.

Hydrogen exchange in proteins

In aqueous solutions the hydrogen atom bound to nitrogen, oxygen and sulphur is labile and exchange with the hydrogen of water. In proteins this is the case for the peptide group, the amino group, the imidazol group, the indole group, the guanidine group, the hydroxy group and the thiol group. For proteins where these groups may engage in hydrogen bonds it is required, that the hydrogen bond is broken for the exchange to take place. The rate of the hydrogen exchange will depend of the frequency by which the hydrogen bond is opening.

Hydrogen exchange can be studied by NMR spectroscopy. By replacing water with deuterium oxide the exchange process will lead to the replacement of hydrogen with deuterium. ^1H NMR can study the exchange process, since the hydrogen giving rise to a ^1H NMR signal is exchanged by deuterium, which is not observable by NMR at the ^1H NMR frequency.

The ^1H NMR signals of the peptide hydrogen atoms in a protein are normally very easily observed. These hydrogen atoms are engaged in hydrogen bond formation in the secondary structures of proteins. By recording the hydrogen exchange rates the stability of the individual hydrogen bonds can be measured. This can be done for every single peptide group of a protein and be used to describe the dynamic properties in the neighbourhood of the amide group, which is being monitored, Figure 32.

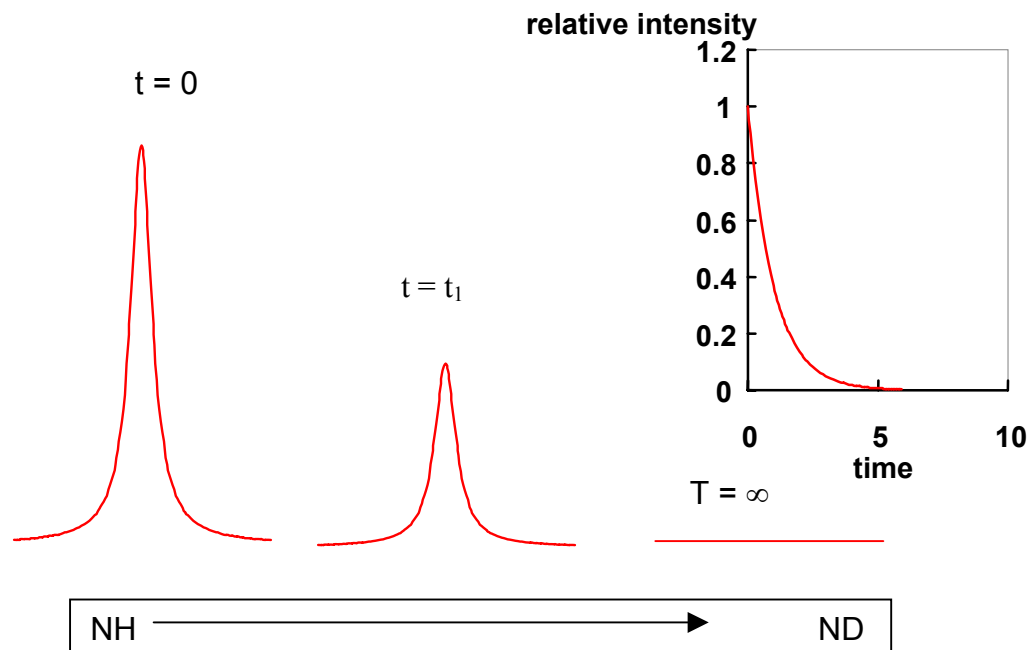


Figure 32. The decrease of the NH NMR signal with time in the process of the proton being exchanged with deuteron, which is not detectable by NMR at the protein frequency.