

6 NMR von Proteinen

6.1 Proteine

6.1.1 Motivation und Grundlagen

Proteine gehören zu den wichtigsten Bestandteilen aller lebenden Organismen. Sie stellen den mengenmäßig größten Anteil der Biomoleküle und spielen eine wichtige Rolle unter anderem für die Struktur sowie für sämtliche chemischen Reaktionen, welche durch Enzyme gesteuert werden. Die Struktur und Funktion von Proteinen zu verstehen ist deshalb zu einer der wichtigsten Ziele von Biologie und Medizin geworden.

Chemisch sind Proteine Makromoleküle, welche aus einer Sequenz von Aminosäuren bestehen. Diese können jeweils durch die Formel $\text{NH}-\text{CHR}_i-\text{CO}$ beschrieben werden, wobei der "Rest", resp. die Seitenkette R_i für die jeweilige Aminosäure charakteristisch ist. Die Sequenz der Aminosäuren ist in den Genen codiert.

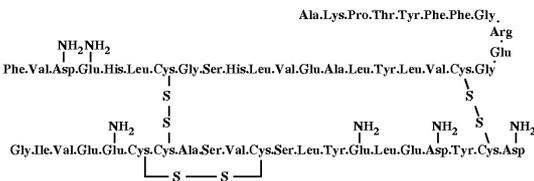


Abbildung 6.1: Aminosäuresequenz von Insulin.

Der erste Schritt für die Charakterisierung von Proteinen ist die Bestimmung der Aminosäuresequenz. Als erstem gelang es Sanger 1955, die Sequenz eines Proteins zu bestimmen (Insulin). Er erhielt dafür 1958 den Nobelpreis. Das Sanger Verfahren wurde später weiter entwickelt und teilweise automatisiert. Da die Sequenz in der DNA codiert ist kann sie auch über die Sequenzierung der DNA bestimmt werden.

6.1.2 Strukturbestimmung

Allerdings ist die Sequenz noch nicht direkt für die Funktion zuständig. Die Ketten der Aminosäuren falten sich und erst die daraus entstehende dreidimensionale Struktur bestimmt die biologische Funktion. Die erste dreidimensionale Struktur eines Proteins wurde 1958 bestimmt (Myoglobin und Hämoglobin).

Die Funktionsweise und damit die Struktur der Proteine ist eine essentielle Grundlage für das Verständnis von vielen biologischen und medizinischen Prozessen. Man versucht deshalb schon seit vielen Jahren, die Struktur von Proteinen zu bestimmen. Dies ist allerdings keine einfache Aufgabe, und erst in den letzten Jahren gelang das im großen Stil.

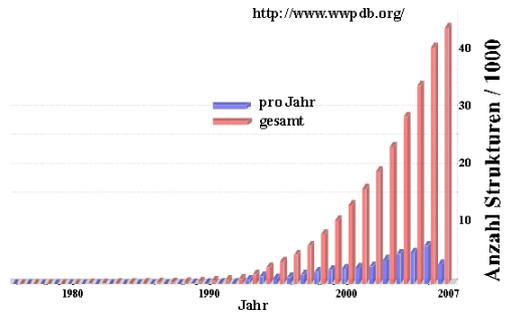


Abbildung 6.2: Anzahl der Proteinstrukturen in der "Protein Data Bank".

Inzwischen wächst die Zahl der Strukturen exponentiell an. Im letzten Jahr wurden mehr als 7000 Strukturen neu in der "protein data bank" deponiert.

Die meisten dieser Strukturen wurden mit Röntgenbeugung an Kristallen bestimmt, aber ein zunehmend größerer Teil wird mit Hilfe von NMR bestimmt. Strukturbestimmung mit Hilfe der NMR ist zwar aufwändiger, hat aber auch gewisse Vorteile gegenüber der Röntgenmethode: Damit kann das Protein in wässriger Lösung, d.h. unter einigermaßen physiologischen Bedingungen untersucht

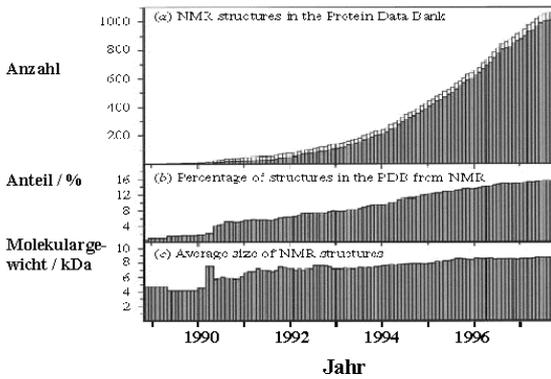


Abbildung 6.3: Gesamtzahl der Proteinstrukturen und Anteil der Strukturen, die mit NMR bestimmt wurden.

werden, während für Untersuchungen mit Hilfe der Röntgenbeugung (gute!) Einkristalle benötigt werden. Da nicht alle Proteine kristallisiert werden können ist häufig die NMR die einzig mögliche Methode.

Neben der Struktur erlaubt die NMR auch in vielen Fällen Untersuchungen zur Funktion. So kann man (in geeigneten Systemen) feststellen, welcher Teil eines Proteins an ein anderes Molekül ankoppelt. Entscheidend dafür ist, dass NMR Untersuchungen in Flüssigkeit, d. h. unter (nahezu) physiologischen Bedingungen durchgeführt werden können. Die Untersuchung von Struktur und Funktion von Proteinen gehört inzwischen zu den wichtigsten Anwendungen der magnetischen Resonanz.

6.1.3 Aminosäuren und Peptide

Proteine sind natürliche Polymere, also Makromoleküle, die aus einer Sequenz von kleineren Molekülen, den Monomeren, zusammengesetzt sind. Die Grundbestandteile der Proteine sind die **Aminosäuren**. Alle Aminosäuren, außer der einfachsten, dem Glycin, sind **chiral** und daher optisch aktiv.

Aminosäuren werden über **Peptidbindungen** miteinander verknüpft: Dabei wird ein Wassermolekül eliminiert und es entsteht eine C-N Bindung von C=O nach N-H.

Auf diese Weise entstehen Ketten von Aminosäuren.

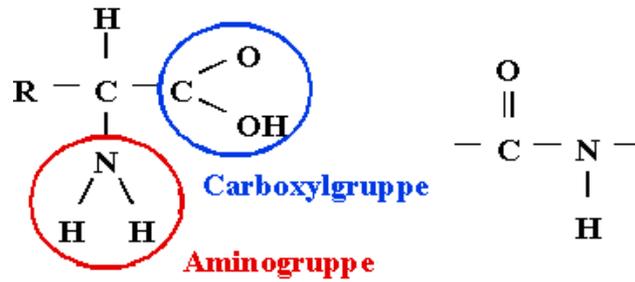


Abbildung 6.4: Aminosäuren (links) und die sie verbindenden Peptidbindungen (rechts).

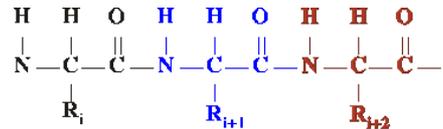


Abbildung 6.5: Sequenz von Aminosäuren.

Kurze Ketten werden als Peptide bezeichnet, lange Ketten als Proteine.

Name	Abk.	Mol.Gew.	Name	Abk.	Mol.Gew.
Alanine	A Ala	89.09404	Methionine	M Met	149.20784
Cysteine	C Cys	121.15404	Asparagine	N Asn	132.11904
Aspartic acid	D Asp	133.10384	Proline	P Pro	115.13194
Glutamic acid	E Glu	147.13074	Glutamine	Q Gln	146.14594
Phenylalanine	F Phe	165.19184	Arginine	R Arg	174.20274
Glycine	G Gly	75.06714	Serine	S Ser	105.09344
Histidine	H His	155.15634	Threonine	T Thr	119.12034
Isoleucine	I Ile	131.17464	Valine	V Val	117.14784
Lysine	K Lys	146.18934	Tryptophan	W Trp	204.22844
Leucine	L Leu	131.17464	Tyrosine	Y Tyr	181.19124

Abbildung 6.6: Aminosäuren

Von allen in der Natur vorkommenden Aminosäuren werden nur 20 ribosomal in Proteine eingebaut. Auch damit ist eine sehr große Strukturvielfalt möglich: für ein Protein mit N Aminosäuren gibt es N^{20} Kombinationsmöglichkeiten. Damit ist schon bei einer Kettenlänge von weniger als 80 Aminosäuren im Universum nicht genügend Masse vorhanden, um alle möglichen Kombinationen zu realisieren.

Die Sequenz der Aminosäuren wird mit einer abgekürzten Notation gelistet. Dabei beginnt man jeweils am N-Terminus, d.h. an demjenigen Ende der Aminosäurenkette, an der sich die NH_2 -Gruppe befindet.

6.1.4 Strukturen

Die Funktion der Proteine ist stark mit ihrer räumlichen Struktur verknüpft. Man unterscheidet unterschiedliche Arten von Strukturen:

Die **Primärstruktur** besteht aus der Aneinanderreihung von Aminosäuren, der Aminosäuresequenz. Das Protein ist sehr langgestreckt, wobei dieser lange Faden sehr flexibel ist: die einzelnen Molekülteile können jeweils um die $N-C_\alpha$ und $C_\alpha-CO$ Bindungen praktisch frei rotieren. Dadurch ergeben sich sehr viele unterschiedliche Möglichkeiten für die räumliche Anordnung. Es ist deshalb eigentlich erstaunlich, dass in Wirklichkeit der überwiegende Teil der Proteinmoleküle eine relativ wohl definierte Struktur besitzt. Diese bildet sich zunächst lokal; man spricht hier von der **Sekundärstruktur**. Diese entsteht durch Wechselwirkungen zwischen den Aminosäuren und wird meist durch Wasserstoffbrücken stabilisiert. Es entstehen Helizes, Faltblätter, Schleifen und anderen Strukturen. Die beiden wichtigsten Strukturelemente sind die α -Helix und das β -Faltblatt.

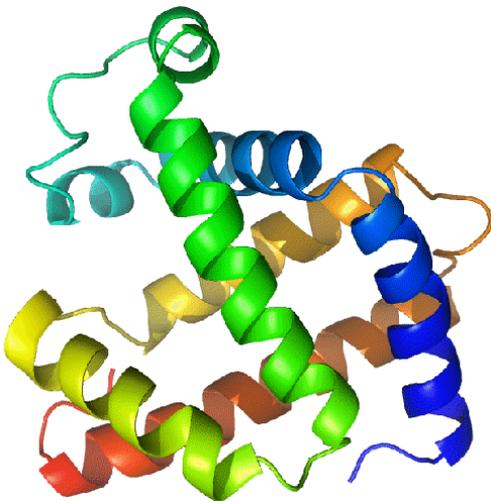


Abbildung 6.8: α -Helices in Myoglobin

Schöne Beispiele für α -Helices findet man in Myoglobin. Dies war auch das erste Protein, dessen Struktur mit Hilfe von Röntgenbeugung bestimmt wurde. In diesen beiden Strukturelementen können die Aminosäuren ihre möglichen Wasserstoffbrücken bilden.

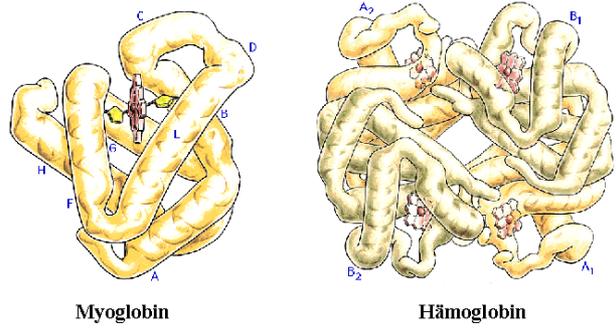


Abbildung 6.9: Tertiär- und Quartärstruktur von Proteinen

Die **Tertiärstruktur** ist die dreidimensionale Anordnung der Sekundärstrukturen und ist häufig die biologisch aktive Struktur des Proteins. Das Protein bildet diese Struktur aufgrund von Wechselwirkungen zwischen den einzelnen Aminosäuren. Die wichtigsten Wechselwirkungen sind Wasserstoff- und Disulphid Brücken, sowie van der Waals Wechselwirkungen. Außerdem spielen hydrophobe Wechselwirkungen eine wichtige Rolle: die Aminosäuren sind teilweise hydrophob, d.h. sie werden vom Wasser gemieden. Hydrophobe Teile eines Proteins sind meist entweder im Inneren der Struktur oder im Inneren einer Lipidmembran. Geladene oder polare Aminosäuren sind dagegen hydrophil.

Die Tertiärstruktur bildet sich spontan, aber nur bei geeigneten Bedingungen. So kann sie z.B. durch Zugabe von Harnstoff aufgelöst werden. Die Funktion eines Proteins läuft häufig über Änderungen dieser Konformation.

Die Aneinanderreihung von Proteinmolekülen in Oligomere wird als **Quartärstruktur** bezeichnet.

6.1.5 Funktion

Die Aufgaben der Proteine sind vielfältig:

- **Strukturbildung und -erhaltung.** Die mechanische Stabilität von Zellen ist dafür ein Beispiel (Collagen).
- **Bewegung.** Die Muskelbewegung entsteht durch ein Zusammenspiel von Actin und Myosin.

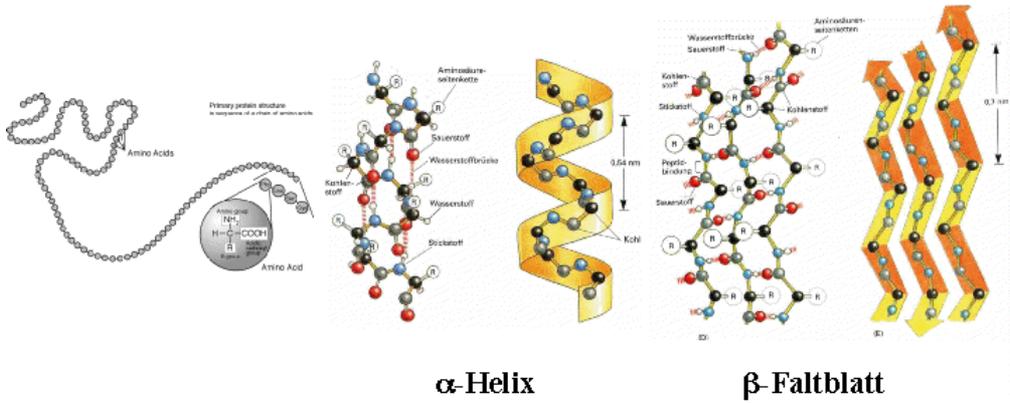


Abbildung 6.7: Primär- und Sekundärstruktur von Proteinen

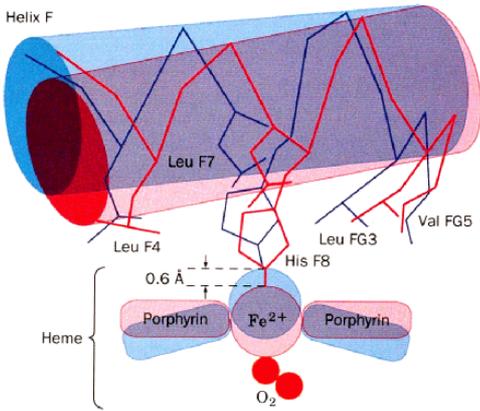


Abbildung 6.10: Bindung von Sauerstoff an das Hämoglobin verändert die Struktur des Komplexes.

lysiert

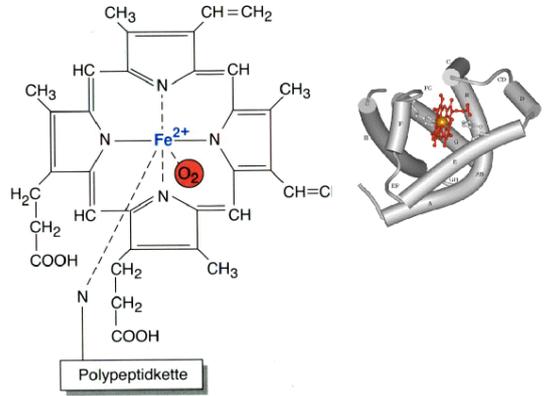


Abbildung 6.11: Kofaktor des Hämoglobins

- **Transport.** Durch Proteine werden Stoffe durch die Blutbahn transportiert. Hämoglobin transportiert Sauerstoff und Kohlendioxid zwischen Lunge und Gewebe. Ein weiteres Beispiel sind Ionenkanäle in Zellmembranen.
- **Schutz und Abwehr.** Im Immunsystem gibt es Immunglobuline und Antikörper für die Immunabwehr.
- **Regulation und Steuerung von Zellvorgängen.** Jeder hormonale Prozeß verläuft zum Beispiel über Proteine als **Rezeptoren**, sie dienen auch häufig der Informationsweiterleitung.
- **Katalyse.** Alle im Organismus ablaufende chemischen Prozesse werden durch Enzyme katalysiert.

In vielen Fällen führen die Proteine diese Funktionen nicht alleine aus, sondern benötigen dazu weitere Moleküle, sogenannte Kofaktoren.

6.2 Methodik der Strukturaufklärung

6.2.1 NMR von Proteinen

Proteine enthalten in erster Linie C, N, O und H Atome. Während die H Atome direkt für die NMR verwendet werden können, sind bei C fast 99% der Kerne ohne Spin, bei O mehr als 99%. Im Falle von N besitzen beide natürlich vorkommenden Isotope

(^{14}N , $I=1$ und ^{15}N , $I=1/2$) einen Spin. Allerdings ist ^{14}N für die NMR von Proteinen nicht geeignet: Der Spin besitzt nicht nur ein magnetisches Dipolmoment, sondern auch ein elektrisches Quadrupolmoment. Dieses sorgt für relativ effiziente Relaxation und damit breite Resonanzlinien. Für hochaufgelöste Spektren muss deshalb Stickstoff mit ^{15}N angereichert werden. Beim Kohlenstoff kann man ^{13}C anreichern. ^{17}O Anreicherung des Sauerstoffs wird nur selten verwendet.

Proteine sind relativ komplexe Moleküle, die eine große Zahl von nicht äquivalenten Spins enthalten. Entsprechend komplex wird das Spektrum.

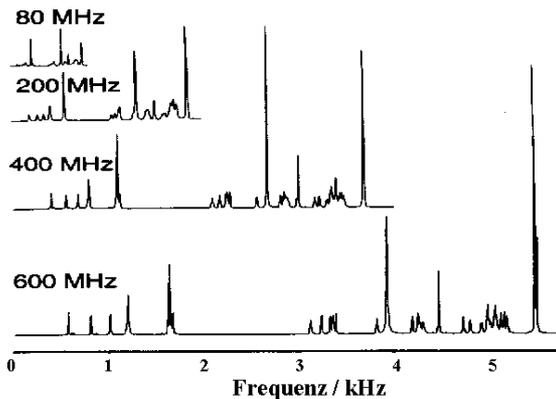


Abbildung 6.12: ^1H -Spektren des Peptides Gramicidin bei unterschiedlichen Feldstärken.

NMR von Proteinen wurde erst möglich, als Magneten mit genügend hohen Feldstärken und hoher Auflösung zur Verfügung standen. In der Figur ist gezeigt, wie die Auflösung eines Spektrums für ein einfaches Protein mit der Feldstärke besser wird. Größere Proteine werden nur noch bei 800, 900 oder 950 MHz ^1H Frequenz gemessen.

Bei einem relativ kleinen Protein wie Ubiquitin kann man bei 500 MHz noch einige der relevanten Bereiche erkennen. Eine Zuordnung ist aufgrund des starken Überlapps aber nicht möglich.

6.2.2 Empfindlichkeit und Linienbreite

Typische Proben bestehen aus etwa 0.5 ml Lösung, in der das Protein eine Konzentration von 0.1 - 3 mM

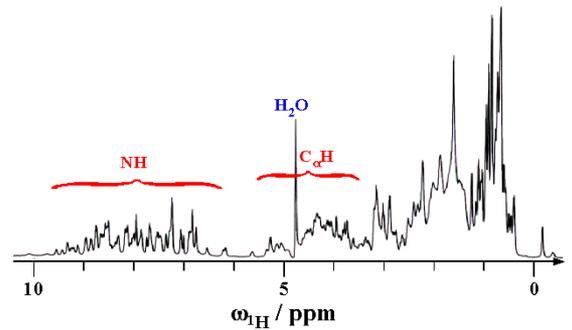


Abbildung 6.13: ^1H NMR Spektrum von Ubiquitin.

vorhanden sind. Bei einer Konzentration von 1 mM befinden sich somit etwa

$$10^{-3} \cdot 0.5 \cdot 10^{-3} \cdot 6 \cdot 10^{23} = 3 \cdot 10^{17}$$

Moleküle in der Probe.

Ein relativ kleines Protein mit 100 Aminosäuren enthält z.B. 100 Resonanzen im Bereich der NH-Protonen (≈ 6.5 - 10.5 ppm) und im Bereich der C_α Protonen (≈ 4 - 6 ppm) und ein Vielfaches davon im Bereich der Seitenketten - Aufspaltungen durch Kopplungen jeweils nicht mitgerechnet. Bei einer ^1H Resonanzfrequenz von 500 MHz entspricht ein Bereich von 2 ppm jeweils 1 kHz. Somit wäre es auch unter idealen Bedingungen nur dann möglich, alle Resonanzlinien getrennt zu beobachten, wenn die Linienbreiten weniger als 5 Hz betragen. Eine gute Auflösung ist deshalb sehr wichtig.

Ein weiteres Problem ist die Empfindlichkeit - absolut sowie relativ zum Lösungsmittel. So enthält reines Wasser 55 Mol/l, also mindestens 10^4 mal Wasser-Moleküle pro Protein-Molekül. Damit das Signal nicht vom Wassersignal überdeckt wird häufig in D_2O gearbeitet. Allerdings geht dabei auch ein Teil der NH-Protonen verloren.

Außerdem wurden experimentelle Techniken zur Unterdrückung des Wassersignals entwickelt. Allerdings führen auch diese häufig zu einer Abschwächung z.B. der NH-Signale.

Je größer ein Protein, desto langsamer bewegt es sich. Gemäß Stokes Gesetz hat ein sphärisches Molekül mit (effektivem hydrodynamischem) Radius

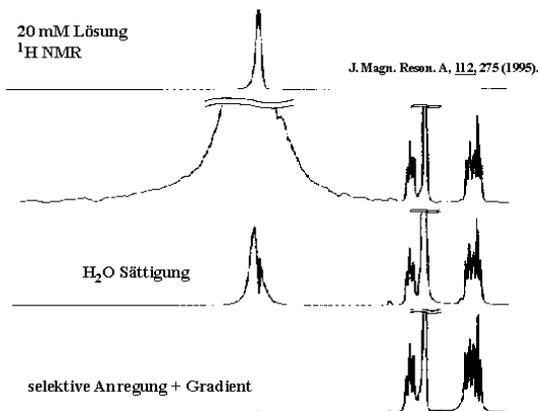


Abbildung 6.14: a) Normales ^1H NMR Spektrum einer 50 mM Lösung. b) 1000 mal vergrößert. c) Wasser vor der eigentlichen Anregung durch selektiven Puls gesättigt. d) Spektrum mit selektiver Anregung (aus J. Magn. Reson. A, 112, 275 (1995)).

r_H die Korrelationszeit

$$\tau_c = \frac{4\pi\eta r_H^3}{3k_B T}$$

Hier stellt η die Viskosität des Lösungsmittels dar. Bei einem relativ kleinen Protein wie Ubiquitin ($M_r = 8400$) beträgt der hydrodynamische Radius $r_H \approx 16.5 \text{ \AA}$ und die Korrelationszeit $\tau_c = 3.8 \text{ ns}$. Dies ist bereits deutlich länger als die inverse Larmorfrequenz (14 T : $\omega_L = 3,8 \cdot 10^9 \text{ s}^{-1}$, d.h. $\tau_c \omega_L \approx 14$). Mit zunehmender Größe des Proteins nimmt die Korrelationszeit zu, die Effizienz des Mittelungsprozesses durch die Rotation nimmt ab und damit die Breite der Resonanzlinien zu.

Die Linienbreiten in den Spektren hängen u.a. auch von der Zubereitung der Proben ab: In welchem Medium und bei welcher Temperatur das Protein die besten Spektren ergibt und außerdem lange genug erhalten bleibt bis die Messdaten aufgenommen sind, muss praktisch für alle Proben neu bestimmt werden. Eine höhere Temperatur ergibt kürzere Korrelationszeiten; allerdings denaturieren alle Proteine bei zu hohen Temperaturen. Ein wichtiger Parameter ist der pH der Lösung: viele Proteine verhalten sich unterschiedlich in saurem, resp. basischen Medium.

6.2.3 Zuordnung der Resonanzen

Der erste Schritt der Strukturaufklärung besteht in der Zuordnung der Resonanzen. Erste Hinweise dafür erhält man aus der chemischen Verschiebung. Wie bereits erwähnt muss man dafür jedoch meistens mehrdimensionale Spektren verwenden.

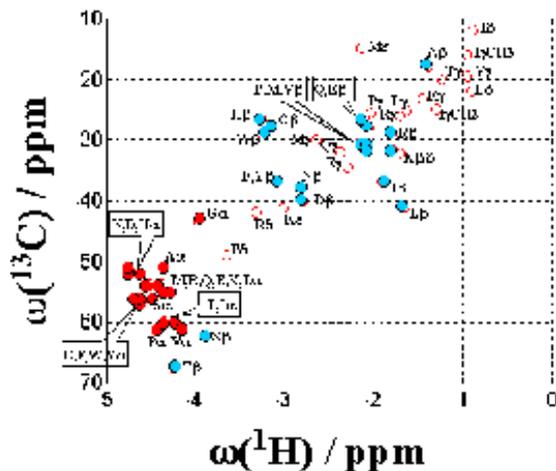


Abbildung 6.15: Chemische Verschiebungen der Aminosäuren im ^{13}C und ^1H NMR Spektrum. Die einzelnen Aminosäuren sind durch ihren 1-Buchstaben Code markiert.

In der Figur sind Richtwerte für die ^1H und ^{13}C Frequenzen für die einzelnen Aminosäuren aufgetragen. Im Idealfall kann man damit die Resonanzen einem bestimmten Aminosäuretyp zuordnen. Allerdings enthält ein Protein im Allgemeinen mehrere Aminosäuren des gleichen Typs an unterschiedlichen Positionen. Außerdem sind die unterschiedlichen Aminosäuren nicht klar getrennt.

Sehr nützlich für die Zuordnung ist z.B. der $\text{NH-C}_\alpha\text{H}$ Bereich des COSY Spektrums. Er ist relativ gut aufgelöst und ergibt von den meisten Aminosäuren genau eine Resonanzlinie. Ausnahmen sind Glyzin (1 C_αH Protonen) und Prolin (kein C_αH Proton).

6.2.4 Seitenketten

Die einzelnen Arten von Aminosäuren unterscheiden sich durch ihre Seitenketten. Es ist deshalb wich-

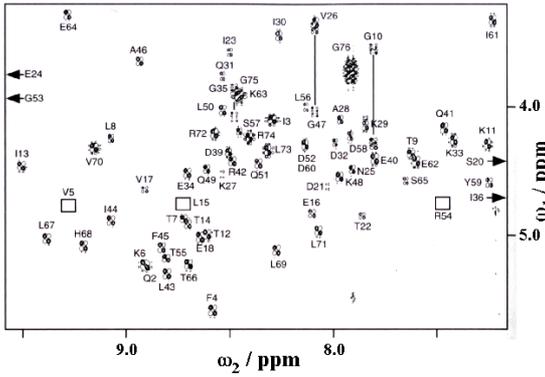


Abbildung 6.16: NH-C α H Bereich des COSY Spektrums von Ubiquitin. Die meisten der erwarteten 78 Linien sind zugeordnet.

tig (un meistens notwendig), diese in die Identifikation einzubeziehen. Da COSY Spektren die chemischen Bindungen zwischen den Kernen reflektieren kann man sie z.B. benutzen, um unterschiedliche Aminosäuren zu unterscheiden. Die meisten Kopplungen, die man in COSY Spektren verwendet, sind Kopplungen über 3 Bindungen, also z.B. H-C-C-H. Somit erhält man vor allem Kreuzpeaks zwischen Protonen, die an den gleichen, oder einen benachbarten Kohlenstoff (oder Stickstoff) gebunden sind.

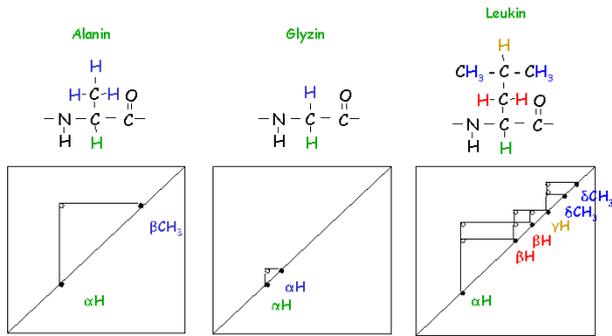


Abbildung 6.17: COSY Muster unterschiedlicher Aminosäuren.

In Figur 6.17 sind die resultierenden Muster für 3 unterschiedliche Aminosäuren gezeigt. Ausgangspunkt für die Zuordnung ist das α -CH Proton, welches entweder auf Grund seiner chemischen Verschiebung oder auf Grund einer Kopplung an ein N-H Proton identifiziert werden kann. Im Alanin ist das α -CH Proton an eine Methylgruppe gekoppelt, im Gly-

zin an ein weiteres α -CH Proton. Im dritten Beispiel (Leukin) ist die Seitenkette relativ lang. Das α -CH Proton ist hier an zwei unterscheidbare Protonen der β -CH₂ Gruppe gekoppelt, diese wiederum an ein einzelnes γ -CH und dieses wiederum an 2 CH₃ Gruppen.

Auch das NOESY Spektrum kann zur Identifikation der Linien genutzt werden.

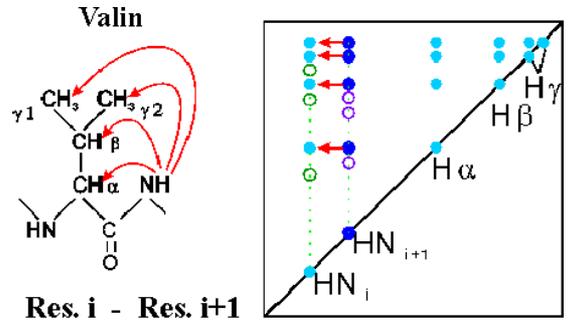


Abbildung 6.18: Identifizierung von Aminosäuren und Nachbarschaften mittels NOESY Kreuzpeaks.

In diesem Beispiel findet man Kontakte von einem NH Proton zu den verschiedenen Protonen im Seitenast der vorangehenden Aminosäure koppelt. Die Struktur (1 α , 1 β , 6 CH₃) zeigt dass es sich um ein Valin handelt.

6.2.5 Sequenzierung

Um die Sequenz der Aminosäuren zu verifizieren und die zugehörigen Resonanzen zu identifizieren kombiniert man die Informationen aus COSY und NOESY Spektren.

Kopplungen

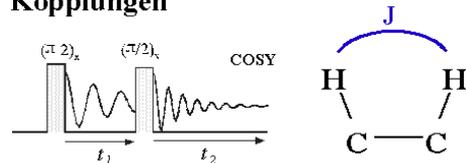


Abbildung 6.19: COSY Spektren liefern Informationen über Kopplungen.

Aus dem COSY Spektrum erhält man die Informationen über "chemisch" benachbarte Protonen,

d.h. über Protonen, welche durch nicht mehr als 3 Bindungen (meist H-C-C-H oder H-C-N-H) getrennt sind. Entlang der Hauptkette des Proteins sind dies die N-H Protonen der Amidgruppe, welche mit dem $\text{C}_\alpha\text{-H}$ Proton der gleichen Aminosäure koppeln.

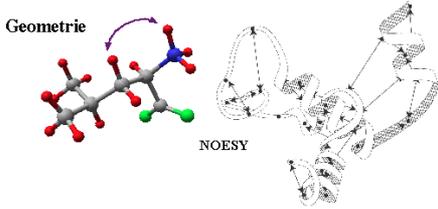


Abbildung 6.20: Geometrische Information aus dem NOESY Spektrum.

Aus dem NOESY Spektrum erhält man hingegen Informationen über die räumliche Distanz zwischen den Protonen. Dies betrifft einerseits die Distanz zwischen dem $\text{C}_\alpha\text{-H}$ Proton einer Aminosäure und dem N-H Proton der nächsten Aminosäure, andererseits die Distanz zwischen unterschiedlichen Bereichen der Kette, welche durch die räumliche Faltung benachbart sind.

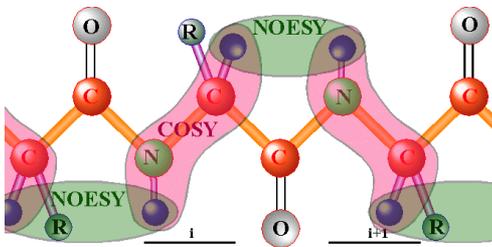


Abbildung 6.21: Sequenzierung durch Kombination von COSY (innerhalb einer AS) und NOESY (aufeinanderfolgende AS).

Aufgrund der räumlichen Nähe zwischen dem NH Proton einer Aminosäure mit der vorangehenden Aminosäure findet man meist einen Kreuzpeak im NOESY Spektrum. Damit ist es möglich, eine Verbindung zwischen zwei aufeinander folgenden Aminosäuren zu erstellen. Innerhalb der Aminosäure kann man wiederum mit Hilfe des COSY Spektrums die zusammen gehörigen CH und NH Protonen korrelieren.

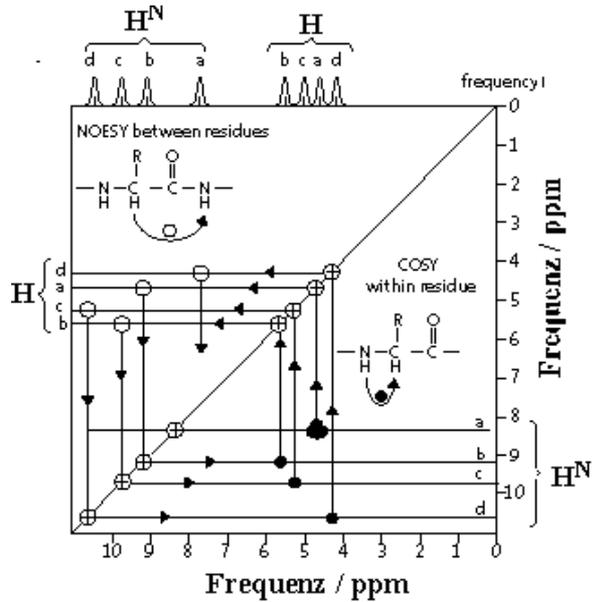


Abbildung 6.22: Kombination von NOESY und COSY Spektrum. Damit ist es prinzipiell möglich, die gesamte Aminosäuresequenz zu bestimmen.

Kombiniert man die Information eines NOESY Spektrums mit dem eines COSY Spektrums, so ist es prinzipiell möglich, die gesamte Aminosäuresequenz zuzuordnen: Das COSY Spektrum erlaubt die Messung der Kopplungen zwischen einem $\alpha\text{-CH}$ und dem NH in der gleichen Aminosäure. Das NOESY Spektrum verbindet das $\alpha\text{-CH}$ mit dem NH der nächsten Aminosäure. Diese beiden Protonen sind sich räumlich benachbart und ergeben deshalb einen Kreuzpeak im NOESY Signal. Wegen der dazwischen liegenden C=O Gruppe findet man jedoch keine messbare Kopplung, so dass im COSY Spektrum kein Signal erscheint.

Da die Spektren häufig zu voll sind verwendet man unterschiedliche Techniken, um den Überlapp der Signale zu reduzieren. Eine wichtige Technik, welche nur Signale von Gruppen von Spins (im Gegensatz zu isolierten Spins) übrig lässt, ist das Doppelquantenfilter. Wie hier am Beispiel von BPTI gezeigt werden damit die einzelnen Resonanzlinien getrennt sichtbar. Dargestellt ist der Bereich der NH-Protonen von BPTI.

Neben der Sequenz interessiert man sich vor al-

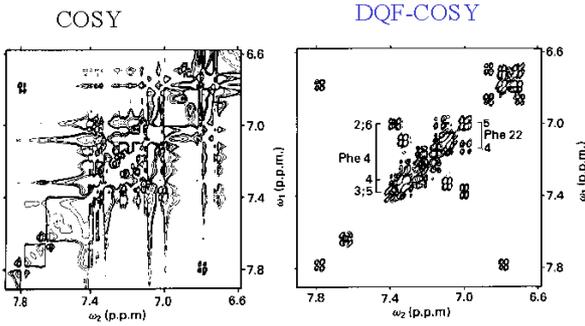


Abbildung 6.23: Vergleich eines gewöhnlichen COSY Spektrums (links) mit einem Doppelquanten-gefilterten COSY Spektrum (rechts).

lem für die Konformation der Kette. Diese wird im Wesentlichen durch die Diederwinkel bestimmt. Eine Möglichkeit, diese zu messen bieten die Kopplungskonstanten. Der Betrag der Kopplungskonstanten hängt ab vom Diederwinkel.

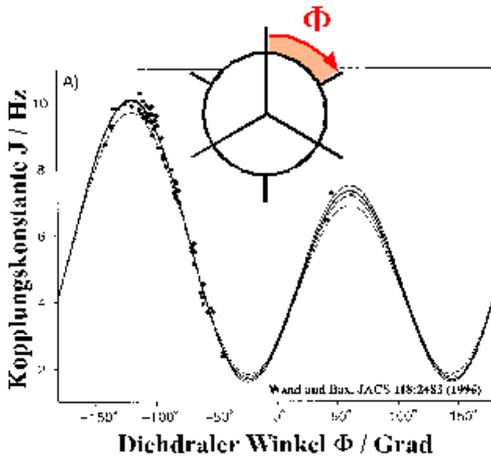


Abbildung 6.24: Kopplungskonstanten als Funktion des dihedralen Winkels in Ubiquitin.

Relevant sind vor allem Kopplungen zwischen 2 Protonen, welche über 3 Bindungen, in einem Fragment HCCH miteinander verbunden sind. Die nach Karplus benannte Beziehung sagt, dass die Kopplungskonstante vom Diederwinkel ϕ wie

$$J \approx A + B \cos \phi + C \cos 2\phi$$

abhängt. Offensichtlich ermöglicht dies i.A. kei-

ne eindeutige Bestimmung des Winkels. Es ist jedoch eine sehr nützliche Information, z.B. um mögliche Strukturen auszuschließen. Allgemein werden Strukturen nicht durch eine einzelne Größe bestimmt.

6.2.6 Heteronukleare Experimente: Empfindlichkeit

Bei größeren Proteinen ist es nicht möglich, die Resonanzen allein aufgrund der ^1H NMR Spektren zuzuordnen, da die Zahl der Resonanzlinien wie auch ihre Breite mit der Größe des Proteins zunimmt. Da Heterokerne einen wesentlich größeren Bereich an chemischen Verschiebungen aufweisen gelingt die Zuordnung mit Hilfe heteronuklearer Experimente noch für deutlich größere Moleküle.

Die interessantesten Heterokerne für Proteine sind ^{13}C und ^{15}N . Die natürliche Häufigkeit von ^{13}C beträgt rund 1.1 %, diejenige von ^{15}N 0.37 %. Damit wäre die Empfindlichkeit für große Moleküle deutlich zu klein. Man arbeitet deshalb ausschließlich mit markierten Molekülen, d.h. mit Molekülen, in denen die entsprechenden Isotope angereichert wurden. Dies trägt natürlich wesentlich zum Aufwand bei der Probenherstellung bei.

Für eine weitere Erhöhung der Empfindlichkeit wird meistens zuerst Magnetisierung von den ^1H auf die Heterokerne übertragen und die Detektion wiederum auf den Protonen durchgeführt. Die Polarisation der Kernspins ist proportional zum gyromagnetischen Verhältnis. Gegenüber ^{13}C ist damit die Polarisation der Protonen um einen Faktor 4 höher, gegenüber ^{15}N sogar um den Faktor 10. Bei der Detektion ist die Empfindlichkeit in etwa proportional zu $\gamma^{3/2}$. Damit gewinnt man bei ^{13}C etwa einen Faktor 8, bei ^{15}N etwa einen Faktor 30.

Das Standardexperiment im heteronuklearen Bereich ist das HSQC, welches vor allem verwendet wird, um NH-Protonen mit den entsprechenden ^{15}N Spins zu korrelieren.

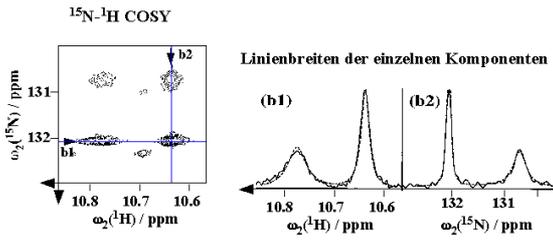


Abbildung 6.25: Unterschiedliche Linienbreiten in einem ^{15}N - ^1H COSY Experiment.

6.2.7 Auflösung und Linienbreite

Nicht alle Linien im Spektrum haben die gleiche Linienbreite. Es kann sich deshalb auszahlen, spezifisch nur eine bestimmte Linie in einem Multiplett anzuregen, um damit die Linienbreite zu minimieren, resp. die Auflösung zu maximieren. Das entsprechende Experiment hat das Akronym TROSY. Es basiert auf der Interferenz unterschiedlicher Relaxationsmechanismen, in diesem Fall Relaxation aufgrund der chemischen Verschiebungsanisotropie (CSA) und der Dipol-Dipol Wechselwirkung. Beides sind orientierungsabhängige Wechselwirkungen, welche durch die Molekülrotation im Mittel verschwinden, aber zur Relaxation beitragen. In einer Linie eines Dubletts addieren sich die beiden Wechselwirkungen, in der anderen heben sie sich gegenseitig auf. Im TROSY Experiment wird deshalb spezifisch die schmalste Linie angeregt, also diejenige, bei der die Beiträge zur Relaxation sich aufheben.

6.2.8 3D Experimente

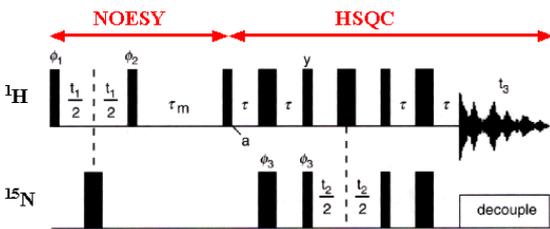


Abbildung 6.26: Pulsfolge für das 3D NOESY-HSQC Experiment.

Ein wichtiges 3D Experiment ist die Kombination NOESY-HSQC: Hier wird die Distanz zwischen

Protonen gemessen, wobei einer der beiden Partner durch die Korrelation mit einem Heterokern (z.B. ^{15}N) identifiziert, resp. die Linie aufgelöst wird.

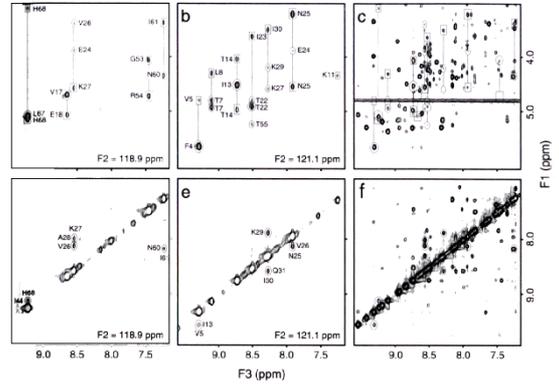


Abbildung 6.27: 2D Ausschnitte aus einem 3D NOESY-HSQC Spektrum von ubiquitin. Gezeigt sind die beiden ^1H Dimensionen, jeweils zu 2 unterschiedlichen ^{15}N Frequenzen. Rechts zum Vergleich das normale NOESY Spektrum.

Die entsprechenden Spektren zeigen eine deutlich bessere Auflösung und geringeren spektralen Überlapp. Insbesondere fehlt in den heteronuklearen Spektren die Wasserlinie, welche im NOESY Spektrum quer durch das Spektrum läuft.

Um ein 3D Spektrum aufzunehmen muss man pro indirekter Dimension eine große Zahl von einzelnen Experimenten durchführen. Dabei regt man meist nicht das ganze Spektrum, sondern nur einen Teilbereich an. Würde man pro indirekter Dimension z.B. 1024 Datenpunkte aufnehmen, so würde ein 3D Experiment mehrere Wochen dauern.

In der Figur sind einige Schemata für 3D Experimente zusammengestellt. Sie zeigen, welche Information aus den einzelnen Spektren gewonnen werden kann.

6.2.9 Dipolkopplungen

Die Dipolkopplung zwischen 2 Kernspins hängt ab von der Orientierung der Kern-Kern Verbindungs-

Experiment	Correlations observed	Magnetization transfer	J couplings ^a
HNCA	$^1\text{H}_i^N - ^{15}\text{N}_i - ^{13}\text{C}_i^\alpha$ $^1\text{H}_i^N - ^{15}\text{N}_i - ^{13}\text{C}_{i-1}^\alpha$		$^1J_{\text{NH}}$ $^1J_{\text{SC}^\alpha}$ $^2J_{\text{SC}^\alpha}$
HN(CO)CA	$^1\text{H}_i^N - ^{15}\text{N}_i - ^{13}\text{C}_{i-1}^\alpha$		$^1J_{\text{NH}}$ $^1J_{\text{SCO}}$ $^1J_{\text{C}^\alpha\text{CO}}$
H(CA)NH	$^1\text{H}_i^\alpha - ^{15}\text{N}_i - ^1\text{H}_i^N$ $^1\text{H}_i^\alpha - ^{15}\text{N}_{i+1} - ^1\text{H}_{i+1}^N$		$^1J_{\text{C}^\alpha\text{H}^\alpha}$ $^1J_{\text{NC}^\alpha}$ $^2J_{\text{NH}}$

Abbildung 6.28: Übersicht über einige 3D Experimente.

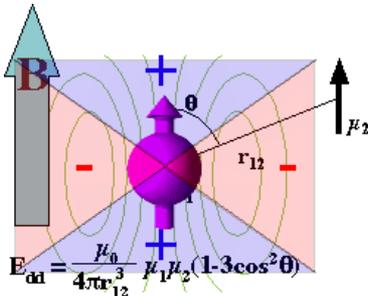


Abbildung 6.29: Abhängigkeit der Dipolkopplung von der Orientierung des internuklearen Vektors.

vektoren:

$$d = d_0(1 - 3\cos^2\theta).$$

Hier bezeichnet θ den Winkel des Kern-Kern Verbindungsvektors mit der z-Achse (der Magnetfeldrichtung).

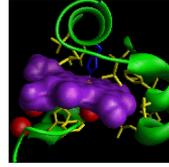
In Flüssigkeiten bewegen sich die Moleküle so schnell, dass man nur den Mittelwert dieser Größe beobachtet. In gewöhnlichen Flüssigkeiten ist die molekulare Bewegung im Normalfall isotrop, d.h. es wird über alle Orientierungen gemittelt. Da der isotrope Mittelwert

$$\int_0^\pi \sin\theta d\theta (1 - 3\cos^2\theta) = 0$$

verschwindet beobachtet man in isotropen Flüssigkeiten normalerweise keine Dipolkopplungen.

Nicht alle Proteine sind jedoch vollständig isotrop. Insbesondere paramagnetische Moleküle können eine erhebliche magnetische Anisotropie aufweisen,

magnetische Anisotropie des Proteins



anisotropes Medium



Abbildung 6.30: Anisotrope Mittelwerte können auf Grund einer anisotropen magnetischen Suszeptibilität des Moleküls oder seiner Umgebung entstehen.

welche dazu führt, dass das Molekül sich zwar frei bewegt, dass diese Bewegung im Mittelwert jedoch nicht vollständig isotrop ist. In diesem Fall tragen die Dipolkopplungen zu den im Spektrum gemessenen Aufspaltungen bei. Man kann diesen Effekt auch bei Proteinen mit einer geringen magnetischen Anisotropie erreichen wenn man sie in ein anisotropes Medium gibt. Dafür kann man Flüssigkristalle oder Viren (Phagen) verwenden.

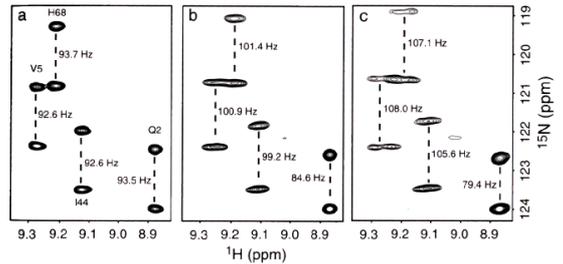


Abbildung 6.31: ^1H - ^{15}N HSQC Spektren von Ubiquitin. Links: Isotropes Lösungsmittel; Mitte, rechts: zunehmend anisotropes LM.

Misst man das Spektrum in einem anisotropen Lösungsmittel, so findet man, dass die gemessenen Kopplungen von der Stärke der Anisotropie abhängen. Man kann die Aufspaltungen als Summe der isotropen Kopplungskonstanten J und der Dipol-Dipol Kopplungskonstanten d betrachten:

$$\delta_{obs} = J + d.$$

Dabei muss in beiden Fällen das Vorzeichen berücksichtigt werden, welches positiv oder negativ sein kann. Die gemittelte Dipolkopplung ist gegeben

durch das Produkt der statischen Dipolkopplungskonstante mit dem Zeitmittelwert der Orientierungsabhängigkeit,

$$d = d_0 \langle 1 - 3 \cos^2 \theta \rangle.$$

Hierbei spielt nicht nur die Stärke der Anisotropie eine Rolle, sondern auch die Lage der relevanten Bindung im Molekül. Aus der Messung der "residual dipolar couplings" (RDCs) kann man deshalb sehr viel Information über die Struktur und Dynamik des Moleküls gewinnen.

Damit gelingt es einem prinzipiell, die Sekundärstruktur des Proteins zu bestimmen. Für die Tertiärstruktur (also die Faltung) benötigt man weitere NOEs, diesmal nicht zwischen aufeinanderfolgenden Aminosäuren, sondern Querverbindungen zu (in der Sequenz) weiter entfernten.

6.3 Anwendungsbeispiele

6.3.1 BPTI: Primär- und Sekundärstruktur

Die ersten Experimente zur Strukturbestimmung von Proteinen mit NMR wurden zu einem wesentlichen Teil an BPTI durchgeführt. Das Molekül hat den Namen Aprotinin oder bovine pancreatic trypsin inhibitor (=BPTI), welches als Medikament verwendet wird (um Blutungen bei Operationen zu mindern). Es besteht aus 58 Aminosäuren, hat die Summenformel $C_{284}H_{432}N_{84}O_{79}S_7$ und damit ein Molekulargewicht von 6511.

In Figur 6.32 ist als Beispiel die Zuordnung der Resonanzen von BPTI dargestellt. Dafür wurden das COSY und das NOESY Spektrum in ein Spektrum kombiniert, wobei das linke obere Dreieck das NOESY Spektrum darstellt, das rechte untere das COSY Spektrum. Diese Darstellungsweise erlaubt eine relativ einfache Sequenzanalyse: wie durch die Linien markiert ist gelangt man jeweils über ein Viereck von einer Aminosäure zur nächsten.

Im Bereich von 7-10 ppm sind die Amid-Protonen im Spektrum zu finden, während die CH-Protonen im Bereich von 4-5 ppm zu finden sind. Man kann

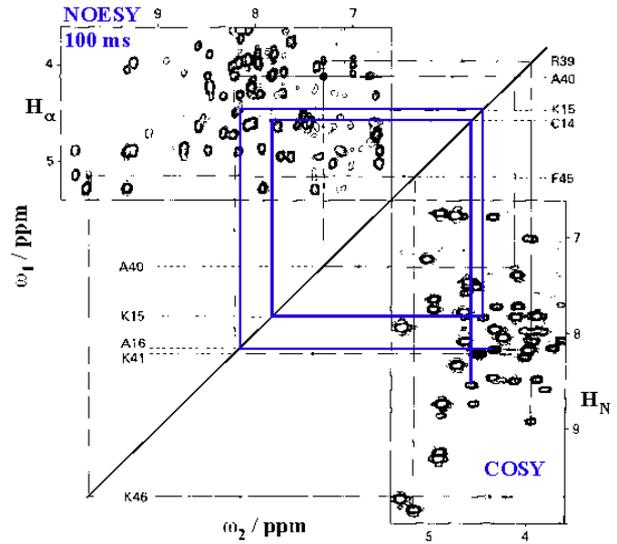


Abbildung 6.32: Kombinationsspektrum (NOESY + COSY) von BPTI, mit Markierung einiger Zuordnungen.

z.B. bei 8.2 ppm mit dem NH Proton der AS 16 beginnen (A16 = Alanin an Pos. 16). Der Kreuzpeak im NOESY Spektrum verbindet diesen zum C_{α} Proton von K15 (Lysin) bei 4.4 ppm und das COSY Spektrum wiederum zum zugehörigen NH bei 7.8 ppm. Das NOESY führt uns dann wieder zu C14 (C=Cystein), usw.

Das NOESY Spektrum kann auch für semi-quantitative Distanzmessungen verwendet werden. Im Kapitel 5 hatten wir für Austauschexperimente allgemein die Zeitabhängigkeit über die Differentialgleichung

$$\frac{d}{dt}A_z = -k(A_z - X_z) - A_z/T_1$$

beschrieben, wobei wir ein Zweispinsystem angenommen hatten. Im Fall des NOESY Experimentes ist $k \propto r^{-6}$, d.h. invers proportional zur 6ten Potenz des Abstandes. Allerdings haben wir es in Proteinen mit weit mehr als 2 Spins zu tun. Diese Gleichung ist somit nur Teil eines Systems von mehreren hundert Differentialgleichungen.

In Figur 6.33 sind die gemessenen Zeitabhängigkeiten für einige Aminosäuren von BPTI dargestellt. Ganz Links sind die Signale für das N-H Proton von

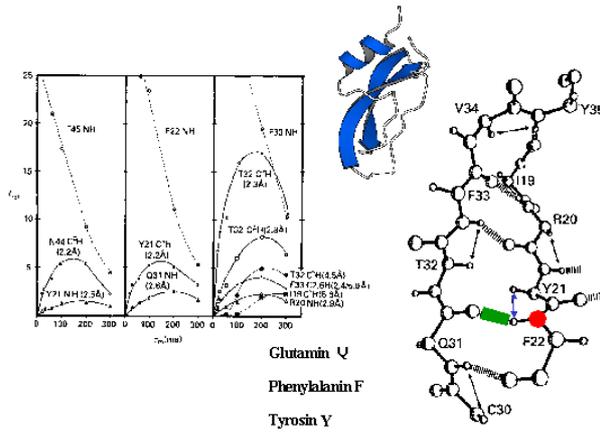


Abbildung 6.33: β -Faltblatt von BPTI: Zeitabhängigkeit einiger Kreuzpeaks mit zugehörigen Distanzen. (Aus Ernst et al. [1], Fig. 9.7.7.)

F22 (in der Struktur rot markiert) und rechts daneben für F33 (ebenfalls rot markiert) aufgetragen. In der Struktur sind für jeweils 2 der Kurven die Verbindungen markiert. Die Diagonalpeaks haben die maximale Amplitude zu Beginn der Mischzeit, während die Kreuzpeaks mit Amplitude 0 beginnen, aufgrund der Kreuzrelaxation wachsen und schließlich aufgrund der Relaxation wieder verschwinden. Man erkennt zwei Arten von Kreuzpeaks: diejenigen welche für niedrige Mischzeiten linear anwachsen und andere die zunächst eine verschwindende Steigung aufweisen. Hierbei handelt es sich um Transferprozesse höherer Ordnung, d.h. mehrstufige Transferprozesse.

Neben den Resonanzen, welche benachbarte Aminosäuren verbinden findet man auch einige Linien, welche zwei weiter auseinander liegende Aminosäuren verbinden. In diesem Fall lassen sie sich erklären wenn man annimmt, dass dieser Teil des Proteins ein β -Faltblatt bildet.

6.3.2 Tertiärstruktur

Die Tertiärstruktur erhält man im Allgemeinen erst wenn eine sehr große Zahl von NOE Kontakten gefunden wurden.

In diesem Beispiel sind einige der insgesamt 3980 Di-

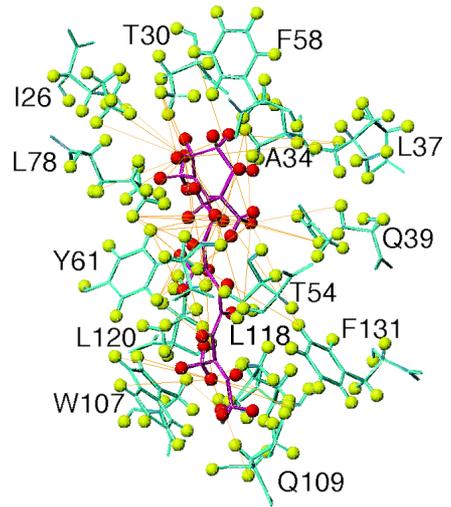


Abbildung 6.34: Struktur von "retinol binding protein" mit NOE distance constraints.

stanzen eingezeichnet, welche aus NOESY Experimenten bestimmt wurden. Es handelt sich um das "retinol binding protein II"; dargestellt ist der Teil des Proteins, der für die Bindung verantwortlich ist, zusammen mit dem Liganden Retinol (Vitamin A). Dargestellt sind diejenigen Distanzen, welche zwischen dem Protein (gelb) und dem Liganden (rot) bestimmt wurden. Hier wurde somit nicht nur die Tertiärstruktur des Proteins selber bestimmt, sondern die gesamte Struktur des Proteins mit dem gebundenen Liganden.

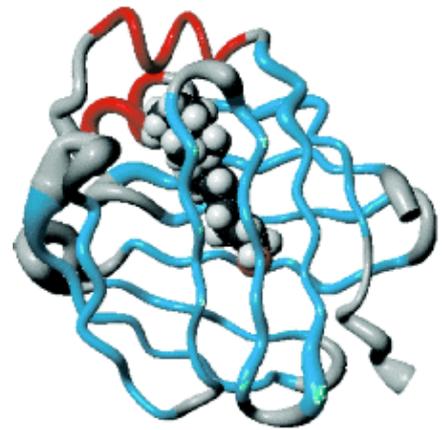


Abbildung 6.35: Gesamtstruktur des RBP.

Die zweite Figur stellt das gesamte Protein dar. Der blaue Strang stellt das Rückgrat des Proteins dar, das

Retinol befindet sich im Inneren und wird von insgesamt zehn Strängen eines β -Faltblattes umgeben.

Die NMR Messungen werden immer mit Simulationsrechnungen (v.a. Molekulardynamik) kombiniert, um zuverlässige Aussagen zu erhalten. Dafür werden die aus der NMR ermittelten Distanzen als zusätzliches Pseudopotenzial in die Rechnung gesteckt.

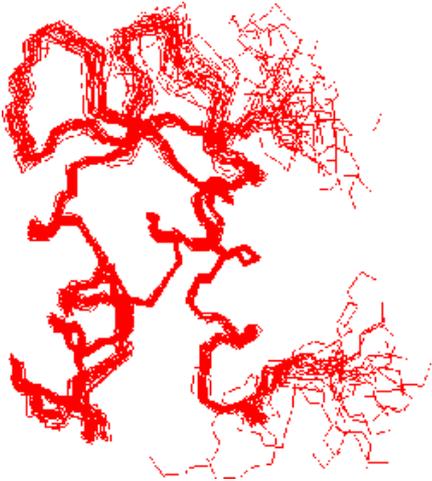


Abbildung 6.36: Überlagerung verschiedener möglicher Konformationen.

In der Figur ist der Verlauf der Hauptkette als Linie dargestellt. Jede einzelne Linie entspricht dem Resultat einer Simulationsrechnung. In vielen Fällen stellt man dann fest, dass nicht für das gesamte Molekül eine eindeutige Lösung erhalten wird. Dies deutet darauf hin, dass in der Lösung Teile des Moleküls relativ flexibel sind.

Neben den rein statischen Informationen sind auch dynamische Informationen sehr wichtig. So kann man nicht nur die Struktur untersuchen, sondern auch Bewegungsprozesse, z.B. als Antwort auf externe Stimuli wie z.B. chemische Substanzen oder Licht. Auch die Wechselwirkung zwischen unterschiedlichen Molekülen, z.B. zwischen Enzym und Substrat liefert wertvolle Informationen.

6.4 Festkörper

6.4.1 Grundlagen

Die meisten NMR Experimente an Proteinen werden in Lösung durchgeführt, weil hier die spektrale Auflösung optimiert wird. Außerdem ist Wasser für die meisten interessanten Proteine die natürliche Umgebung.

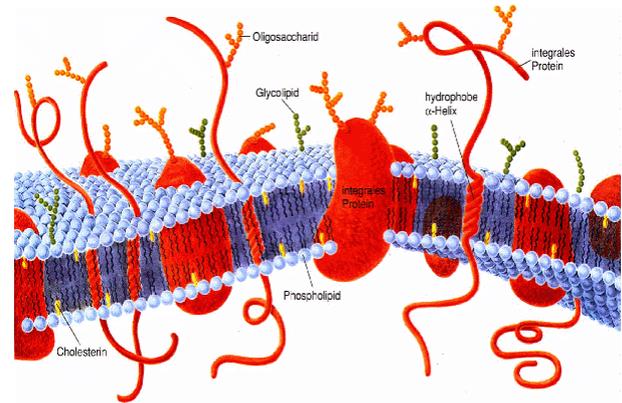


Abbildung 6.37: Membranproteine.

Dies trifft jedoch nicht für alle Proteine zu. Insbesondere Transmembran-Proteine, also Proteine, welche in eine biologische Membran eingebettet sind. Diese sind stark lipophil und lösen sich deshalb in Wasser nicht.

Moleküle in einem Festkörper geben meist sehr breite NMR Spektren (mehrere kHz bis mehrere 10 kHz) so dass es meist schwierig ist, die Resonanzen von einzelnen Kernen zu identifizieren. Dies ist auf die starke Dipol-Dipol Kopplung zurückzuführen, sowie auf die Überlagerung von vielen unterschiedlichen Orientierungen in einem Pulver.

Die chemische Verschiebung ist grundsätzlich anisotrop, d.h. sie hängt ab von der Orientierung des Moleküls bezüglich dem Magnetfeld:

$$\delta\omega_L = a(3\cos^2\theta - 1).$$

Hier bezeichnet θ den Winkel zwischen der Hauptachse des CSA Tensors (CSA = chemical shift anisotropy) und dem Magnetfeld. In einem Pulver kommen alle Orientierungen vor und das beobachtete

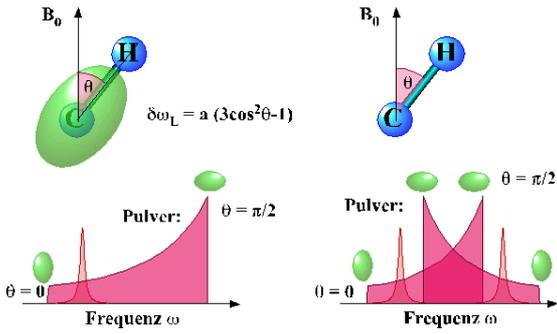


Abbildung 6.38: Pulverspektren für chemische Verschiebungsanisotropie (links) und Dipol-Kopplung (rechts).

Spektrum entspricht einer Mittelung über alle Orientierungen.

Das gleiche gilt für die Dipol-Dipol Wechselwirkung: Sie hat die gleiche Orientierungsabhängigkeit. Das Pulvermittel für die beiden Linien eines Dipol-aufgespaltenen Spektrums entspricht einer Überlagerung von zwei CSA-Mustern.

6.4.2 Magisch Winkel Rotation (MAS)

In einer Flüssigkeit beobachtet man diese Verbreiterungen nicht, weil die Moleküle schnell rotieren und dadurch ein Bewegungsmittelwert

$$\int_0^\pi d\theta \sin\theta (3\cos^2\theta - 1) \int_0^{2\pi} d\phi = 0$$

gebildet wird. Eine isotrope Mittelung technisch durchzuführen ist schwierig. Es reicht aber, wenn man statt dessen die Probe um eine geeignete Achse rotiert. Bei einer Rotation um eine Achse, die gegenüber dem Magnetfeld um den Winkel θ_1 geneigt ist, wird der Orientierungsterm auf den Betrag

$$\langle 3\cos^2\theta - 1 \rangle = \frac{1}{2}(3\cos^2\theta_1 - 1)(3\cos^2\theta_2 - 1)$$

gemittelt. Hier stellt θ_2 den Winkel zwischen der Tensorhauptachse des Spins und der Rotationsachse dar. Dieser Ausdruck verschwindet offensichtlich dann, wenn

$$\theta_1 = \cos^{-1} \frac{1}{\sqrt{3}} = \theta_m \approx 54.7^\circ.$$

Rotiert man die Probe um diesen Winkel, so verschwinden alle anisotropen Wechselwirkungen. Man nennt ihn deshalb den magischen Winkel. Er entspricht der Raumdiagonalen in einem Würfel.

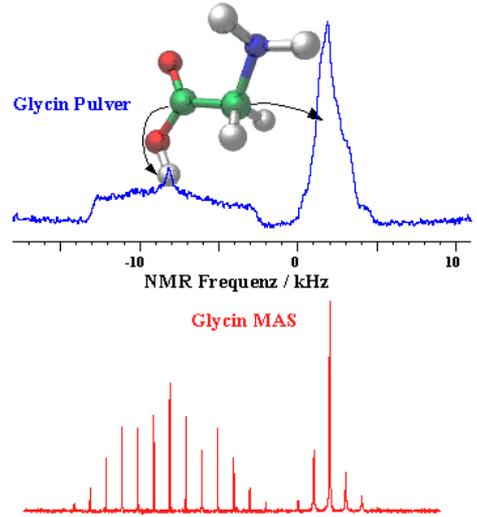


Abbildung 6.39: Statisches Pulverspektrum von Glycin (oben) und MAS Spektrum (unten).

Wie in der Figur gezeigt erhält man bei einer Rotation um den magische Winkel (Magic Angle Spinning, MAS) sehr schmale Linien, im Idealfall vergleichbar mit den Linien in einer Flüssigkeit. Dafür muss die Rotationsgeschwindigkeit genügend hoch sein. Typische Werte sind heute 10-30 kHz, mit Spitzenwerten bis 80 kHz. Die erreicht man, indem man die Probe in einen Rotor mit Turbinenschaufel füllt, welche von einem Luft- oder Stickstoffstrom mit hoher Geschwindigkeit angetrieben wird.

Die Entwicklung der MAS Technik erhöhte für die Festkörper-NMR die Auflösung ebenso drastisch wie die Einführung höherer Felstärken.

Damit ist es heute möglich, Festkörper NMR Spektren zu erhalten, welche eine ähnlich hohe Auflösung aufweisen wie entsprechende Flüssigkeitsspektren. Für die Untersuchung von Proteinen in Festkörpern wird ausschließlich MAS NMR verwendet.

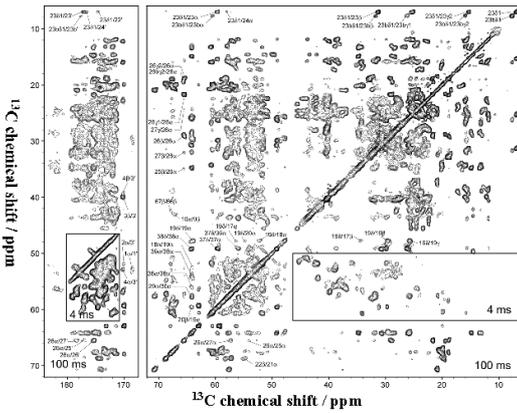


Abbildung 6.40: Festkörperspektrum mit Hilfe von MAS: ^{13}C - ^{13}C -Spindiffusionsspektrum.

6.4.3 Seide

Es gibt Proteine, welche natürlicherweise in fester Form vorliegen. Dazu gehört z.B. Seide, welche als halb-kristallines Material vorliegt. Seide ist eines der Materialien mit der größten Reißfestigkeit - höher als Stahl. Um diese besonderen Eigenschaften besser zu verstehen werden Daten über die Struktur benötigt.

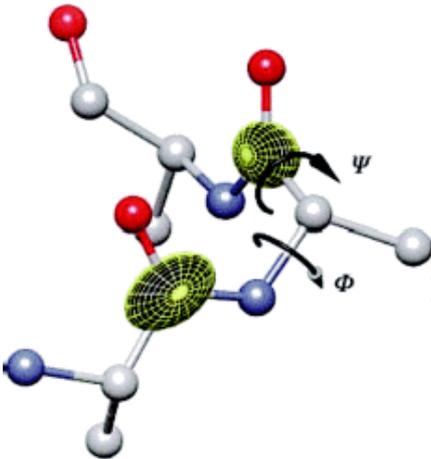


Abbildung 6.41: Aminosäuren in Seide

Die Sequenz der Seide ist bekannt, nicht aber ihre sekundäre und tertiäre Struktur. Spinnen erzeugen aus dem gleichen Ausgangsmaterial unterschiedliche Arten von Seide, z.B. besonders reissfeste oder besonders elastische Fäden.

Der Unterschied zwischen diesen Arten liegt offenbar in der Anordnung der Proteinfäden im Festkörper. Somit wird wesentlich durch die Rotation um C-C und C-N Bindungen bestimmt. Sie kann z.B. bestimmt werden, wenn man die relative Orientierung von benachbarten C=O Kohlenstoffen bestimmt. Beugungsmethoden helfen dabei kaum weiter, da die langreichweitige Ordnung sehr gering ist.

Hingegen ist es möglich, mit NMR die relative Orientierung der CSA-Tensoren zu bestimmen (J.D.v. Beek, L. Beaulieu, H. Schäfer, M. Demura, T. Asakura, and B.H. Meier, Nature 405, 1077 (2000)). Da diese immer an die chemischen Bindungen gekoppelt sind erhält man daraus die gewünschte Information.

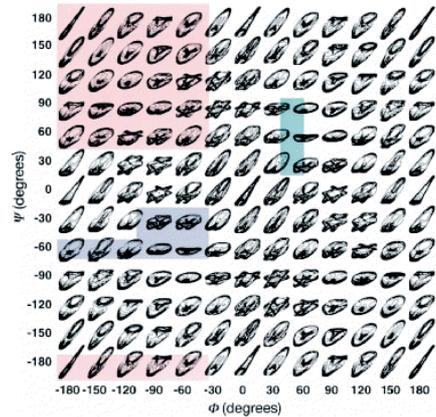


Abbildung 6.42: Mögliche 2D Spektren für unterschiedliche Orientierungen der beiden dihedralen Winkel ψ und ϕ .

Die Figur zeigt die erwarteten Pulverspektren als Funktion der beiden Winkel ψ und ϕ . Die farbige unterlegte Bereiche sind energetisch günstig (gemäß MD-Simulationen).

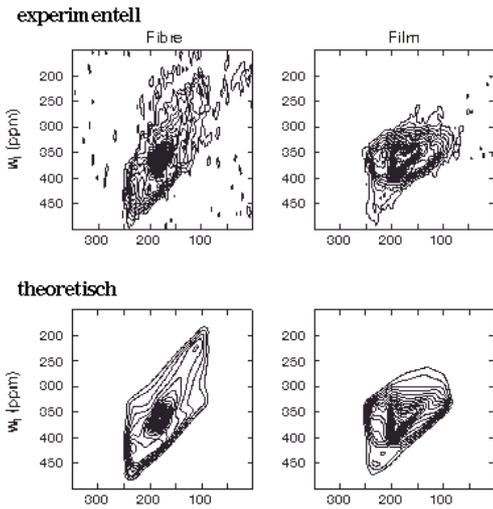


Abbildung 6.43: Vergleich von gerechneten und gemessenen DOQSY Spektren von Spinnenseide.

Vergleicht man die gemessenen Spektren (DOQSY = Double Quantum Spectroscopy) mit Spektren, welche für bestimmte dihedrale Winkel gerechnet wurden, so kann man den Bereich der möglichen Winkel stark einschränken. In Kombination mit Kraftfeldrechnungen kann die vorliegende Orientierungsverteilung bestimmt werden.

Die beste Übereinstimmung wird für die Wertepaare $(-140^\circ, 35^\circ)$ und $(-57^\circ, -47^\circ)$ erhalten.

6.4.4 Lichtinduzierte Konformationsänderung

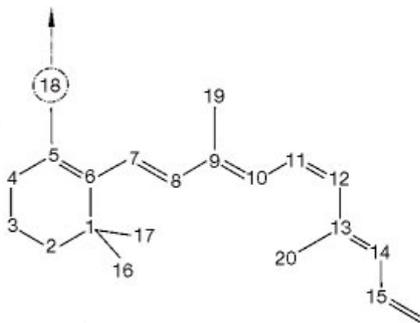


Abbildung 6.44: Struktur von Retinal

Ein weiteres Beispiel für die Möglichkeiten der Festkörper NMR für die Untersuchung von großen Bio-

molekülen ist die Untersuchung der Aktivierung von Retinal im photosynthetischen Komplex Rhodopsin. Hier wird nicht das Protein, sondern ein darin eingelagertes kleines Molekül, welches für die Funktion des Proteins eine entscheidende Bedeutung hat, untersucht. Retinal weist eine Reihe von benachbarten Doppelbindungen auf, welche die planare Struktur stabilisieren und die Absorption von Licht ermöglichen. Es war bereits bekannt, dass sich das Molekül bei der Absorption bewegt.

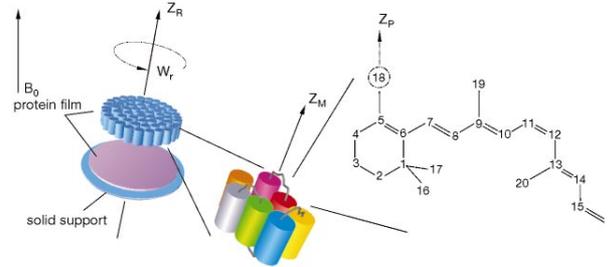


Abbildung 6.45: Orientierte Proben

Um die Art der Umordnung zu bestimmen wurden orientierte Proben hergestellt, in der jeweils eine der Methylgruppen 18, 19 oder 20 deuteriert waren.

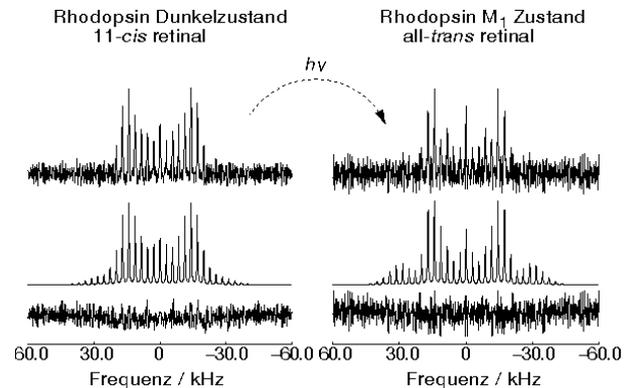


Abbildung 6.46: MAS Spektren von Retinal.

Das Muster der MAS Seitenbanden hängt von der Orientierung der C-C(2H)3 Bindung bezüglich der Rotationsachse ab. Die gemessenen Spektren wurden verglichen mit Rechnungen.

Aus dem Vergleich der gerechneten mit den gemessenen Spektren können die Winkel bestimmt werden. In der Figur ist dargestellt wie die Spitze des Moleküls sich während der Belichtung bewegt.

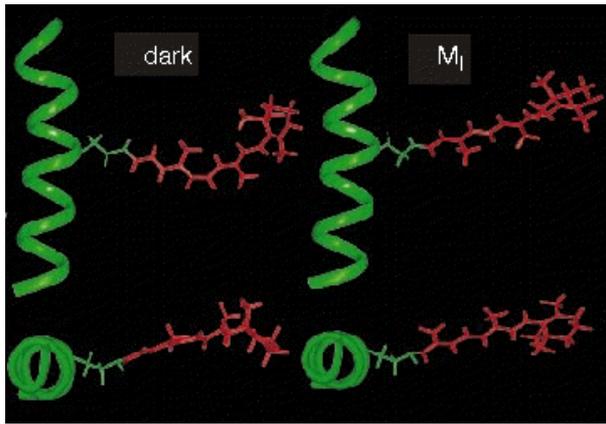


Abbildung 6.47: Struktur des Dunkelzustandes (links) und des aktivierten Zustandes (rechts) aus zwei Blickrichtungen.