

# Sample preparation of membrane proteins suitable for solid-state MAS NMR and development of assignment strategies

Dissertation zur Erlangung des akademischen Grades des  
Doktors der Naturwissenschaften (Dr. rer. Nat.)

eingereicht an der Mathematisch-Naturwissenschaftliche Fakultät  
der Universität Potsdam  
Institut für Biologie und Biochemie

vorgelegt von Matthias Hiller

Diplom-Ingenieur (FH) für Biotechnologie

aus Bautzen

April 2009

Published online at the  
Institutional Repository of the University of Potsdam:  
URL <http://opus.kobv.de/ubp/volltexte/2009/3724/>  
URN urn:nbn:de:kobv:517-opus-37246  
[<http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-37246>]

## Abbreviations

## Summary

## Zusammenfassung

<b>1 General Introduction</b>	<b>1</b>
1.1 Approaches for structural studies of membrane proteins	2
1.1.1 Membrane protein sample preparation	2
1.1.2 Three dimensional (3D) crystallisation and X-ray crystallography	4
1.1.3 Two-dimensional (2D) crystallisation and electron microscopy (EM)	5
1.1.4 Solution- and solid-state NMR	8
1.2 Fundamentals of nuclear magnetic resonance (NMR)	12
1.2.1 Solid-state MAS NMR	15
1.2.2 Basic solid-state MAS NMR techniques	17
1.2.3 Structure determination	21
1.3 The outer membrane protein G (OmpG) from <i>E.coli</i>	24
1.4 Overview of the thesis	30
<b>2 Material and Methods</b>	<b>32</b>
2.1 Materials	32
2.1.1 <i>E. coli</i> expression and cloning strains	32
2.1.2 Expression-plasmid	32
2.1.3 Isotopically labelled substances	33
2.2 Methods	34
2.2.1 Expression of labelled and unlabelled OmpG in <i>E. coli</i> BL21(DE3)	34
2.2.2 Protein purification from inclusion bodies	37
2.2.3 OmpG refolding	37
2.2.4 Reconstitution into lipid bilayers and two dimensional crystallisation	39
2.2.5 Protein PEG-precipitation	39
2.2.6 Circular dichroism (CD) spectroscopy	40
2.2.7 Electron microscopy	40
2.2.8 Solid-state MAS NMR spectroscopy	40

<b>3 Results and Discussion</b>	<b>44</b>
3.1 Expression and purification of isotopically labelled OmpG	44
3.1.1 Expression and purification of OmpG from inclusion bodies	44
3.1.2 Refolding of OmpG	46
3.1.3 Improvement of recombinant OmpG expression	49
3.1.4 Modified expression protocols or HCDF can save costs for isotopically labelled material	53
3.2 Preparation of OmpG samples suitable for solid-state MAS NMR investigations	55
3.2.1 2D crystals of OmpG	55
3.2.2 Proteoliposomes of OmpG	59
3.2.3 PEG precipitate	61
3.2.4 Is there a perfect sample-preparation technique?	64
3.3 Assignment of solid-state MAS NMR spectra of OmpG	69
3.3.1 [2,3- <sup>13</sup> C]-labelling of aromatic residues – getting a head start in the assignment of OmpG	69
3.3.2 Assignment using spectral editing techniques: selection of methyl groups	82
3.3.3 Enhancements of the assignment using other specifically labelled samples	85
3.3.4 Secondary structure elements calculated from chemical shift data	93
<b>4 Conclusions and Outlook</b>	<b>96</b>
<b>5 References</b>	<b>101</b>
<b>6 Appendix</b>	<b>112</b>
6.1 <sup>13</sup> C and <sup>15</sup> N chemical shifts of OmpG derived from solid-state MAS NMR experiments	112
6.2 Publications	115
6.3 Curriculum vitae	117
<b>Danksagung</b>	<b>118</b>

## Abbreviations

1,2 and 3D	one, two and three dimensional
Å	Ångström
ADP/ATP	adenosine diphosphate / adenosine triphosphate
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
cmc	critical micellar concentration
CP	cross polarisation
CSA	chemical shift anisotropy
DARR	dipolar assisted rotational resonance
DDM	dodecyl- $\beta$ -D-maltoside
<i>E. coli</i>	<i>Escherichia coli</i>
EM	electron microscopy
ER	endoplasmatic reticulum
FID	free induction decay
HCDF	high cell density fermentation
Hz	Hertz
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
K	Kelvin
kDa	kilo dalton
LDAO	Lauryl-N,N-dimethyldodecylamine-N-oxide
LPR	lipid to protein ratio
MAS	magic angle spinning
MP	membrane protein
nAChR	nicotinic acetylcholine receptor
NMR	nuclear magnetic resonance
OD <sub>600</sub>	optical density measured at 600 nm
OG	octyl- $\beta$ -D-glycopyranoside
OmpG	outer membrane protein G
Omps	Outer membrane porins
PCR	polymerase chain reaction
PDB	Protein Data Bank
PDSF	Proton driven spin diffusion

PEG	polyethylene glycol
ppm	parts per million
REDOR	rotational-echo double resonance
Rf	radio frequency
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SH3	Src-homology 3
SPINAL	small phase incremental alternation
SRP	Signal recognition particle
TEDOR	transferred-echo double resonance
TPPM	two pulse phase modulation
TROSY	transverse relaxation optimised spectroscopy

### Single- and three-letter amino acid code

Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Metionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophane	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

## Summary

Although the basic structure of biological membranes is provided by the lipid bilayer, most of the specific functions are carried out by membrane proteins (MPs) such as channels, ion-pumps and receptors. Additionally, it is known, that mutations in MPs are directly or indirectly involved in many diseases. Thus, structure determination of MPs is of major interest not only in structural biology but also in pharmacology, especially for drug development. Advances in structural biology of membrane proteins (MPs) have been strongly supported by the success of three leading techniques: X-ray crystallography, electron microscopy and solution NMR spectroscopy. However, X-ray crystallography and electron microscopy, require highly diffracting 3D or 2D crystals, respectively. Today, structure determination of non-crystalline solid protein preparations has been made possible through rapid progress of solid-state MAS NMR methodology for biological systems. Castellani *et. al.* solved and refined the first structure of a microcrystalline protein using only solid-state MAS NMR spectroscopy. These successful application open up perspectives to access systems that are difficult to crystallise or that form large heterogeneous complexes and insoluble aggregates, for example ligands bound to a MP-receptor, protein fibrils and heterogeneous proteins aggregates. Solid-state MAS NMR spectroscopy is in principle well suited to study MP at atomic resolution.

In this thesis, different types of MP preparations were tested for their suitability to be studied by solid-state MAS NMR. Proteoliposomes, poorly diffracting 2D crystals and a PEG precipitate of the outer membrane protein G (OmpG) were prepared as a model system for large MPs. Results from this work, combined with data found in the literature, show that highly diffracting crystalline material is not a prerequisite for structural analysis of MPs by solid-state MAS NMR. Instead, it is possible to use non-diffracting 3D crystals, MP precipitates, poorly diffracting 2D crystals and proteoliposomes. For the latter two types of preparations, the MP is reconstituted into a lipid bilayer, which thus allows the structural investigation in a quasi-native environment. In addition, to prepare a MP sample for solid-state MAS NMR it is possible to use screening methods, that are well established for 3D and 2D crystallisation of MPs. Hopefully, these findings will open a fourth method for structural investigation of MP.

The prerequisite for structural studies by NMR in general, and the most time consuming step, is always the assignment of resonances to specific nuclei within the protein. Since the last few years an ever-increasing number of assignments from solid-state MAS NMR of uniformly carbon and nitrogen labelled samples is being reported, mostly for small proteins of up to around 150 amino acids in length. However, the complexity of the spectra increases with increasing molecular weight of the protein. Thus the conventional assignment strategies developed for small proteins do not yield a sufficiently high degree of assignment for the large MP OmpG (281 amino acids).

Therefore, a new assignment strategy to find starting points for large MPs was devised. The assignment procedure is based on a sample with [2,3- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labelled Tyr and Phe and uniformly labelled alanine and glycine. This labelling pattern reduces the spectral overlap as well as the number of assignment possibilities. In order to extend the assignment, four other specifically labelled OmpG samples were used. The assignment procedure starts with the identification of the spin systems of each labelled amino acid using 2D  $^{13}\text{C}$ - $^{13}\text{C}$  and 3D NCACX correlation experiments. In a second step, 2D and 3D NCOCX type experiments are used for the sequential assignment of the observed resonances to specific nuclei in the OmpG amino acid sequence. Additionally, it was shown in this work, that biosynthetically site directed labelled samples, which are normally used to observe long-range correlations, were helpful to confirm the assignment.

Another approach to find assignment starting points in large protein systems, is the use of spectroscopic filtering techniques. A filtering block that selects methyl resonances was used to find further assignment starting points for OmpG.

Combining all these techniques, it was possible to assign nearly 50 % of the observed signals to the OmpG sequence. Using this information, a prediction of the secondary structure elements of OmpG was possible. Most of the calculated motifs were in good agreement with the crystal structures of OmpG.

The approaches presented here should be applicable to a wide variety of MPs and MP-complexes and should thus open a new avenue for the structural biology of MPs.



## Zusammenfassung

Obwohl biologische Membranen hauptsächlich aus Lipiden bestehen, wird deren biologische Funktion durch die in sie eingebetteten Membranproteine (z.B. Kanäle, Ionenpumpen und Rezeptoren) bestimmt. Mutationen in dieser Proteinklasse können direkt oder indirekt zum Auftreten verschiedener Krankheitsbilder führen. Die Untersuchung der dreidimensionalen Struktur von Membranproteinen ist daher nicht nur strukturbiochemisch, sondern auch von pharmakologischen Interesse. Die Fortschritte in der Strukturaufklärung von Membranproteinen sind im wesentlichen auf die Anwendung dreier Techniken zurückzuführen, der Röntgenkristallographie, der Elektronen-mikroskopie und der Lösungs-NMR. Für die Strukturuntersuchung mittels Röntgenkristallographie und Elektronenmikroskopie sind jedoch hochaufgelöste 3 bzw. 2 dimensional Kristalle erforderlich. In den letzten Jahren wurde eine weitere Methode, die Festkörper NMR Spektroskopie, für Strukturuntersuchungen an Proteinproben im festen Aggregatzustand entwickelt. Castellani *et. al.* haben als erste die Struktur eines mikrokristallinen Proteins mit Hilfe der Festkörper NMR Spektroskopie gelöst. Diese Technik ist vor allem deshalb so erfolgreich, weil sie Zugang zu Proteinsystemen ermöglicht, die nicht kristallisieren, große heterogene Komplexe oder unlösliche Aggregate bilden. Geeignete Systeme für Festkörper NMR Untersuchungen sind beispielsweise an Rezeptoren gebundene Liganden, Protein-Fibrillen oder heterogene Biomoleküle. Diese Methodik damit im Prinzip auch für die Strukturuntersuchung von Membranproteinen auf atomarer Ebene nutzbar.

Dieser Arbeit beschäftigt sich mit der Herstellung verschiedener Membranproteinpräparationen, welche eine Aufnahme von hochaufgelösten Festkörper NMR Spektren erlauben. Dafür wurde das Protein G der äußeren Membran (outer membrane protein G, OmpG) von *Escherichia coli* als Modellsystem gewählt. Für Festkörper NMR spektroskopische Untersuchungen wurde OmpG in Proteoliposomen und schlecht aufgelösten 2D Kristallen rekonstituiert. Darüber hinaus wurde ein OmpG PEG Präzipitat erzeugt. Die Ergebnisse dieser Arbeit und Literaturdaten zeigten, dass zur Strukturuntersuchung von Membranproteinen mittels Festkörper NMR keine geordneten Kristalle notwendig sind. Bei der Herstellung von Proteoliposomen und 2D Kristallen werden Membranproteine in Membranen rekonstituiert. Diese Präparationsart bietet daher die Möglichkeit, Membranproteine quasi in ihrer natürlichen Umgebung mittels Festkörper NMR zu untersuchen. Darüber hinaus können zur Herstellung dieser

Proben schon bereits in der Strukturbiologie (Röntgenkristallographie und Elektronenmikroskopie) etablierte Kristallisationsmethoden eingesetzt werden. Der Vorteil, dass zur Strukturuntersuchung mittels Festkörper NMR Spektroskopie keine hochaufgelösten Kristalle erforderlich sind, eröffnet daher einen neuen Weg zur Untersuchung von Membranproteinen.

Eine wichtige Voraussetzung, um aus NMR Daten eine Proteinstruktur berechnen zu können, ist die Zuordnung der einzelnen Signale in den Spektren zur jeweiligen Aminosäure in der Proteinsequenz. In den letzten Jahren konnte mit verschiedenen Zuordnungsstrategien gezeigt werden, dass dies bei kleinen uniform  $^{13}\text{C}$ ,  $^{15}\text{N}$ -markierten Proteinen (bis zu 150 Aminosäuren) auch möglich ist. Es wurde jedoch auch klar, dass die Komplexität der NMR Spektren mit dem Molekulargewicht des zu untersuchenden Proteins stark zunimmt. Daher ist die Anwendung konventioneller Zuordnungsstrategien für große Membranproteine wie OmpG (281 Aminosäuren) äußerst schwierig. Aus diesem Grund wurde in dieser Arbeit eine Methode entwickelt, welche auf einfache Weise das Auffinden von Startpunkten für die sequentielle Zuordnung erlaubt. Diese Methode basiert auf der Verwendung einer spezifisch markierten Proteinprobe, in der die aromatischen Aminosäuren Tyrosin und Phenylalanine  $^{15}\text{N}$  markiert und in den Positionen 2 ( $\text{C}\alpha$  Atom) und 3 ( $\text{C}\beta$  Atom)  $^{13}\text{C}$  markiert sind. Darüber hinaus sind in dieser Probe die beiden kurzkettigen Aminosäuren Ala und Gly uniform  $^{13}\text{C}$  und  $^{15}\text{N}$  markiert. Mit dieser Probe konnte gezeigt werden, dass dieses Markierungsmuster einerseits die spektrale Signalüberlagerung und andererseits die Anzahl von Zuordnungsmöglichkeiten stark verringert. Um die sequentielle Zuordnung weiter voranzutreiben wurden vier weitere spezifisch markierte OmpG Proben hergestellt. Für die Zuordnung der NMR Signale wurden zunächst 2D  $^{13}\text{C}$ - $^{13}\text{C}$  and 3D NCACX Korrelationsexperimente mit kurzen  $^{13}\text{C}$ - $^{13}\text{C}$  Mischzeiten aufgenommen. Diese Spektren erlauben zunächst die Zuordnung der einzelnen Spinsysteme jeder markierten Aminosäure. In einem zweiten Schritt wurden 2D and 3D NCOCX Korrelationsexperimente verwendet, um die einzelnen NMR Signale einer bestimmten Aminosäure in der OmpG Sequenz zuzuordnen.

Zwei weitere Markierungsmuster wurde durch rekombinante Expression in einem Nährmedium, welches 1,3- $^{13}\text{C}$  oder 2- $^{13}\text{C}$ -markiertes Glycerol als einzige Kohlenstoffquelle beinhaltet, hergestellt. Proben dieser Art wurden bisher bei kleineren Proteinen zur Gewinnung von Abstandsinformationen genutzt. In dieser Arbeit konnte

jedoch gezeigt werden, dass diese Methode sehr nützlich sein kann um die bereits erhaltene Zuordnung zu bestätigen.

Eine weitere Möglichkeit zur Identifikation von Startpunkten für die Zuordnung größerer Proteine ist die Anwendung spektroskopischer Filtermethoden. Ein für Methylgruppen selektiver Filterblock wurde hierfür in ein  $^{13}\text{C}$ - $^{15}\text{N}$  heteronukleares Experiment eingebaut.

Durch Nutzung der hier vorgestellten Methoden war es bisher möglich 50 % der NMR Signale der OmpG Proteinsequenz zuzuordnen. Zur Überprüfung der gewonnenen Daten wurden diese zur Vorhersage von Sekundärstrukturelementen genutzt. Es konnte gezeigt werden, dass die berechneten Struktur motive in guter Übereinstimmung zu den bisher veröffentlichten OmpG Strukturen liegen.

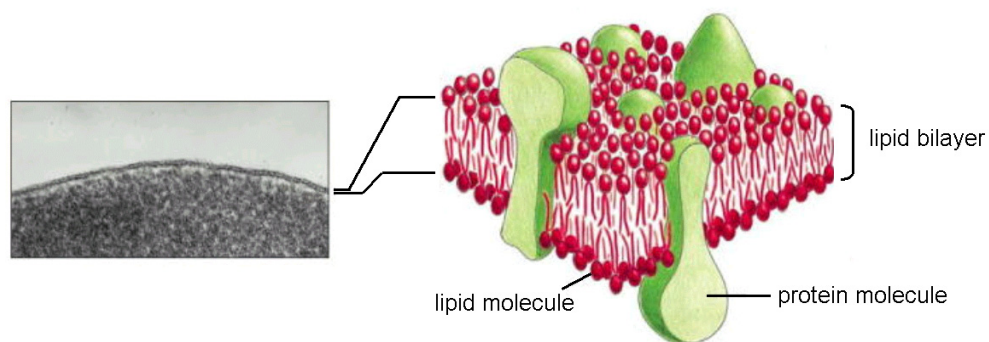
Die in dieser Arbeit angewendeten Methoden sollten auf eine Vielzahl anderer Membranproteine anwendbar sein und somit einen neuen Weg zur strukturbiologischen Untersuchung von Membranproteinen eröffnen.



# 1 General Introduction

Complex reactions, like the neuronal signal-transduction in nerve cells, the synthesis of energy (ATP) or the photosynthesis and the photorespiration in plants, occur simultaneously in different cells and/or sub-cellular compartments. To ensure, that all these processes can proceed without problems, compartmentalisation by membranes is one of the most important necessities in life. Membranes isolate cells from each other and maintain the essential differences between the cytosol and the extracellular environment. Inside one cell, sub-cellular organelles, like the endoplasmic reticulum (ER), golgi apparatus, mitochondria and chloroplasts, are also surrounded by membranes to separate the different contents of each organelle from the cytosol.

Biological membranes are very thin films of lipid- and membrane protein molecules (Figure 1.1). Lipid molecules are arranged as a continuous double layer (lipid-bilayer) formed by three major classes of lipids: phospholipids, cholesterol and glycolipids. However, different mixtures of lipids are found in membranes of different cells and sub-cellular compartments.



**Figure 1.1 View of a cell membrane.** Left, electron micrograph of the plasma membrane from human red blood cells. Right, three dimensional model of a cell membrane. The figure was taken from the book “Molecular biology of the Cell” [1].

By means of these properties, membranes have four major functions. First, they serve as a permeability barrier for most water-soluble molecules and prevent that unwanted material passes from the extracellular milieu into the cell. Second, proteins in the plasma membrane anchor the membrane to the intracellular cytoskeleton, thus membranes serve as a scaffold. Third, electrical charges can be separated across the membrane, thus the lipid bilayer serves as a capacitor. Fourthly, membranes form the

matrix for membrane proteins (MPs). Although the basic structure of biological membranes is provided by the lipid bilayer, most of the specific functions are carried out by MPs such as channels, ion-pumps and receptors.

It is known, that mutations in MPs are directly or indirectly involved in many diseases, such as cystic fibrosis, epilepsy, retinitis pigmentosa and nephrogenic diabetes insipidus [2,3]. Thus, it is not surprising that 90 % of all drug targets are MPs. However, only 428 entries and 182 unique structures of MPs have been deposited in the Protein Data Bank (PDB) by February 2009 according to the list held by Prof. Stephen White on his website [4]. The number of solved MP structures is still far behind the number of solved structures of soluble proteins (~20,000 structures of non-redundant protein sequences) deposited in the PDB since 1976 [5]. Taking into account, that the human genome codes for ~36,000 structural genes and it is estimated that ~30% of all eukaryotic proteins are MPs, less than 1 % of the known protein structures are MP structures. Therefore, structure determination of MPs is of major interest not only in structural biology but also in pharmacology, especially for drug development.

## **1.1 Approaches for structural studies of membrane proteins**

The increasing availability of genome information, a better biochemical and biophysical understanding of MPs and new technical developments in the field of structure determination has increased the number of MP structures deposited in the PDB. X-ray crystallography, electron microscopy (EM) and nuclear magnetic resonance (NMR) spectroscopy are the three major techniques that are used to obtain structural details of MPs. A challenge for all these methods is the expression, isolation and purification of large amounts of the MP of interest (tens or even hundreds of milligrams), as well as finding suitable sample conditions.

### **1.1.1 Membrane protein sample preparation**

In early studies, MPs such as the bacterial photosynthetic reaction centre, the visual pigment rhodopsin, the P-type  $\text{Ca}^{2+}$ -ATPase, the light-harvesting complex or the nicotinic acetylcholine receptor (nAChR), for example, were purified from natural sources [6-10]. The natural abundance of most other MPs, however, is usually too low

to isolate sufficient material for structural studies. Thus, an efficient expression system for the heterologous over-expression of MPs is required.

In prokaryotes, the ribosome-bound nascent chain of an MP interacts with the signal recognition particle (SRP) and the SRP-receptor, which targets the complex to the Sec-translocase of the cytosolic membrane. In the second step, the MP is translated further and inserted into the membrane by the Sec-translocase [11]. In eukaryotes the ribosome-bound nascent chain is typically translated and inserted into the ER membrane via a similar mechanism. Subsequently, the MP undergoes different modifications (e.g. glycosylation and folding) in the ER and the golgi apparatus before it is transferred and inserted into the plasma membrane via exocytosis [12]. However, when over-expressing recombinant MPs, prokaryotic and eukaryotic systems can have difficulties in assisting the targeting, insertion and folding of over-expressed MPs for the following reasons. First, in both systems, MP folding is assisted by chaperones. Thus, it is possible that the amount of chaperones is too low or the chaperones are completely absent, which often causes protein aggregation. Second, membranes from prokaryotes or eukaryotes differ in their lipid composition, which can have a significant effect on the insertion, folding and functioning of heterologous over-expressed MPs. Additionally, the accumulation of non-native structures in a membrane can induce stress responses and activate proteolytic systems of the expression host. Third, the membrane space required to accommodate heterologous over-expressed MPs represents another potential bottleneck. The knowledge about these bottlenecks, leads to the development of various expression systems, including the *Escherichia coli* (*E.coli*) strains C41(DE3) and C43(DE3), the yeasts *Pichia pastoris* and *Saccharomyces cerevisiae*, the baculovirus/insect- and several mammalian cell systems, that improve the heterologous over-expression of functional MPs [13-16]. Additionally, cell-free expression systems facilitating high yields of functional MPs were developed [17]. In addition to the optimisation of expression systems and/or conditions, the MP itself can be modified to improve its functional over-expression yields. For example, N-terminal fusion to a signal sequence facilitates the targeting to the cytoplasmic or ER membrane. Additionally, C-terminal fusion can protect the MP against attacks by proteases. Comprehensive reviews are given by Grisshammer and Wagner [18,19].

All these expression strategies can be used to over-express MPs for X-ray crystallography and EM studies. However, structural studies by NMR spectroscopy are a special case, because the sample must be labelled with stable isotopes such as  $^{13}\text{C}$

and/or  $^{15}\text{N}$ . *E.coli* is therefore considered as the primary expression system, since the costs for isotopically labelled sole sources are much lower, cell cultures growth rapidly and expression protocols can be scaled up more easily [20]. However, the over-expression of MP in *E. coli* often results in the formation of aggregates (inclusion bodies). In the case of  $\beta$ -barrel MPs (termed porins), those inclusion bodies can be isolated and subsequently refolded [21]. In contrast to that, there are very few examples of  $\alpha$ -helical MPs that have been successfully refolded after isolation from inclusion bodies [22].

Therefore, in the last few years, isotopically labelled growth media for yeasts, the baculovirus/insect and mammalian cell-systems have been developed by Cambridge Isotope Laboratories, Inc. in cooperation with scientific groups [23,24]. Additionally, the cell-free expression systems have been used to produce functional isotopically labelled MPs in the presence of detergents [17,25].

Even if the yield of the recombinant MP is satisfactory, the second hurdle is the purification of the protein to near homogeneity. Prior to purification, the MP must be in a detergent-solubilised state that is no longer associated with the lipid bilayer. There are a large number of detergents available and choosing the right one is often a process of trial and error. Therefore, it is necessary to screen a number of detergents to optimise the recovery of soluble and functional material. The most commonly used detergents are polyoxyethylene glycols, non-ionic alkyl-sugars with a  $\text{C}_8$ - $\text{C}_{12}$  alkyl chains and the zwitterionic detergents CHAPS and LDAO. Comprehensive reviews are given by le Maire and Prive [26,27]. After successful solubilisation the MP can be purified from other impurities using standard chromatography techniques (e.g. ion-exchange, affinity and size-exclusion chromatography).

### **1.1.2 Three dimensional (3D) crystallisation and X-ray crystallography**

Structural data obtained by X-ray crystallography combined with additional biophysical investigations have given us insight into the molecular basis of different physiological processes, like the light induced electron transfer and the resulting ATP-synthesis in chloroplast, the detection of photons at the photoreceptor cells, the reduction of the cytosolic  $\text{Ca}^{2+}$  concentration driven by the  $\text{Ca}^{2+}$ -ATPase (SERCA1)



which causes the relaxation of muscle cells, and the principles of ion transport over the plasma membrane by channels and pumps [7,8,28,29].

X-ray crystallography provides high-resolution atomic structures of MPs (1-4 Å) and so far it is the most successful technique used in this field. However, X-ray crystallography is an all or nothing method, which requires highly diffracting 3D crystals. If the crystalline material does not show any diffraction, there is no possibility of solving the structure of the MP.

MPs solubilised in detergent micelles are flexible and dynamic complexes. However, any degree of flexibility in the crystal lattice is detrimental for growing high-resolution 3D crystals. The flexibility of a detergent micelle is strongly influenced by the polarity of the head group and the length of the alkyl-chain of the detergent molecules [30]. Therefore, searching for the most suitable detergent is a critical step on the one hand. On the other hand, heterogeneity in the MP sample due to the presence of flexible loops also hampers the crystallisation process. This was nicely illustrated for the crystallisation of the  $\beta_2$ -adrenergic receptor [31,32]. Based on prior biophysical studies, the third intracellular loop and the C-terminus of this receptor was found to be the most flexible region. Therefore, the receptor was expressed as a C-terminal truncated version and most of the third intracellular loop was replaced by T4 lysozyme. The authors could show that this modified receptor retains near-native properties and crystal formation was improved.

Over the last few years, different techniques to enhance the resolution of 3D crystals have been developed. These include the use of lipids to stabilise MPs, the use of antibody fragments to increase the number of protein-protein contacts, the use of lipid cubic phases and bicelles [33-36]. Additionally, to test different buffer conditions and additives, crystallisation screens (e.g. MembFac from Hampton research) are commercially available.

### **1.1.3 Two-dimensional (2D) crystallisation and electron microscopy (EM)**

X-ray crystallography is the most frequently used method in structure determination of MPs. However, detergent micelles are an appropriate but imperfect substitution for the lipids surrounding MPs in situ. In particular, the structural integrity of a MP may depend on essential lipids. In this difficult context, reconstitution of MPs into artificial

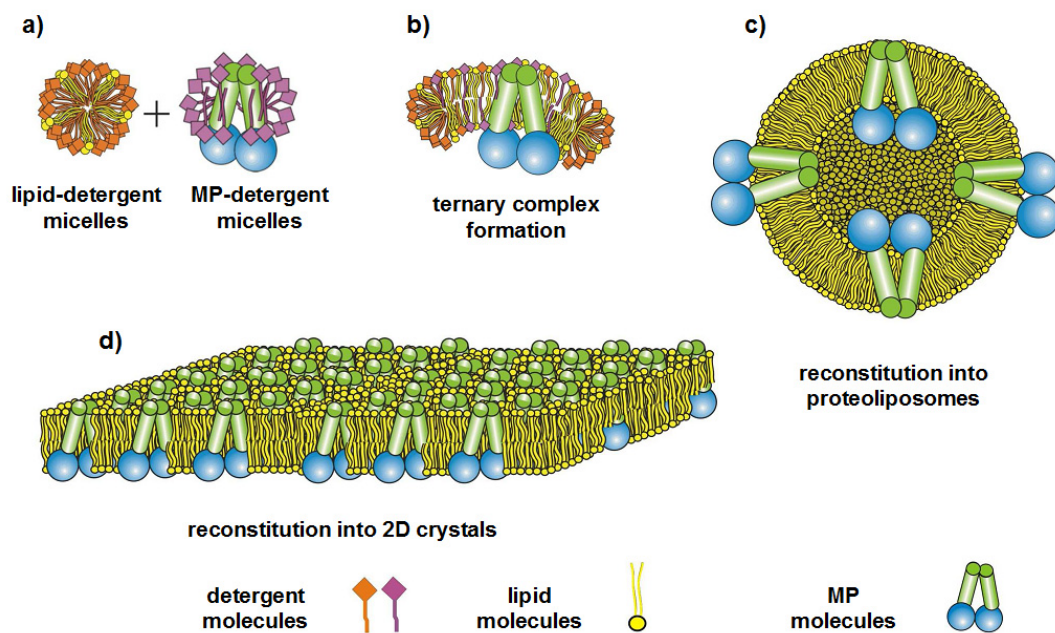
membranes to form two dimensional (2D) crystals and their structural analysis by EM has opened a new way to solve structures of MPs in a native-like environment. The resolution of MP structures obtained by EM lies in the range of 3 to 6 Å and is thus significantly lower than for X-ray crystallography. However, EM of MP in 2D crystals is not an all or nothing method like X-ray crystallography. Using poorly diffracting 2D crystals, it is at least possible to generate a molecular shape of the MP of interest. Structures of MPs determined by EM have given us a fundamental knowledge of the molecular function, for example, of the light-harvesting complex in plants, the selective water channel aquaporin 1 in the plasma membrane and the nAChR in nerve cells [10,37,38].

It has been shown that EM can also be used as a control for MP structures determined by X-ray crystallography and to assess structural differences observed in detergent micelles and in native lipid bilayers. The structure of the multidrug transporter EmrE obtained by EM and the first X-ray structure, for example, differ significantly [39]. Thus, X-ray data were proven and recalculation of the electron density maps results in a structure, which closely matches the structure obtained by EM [40]. A second example is the mitochondrial ADP/ATP carrier, which shows a monomeric structure in detergent micelles. However, the same protein reconstituted into lipid bilayers exhibits a dimeric structure [41,42].

EM of MPs reconstituted into lipid bilayers offers an additional advantage over X-ray crystallography, since the structural studies are performed on the same system (lipid bilayers) used for functional studies. This was nicely illustrated in an investigation of the opening-mechanism of the nAChR [43]. The open state of the channel was captured by spraying acetylcholine on tubular crystals and rapid freezing of the samples within less than 5 ms.

2D crystals of MPs can be produced in two different ways. The first method involves the induction of regular packing of a highly abundant MP in its native lipid bilayers. This can be achieved by incubating the isolated membranes with phospholipase A<sub>2</sub> to remove excess lipids [44]. The second method involves reconstitution of a detergent solubilised MP into lipid bilayers at high protein concentrations. The detergent solubilised MP is mixed with solubilised lipids to form homogenous MP-detergent-lipid micelles (ternary complex) (Figure 1.2). Detergent removal may then, in the ideal case, result in the formation of proteoliposomes with large 2D crystalline regions. This process is determined by factors like the lipid-protein

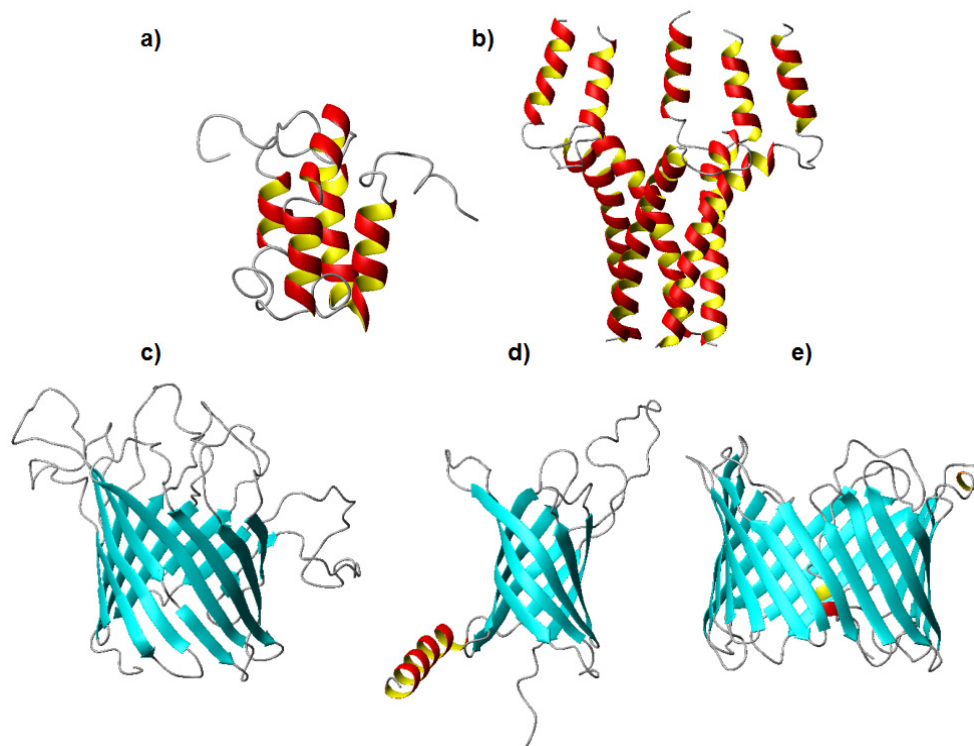
ratio (LPR), type of lipids, detergent concentration, temperature, pH and ionic strength. In particular, the chosen LPR and the speed of detergent removal are critical. With an excess of lipid over protein, the protein is mainly incorporated into liposomes, similar to its native state (Figure 1.2). For crystal packing, the LPR must be as low as possible to ensure close packing (Figure 1.2). However, an excess of protein over lipid can also produce amorphous aggregates. Additionally, slow detergent removal results in the formation of large flat 2D crystals. Those crystals are often composed of a stack of 2D crystalline lamellae, which hamper the effort to collect images by EM and thus the construction of a 3D map of the MP. In contrast to that, rapid detergent removal promotes the formation of much smaller unilamellar 2D crystals, which are much better for the collection of images by EM. Frequently used methods to remove the detergent include dialysis, the hydrophobic adsorption onto polystyrene beads (also termed as Biobeads®) and dilution of the ternary complex. Comprehensive reviews on 2D crystallisation of MPs are given by Kühlbrandt, Hasler and Rigaud [37,44,45].



**Figure 1.2 Reconstitution of MP into proteoliposomes and two-dimensional crystallization [137].** (a) The starting mixture is composed of lipid-detergent micelles and MP-detergent micelles, (b) ternary micelles are formed upon equilibration in the starting mixture. After subsequent detergent removal different structures are formed: (c) at a higher lipid-to-protein ratios (LPR) the membrane protein is reconstituted into proteoliposomes. (d) at a sufficiently low LPR 2D crystals can assemble.

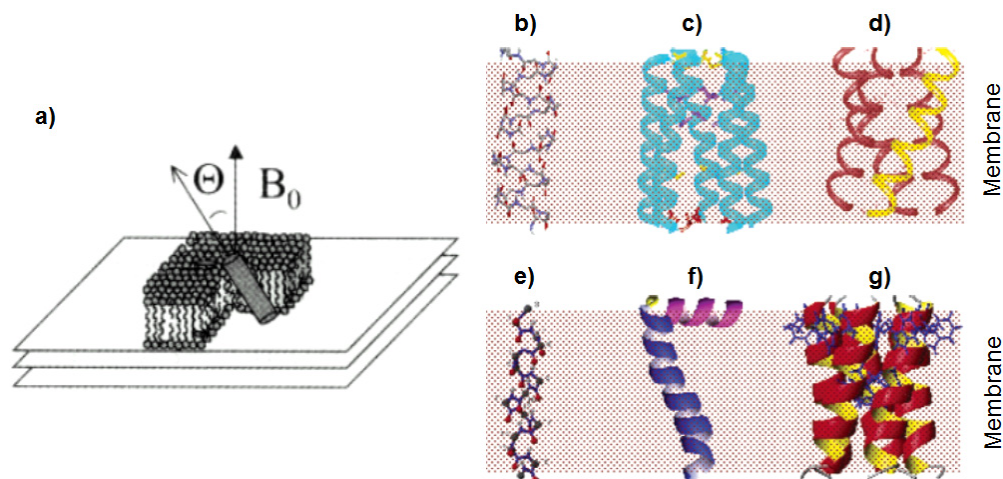
### 1.1.4 Solution- and solid-state NMR

Solution NMR has been established as a major method to determine structures of proteins and protein complexes that are readily soluble in aqueous solution [46]. In addition to the elucidation of their structures, NMR also offers unique opportunities to probe dynamic processes of such proteins and complexes. Membrane proteins, when embedded in lipid bilayers, are not amenable to solution NMR techniques because the tumbling of those systems is slow and highly anisotropic, which leads to unfavourable relaxation and hence very broad resonance lines. An alternative approach is the use of MPs in detergent- or lipid micelles, which tumble fast enough on the NMR time-scale. In combination with the use of perdeuterated  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled samples and TROSY-based triple resonance techniques, spectra of detergent solubilised MPs with reasonable line widths can be obtained by solution NMR [47,48]. Complete or nearly complete backbone assignments have been obtained for several  $\beta$ -barrel MPs, e.g. outer membrane protein G (OmpG), and PagP from *E. coli* and the human anion channel VDAC1, and for four  $\alpha$ -helical MPs, mistic from *Bacillus subtilis* and human pentameric phospholamban (PLB) [49-53] (Figure 1.3).



**Figure 1.3** solution NMR structures of MPs in detergent micelles. (a) Mistic (PDB code 1YGM), (b) Phospholamban (PDB code 1ZLL), (c) OmpG (PDB code 1JQY), (d) PagP (PDB accession code 1MM5) and (e) VDAC1 (PDB code 1K4T).

Over the past two decades, solid-state NMR has emerged as an effective method for structural studies of quasi-immobilized biomolecules, such as amyloid systems, MP-bound ligands and MPs. These samples are characterised by anisotropic interactions like chemical shift anisotropy (CSA), dipolar interactions and quadrupolar couplings, which lead to resonance broadening (see chapter 1.2). To overcome this problem, two general approaches for obtaining high-resolution solid-state NMR spectra have been developed. The first approach is specific to MPs and relies on macroscopically aligned membranes on glass plates (Figure 1.4a). Orientational restraints of membrane integrated peptides or proteins are derived from dipolar coupling and chemical shift measurements, so-called PISEMA experiments [54]. The values of both depend on the relative alignment of the molecule to the magnetic field ( $B_0$ ). To align lipid bilayers on glass plates, lipids are dissolved in organic solvents followed by evaporation and lipid hydration [55]. Additionally, bicelles have also been used to study aligned MPs by solid-state NMR [56,57]. Using this technique, structures of membrane embedded peptides, for example Gramicidin, the M2 transmembrane segment of the nAChR, the M2 transmembrane domain from the M2 proton channel of the influenza A viral coat, the M6 transmembrane domain of the  $\alpha$ -factor receptor from *Saccharomyces cerevisiae*, the Fd bacteriophage coat protein and the M2 transmembrane domain of the HIV virus protein U (Vpu) were solved by solid-state NMR (Figure 1.4b-g) [55].



**Figure 1.4 Structures determined by solid-state NMR in oriented lipid bilayers.** a) Oriented lipid bilayers put on solid glass-plates [58]. b) Gramicidin (PDB code: 1MAG), c) the M2 segment of the nAChR (PDB code 1CEK), d) the closed-state structure of M2 proton-channel (PDB code 1NYJ), e) the 6<sup>th</sup> transmembrane domain of the  $\alpha$ -factor receptor (PDB code 1PJD), f) the Fd bacteriophage coat protein (PDB code 1MZT) and g) the virus protein U (PDB code 1PJE) from the HIV-1 virus [55].

The second approach involves the **Magic Angle Spinning (MAS)** technique where a randomly oriented sample rapidly spins around the so-called magic angle ( $\theta_m=54.7^\circ$ ) in the static magnetic field ( $B_0$ ) (see also chapter 1.2.4). MAS applied to randomly oriented samples produces sufficiently resolved spectra and allows the detection of the isotropic chemical shifts. This has important advantages for structural investigations by solid-state MAS NMR. Well-established strategies for resonance assignment known from solution NMR can be implemented and chemical shift databases compiled from solution NMR studies can be accessed to identify amino-acids by their characteristic chemical shift pattern (see chapter 1.2.5). Thus, solid-state MAS NMR spectra of solid protein-preparations of at least seven soluble proteins (including the basic pancreatic trypsin inhibitor (BPTI) [59], the  $\alpha$ -spectrin SH3 domain [60,61], the regulatory protein Crh [62], human ubiquitin [63,64], thioredoxin [65], the scorpion venom kaliotoxin [66] and the  $\beta$ -immunoglobulin binding domain of protein G (GB1) [67]), three amyloid fibrils systems (including the Alzheimer  $\beta$ -amyloid fibrils [68], the second WW domain of the human transcription activator [69] and the fungal prion protein HET-s [70]) and two MP-bound ligands (kaliotoxin bound to the KcsA-Kv1.3 channel [71] and neurotoxin II bound to the nAChR [72]) have been fully or partially assigned.

These studies were performed with numerous different sample preparations, for example frozen liquids, gels, lyophilised powders, micro- and nano crystalline material and proteoliposomes. However, it has been shown that the quality of the spectra is mostly dependent on the art of sample preparation. In the case of the  $\alpha$ -spectrin SH3 domain, spectral resolution was critically compared for different preparation types [60]. Lyophilisation was not found to provide sufficient resolution, whereas a micro-crystalline ammonium sulphate precipitated sample gave highly resolved NMR spectra. This is due to the fact that proteins in a lyophilised powder might be properly folded, however, the conformation of the side chains will be quite variable and each protein is in a different environment. Each molecule will then have a slightly different chemical shift for each side-chain atom. Additionally, it was shown that nano-crystalline material far too small to be useful in X-ray diffraction studies is suitable for structural studies by solid-state MAS NMR [73]. In the case of amyloid fibrils, solid-state MAS experiments have been carried out on lyophilised [74] and on fully hydrated amyloid fibrils [75,76] or by re-hydrating the fibrils after lyophilisation [68]. Hydrated samples exhibit narrower line widths, in particular for side chain resonances. This is due to the fact that inhomogeneous broadening arising from static disorder in lyophilised samples is

averaged out by molecular motions in hydrated samples [77]. The high spectral resolution of a ligand bound to a receptor or channel can be attributed to the defined ligand-binding pocket [71,72].

Using randomly-oriented MP samples, bacteriorhodopsin and rhodopsin probably are the most extensively studied MPs by solid-state MAS NMR [78,79]. These studies were done mostly using membrane preparations from natural sources. Additionally, specifically labelled MP-samples for measurement of helix-helix distances and specifically labelled chromophores for the determination of the conformation of light cycling intermediates were used. In the beginning of this work, only one example of a fully labelled, integral MP has been documented. This was the LH2 light harvesting complex from *Rhodospseudomonas acidophila*, which yielded assignments of the  $^{15}\text{N}$ ,  $^{13}\text{C}$ -signals for the membrane-spanning portion. The high resolution obtained for this sample is attributed to the intrinsic rotational symmetry of the homo-nonameric membrane protein complex [80,81].

All these studies illustrate that solid-state MAS NMR benefit from homogeneous samples with local order, but without needing of highly ordered crystalline samples required for X-ray crystallography. Therefore, experiments on MPs can theoretically be performed with numerous different sample preparation types, for example, detergent solubilised MPs, nano-crystalline material originating from 3D crystallisation screens, MPs reconstituted into lipid bilayers (proteoliposomes) or 2D crystals. Additionally, the latter two types allow the structural investigation of MPs in a near-native environment. Many questions remain in this field. **First**, to what extent does the degree of local order determine the resolution of solid-state MAS NMR spectra of uniformly labelled MPs? **Second**, is it possible to assign spectra of larger MPs using the assignment procedures used for the examples presented above?

## 1.2 Fundamentals of nuclear magnetic resonance (NMR)

In the following section the physical basis of NMR and the NMR techniques used in this work are briefly introduced.

NMR spectroscopy probes the interaction between intrinsic magnetic moments of nuclei and an external magnetic field. The elementary particles, neutrons and protons have the intrinsic quantum mechanical property of a spin. The overall spin of the nucleus is determined by the spin quantum number ( $I$ ). If a nucleus has the same number of protons and neutrons, there is no overall spin ( $I = 0$ ). Nuclei with a spin quantum number  $I > \frac{1}{2}$ , have also an electrical quadrupolar moment. The resulting quadrupolar interaction is very strong, limiting the usability of these nuclei.

Therefore, in biomolecular NMR the most commonly measured nuclei are those with a spin quantum number of  $\frac{1}{2}$ , such as  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$  and  $^{31}\text{P}$  (Table 1).

*Table 1 Selection of nuclear isotopes and their characteristics.*

Isotope	spin quantum number $I$	natural abundance [%]	gyromagnetic ratio $\gamma$ [ $10^6 \text{ rad s}^{-1} \text{ T}^{-1}$ ]
$^1\text{H}$	$\frac{1}{2}$	99,98	267.522
$^2\text{H}$	1	0,016	41.066
$^{13}\text{C}$	$\frac{1}{2}$	1,108	67.283
$^{15}\text{N}$	$\frac{1}{2}$	0,37	-27.12
$^{31}\text{P}$	$\frac{1}{2}$	100	108.394

Nuclei with spin  $\frac{1}{2}$  have two possible spin states:  $1/2$  or  $-1/2$  (also referred to as “up” and “down” or “ $\alpha$ ” and “ $\beta$ ”, respectively). The energy of both spin states in absence of an external magnetic field is the same (Figure 1.5).

In a strong magnetic field ( $B_0$ ) the spins precess around  $B_0$  with their Larmor frequency  $\omega_0$ , which is related to  $B_0$  by the following equation:

$$\omega_0 = -\gamma B_0 \quad [\text{Eq. 1}]$$



The two spin states correspond to two different energy levels, with the z component either parallel (low energy) or antiparallel (high energy) to  $B_0$ . The energy difference  $\Delta E$  corresponds to the Larmor frequency, with which the NMR phenomenon can be described [Equation. 2] (Figure 1.5).

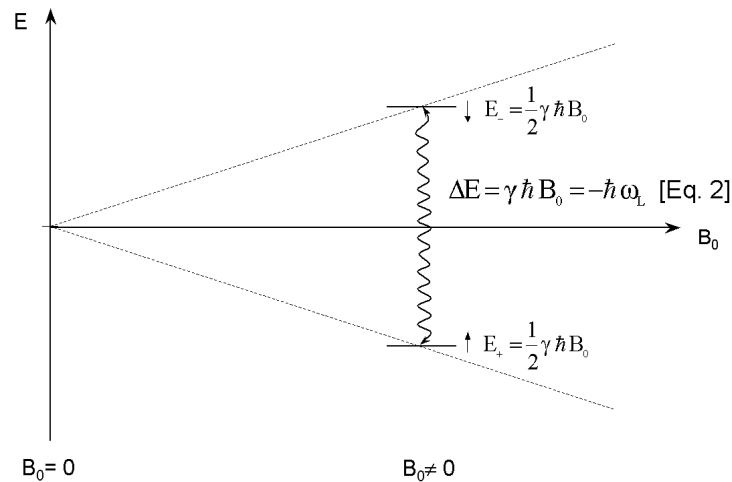


Figure 1.5 Phenomenon of nuclear magnetic resonance.

For an easier discussion it is useful to imagine the spins being situated in a Cartesian coordinate system (Figure 1.6). By convention, the orientation of  $B_0$  is always along the z-axis. If a sample is placed into a strong magnetic field, a macroscopic nuclear net magnetisation ( $M_0$ ) gradually builds up until a thermal equilibrium is reached (Figure 1.6a). However, the experimental study of  $M_0$  is impractical, since the external magnetic field (MHz) is much stronger than  $M_0$  (Hz). Therefore, a radio frequency (rf) pulse (with a frequency close to the Larmor frequency  $\omega_0$  of the nuclei) is applied in such a way that the nuclear spin is rotated into the  $XY$ -plane. As a result, the macroscopic nuclear magnetisation rotates perpendicular to  $B_0$  (Figure 1.6c above). This magnetisation is called transverse magnetisation ( $M_T$ ). Now suppose that the rf-pulse is turned off. The spin precesses around the Z-axis and will slowly relax to the thermal equilibrium (Figure 1.6c, below). The rotating magnetisation generates a magnetic field, which is associated with an electric field. If a coil is near the sample the electric field induces an oscillating current into the coil. Note the geometry of the coil is arranged in such a way that the winding axis of the coil is perpendicular to  $B_0$  in order to detect the precessing  $M_T$ . The resulting current can be detected by using a sensitive radio frequency detector and thus an NMR signal or *free-induction decay* (FID) can be

detected. The detected signal is composed of  $X$ - and  $Y$ -magnetisation yielding simultaneously the amplitude as well as the phase of the signal (Figure 1.6d, above). The NMR spectrum is calculated by Fourier transformation of the FID (Figure 1.6d, below).

It might appear from the above that all nuclei of the same isotope (and hence the same  $\gamma$ ) would resonate at the same frequency (also termed as chemical shift). This is not the case. The most important perturbation of the NMR frequency is the 'shielding' effect of the surrounding electrons. As an example a carboxyl group is more de-shielded, whereas a methyl group is more shielded by the surrounding electrons. Thus, the chemical shift of the carboxyl group is shifted more up-field and that of the methyl-group shifted more down-field due to the chemical environment. This explains why NMR is a direct probe of chemical structure.

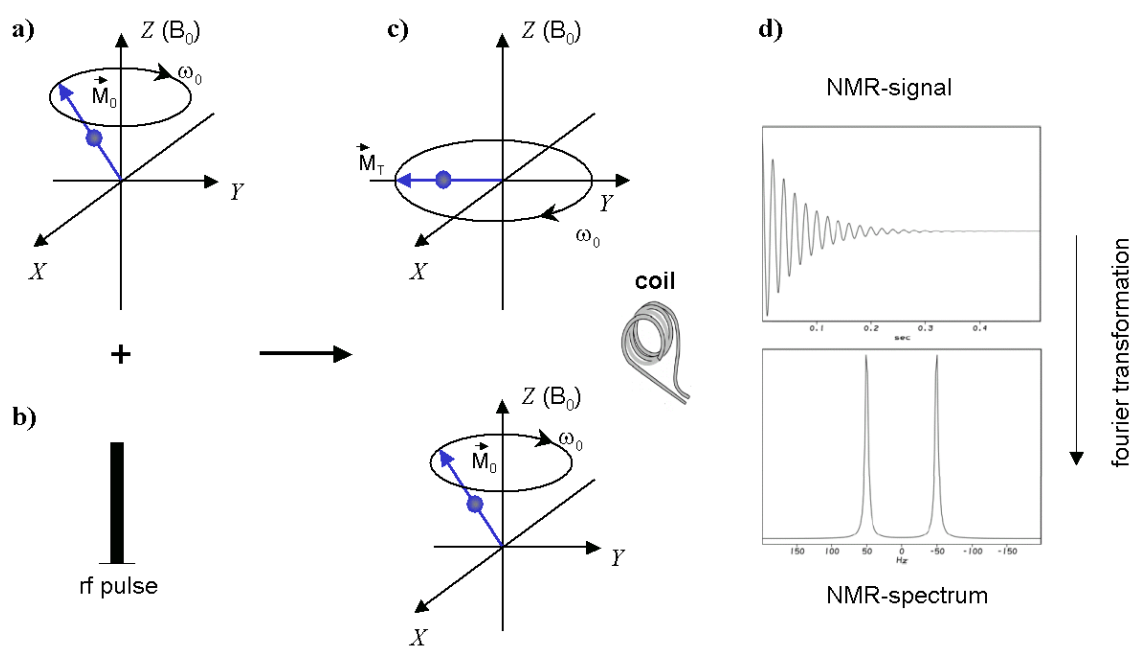


Figure 1.6 Pulsed Fourier-transformed NMR (for more details see text above).

### 1.2.1 Solid-state MAS NMR

Solid-state MAS NMR has emerged as an effective method for structural studies of quasi-immobilized biomolecules. In most solids, fast molecular tumbling is absent and anisotropic interactions like homo- and heteronuclear dipolar couplings, the chemical shift anisotropy (CSA) and, for spins  $> 1/2$ , quadrupolar couplings, lead to resonance broadening.

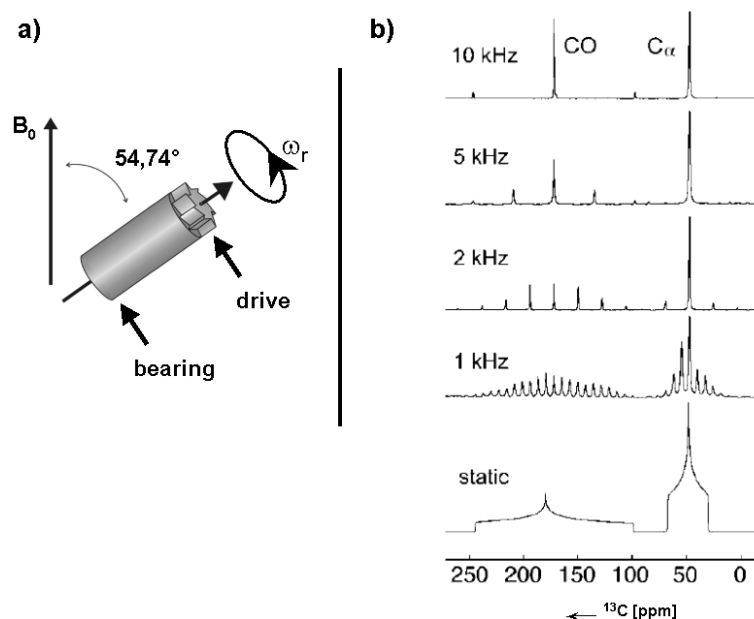
The dipolar coupling arises from the through-space interaction (or coupling) of two dipoles. The strength of the heteronuclear dipolar coupling is represented by the following Hamiltonian [3]:

$$H_D = -d \frac{1}{r^3} (3 \cos^2 \theta - 1) I_z S_z \quad [\text{Eq. 3}]$$

where  $d$  is the dipolar coupling constant,  $r$  is the distance between the interacting spins,  $I_z$  and  $S_z$  are the transverse components of the local magnetic field due to spins  $I$  and  $S$ .  $\theta$  is the angle between the internuclear vector and the static magnetic field  $B_0$ . The  $r^{-3}$  distance dependency allows the detection of distance-dependent restraints in solid-state NMR. The orientational dependence of the dipolar coupling means that it is not active in liquid-state NMR spectroscopy: the reorientation time of a molecule in solution is much faster than the time during which the dipolar coupling evolves, and thus the term  $(3\cos^2\theta-1)$  in Equation 3 is averaged to zero. In contrast to that, a solid sample consists of static crystallites, which are randomly oriented with respect to  $B_0$  and so the dipolar coupling will be active. For each coupling, a spin's resonance line will be broadened. Given the typical size of the dipolar coupling (at a distance of 1 Å, 120 kHz for H-H, 30 kHz for C-H and 12 kHz for N-H), the distance over which it is active (up to 10 Å) and the number of  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei present in a protein sample, each spin will be subject to a large number of dipolar interactions and thus its resonance lines will be broadened. In order to remove the dipolar coupling and restore narrow resonance lines, the  $(3\cos^2\theta-1)$  term has to be effectively set to zero. This can be done by fast spinning the sample around an axis inclined at an angle of  $\theta = 54.7^\circ$  with respect to the  $z$ -axis (the so-called magic angle). This results in a time average of each single  $\theta$  that equals the magic angle and hence averages out this interaction.

An additional method to remove  $^1\text{H}$ - $^{13}\text{C}$  dipolar couplings is to use rf-pulses. By continually applying rf-pulses that rotate the proton nuclear spins between their spin-up and spin-down states the orientation of their magnetic moments is averaged and dipolar couplings are reduced. This technique is called proton decoupling. Today different rf-pulse schemes, such as **continuous-wave (CW) irradiation**, **two pulse phase modulation (TPPM)** or **small phase incremental alternation (SPINAL)** can be used for proton decoupling [82-84]. However, the use of high power proton decoupling schemes is restricted by the amount of power the probe and the sample can sustain without damage. Special care has to be taken when investigating protein samples because, due to their high water content, severe sample heating is often observed.

The second anisotropic effect, which has a bearing on solid-, but not solution-state samples, is the chemical shift anisotropy. As described in section 1.2, the chemical shift of a nucleus is strongly dependent on the orientation of the surrounding electron cloud. In liquids, molecules randomly and rapidly sample the full range of orientations, so that even a strongly asymmetric electron distribution will appear spherical on the NMR time scale. This phenomenon averages out the anisotropic part of the chemical shift in solution NMR. In contrast to that, in randomly oriented crystallites, the orientation of the electron cloud around a particular atomic site differs from molecule to molecule. For example, the resonance frequency of a carbonyl carbon atom can differ by more than 120 ppm, depending on the orientation of the C=O moiety. Thinking about solution NMR the effect of the CSA can be eliminated in solid-state NMR by rotating the sample at a high spinning speed ( $\omega_r = 1000\text{-}50000$  Hz) around the magic angle ( $\theta = 54.7^\circ$ ) (Figure 1.7a). This so-called solid-state magic-angle spinning (MAS) NMR technique averages out the anisotropic part of the chemical shift and only the isotropic part will be measured (Figure 1.7b).



**Figure 1.7 The solid-state MAS NMR technique.** a) The rotor, that contains the sample, is tilted at  $54.7^\circ$ . By the combined application of bearing and driving pressures, the sample rotates around itself, at the desired frequency. b) Solid-state MAS  $^{13}\text{C}$ -NMR spectra of uniformly  $^{13}\text{C}$ -labelled glycine powder sample at different spinning speeds. The Figure was taken from [85].

## 1.2.2 Basic solid-state MAS NMR techniques

In solution NMR, detection of protons is most popular due to their high natural abundance in biological samples. Additionally, protons are the most sensitive nuclei as they have the largest gyromagnetic ratio  $\gamma$  (Table 1). To increase resolution of overlapping  $^1\text{H}$  resonances,  $^{13}\text{C}$  and  $^{15}\text{N}$  are used in multidimensional NMR techniques. However, in solid-state MAS NMR spectroscopy the usage of  $^1\text{H}$  nuclei is very difficult due to the dipolar couplings (see above). Thus, in solid-state MAS NMR mainly  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei are used for detection. The drawbacks of directly detecting low  $\gamma$  nuclei, like  $^{13}\text{C}$  or  $^{15}\text{N}$ , are their low isotopic abundances and lower spin polarization and hence lower signal intensity (Table 1). Using a technique, which is called cross polarization, some of these problems can be circumvented.

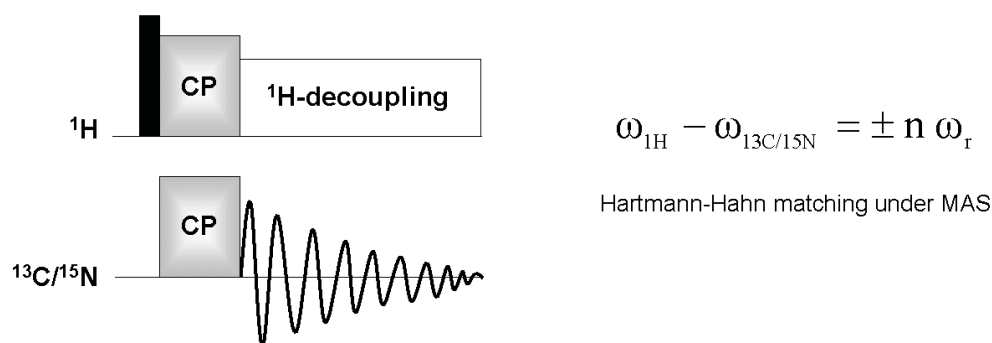
### 1.2.2.1 Cross polarisation

To enhance the signals from  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei, most solid-state MAS NMR experiments routinely involve the transfer of polarisation from highly abundant protons by using the cross polarisation (CP) technique [86,87].

The process of CP occurs through the tendency of the magnetisation to flow from highly polarised nuclei to nuclei with lower polarisations when the two are brought into contact.

To establish a dipolar contact between two different spin systems, here  $^1\text{H}$  and  $^{13}\text{C}$ , a particularly effective approach is that of Hartmann and Hahn [85]. This approach requires the simultaneous application of two continuous rf-pulses, one at the resonance frequency of the  $^1\text{H}$  spin and one at the resonance frequency of the  $^{13}\text{C}$  spin. Thus, both spins rotate independently around a particular axis at rates determined by the amplitude of the two rf-pulses. When the frequencies of the  $^1\text{H}$  and  $^{13}\text{C}$  spin are equal, an energy conserving dipolar contact between the two spins is created. The experimental implementation of this concept is shown in Figure 1.8. First, the proton magnetisation is brought into the  $XY$  plane by a  $\pi/2$  pulse. Rf-pulses are then applied to the  $^1\text{H}$  and  $^{13}\text{C}$  spins for a period, causing the magnetisation to be exchanged between the  $^1\text{H}$  spins and the  $^{13}\text{C}$  spins. Finally, the  $^{13}\text{C}$  spins are detected while the  $^1\text{H}$  spins are decoupled. Under MAS, CP occurs at the so-called side band matching condition.

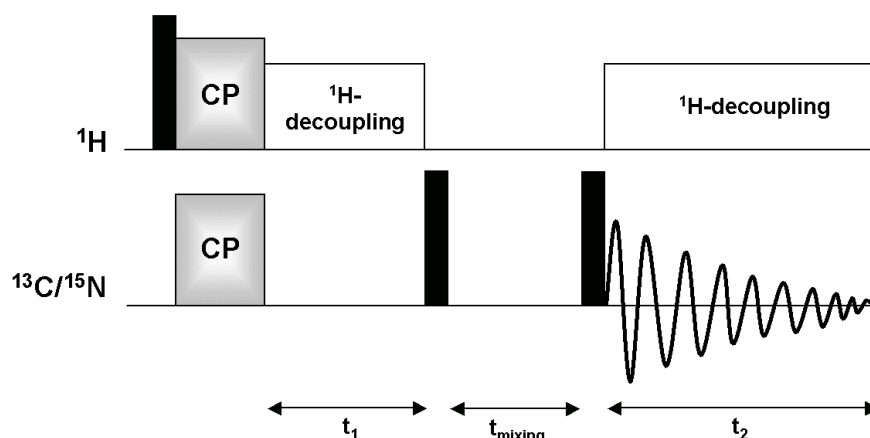
These match conditions are generally quite narrow, but in practice they are broadened by the application of a ramped CP-pulse [88]. This pulse sequence in combination with MAS forms the basis of proton enhanced solid-state MAS NMR spectroscopy.



**Figure 1.8** A simple 1D  $^{13}\text{C}/\text{CP}$  MAS pulse sequence (left) and Hartmann-Hahn matching (right). Following  $^1\text{H}$  excitation, CP between  $^1\text{H}$  and  $^{15}\text{N}$  or  $^{13}\text{C}$  creates the initial  $^{15}\text{N}/^{13}\text{C}$  magnetisation transfer. During the acquisition period, proton decoupling is usually applied.

### 1.2.2.2 Spin diffusion

Dipolar coupled spins exchange longitudinal magnetisation, which is therefore distributed among the coupled spins in a spin system. This process is termed spin diffusion. Under MAS it is generally attenuated and does not occur among  $^{13}\text{C}$  or  $^{15}\text{N}$  nuclei as their dipolar couplings are averaged to zero. However, strong couplings to a proton spin network broaden the signals of low  $\gamma$  nuclei and enable the exchange of magnetisation in an energy conserving way. This process is referred to as **proton-driven spin-diffusion** (PDS, Figure 1.9) [89]. By irradiating the protons at the rotary resonance condition (one or two times the MAS frequency) the magnetisation exchange can be accelerated. This technique is referred to as **dipolar-assisted rotational-resonance** (DARR) [90]. The efficiency of the magnetisation transfer depends on the distance of the nuclei involved and can be used to correlate nuclei, which are close in space (e.g. up to 7-8 Å for  $^{13}\text{C}$ ). We will see in section 1.2.5, that spectra recorded in such a way, forms the basis for the identification of different amino acid types.



**Figure 1.9 Pulse program of a standard 2D homonuclear PDS dipolar correlation experiment, applicable in the case of  $^{13}\text{C}$ - $^{13}\text{C}$  or  $^{15}\text{N}$ - $^{15}\text{N}$  correlation spectra.** Following  $^1\text{H}$  excitation, cross-polarization between  $^1\text{H}$  and  $^{15}\text{N}$  or  $^{13}\text{C}$  creates the initial  $^{15}\text{N}/^{13}\text{C}$  magnetization. Following the nitrogen or carbon evolution, a  $90^\circ$  pulse on the low- $\gamma$  nucleus brings back the magnetisation along the z-axis. During the mixing time  $t_{\text{mixing}}$ , proton decoupling is switched off and the polarisation-transfer between  $^{15}\text{N}$  or  $^{13}\text{C}$  spins occur in the presence of  $^1\text{H}$ - $^1\text{H}$  and  $^1\text{H}$ - $^{15}\text{N}$  (or  $^1\text{H}$ - $^{13}\text{C}$ ) dipolar couplings. During all evolution periods ( $t_1$  and  $t_2$ ), proton decoupling is usually applied.

### 1.2.2.3 Specific cross polarisation experiments

Spectral assignment requires heteronuclear transfer steps that allow a direct magnetisation transfer from one spin to the next in the polypeptide chain, to obtain the sequential assignment. To achieve heteronuclear transfer in polypeptides, the conventional CP approach has been modified for polarisation transfer between  $^{15}\text{N}$ - $^{13}\text{C}$  spins in a chemical shift selective manner [91]. Selective magnetization transfer between the backbone nitrogen atom and the  $\text{C}\alpha$  or  $\text{C}'$  carbon atom can be established using specific  $^{15}\text{N}$ - $^{13}\text{C}$  CP (specific-CP) experiments, where off-resonance rf-pulses (low-power) are applied to spin lock the  $^{15}\text{N}$  and  $^{13}\text{C}$  spins. The spin locking has to fulfil the specific-CP matching condition,

$$\sqrt{\Omega_{^{13}\text{C}}^2 + \omega_{^{13}\text{C}}^2} - \sqrt{\Omega_{^{15}\text{N}}^2 + \omega_{^{15}\text{N}}^2} = n\omega_r \quad [\text{Eq. 4}]$$

where  $\omega_r$  is the sample spinning frequency,  $\Omega_{^{13}\text{C}/^{15}\text{N}}$  is the frequency offset and  $\omega_{^{13}\text{C}/^{15}\text{N}}$  is the rf pulse strength for the  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei. The low rf-powers ensure that the CP is band-selective, while the off-resonance components can be adjusted to fulfil the specific-CP conditions for NCA or NCO spectra. Additionally, a gain in sensitivity for such an experiment was observed by applying an adiabatic pulse shape to one of the CP pulses [92].

NCA and NCO experiments can be extended by a DREAM scheme, yielding N(CA)CB, N(CO)CA and N(CA)CO experiments which can also be run in a three dimensional fashion [93]. Alternatively, a  $^{13}\text{C}$ - $^{13}\text{C}$  PDS or DARR step can be implemented, to achieve chemical shift information of carbon atoms in the side chains [61]. In ideal cases, these schemes can be extended to record a four dimensional CANCOX experiment [94]. However, for such experimental schemes relatively high signal-to-noise ratios are necessary.

### 1.2.2.4 Recoupling techniques

The dipolar couplings that are removed by MAS can be reintroduced by certain pulse schemes. Thereby a train of rf-pulses is used which inverts the magnetisation every half rotor period. This results in an incomplete averaging of the dipolar couplings and drives a through space magnetisation transfer between the coupled spins. This can



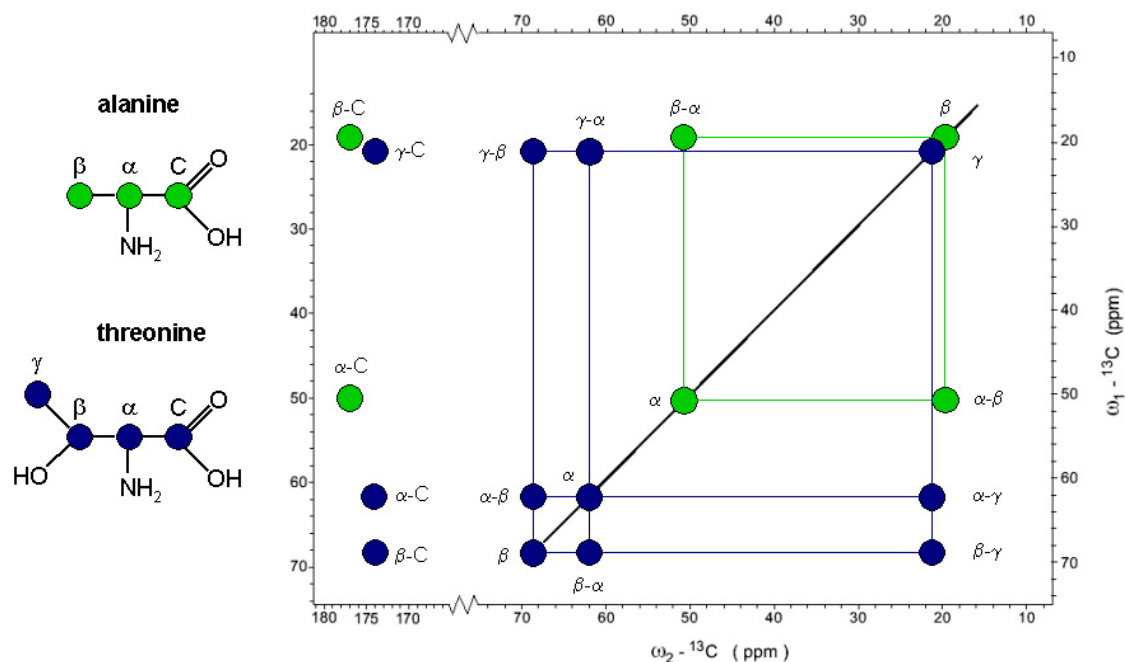
either be done in a homo- or a heteronuclear fashion and is referred to as radio-frequency driven recoupling (RFDR) or transferred/rotational-echo double resonance (TEDOR/ REDOR), respectively [95-97]. This method is particularly useful for generating correlations between  $^{13}\text{C}$  and  $^{15}\text{N}$  nuclei, which are needed for sequence-specific assignment.

### **1.2.3 Structure determination**

NMR spectroscopy is well suited to collect structural information at atomic resolution on complex systems, such as proteins, polymers, or inorganic materials. In the following section the assignment and structure determination process of proteins will be described.

#### **1.2.3.1 Assignment of solid-state MAS NMR spectra**

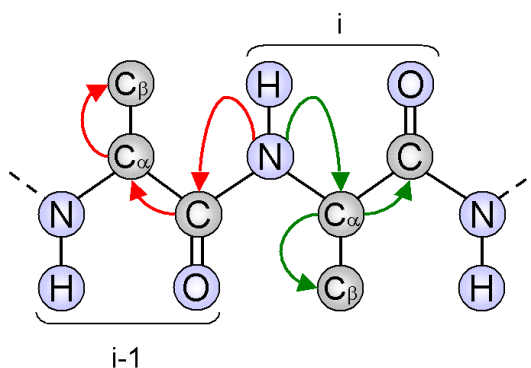
To obtain site-specific information from NMR spectra, the observed isotropic chemical shifts have to be assigned to the nuclei in the protein. In solid-state MAS NMR spectroscopy complete assignment of small proteins has become possible only during the last decade, and is still not routine. The established sequential assignment procedures are based on  $^{13}\text{C}$ - $^{13}\text{C}$  (see 1.2.4.2) and  $^{13}\text{C}$ - $^{15}\text{N}$  (see 1.2.4.3) correlation experiments. In the first step of resonance assignment,  $^{13}\text{C}$ -homonuclear correlation experiments that allow the identification of different amino acid types based on their characteristic chemical shift correlation patterns are recorded. In Figure 1.10, a schematic  $^{13}\text{C}$ - $^{13}\text{C}$  PDSM spectrum with the typical correlation patterns for alanine and threonine residues is illustrated. All cross peaks reflect intra-residual correlation networks in a given protein. In order to assign the signals to specific amino acid types, the averaged chemical shift values of  $\text{C}'$ ,  $\text{C}\alpha$ ,  $\text{C}\beta$  and  $\text{C}\gamma$  carbon atoms from the BioMagResBank (BMRB) can be used [98].



**Figure 1.10 Illustration of the assignment of amino acid types in a protein.** The illustrated schematic  $^{13}\text{C}$ - $^{13}\text{C}$  spectrum shows the typical correlation patterns of the amino acids alanine and threonine. The averaged chemical shift values for the  $\text{C}'$ ,  $\text{C}\alpha$ ,  $\text{C}\beta$  and  $\text{C}\gamma$  carbon atoms were taken from the BioMagResBank (BMRB) [98].

In the second step, the amino acid spin systems need to be linked sequentially into a chain, which can then be matched to the protein sequence to yield sequence-specific assignments. This can be achieved using heteronuclear  $^{13}\text{C}$ ,  $^{15}\text{N}$  correlation experiments, that correlate the backbone nitrogen either to the neighbouring  $\text{C}\alpha$  carbon atom of the same residue (i) or to the carbonyl carbon atom of the previous one (i-1) (Figure 1.11). These experiments are referred to as NCA and NCO, which can be further extended to provide links not only to the  $\text{C}\alpha$  and  $\text{C}'$  atoms (see 1.2.4.3), but also to the whole side chain in so-called NCACX and NCOCX experiments [61].

Using this assignment procedure, near full assignments have been obtained for a number of small soluble proteins and amyloid fibrils, like the  $\alpha$ -spectrin SH3 domain, ubiquitin, the catabolite repression histidine-containing phosphocarrier protein (Crh) from *Bacillus subtilis*, the  $\beta$ 1 immunoglobulin binding domain of protein G (GB1), the non-crystalline amyloid fibrils of the HET-s prion protein and the amyloid fibrils of human CA150.WW2 [61-63,67,69,70].



**Figure 1.11 Illustration of the magnetisation transfer during an NACX (green) and NCOCX (red) experiment.** First, proton rf-fields are used to prepare  $^{15}\text{N}$  polarisation via a broad band ( $^1\text{H}$ ,  $^{15}\text{N}$ ) CP transfer. Second, a band selective  $^{15}\text{N}$ - $^{13}\text{C}$  CP step is used to transfer the magnetisation from  $^{15}\text{N}$  to  $\text{C}\alpha$  or CO. Subsequently a  $^{13}\text{C}$ - $^{13}\text{C}$  PDSM mixing scheme can be used to transfer the magnetisation between  $\text{C}\alpha$  or CO and the side chain carbon atoms (CX).

### 1.2.3.2 Structural restraints and structure calculation

Protein structure determination from solution NMR data is well established and many methods can be used or adapted for the calculation of protein structures from solid-state MAS NMR data. In solution NMR distance restraints (from NOESY experiments), dihedral restraints (from J-couplings or chemical shifts) and residual dipolar couplings are used to calculate a protein structure.

In solid-state MAS NMR, distance restraints can be collected from PDSM and DARR experiments, which reveal distance dependent homonuclear magnetisation transfers between carbons or nitrogens [89,90,99]. Spin diffusion across hydrogen bonds is probed, using nitrogen and carbon atoms for detection, in CHHC or NHHC experiments [100]. Heteronuclear distances can be obtained by using heteronuclear recoupling experiments such as REDOR [96,101]. By adjusting the length of the two REDOR mixing periods, it is possible to selectively re-couple N-C spin pairs separated by different distances.

Dihedral restraints for the  $\Phi$  and  $\Psi$  angles can be predicted in a manner analogous to solution NMR from the chemical shift values determined for the  $\text{C}\alpha$ ,  $\text{C}\beta$ ,  $\text{C}'$  carbon and the backbone nitrogen atoms using the software TALOS [102].

The protein structure calculation can be carried out in the same manner as when using restraints extracted from solution NMR data. The restraints are used in conjunction with a force field to carry out a molecular-dynamics-based simulated annealing protocol (effectively heating and cooling down the protein structure while imposing the experimental restraints and force field upon it). An ensemble of structures

that are all in agreement with the experimental data is generated and should represent the real structure of the protein.

Semi-automated approaches, which are used for solution NMR data such as the iterative assignment of NOESY peaks, can be adapted to solid-state MAS NMR data. The NOESY spectra are simply replaced by solid-state MAS NMR PDSO experiments such as 2D  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra and 3D NCACX and NCOCX spectra [103].

### **1.3 The outer membrane protein G (OmpG) from *E.coli***

The outer membrane of Gram-negative bacteria acts as a selective permeability barrier and prevents uncontrolled exchange of solutes and nutrients such as sugars, nucleotides, amino acids and ions between the cell exterior and interior. The uptake of these substances from the medium to the periplasm occurs through channel-forming integral MPs known as outer-membrane porins (Omps), which function as molecular filters. Omps are membrane-spanning  $\beta$ -barrels (surrounding an aqueous pore) with short turns extending into the periplasmic space and long flexible loops extending in the extra cellular space (Figure 1.12). Classical porins conform to a common pattern whereby three monomers associate into trimers. This is their oligomeric state in 2D and 3D crystals, and also their functional state in lipid bilayers [104].

However, Omps also serve as membrane anchors (OmpA), or promote bacterial adhesion to mammalian cells (OmpX) [105,106]. Others are membrane-bound enzymes, such as the protease OmpT or the phospholipase OmpLA [107,108]. Yet others are large barrels with a plug domain, and are specifically adapted to the sequestering of essential compounds such as siderophores for iron or vitamin B12 uptake (FhuA and BtuB) [109,110]. Additionally, similar  $\beta$ -barrel-forming membrane proteins are thought to participate in solute exchange and protein translocation in the outer membranes of chloroplasts and mitochondria [111,112]. All these examples demonstrate the great versatility of  $\beta$ -barrel Omps.

In *E. coli* the main porins for sugar uptake are LamB and ScrY [113,114]. Thus, porin proteins are critical for bacterial growth. However, it was shown that porin-deficient strains grow slowly and accumulate mutations that permit expression of new porin proteins. This finding led to the isolation of *E. coli* mutant strains that acquired the ability to synthesise new porin proteins [104]. The porin OmpG, for example, was

discovered among mutants (*E.coli* RAM 194) that gained increased outer membrane permeability in the absence of the classical porins OmpF and LamB [115]. The novel porin is produced in response to a chromosomal mutation termed *cog-192*. Molecular characterisation of *cog-192* revealed that it is a large chromosomal deletion that placed the OmpG expression under the control of the *pspA* promoter [116]. Growth and uptake experiments showed that OmpG displays a porin-like activity. Thus, OmpG seems to function as a rescue porin.

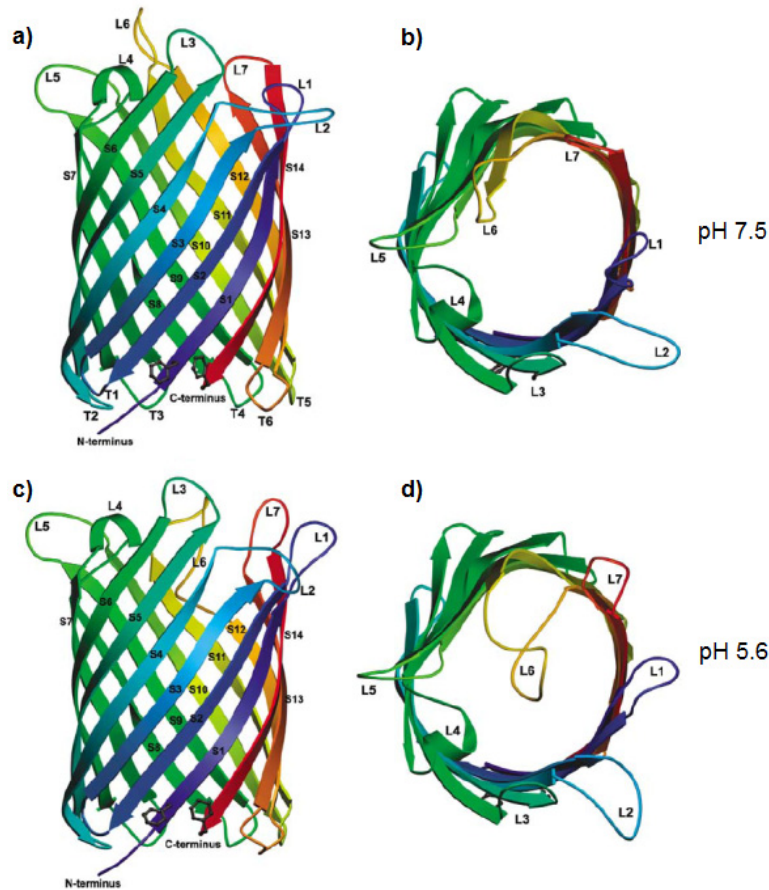
The gene for OmpG encodes a polypeptide with 301 amino acids, which show all features of an Omp: a signal sequence of 21 amino acids at its N-terminus, which is cleaved off during export into the outer membrane, absence of long hydrophobic stretches, lack of cysteine residues and a C-terminal phenylalanine, which is important for membrane insertion. Circular dichroism spectroscopy of detergent-solubilised OmpG indicated that, like other outer-membrane porins, it consists largely of  $\beta$ -sheet secondary structure with little  $\alpha$ -helix content. Proteoliposome swelling assays have shown that OmpG is a non-selective channel for mono-, di- and trisaccharides, with an unusually large limiting pore diameter of around 2 nm. In contrast to the classical Omps, chemical cross-linking and two-dimensional electrophoresis suggested that OmpG is a monomer [116].

A projection map at 6 Å resolution obtained by cryo-EM of 2D crystals revealed a ring-shaped density indicative of a monomeric  $\beta$ -barrel, with no evidence of oligomer formation [117]. Initial secondary structure predictions of OmpG suggested 16  $\beta$ -strands [116]. However, more recently this number was estimated to be 14, which agrees better with the diameter of the  $\beta$ -barrel in the projection map [117,118].

The crystal structure of OmpG was solved recently by two groups [119,120]. For one of these studies OmpG was extracted from native membranes with lauryl-*N,N*-dimethyldodecylamine-*N*-oxide (LDAO). During the purification step the detergent was exchanged against tetraoxy-ethylene-mono-*n*-octylether (C8E4) and the subsequent crystallisation was performed at pH 5.5 [119]. In contrast to that, Yildiz et al. expressed OmpG without its signal sequence, extracted and purified the protein from inclusion bodies and subsequently refolded the protein into detergent micelles. Two different crystal forms at pH 5.6 and pH 7.5 were obtained using octyl- $\beta$ -D-glycopyranoside (OG) and LDAO as detergent, respectively [120].

The three structures show similar structural features typical for  $\beta$ -barrel porins (Figure 1.12a and b). OmpG forms a monomeric 14-stranded ( $\beta$ 1- $\beta$ 14)  $\beta$ -barrel with six short

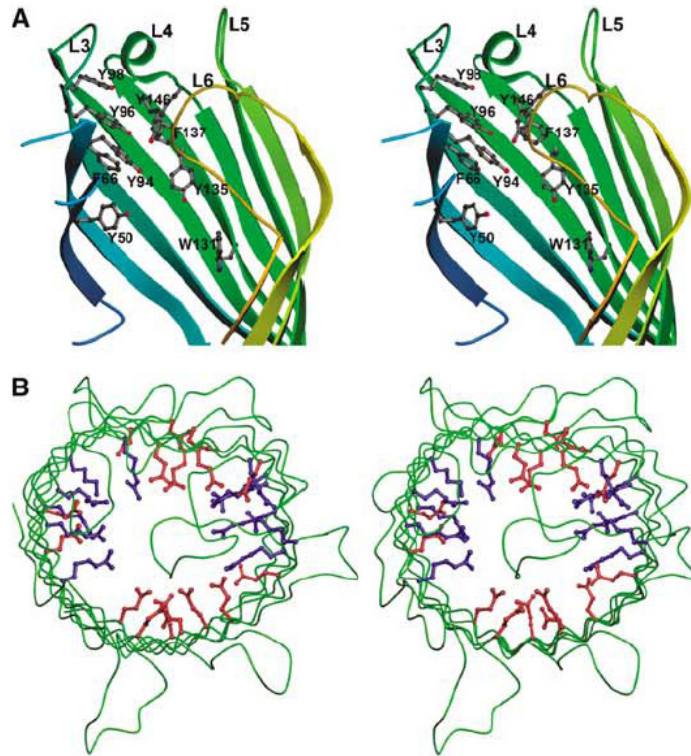
turns (T1-T6) extending into the periplasmic space and seven large loops (L1-L7) extending into the extracellular space. Loop 4 contains a short  $\alpha$ -helix. The averaged angle between the  $\beta$ -strands and the membrane plane is around  $43^\circ - 60^\circ$ .



**Figure 1.12** Crystal structures of OmpG in the open and closed state. a and b) OmpG in the open state (PDB code 2IWW). c and d) OmpG in the closed state at pH 5.6 (PDB code 2IWW). Figure taken from [120].

In all crystal structures the side chain densities were well defined except for the loop residues 20-27, 58-62, 220-231 and 261-266, indicating a degree of flexibility in L1, L2, L6 and L7. The pore diameter measured between C $\alpha$  atoms of opposite  $\beta$ -strands is  $\sim 25 \text{ \AA}$  or  $\sim 13 \text{ \AA}$  when taking side chains into account. The entry of disaccharides into the pore is facilitated by a line of aromatic residues (also termed an aromatic slide) adjusted parallel to the barrel axis (Figure 1.13) [119,120]. The authors suggested that the OH groups in the aromatic slide may function as a guide rail for incoming substrates. Additionally, the pore is lined by charged and polar residues (mainly Glu and Arg), which might create an electrostatic barrier that prevents ions and charged compounds from entering the channel. These data are consistent with the assumption that OmpG could be needed for the transport of large oligosaccharides under very specific growth

conditions [116]. As mentioned above, Yildiz et al. obtained two different crystal structures by crystallising OmpG at pH 7.5 and pH 5.6 [120].



**Figure 1.13 Stereo views of the arrangement of aromatic and positively and negatively charged residues in the OmpG pore.** (A) Lines of aromatic residues ('aromatic slides') leading from the extracellular channel entrance to the pore interior. (B) Top view of positive charged side chains in the pore (mostly Arg, blue) and negatively charged residues (mostly Glu, red) on opposite sides of the barrel wall. The figure was taken from [120].

The latter, shows a closed conformation in which loop 6 folds into the pore (Figure 1.12c and d). The authors postulate that the pH-induced conformational change of loop 6 is induced by the protonation of two solvent exposed histidine residues (His 231 and His 261) located at the C-terminal end of loop 6 and the N-terminal end of loop 7. At pH 7.5 the distance of the imidazol rings is 4.9 Å, whereas at pH 5.6, this distance increases to 13.7 Å. The unzipping of the hydrogen bond between His 231 and His 261 results in an increase in the length of loops 6 and 7. Loop 7 and the adjacent loop 1 move outward to make room for the reorientation of loop 6 into the pore [120].

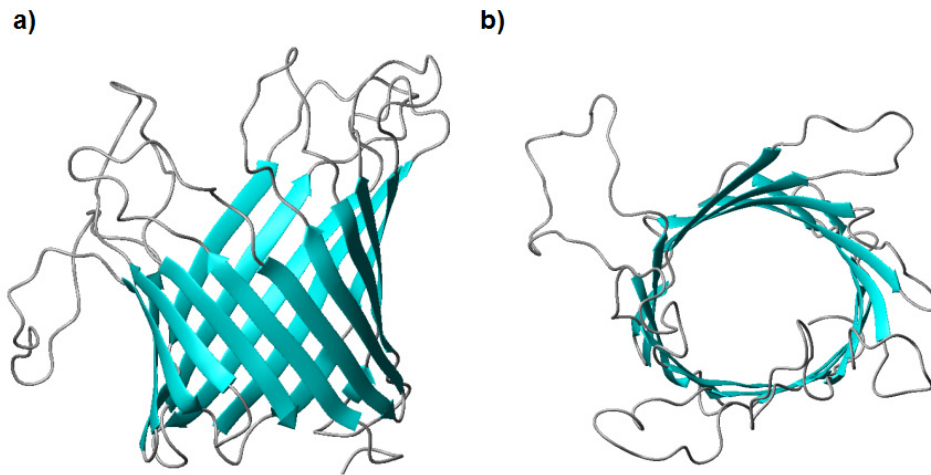
These findings are consistent with electrophysiological channel gating measurements with OmpG [118]. At neutral and basic pH the OmpG channel was open, but it was closed at acidic pH below pH 5.0. At pH 6.0, OmpG fluctuated between the open and closed state. Additionally, the projection map obtained from 2D crystals also shows restricting loops in the middle of the pore. This might be due to the fact that the 2D

crystals were stained with uranylacetate (pH ~ 4.0) immediately before the EM investigation. So far, OmpG is the only Omp for which pH depending gating has been demonstrated and structurally validated by EM and X-ray crystallography [117,118,120]. Surprisingly, the Subbarao crystal structure resembles the open state and not the closed state of the Yildiz structure, even though both crystals were formed at nearly the same pH (pH = 5.5 and 5.6, respectively) [119,120]. The reason for this discrepancy is presently unresolved.

More recently, a third OmpG structure determined by solution NMR was published [49]. Deuterated  $^{13}\text{C}$ ,  $^{15}\text{N}$  labelled OmpG sample was expressed as inclusion bodies. The protein was purified under denaturing conditions (8M urea) and subsequently refolded into OG detergent micelles. In a last step, the detergent was exchanged against dodecylphosphocholine (DPC) and NMR datasets were recorded at pH 6.3. The structure was determined with an rmsd of  $1.43 \pm 0.3 \text{ \AA}$  for the backbone of the  $\beta$ -barrel residues,  $1.67 \pm 0.29 \text{ \AA}$  for the backbone of the turn residues and  $7.37 \pm 1.29 \text{ \AA}$  for the backbone of the residues located in the loops (Figure 1.14). The structure is in good agreement with both crystal structures in the  $\beta$ -barrel region, although the barrel is a little shorter. Additionally, the extra cellular loop region indicates high flexibility. This might be due to the pH value (pH 6.3) at which the protein may undergo conformational exchange between the open and closed conformation on the microsecond to millisecond time scale.

Because of their robustness, pore forming  $\beta$ -barrel proteins have considerable potential in biotechnology as stochastic sensors for single molecule detection [121,122]. The ionic current flowing through a pore under an applied potential is altered when an analyte binds to the lumen of the pore. Measurement of the frequency of occurrence of binding events allows the determination of the concentration of the analyte. To date, studies of stochastic sensor elements have mostly focused on staphylococcal  $\alpha$ -hemolysin, a heptameric  $\beta$ -barrel pore-forming toxin [122-124]. However, the heptameric character of this pore makes it difficult to fine-tune the properties of the pore.





**Figure 1.14 Solution-structure of OmpG at pH 6.8.** Lowest-energy conformer of the 10 NMR structure ensemble calculated from NOE, dihedral angle, and hydrogen bond restraints.

Recently, it was shown by Chen et al., that the monomeric OmpG is well suited as a stochastic sensing pore [125]. However, it is known that OmpG exhibits spontaneous gating behaviour (switching between the open and closed state at neutral pH), which interferes with the application of the pore as a biosensor. Chen et al. found out by molecular dynamics simulations, that the flexibility of L6 and  $\beta$ 12 are responsible for most of the spontaneous gating activity. In a second step an OmpG mutant was generated to stabilise the position of  $\beta$ 12 and L6 in the open state. First, a disulfide bond was created between  $\beta$ 12 and  $\beta$ 13 by introducing cysteine residues at position G231 and D262, which are located at the extracellular ends of  $\beta$ 12 and  $\beta$ 13 (Figure 1.3). Second, D215 was deleted to optimise the hydrogen bonding between  $\beta$ 11 and  $\beta$ 12. It was possible to detect ADP by using this OmpG double mutant in combination with a molecular adapter [125]. The authors intend to expand the utility of OmpG as a sensor protein in the future.

## **1.4 Overview of the thesis**

This thesis describes the development of methods for the structure determination of large MPs by solid-state MAS NMR. Their potential is demonstrated by the partial resonance assignment and analysis of the secondary structure of the 281-residue OmpG protein, used as a model system. These methods should be widely applicable for structural investigation of those membrane proteins, which do not easily form 3D or 2D crystals or are not accessible to solution NMR.

The first step in a structural study by NMR is the production of sufficient amounts (> 10mg) of isotopically  $^{13}\text{C}$  and/or  $^{15}\text{N}$  labelled protein samples. In **chapter 3.1** the use of two *E. coli* expression protocols that allow the production of high amounts of uniformly- or selectively-labelled protein samples are described and compared. The protocols significantly reduce the costs of isotopically labelled sources and production time. Additionally, the protocols demonstrate that isotopically labelled cell mass is not a requirement for high-level  $^{13}\text{C}$ ,  $^{15}\text{N}$  incorporation during recombinant protein expression.

In **chapter 3.2** the preparation of different isotopically labelled OmpG sample types, including poorly diffracting 2D crystals, proteoliposomes and a PEG precipitate, is described. The spectral resolution of these samples was compared in 1D and 2D solid-state MAS NMR spectra. The results demonstrate that for solid-state MAS NMR a homogeneous sample with a high degree of local order is required. In contrast to that, a crystalline sample, which contains long-range order, is necessary for high-resolution structural investigations by X-ray crystallography or EM. Thus, solid-state MAS NMR can be a helpful tool for the structural investigation of MPs, which do not easily form 3D or 2D crystals.

In order to obtain structural information from NMR spectra, the major prerequisite and the most time consuming step, is always the assignment of resonances to specific nuclei within the protein. In **chapter 3.3** an assignment strategy that allows the assignment of large MPs is introduced. For this purpose multidimensional NMR experiments with different specifically labelled samples were recorded. The resulting spectra were used to find starting points for the assignment and then connect the assigned fragments to each other. Additionally, the extracted chemical shift information is used to calculate secondary chemical shifts, which gives a first insight into the secondary structure of OmpG. A comparison with the published X-ray structure of OmpG shows high

similarity, which confirms the usefulness of solid-state MAS NMR for the structural investigation of MPs.

## 2 Material and Methods

### 2.1 Materials

All chemicals, if not indicated otherwise, were purchased from VWR International GmbH (Germany) or Carl Roth GmbH & Co. KG (Germany).

#### 2.1.1 *E. coli* expression and cloning strains

Strain	genotype	company
BL21 (DE3)	F <sup>-</sup> , <i>ompT</i> , <i>hsdS<sub>B</sub></i> (r <sub>B</sub> <sup>-</sup> m <sub>B</sub> <sup>-</sup> ), <i>gal</i> , <i>dcm</i> (DE3)	Merck Biosciences Germany
C43 (DE3)	derivate of BL21(DE3) [13]	Imaxio /Biopole Clermont-Limagne France
DH5α	F <sup>-</sup> , φ80 <i>lacZ</i> ΔM15, Δ( <i>lacZYA-argF</i> )U169, <i>recA1</i> , <i>endA1</i> , <i>hsdR17</i> (r <sub>k</sub> <sup>-</sup> , m <sub>k</sub> <sup>+</sup> ), <i>phoA</i> , <i>supE44</i> , <i>thi-1</i> , <i>gyrA96</i> , <i>relA1</i> , λ <sup>-</sup>	Invitrogen, Germany

#### 2.1.2 Expression-plasmid

The pET-OmpG plasmid was a gift from Prof. W. Kühlbrandt from the MPI for Biophysics (Frankfurt am Main, Germany). The DNA of the mature form of OmpG was cloned into the pET-26b plasmid (Novagen, Merck Biosciences, Germany) using the *NdeI* and *SacI* restriction sites. The coding sequence for OmpG was amplified from chromosomal DNA (D2001) of *E. coli* strain B (SIGMA, Deisenhofen, Germany) using the following primers: forward, 5'-GATCTCGGTTGGGCTGGCTTCTGTC TCCCT; reverse, 5'-CCGACGCAGGA GTTAGGTCAACAAAGCTGCG. The product was used in a second PCR (primers: forward, 5'-GGCCTGCGCACATA TGGAGGAAAGGAACGAC; reverse 5'-CGGATAAGGAGCTCGCGCCG CATCC)

to introduce the restriction sites for *NdeI* and *SacI*. The amplified DNA was digested with *NdeI* and *SacI* and cloned into a pET-26b expression plasmid from Novagen (Bad Soden, Germany). The resulting construct, termed pET-OmpG, contained the mature form of OmpG (280 amino acids) without the signal sequence, but with an N-terminal methionine.

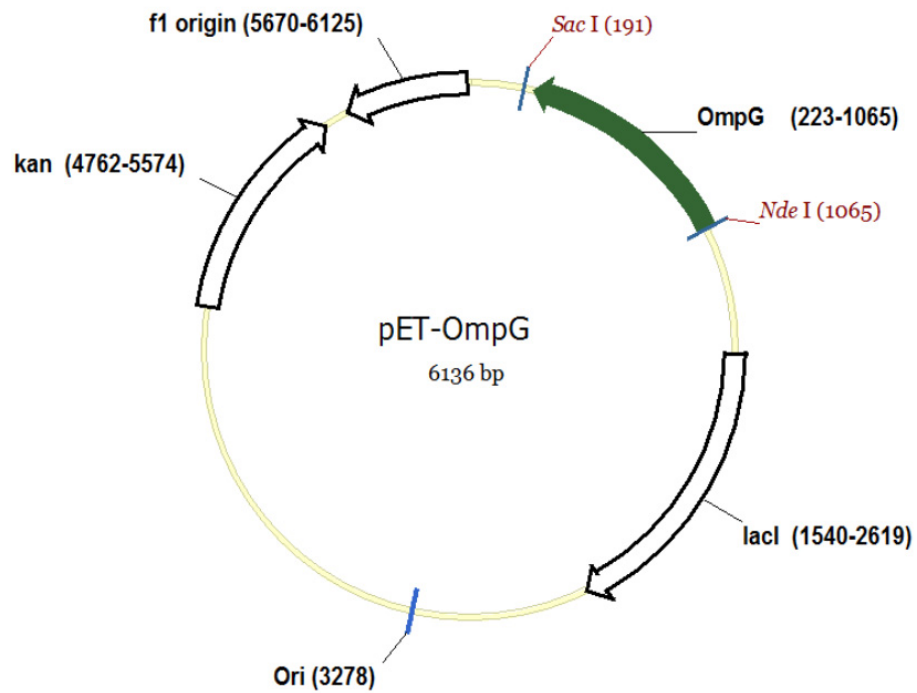


Figure 2.1. Plasmid-map of the pET-OmpG construct.

### 2.1.3 Isotopically labelled substances

All isotopically labelled substances were purchased from Cambridge Isotope Laboratories, Inc. (Andover, USA).

## 2.2 Methods

Basic techniques for handling of DNA and cells were essentially followed as described in the manual “Molecular Cloning: A Laboratory Manual” by Sambrook and Russell [126].

### 2.2.1 Expression of labelled and unlabelled OmpG in *E. coli* BL21(DE3)

#### 2.2.1.1 Isotopically labelled supplements for different labelling-schemes

Unlabelled OmpG was expressed in LB-medium. Expression of labelled OmpG was conducted in standard M9 minimal medium. The standard expression protocol was done using shaking flasks. The amounts of sole carbon and nitrogen sources for uniformly- and site-specifically labelled OmpG samples are listed below (Table 2.1).

**Table 2.1** Amounts of sole carbon and nitrogen sources for labelled growth media.

labelling scheme	Sample	labelled source	amount [g/l culture]
uniformly $^{13}\text{C}$ - and $^{15}\text{N}$ -labelling	u- $^{13}\text{C}$ , $^{15}\text{N}$ -OmpG	u- $^{13}\text{C}$ -glucose $^{15}\text{N}$ - $\text{NH}_4\text{Cl}$	2 0.5
site specific $^{13}\text{C}$ - and $^{15}\text{N}$ -labelling	[1,3 $^{13}\text{C}$ , $^{15}\text{N}$ ]-OmpG	[1,3 $^{13}\text{C}$ ]-glycerol $^{15}\text{N}$ - $\text{NH}_4\text{Cl}$ $\text{NaHCO}_3$	2 0.5 2
	[2 $^{13}\text{C}$ , $^{15}\text{N}$ ]-OmpG	[2 $^{13}\text{C}$ ]-glycerol $^{15}\text{N}$ - $\text{NH}_4\text{Cl}$ $^{13}\text{C}$ - $\text{NaHCO}_3$	2 0.5 2

Amino-acid selectively labelled OmpG samples were prepared using the modified (2.2.1.3) and high cell-density fermentation (HCDF, 2.2.1.4) procedures by forward- and reverse labelling. The amounts of labelled and unlabelled supplements used are listed separately.

### **2.2.1.2 Standard expression protocol**

An overnight LB culture of BL21(DE3) transformed with pET-OmpG was used to inoculate one litre of LB or M9-minimal medium culture. Cells were shaken at 37 °C and 180 rpm until an optical cell density of ~0.6 measured at 600 nm (OD<sub>600</sub>) was reached. The expression of OmpG was induced by the addition of 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After three hours of expression, cells were harvested by centrifugation (5000 g, 10 min at 8 °C). Cells were resuspended and washed with an ice cold 0.15 M NaCl solution and centrifuged again. The amount of biomass was determined and cells were stored at –80 °C.

### **2.2.1.3 Modified expression method [127]**

A four litre pre-culture (LB-medium) was inoculated with an overnight culture. The pre-culture was shaken at 37 °C and 180 rpm up to an OD<sub>600</sub> of 0.5. Cells were harvested by centrifugation (5000 xg, 10 min at 8 °C) and washed with 200 ml of ice cold 1x M9 salt-solution. Cells were harvested again and subsequently resuspended in one or two litres of M9-minimal medium, depending on the type of carbon source in the medium. Additionally, it was important to use 2x M9-salt solution in the medium in order to stabilise the pH.

If glucose was used as the carbon source, cells from a 4 litre LB-culture were resuspended in one litre of M9-minimal medium supplemented with the indicated amounts of carbon and nitrogen sources (Table 2.1). When glycerol was used as the carbon source, cells from a 4 litre LB culture were resuspended in 2 litres of M9 minimal medium supplemented with the indicated amounts of carbon and nitrogen sources (Table 2.1).

For amino acid selective labelling a forward labelling strategy was used. Cells from a 4 litre LB culture were resuspended in 2 litres of M9 minimal medium containing all amino acids (200 mg) in unlabelled or labelled form, 2 g of glucose and 0.5 g of NH<sub>4</sub>Cl. Cells were then shaken at 37 °C and 180 rpm for one hour for recovery of growth and clearance of unlabelled metabolites. Then protein expression was induced by the addition of 1 mM IPTG. After three hours of incubation, the cells were harvested and stored as described under 2.2.1.2.

#### 2.2.1.4 High cell density fermentation (HCDF)

A fully automated high cell density fermentation (HCDF)-system from DASGIP® was used to produce mainly amino acid selectively labelled OmpG samples. The system allows production of high cell densities and precise process control (e.g control of pH, temperature, O<sub>2</sub> supply and stirring speed and the nutrient feed) simultaneously. The fermentation consists of a batch phase, a biomass accumulating phase and an expression phase [128].

In the following, the process control and the media composition during the fermentation is described. During the entire process the pH-value (7.0), temperature (30 °C), agitator-speed and pO<sub>2</sub>-value were measured and regulated computer-aided. The media composition used here was described by Fiedler *et al.* (2007) [128].

The batch medium was inoculated with an overnight culture up to an OD<sub>600</sub> of around 0.5. When the pO<sub>2</sub> reached a value >30% during the batch phase, the agitation-speed was increased up to 1000 rpm to ensure that all nutrients in the batch were fully consumed. As a result of batch consumption the pO<sub>2</sub>-value exceeded 50%. This was the starting point for the biomass accumulating feed, which was dosed into the reactor until the pO<sub>2</sub>-value dropped to 30%. The feed was stopped and the pO<sub>2</sub> increased again. This regulation scheme was used during the whole process to obtain a constant growth rate and to avoid over-feeding. After four hours of biomass accumulating feed the process automatically switched to the labelled feed for another four to seven hours.

To produce selectively labelled OmpG samples by **forward labelling**, a separate amino-acid feed (140 ml) was used. 130 mg of each amino acid except tyrosine (100 mg) was dissolved in 140 ml of 2x M9-salt solution. At the beginning of the expression phase 35 ml of the amino acid feed were manually added. After 30 min the expression was induced by the addition of 1 mM IPTG (5 ml). The remainder amino acid feed was then dosed using a pump with a rate of 30 ml/.

**Reverse labelling** of selectively labelled OmpG samples was done using an expression feed containing U-[<sup>13</sup>C]-glucose. Amino acids, which were not to be labelled were added in unlabelled form (amino acid feed see above). All other settings were the same as above. At the end of the fermentation, cultures were cooled down to 10 °C. Cells were harvested as described above (2.2.1.2).



## **2.2.2 Protein purification from inclusion bodies**

OmpG purification was carried out in a similar manner to that described by Conlan et al.[13]. Cells (3 g biomass) were resuspended in 30 ml 10 mM Tris-HCl at pH 8.0 (buffer A) containing a protease-inhibitor cocktail (Complete, Roche) and broken by using a French Press. The cell homogenate was centrifuged at 60,000 g and 4 °C for 30 min. The resulting insoluble fraction was washed first in 30 ml buffer A containing 1 % (w/v) Triton X-100, 1 M Urea and then in 30 ml buffer A containing 1 M urea and was centrifuged after each washing step. The resulting inclusion body pellet was resuspended in 30 ml buffer A containing 8 M urea and incubated at 8 °C. After 3 hours the resulting solution was centrifuged at 60,000 g and 4 °C for 30 min. The solubilised inclusion body fraction was then loaded on to a Q-Sepharose Fast-Flow column with a bed volume of 180 ml. The column was washed with three column volumes of buffer A containing urea (8 M). OmpG was eluted with a linear gradient of NaCl (0-1 M). The final concentration was determined by measuring the absorbance at 280 nm in buffer A containing urea (8 M) with an extinction coefficient of 85060 M<sup>-1</sup> cm<sup>-1</sup> [129].

## **2.2.3 OmpG refolding**

### **2.2.3.1 Refolding screen**

OmpG was refolded by rapid dilution in the presence of detergent micelles. A screen using detergents with different head groups and chain lengths was conducted. Typically, denatured protein was diluted into detergent containing buffer up to a protein concentration of 50 µg/ml. The concentration of detergents in buffer A was set to 1-3 times their critical micelle concentration (cmc). The refolding efficiency was monitored by differential migration of the folded (28 kDa) and unfolded (34 kDa) form of OmpG in the SDS-PAGE gel [116]. The detergents used are listed in table 2.2.

Table 2.2 List of detergents used for OmpG refolding.

detergent	cmc [mM]	supplier
<b><math>\beta</math>-D-glycopyranoside</b>		
hexyl-	250	<b>Glycon Biochemicals</b> (Germany)
heptyl-	79	
octyl	20-25	
nonyl-	6.5	
decyl-	2.2	
<b>N-methyl glucamide</b>		
Mega 8	108	<b>Glycon Biochemicals</b> (Germany)
Mega 9	25	
<b><math>\beta</math>-D-maltoside</b>		
nonyl-	6	<b>Glycon Biochemicals</b> (Germany)
decyl-	2.2	
undecyl-	0.59	
dodecyl-	0.15	
<b>Cyclohexyl*-<math>\beta</math>-D-maltoside</b>		
Cymal 6 <sup>(*)</sup>	0.56	<b>Antrace Inc. (USA)</b>
Cymal 7 <sup>(*)</sup>	0.19	
<b>polyoxyethylene</b>		
C <sub>8</sub> E <sub>5</sub>	7.1	<b>BACHEM (Germany)</b>
C <sub>10</sub> E <sub>6</sub>	0.9	
Triton X-100	0.23	<b>Roth GmbH (Germany)</b>
Genapol X-80	0.06-0.15	<b>Sigma-Aldrich (Germany)</b>
<b>zwitterionic</b>		
LDAO	2	<b>Sigma-Aldrich (Germany)</b>
Anergent 3-10	39	<b>Antrace Inc. (USA)</b>
Anergent 3-12	2.8	

### 2.2.3.2 Large-scale-refolding

Large-scale refolding of OmpG, (100 mg denatured protein) was done by rapid dilution and agitation into two litres of buffer A containing 1 mM DDM and 0.6 mM L-arginine at 8 °C using a peristaltic pump with a flow rate of 0.1 ml/min. The refolded protein was concentrated and washed with buffer A containing 1 mM dodecyl- $\beta$ -D-maltoside (DDM) to a final concentration of 1–2 mg/ml in an ultra-filtration chamber (Millipore, Schwalbach, Germany) with a membrane cut-off of 30 kDa.

### 2.2.3.3 Enrichment of folded OmpG

Detergent concentration in the refolded OmpG sample was reduced by using a second anion-exchange chromatography column (6 ml Resource-Q, GE Healthcare, Germany) The protein was bound to the column and subsequently washed with 3 column volumes of buffer A containing 0.4 mM DDM. OmpG was eluted at 300 mM NaCl in buffer A containing 0.4 mM DDM. To enrich the folded form of OmpG all fractions were analysed by SDS-PAGE. Fractions containing folded OmpG were pooled and concentrated using a centrifugal filter device (Ultrafree-15, cut-off 50 kDa, Millipore) to a final concentration of around 3 mg/ml.

### 2.2.4 Reconstitution into lipid bilayers and two dimensional crystallisation

Refolded OmpG was reconstituted into lipid bilayers. For this purpose, an *E. coli* total lipid chloroform extract (20 mg) was dried in a nitrogen stream and subsequently in a vacuum pump. The resulting lipid film was resuspended in buffer A (5 ml) containing 34 mM octyl- $\beta$ -D-glycopyranoside. Aliquots of this lipid solution and refolded OmpG (~3 mg/ml) were mixed to yield a lipid-to-protein ratio (LPR) of 1:2 (w/w). For 2D crystallisation, the detergent was removed by dialysis (dialysis-tube molecular weight cut-off 25 kDa, Roth, Karlsruhe, Germany) at 20 °C against 5 litres of a buffer containing 20 mM Tris-HCl, pH 7.0, 150 mM NaCl, 25 mM MgCl<sub>2</sub>, 3 mM NaN<sub>3</sub>. The dialysis buffer was changed every 5 days. 2D crystals were observed after 7 days. To ensure complete removal of the detergent the crystal solution was dialysed for a further 4 weeks.

### 2.2.5 Protein PEG-precipitation

Crystallisation conditions were used as described by Yildiz *et al* [120]. 3.3 ml protein (3mg/ml) in buffer A containing 0.4 mM DDM was mixed with 3.3 ml crystallisation buffer containing 100 mM HEPES pH 7.5, 200 mM CaCl<sub>2</sub> and 30 % (w/v) PEG 4000 in a sitting drop. The 3D crystalline PEG-precipitate was formed directly during the mixing step. The resulting sample was analysed by microscopy and collected by centrifugation (see 2.2.7 and 2.2.9).

### **2.2.6 Circular dichroism (CD) spectroscopy**

CD spectra were taken on a J720 CD-spectrometer from Jasco (Tokyo, Japan) in a quartz cuvette with 1 mm path length. Unfolded and folded protein (25  $\mu$ M) was measured in buffer A containing urea (8 M) or DDM (0.4 mM), respectively. Each sample was scanned 5 times from 260 to 190 nm with a step size of 1 nm. Data were baseline corrected and converted to molar residue ellipticity.

### **2.2.7 Electron microscopy**

The crystallised samples (2–4  $\mu$ l) were applied to a carbon-coated grid (400 mesh size, copper rhodium) from PLANO (Wetzlar, Germany) and incubated for 20 s. The grid was stained with 2  $\mu$ l of uranyl-acetate solution (2% (w/v)) and residual solution was removed with filter paper. Images were recorded on a CM12 electron microscope (Philips) at an accelerating voltage of 120 kV with an electron dose of 10 e<sup>-</sup>A<sup>-2</sup> on Kodak SO-163 film at different magnifications.

### **2.2.8 Solid-state MAS NMR spectroscopy**

Before using samples for solid-state MAS NMR investigations, the buffer was exchanged to a phosphate buffer (20 mM, pH 7.0) containing 50 mM NaCl. Samples were centrifuged at 60,000 g for 15 min and the resulting pellet was resuspended in phosphate buffer three times. 2D crystals of OmpG were finally sedimented in an ultracentrifuge at 100,000 g for 1 h and at 4 °C.

PEG precipitated samples were finally centrifuged at 60,000 g, 4 °C for 1 h. Afterwards, samples were transferred into a 3.2 mm or 4 mm MAS rotor (Bruker, Karlsruhe, Germany) by centrifugation in a table top centrifuge.

All NMR measurements were performed on AVANCE-900 and AVANCE-400 spectrometers (Bruker), equipped with a 3.2 mm or 4 mm triple-resonance MAS probe (Bruker).

### 2.2.8.1 1D $^{13}\text{C}$ - and $^{15}\text{N}$ -CP/MAS experiments

One-dimensional (1D) CP/MAS NMR spectra were recorded on a 400 MHz wide-bore or on a 900 MHz narrow-bore spectrometer at a MAS frequency ( $\omega_{\text{R}}/2\pi$ ) of 8.0 kHz or 13 kHz, respectively. Magnetisation was transferred from  $^1\text{H}$  to  $^{13}\text{C}$  or  $^1\text{H}$  to  $^{15}\text{N}$  with a ramped CP with a contact time of 1500  $\mu\text{s}$  [88]. Spin lock field strengths of  $\sim 55$  kHz for  $^1\text{H}$ ,  $\sim 70$  kHz for the  $^{13}\text{C}$  and  $\sim 45$  kHz for  $^{15}\text{N}$  were used. During the acquisition time proton decoupling was applied. For this the two-pulse phase modulation (TPPM) or the small phase increment alteration technique (SPINAL 64) with a proton RF field of  $\sim 90$  kHz was used [83,84].

### 2.2.8.2 2D $^{13}\text{C}$ - $^{13}\text{C}$ PDSM experiments

2D  $^{13}\text{C}$ - $^{13}\text{C}$  spectra of OmpG were recorded on a 700 MHz wide-bore or 900 MHz narrow-bore spectrometer at a MAS frequency ( $\omega_{\text{R}}/2\pi$ ) of 10 or 13 kHz, using the proton-driven spin-diffusion (PDSM) mixing scheme [89]. Magnetisation was transferred from  $^1\text{H}$  to  $^{13}\text{C}$  with a ramped CP of 1500  $\mu\text{s}$  and spin lock field strengths of  $\sim 55$  kHz for  $^1\text{H}$  and  $\sim 70$  kHz for the  $^{13}\text{C}$  ramp. PDSM mixing times between 20-700 ms in combination with dipolar assisted rotational resonance (DARR) were chosen [90]. A proton rf field of 90 kHz was applied for SPINAL 64 decoupling. 256 scans per increment were collected, with an effective evolution time of  $\sim 6$  ms in the indirect dimension. The 2D data were thus recorded in 2-5 days, depending on the length of the PDSM mixing time.

### 2.2.8.3 2D NCA/NCO experiments [91]

2D NCA/NCO spectra were recorded on a Bruker 400 MHz wide-bore spectrometer at 280 K. First, magnetisation was transferred from  $^1\text{H}$  to  $^{15}\text{N}$  with a ramped CP of 1500  $\mu\text{s}$  and spin lock field strengths of  $\sim 55$  kHz for  $^1\text{H}$  and  $\sim 45$ -50 kHz for the  $^{15}\text{N}$  ramp. Following the evolution of nitrogen, adiabatic CP was employed to selectively transfer magnetization from  $^{15}\text{N}$  to  $^{13}\text{C}$  with RF powers of  $\sim 15$  and 23 kHz, respectively. The amide signals and the C $\alpha$  or C' signals were irradiated on-resonance. During the evolution periods, proton decoupling was applied using SPINAL64 (90 kHz). 256 scans per increment were collected, with an effective evolution time of 2.6 ms in the indirect dimension. The data set was thus recorded in 12 hours.

#### **2.2.8.4 2D NCACX/NCOCX experiments [61]**

2D NCACX/NCOCX spectra were recorded in the same way as NCA/NCO experiments (2.2.9.3). However, after the  $^{13}\text{C}$ -evolution period a PDSM mixing scheme with different  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times (20-50 ms) for the NCACX and NCOCX experiments, was included. During all evolution periods proton decoupling was applied using SPINAL64 (90 kHz). Data sets were recorded using evolution times of 2.6 ms.

#### **2.2.8.5 3D NCACX/NCOCX experiments**

Spectra were recorded as described under 3.2.9.3, implementing a PDSM mixing scheme of between 20 and 400 ms in combination with DARR. SPINAL64 was applied for proton decoupling (90 kHz). The data sets were recorded using evolution times of 6.8 ms in  $t_1$  and 6.4 ms in  $t_2$ . Each FID was averaged from 128 scans yielding a total measurement time of 8-10 days per spectrum.

#### **2.2.8.6 2D $^{15}\text{N}$ - $^{13}\text{C}$ TEDOR experiments**

A TEDOR transfer scheme was used for the  $^{13}\text{C}$ - $^{15}\text{N}$  correlation experiment on a Bruker 400 MHz wide bore spectrometer at 280 K [101,130]. The REDOR mixing time was 1 ms for one bond transfer at a spinning frequency  $\omega_R/2\pi = 8$  kHz. Typical  $^{15}\text{N}$   $\pi$ -pulse lengths during the REDOR mixing for OmpG were 13  $\mu\text{s}$ . After  $t_1$ -evolution of  $^{15}\text{N}$ , a z-filter was used. During  $t_1$ -evolution, acquisition and REDOR-mixing, SPINAL64 decoupling at 90 kHz was used for proton decoupling. The  $t_1$ -increment was chosen to be  $1/\omega_R = 12.5$   $\mu\text{s}$ . The total indirect acquisition time was 6 ms.

#### **2.2.8.7 2D CHHC experiment**

The 2D CHHC spectrum was recorded on a 900 MHz narrow bore spectrometer at a MAS frequency of 13 kHz and 280 K. The magnetisation was transferred from  $^1\text{H}$  to  $^{13}\text{C}$  with a CP contact time of 1.5 ms. Spin lock field strengths of  $\sim 55$  kHz for  $^1\text{H}$ ,  $\sim 70$  kHz for the  $^{13}\text{C}$  and  $\sim 45$  kHz for  $^{15}\text{N}$  were used. The second and third CP were each 50  $\mu\text{s}$  long to guarantee for one bond  $^1\text{H}$ - $^{13}\text{C}$  transfer. 100  $\mu\text{s}$   $^1\text{H}$  spin diffusion mixing was used for proton-proton recoupling. 384 transients were recorded in  $t_1$  with a total indirect acquisition time of  $\sim 8$  ms. The direct acquisition time was 15 ms.

### 2.2.8.8 Methyl-filtered spectra

Methyl-filtered solid-state NMR spectra of OmpG were recorded at a MAS frequency ( $\omega_R/2\pi$ ) of 8.0 kHz at 280K on a AVANCE 400 wide-bore spectrometer. The spectrometer was equipped with a 4 mm triple-resonance CP/MAS probe.

A ramped CP from  $^1\text{H}$  to  $^{13}\text{C}$  was applied, followed by a squared phase-inversion CP at the  $n = 1$  Hartmann-Hahn condition. This polarisation inversion (PI) period cancelled the CH-signals, whereas  $\text{CH}_2$  groups were inverted to approximately  $-1/3$  of their initial intensity. The PI time was determined experimentally. It was found that 50  $\mu\text{s}$  lead to the desired signal inversion. Following PI, the  $\text{CH}_2$  and residual CH signals were zeroed by applying a short (60  $\mu\text{s}$ ) dipolardephasing period. The methyl (and non-protonated) carbons were weaker coupled to  $^1\text{H}$  and were not dephased. The carbon chemical shift evolution was refocused by a  $\pi$ -pulse in the middle of the dephasing period.

For the  $^{13}\text{C}$ - $^{15}\text{N}$  exchange, a TEDOR transfer scheme was introduced after selection of methyl carbons by CPPI and dipolar dephasing. REDOR mixing times of 2, 3, 6 or 8 ms were applied in these studies, to achieve transfer over distances from  $\sim 2.5 \text{ \AA}$  (short mixing time) up to about 5  $\text{\AA}$  for the longest mixing times<sup>1</sup>. Typical  $^{15}\text{N}$   $\pi$ -pulse lengths during the REDOR mixing were 13.3  $\mu\text{s}$ . After  $t_1$ -evolution of  $^{15}\text{N}$ , a z-filter was used to retain rotor synchronisation by compensating for pulse durations and to remove transverse spin components. During  $t_1$ -evolution, acquisition and REDOR-mixing, two-pulse phase-modulation (TPPM) decoupling at  $\sim 87 \text{ kHz}$  were used for heteronuclear decoupling. The  $t_1$ -increment was chosen to  $1/\omega_R = 125 \mu\text{s}$ . The total indirect acquisition time amounted to 6 ms.

### 2.2.8.9 NMR-Data processing and analysis

All NMR data were processed with TOPSPIN, version 2.0 (Bruker, Germany) and analysed using the Sparky version 3.11 (T. D. Goddard, D. G. Kneller, University of California, USA) or CCPNmr Analysis [131].

## 3 Results and Discussion

### 3.1 Expression and purification of isotopically labelled OmpG

#### 3.1.1 Expression and purification of OmpG from inclusion bodies

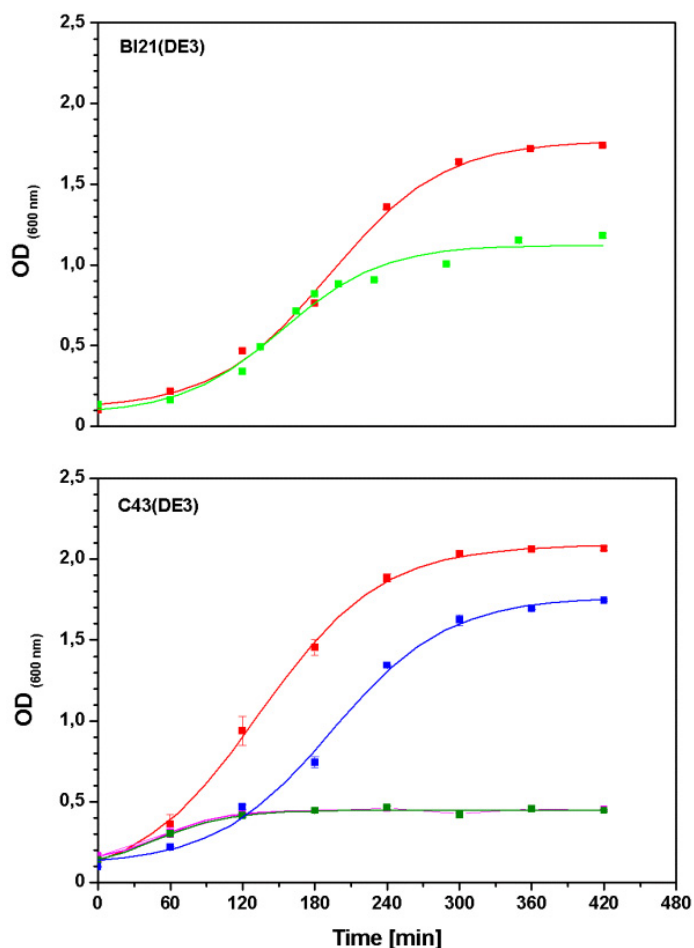
An essential pre-requisite for structural analysis of proteins by NMR is the availability of high amounts (> 10mg) of a homogenous isotopically labelled protein sample both in terms of purity and conformation. However, native expression of OmpG in the *E. coli* strain RAM194 only yields around 1 mg/L culture and co-purifies with a large amount of lipids [116,117]. Alternatively, OmpG was expressed without the leader peptide causing the formation of a non-functional cytoplasmic *inclusion body* precipitate. Subsequent *in vitro* refolding yielded 12.5 mg per litre culture of a functional OmpG sample, as was demonstrated by planar lipid bilayer recordings [118]. Thus, the expression of OmpG in inclusion bodies was favoured for the following reasons:

1. For solid-state MAS NMR measurements the proteins must be either  $^{13}\text{C}$ - and/or  $^{15}\text{N}$ -labelled, which means that cultures must be grown in M9-minimal media with the addition of isotopically labelled carbon and nitrogen sources. In the case of native expression, this would result in high costs for 10 mg of labelled protein.
2. The expression in inclusion bodies is by itself a purification step, since they often almost exclusively contain the over-expressed protein (>95 %).
3. Many other outer membrane porins have been successfully refolded with high yields [21]. Therefore, inclusion body formation provides a straight forward strategy for recombinant protein expression, which can be adapted for OmpG.

OmpG was expressed without the signal peptide but with a N-terminal methionine (starting with M-EERNDWH...) in *E. coli* BL21(DE3) and C43(DE3) [13,132]. The transformed pET-OmpG plasmid was a gift of the group of Prof. W. Kühlbrandt from the MPI of Biophysics (Frankfurt am Main, plasmid map see section 2).



Using LB-medium for cultivation, the C43(DE3) strain achieved slightly higher cell densities than the BL21(DE3) strain (Figure 3.1). In M9 minimal media, however, the C43(DE3) strain had a very short logarithmic growth phase and only achieved low cell densities. Instead the BL21(DE3) strain reached optical cell densities of 1-1.2.

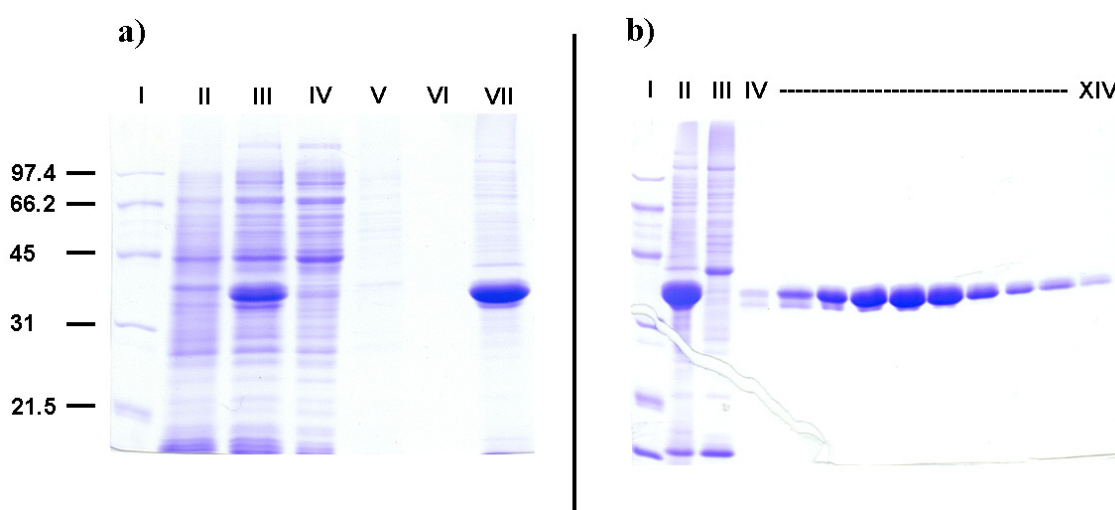


**Figure 3.1. Growth behaviour of *E. coli* BL21(DE3) and C43(DE3) expressing OmpG.** The expression of OmpG was induced at an optical density of 0.6 by the addition of 1mM IPTG. a) Growth curves of BL21(DE3) in LB- (red) and M9-minimal medium (green). b) C43(DE) in LB- (red), M9-minimal (green), in M9-minimal medium without the transformation of the plasmid (magenta) and a mixture (50:50 v/v) of M9-minimal and Silantes OD2 medium (blue).

To test whether the toxicity comes from the expression plasmid, C43(DE3) was grown on M9 minimal medium without the transformation of pET-OmpG. However, the culture showed the same growth behaviour as the transformed culture. The problem could only be overcome by using a mixture (50:50 v/v) of M9-minimal medium and Silantes OD2 medium. The latter is an isotopic labelled rich growth medium optimised to ensure reproducible growth rates, cell densities and high yields of recombinant expressed protein. In this media mixture, the C43(DE3) strain achieved normal growth rates (Figure 3.1). This indicates that growing C43(DE3) in M9 minimal medium is

hampered by a lack of important nutrients. Thus all further protein expression experiments were done using the *E. coli* BL21(DE3) strain.

Recombinant OmpG was found exclusively in inclusion bodies, which were washed and separated from cell debris by several centrifugation steps (Figure 3.2a). Subsequently, OmpG was solubilised in a buffer containing 8 M urea and purified under denaturing conditions by anion exchange chromatography (Source Q Column, Figure 3.2b lane IV to XIV). This procedure yielded up to 25 mg of pure denatured U- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG per litre culture. However, the 25 mg of OmpG are denatured protein and have to be refolded in detergent micelles.



**Figure 3.2. Purification of OmpG.** a) **Separation and solubilisation of the inclusion body fraction.** Lane I, molecular weight marker; Lane II *E. coli* BL21 (DE3) un-induced; Lane III, *E. coli* BL21 (DE3) after 3 hours of induction; Lane IV soluble protein fraction after FRENCH-Press; Lane V supernatant after first inclusion body wash step; Lane VI supernatant after second inclusion body wash step; Lane VII, inclusion body fraction after solubilisation with 8 M urea. b) **OmpG purification by anion exchange chromatography.** Lane I, molecular weight marker; Lane II, inclusion body fraction before loading onto the column; Lane III, eluate at 200 mM NaCl; Lane IV to XIV, eluate at 300 mM NaCl.

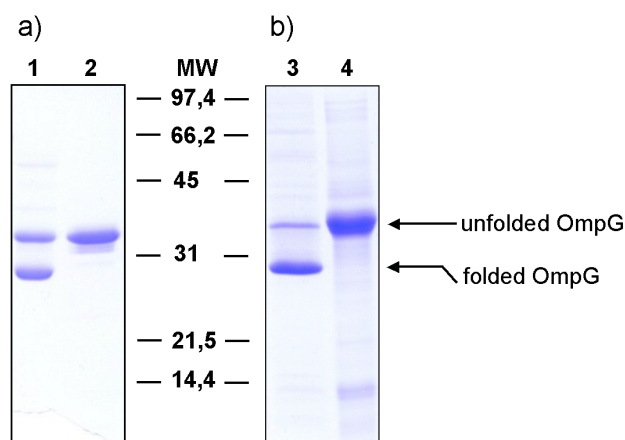
### 3.1.2 Refolding of OmpG

The initial refolding screens were done in co-operation with K.R. Vinothkumar from the group of Prof. W. Kühnbrandt (MPI of Biophysics, Frankfurt am Main). Refolding experiments on the bacterial OmpA showed that refolding is achieved when the detergent concentration is above its critical micelle concentration (cmc)[133]. OmpG refolding was screened by dilution of the denatured form into different classes of detergents with variable chain lengths at concentrations above their cmc. The refolding efficiency was monitored and compared by differential migration of the folded and the

unfolded form of OmpG on SDS-PAGE gels, which depends on whether the sample was heated before analysis or not [134]. The denatured form of OmpG in urea migrates at an apparent molecular weight of 34 kDa, whereas the refolded form runs at around 28 kDa (Figure 3.3) [116]. Therefore, all gels were run at 4 °C to prevent heat denaturing during the gel run. Here, a summary of the results of the initial OmpG refolding screens will be presented.

Using zwitterionic detergents such as LDAO, Anergent 3-10 and Anergent 3-12 a low refolding efficiency was observed and the protein tended to form higher aggregates. Non-ionic detergents belonging to the polyoxyethylene group (Genapol X-80 and Triton X-100) yielded a refolding efficiency of about 50 %. Detergents with glucoside, maltoside and glucamide as their head group were most efficient. For example approximately 70 % refolding was observed for n-dodecyl  $\beta$ -D-maltoside (DDM, Figure 3.3a). In all screens where non-ionic detergents were used, it was observed that a detergent concentration of at least two times the cmc is necessary to start the refolding process.

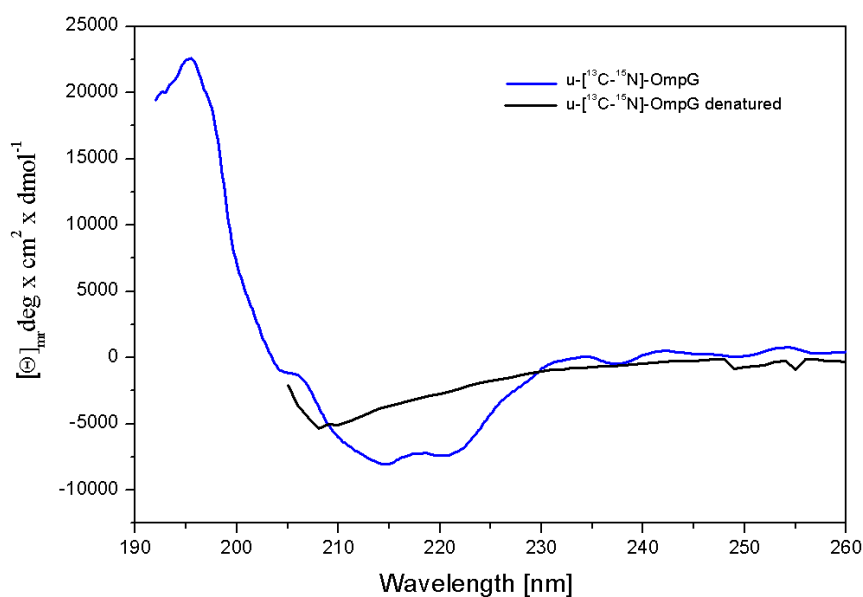
To enhance the refolding efficiency achieved by DDM, L-arginine was used as an additive in the refolding buffer. The addition of different arginine concentrations to the refolding buffer was tested. The addition of 0.6 M L-arginine improved the refolding efficiency up to more than 90 % (Figure 3.3b). For a long time it was suggested that arginine enhances the refolding reaction like a chaperone. Recently Arakawa *et al.* showed that arginine increases the yield of folded protein by suppressing protein aggregation during the refolding reaction [135].



**Figure 3.3 Increased OmpG refolding efficiency using L-arginine as an additive.** a) Refolding of U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -OmpG in DDM without the addition of L-arginine. Lane 1, refolded OmpG. Lane 2, same as lane 1 except that the sample was boiled. b) Refolding with the addition of 0.6 M L-arginine. Lane 3, refolded U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -OmpG. Lane 4, same as lane 3 except that the sample was boiled.

An excess of detergent is normally employed when a MP is solubilised from native membranes, to ensure complete dissolution of the membrane and to provide a large number of micelles, that only one protein molecule is present per micelle. However, for further physicochemical, biochemical or structural characterization, it is often necessary to remove the unbound detergent molecules. Therefore, a second anion exchange chromatography column was used to reduce the detergent concentration from 1 mM DDM to 0.4 mM DDM (see chapter 2.2.3.3).

The secondary structure of unfolded OmpG in urea and refolded OmpG in a detergent-containing buffer was analysed by *Circular-dichroism* (CD) spectroscopy (Figure 3.4). The refolded protein showed a curve expected for a  $\beta$ -sheet structure with a minimum mean residue ellipticity at 215nm and a maximum mean residue ellipticity at 195nm. In comparison to that, OmpG in urea-containing buffer shows a curve expected for a random-coil structure. The CD spectrum of the refolded OmpG confirms the formation of the correct secondary structure in agreement with the observed migration in the SDS gel, since the refolded OmpG resists the SDS denaturation and migrates faster than the completely unfolded product (Figure 3.3 and 3.4). This procedure yielded up to 16.8 mg of refolded U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -OmpG per litre culture. All labelled samples used in this study were prepared and analysed in the same way. In table 3.1 a summary of the yields for each labelled OmpG sample is presented (Table 3.1).



**Figure 3.4** *Circular dichroism* spectra of refolded and unfolded U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -OmpG. The protein concentration was 25  $\mu\text{M}$ . Samples were scanned in a 1 mm path-length quartz cuvette, as described in the experimental section.

### 3.1.3 Improvement of recombinant OmpG expression

To improve the yield of isotopically labelled protein, two different fermentation techniques were tested. The first technique is a modified expression protocol based on normal shaking flasks [127]. The second technique uses a fully automated bioreactor system that allows the growth of high cell densities, before the protein expression is induced [127,128].

#### 3.1.3.1 A modified expression protocol [127]

Marley *et al.* introduced a procedure for the production of isotopically labelled proteins that permits high yields of labelled protein while significantly reducing the cost in both isotopes and production time (Figure 3.5) [127]. First, the cell mass is predominantly produced in unlabelled rich media (LB-media) allowing fast growth to high cell densities. Following growth in LB-media, cells are exchanged into M9-minimal media with labelled carbon and nitrogen sources. During this step, cells can be concentrated by centrifugation and re-suspension in a smaller volume of isotopically labelled M9-minimal medium. After one-hour of incubation to recover the growth and clearance of unlabelled metabolites, the protein expression can be induced (Figure 3.5).

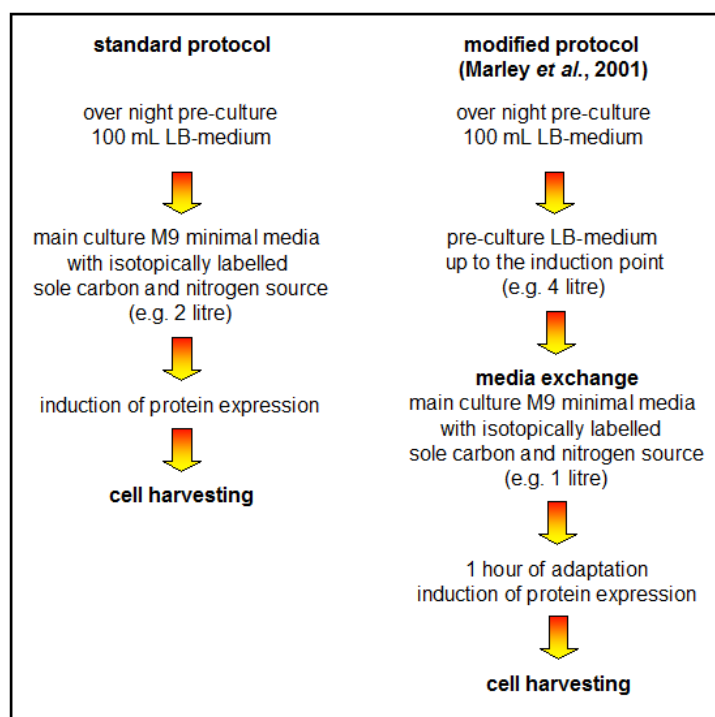


Figure 3.5 Schematic representations of the standard and the modified cultivation protocols.

This method was followed, so as to produce high amounts of recombinant isotopically labelled OmpG. In a first step, in small-scale experiments it was tested how much the cells can be concentrated during the medium exchange from LB medium to the labelled M9 minimal medium. Using glucose as sole carbon sources it was possible to concentrate the cells four-fold. Thus, the biomass obtained was threefold higher and the yield of purified uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled OmpG (U- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG) was increased seven-fold in comparison to the standard protocol (Table 3.1).

Using glycerol as the sole carbon source, however, a four-fold concentration led to a decreased expression level of OmpG. In contrast to that, cells concentrated two fold showed normal OmpG expression yields. Compared to the standard protocol, the achieved biomass was increased 1.7-fold and the yield of purified 1,3- $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled OmpG ([1,3- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG) was increased three-fold (Table 3.1).

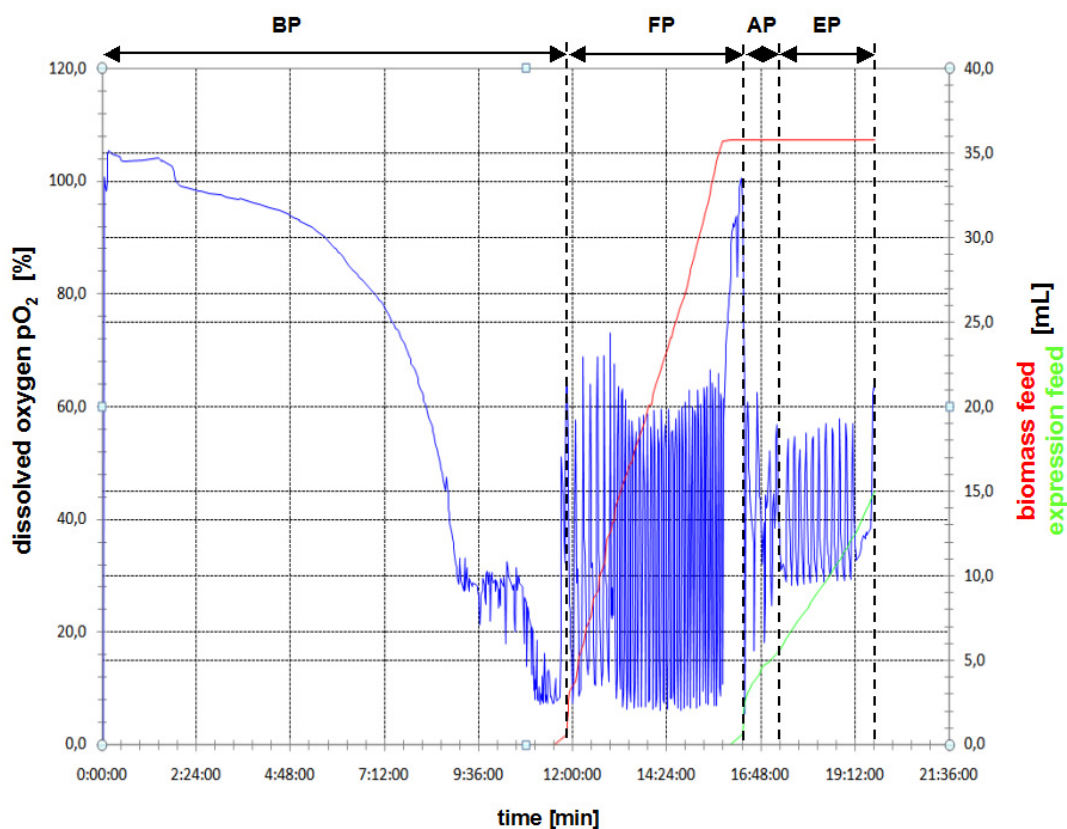
To express specifically labelled OmpG, the modified protocol was used without performing any preliminary small-scale tests. Cells from a 4 litre pre-culture were resuspended in 2 litres of M9-minimal medium containing 200 mg of each labelled and unlabelled amino acid. The results for a [GAFY]-OmpG sample in which only the amino acids alanine (A) and glycine (G) were uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled and the amino acids phenylalanine (F) and tyrosine (Y) were 2,3- $^{13}\text{C}$ - and uniformly  $^{15}\text{N}$ -labelled are shown in table 3.1.

In summary, a number of OmpG samples have been expressed with different sole carbon sources using the modified expression protocol. The improvement of protein yields reduces the costs for isotopically labelled sources. Also, higher cell-densities have no negative influence on the plasmid stability, as the protein yield is increased proportionally to the biomass. It has to be mentioned here that isotopic incorporation rate could not be determined by mass spectrometry for the labelled OmpG samples. This might be due to the strong hydrophobicity of the protein. Marley *et al.* described an isotopic incorporation rate of 95% for  $^{13}\text{C}$  and 90% for  $^{15}\text{N}$  nuclei. As will be shown in section 3.3, such reduced isotope incorporation has no significant influence on the sensitivity of the recorded NMR experiments. Thus, the method offers an alternative to fermenter cultivation that requires additional set-up time and specialized equipment. The time required to prepare isotopically labelled samples using this method is slightly greater than that using the standard expression protocol. However, the initial growth on unlabeled rich media is time saving in comparison to the standard protocol, which

requires significantly longer to reach cell densities suitable for the induction of protein expression.

### **3.1.3.2 High cell density fermentation (HCDF)**

In a second approach a bioreactor system that allows fully automated high cell density fermentation (HCDF) was tested [128]. The bioreactor system is comprised of a temperature controlled 1 litre incubator placed on a magnetic stirrer platform, equipped with pH and pO<sub>2</sub> electrodes, and feed and air supply/removal connections. The fermentation procedure can be divided into four steps: **1.** Batch phase (BP) growth of cells. **2.** Fed phase (FP) in which the culture is grown to high cell densities. **3.** Adaptation phase (AP) in which the medium is switched to a labelled feed for 1 h before induction of protein expression. **4.** Protein expression phase (EP) induced by the addition of IPTG. During each step the consumption of the sole carbon sources can be monitored by the amount of dissolved oxygen (pO<sub>2</sub>) in the medium. The pO<sub>2</sub> behaviour during the expression of a specifically labelled [GAVLS,<sup>15</sup>N]-OmpG sample is shown in figure 3.6 as an example. The batch medium was inoculated with an overnight culture of *E. coli* B121(DE3) transformed with a pET-OmpG plasmid. When the cells are growing the consumption of the sole carbon sources leads to a decrease of the pO<sub>2</sub>-value in the medium. When the pO<sub>2</sub> reached a value >30%, the agitation-speed was increased up to 1000 rpm to ensure that all nutrients in the batch were fully consumed. As a consequence the pO<sub>2</sub>-value increases immediately as no oxygen is used for the metabolism. When the pO<sub>2</sub>-value reaches 50% the biomass feed is turned on and the metabolism of the cells starts again. Due to the metabolism of the cells the pO<sub>2</sub>-value decreases again to less than 30%, at which point the biomass feed is turned off (Figure 3.6). Thus, the pO<sub>2</sub>-value can be used as a monitor of the availability of sole carbon and/or nitrogen sources. This dosing procedure was used for 4 hours to achieve a culture with high cell densities (OD<sub>600</sub>= 12-14). After 4 hours, 35 ml of the expression feed containing labelled sole carbon and nitrogen sources was added by hand. Afterwards, the expression-feed was dosed and added to the medium in the same manner as the biomass feed (Figure 3.6). The protein expression was induced after one hour of dosing to ensure that all unlabelled metabolites had been consumed. After 4 hours of expression the cells were harvested and stored at -80°C.



**Figure 3.6** Graphical representation of the  $pO_2$ -value and feeds during the fermentation process for the [GAVLS,  $^{15}N$ ]-OmpG expression. BP = Batch phase, FP = Fed phase, AP = Adaptation phase, EP = Expression phase.

This fermentation protocol was used to express four specifically labelled OmpG samples (Table 3.1). An expression feed containing a mixture of labelled and unlabelled amino acids and  $^{15}N$ - $NH_4Cl$  were used for the expression of the [GAVLS,  $^{15}N$ ]-, [RIGA,  $^{15}N$ ]- and [GANDSH,  $^{15}N$ ]-OmpG samples. For the fourth sample the reverse labelling strategy was used, where an expression feed containing uniformly  $^{13}C$  labelled glucose and  $^{15}N$ - $NH_4Cl$  was supplemented with unlabelled amino acids (Table 3.1).

In comparison to the standard and the modified flask expression protocols a drastically increased biomass was achieved using the HCDF procedure (Table 3.1). Additionally, the amount of labelled sole carbon sources (amino acids) used for the HCDF could be decreased by a factor of four compared to the modified flask expression protocol. However, the yield of purified OmpG per gram biomass is nearly the same as obtained for the standard flask expression protocol. This means, that many of the cells are losing the ability to express OmpG during the HCDF. Therefore, the presence of the expression plasmid (pET-OmpG) in the cells was tested, by plating them out on LB-agar plates with and without kanamycine, after each HCDF. The results indicate that nearly 40 to 50 % of the cells lose the pET-OmpG plasmid, encoding the kanamycine



resistance gene, during the HCDF. Although the amount of kanamycine used for the HCDF is five times higher than in a standard-flask expression protocol, the high density of cells might lead to kanamycine degradation. Therefore, it should be tested how the plasmid stability can be secured in the future. Nevertheless, the yields of specifically labelled OmpG samples achieved by the HCDF are seven to eight times higher compared to a standard flask expression protocol. Additionally, the costs for isotopically labelled material could be reduced by a factor of four (Table 3.1).

#### **3.1.4 Modified expression protocols or HCDF can save costs for isotopically labelled material**

In this section it was shown that costs for isotopically labelled material could be reduced by a factor of up to four using a modified flask expression protocols or a high cell density fermentation procedure. The protocols were applied to the  $\beta$ -barrel MP OmpG, which can be expressed as a non-functional inclusion body precipitate. An in vitro refolding step yields high amounts of native OmpG per litre culture. In the literature it has been described that refolding  $\beta$ -barrel MPs from inclusion bodies is much easier than for  $\alpha$ -helical MPs. Therefore,  $\alpha$ -helical MPs will be mostly expressed in the cytoplasmic membrane of gram-negative bacteria or in the membrane of yeasts or higher eukaryotic expression systems (see section 1.1.1). The major disadvantage of such expression systems is the low yields (< 1mg/l culture) of the recombinant MPs. Thus, for  $\alpha$ -helical MPs the expression protocols presented here may be of major interest, since HCDF might be adapted to eukaryotic expression systems, like the yeast *Pichia pastoris* and the baculovirus/insect cell expression system. Additionally, different isotopically labelled growth media for these expression systems have been developed in recent years [23,24].

Table 3.1 Final yields the isotopically labelled OmpG samples.

carbon source [g/L]	U-[13C, 15N]- OmpG		[1,3- <sup>13</sup> C, <sup>15</sup> N]- OmpG		[2- <sup>13</sup> C, <sup>15</sup> N]- OmpG		[GAFY]- OmpG		[GAVLS, <sup>15</sup> N]- OmpG		[RIGA, <sup>15</sup> N]- OmpG		[GANDSH, <sup>15</sup> N]- OmpG		[TEMPQANDSG, <sup>15</sup> N]- OmpG	
	S	M	S	M	S	M	M	HCDF	HCDF	HCDF	HCDF	HCDF	HCDF	HCDF	HCDF	HCDF
bm [g/L]	2.9	8.7	1.8	3.2	1.7	3.3	3.9	26	27	30	29					
denat. OmpG [mg/L]	25	143	19,8	62	18	54	118	180	197	212	185					
ref. OmpG [mg/L]	16.8	86	9.9	31	9	27	59	122	124	134	126					

S = standard-flask expression protocol

M = modified-flask expression protocol

HCDF = high cell-density fermentation

bm = biomass

AA = labelled or unlabeled amino acids

Glc = [13C]-labelled or unlabelled glucose

Gly = [1,3-13C]- or [2-13C]-glycerol

denat. OmpG = OmpG in a buffer containing 8M Urea after the first anion-exchange chromatography step

ref. OmpG = refolded OmpG after the second anion-exchange chromatography step

## 3.2 Preparation of OmpG samples suitable for solid-state MAS NMR investigations

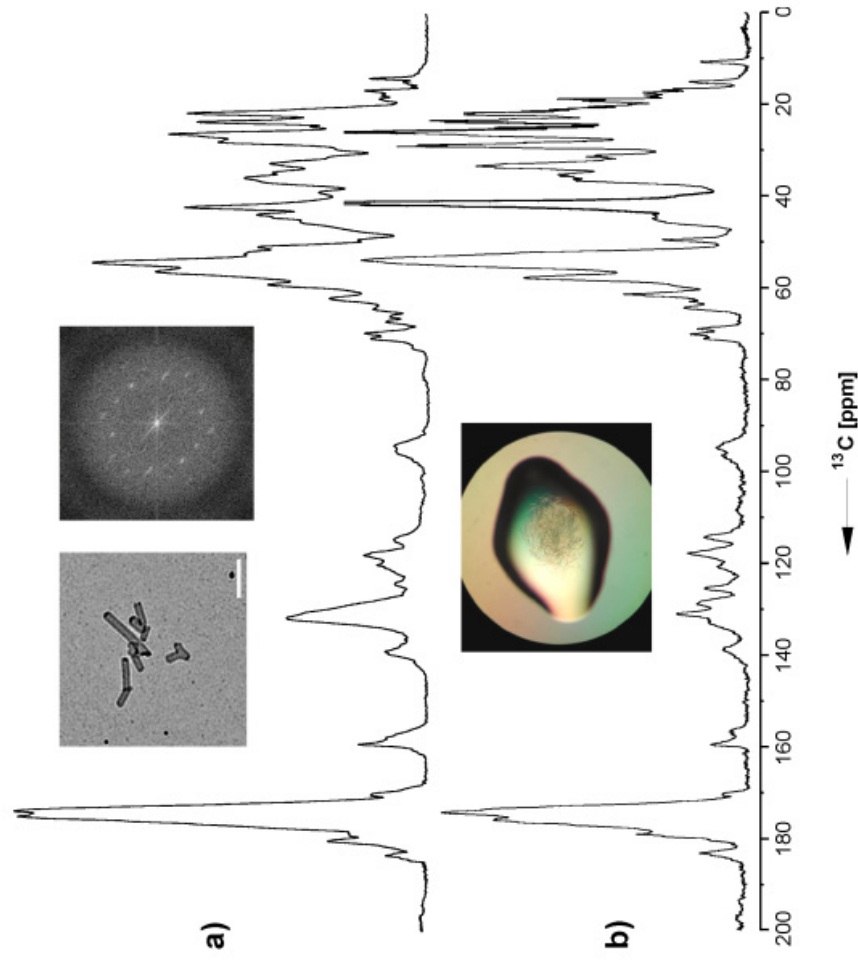
As has been described in chapter 1.1.5, solid-state MAS NMR does not require highly ordered 3D crystalline samples like X-ray crystallography. Experiments can be performed with numerous different solid sample preparations, for example lyophilised powders, or nano- and micro-crystalline precipitates. In the case of MPs, different preparation types are possible. To test the activity of a MP, for example, it is often reconstituted into liposomes (proteoliposomes). Structural investigations of MPs by EM or X-ray crystallography use 2D and 3D crystals, respectively (see sections 1.1.3 and 1.14). In solution NMR, MPs are studied in detergent micelles, which mimic their natural environment (see chapter 1.1.5). To prepare an OmpG sample suitable for solid-state MAS NMR spectroscopy, different preparation types including 2D crystals, proteoliposomes and a nano-crystalline precipitate were analysed by EM and solid-state MAS NMR. Initial solid-state MAS NMR experiments were performed using an uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labelled OmpG sample (U-[ $^{13}\text{C}$ , $^{15}\text{N}$ ]-OmpG). Further spectra were recorded using a specifically labelled [GAFY]-OmpG sample. Such a sample allows a better comparison of the resulting spectra, since only the amino acids alanine, glycine, phenylalanine and tyrosine are  $^{13}\text{C}$  labelled, which reduces spectral overlapping in the NMR data.

### 3.2.1 2D crystals of OmpG

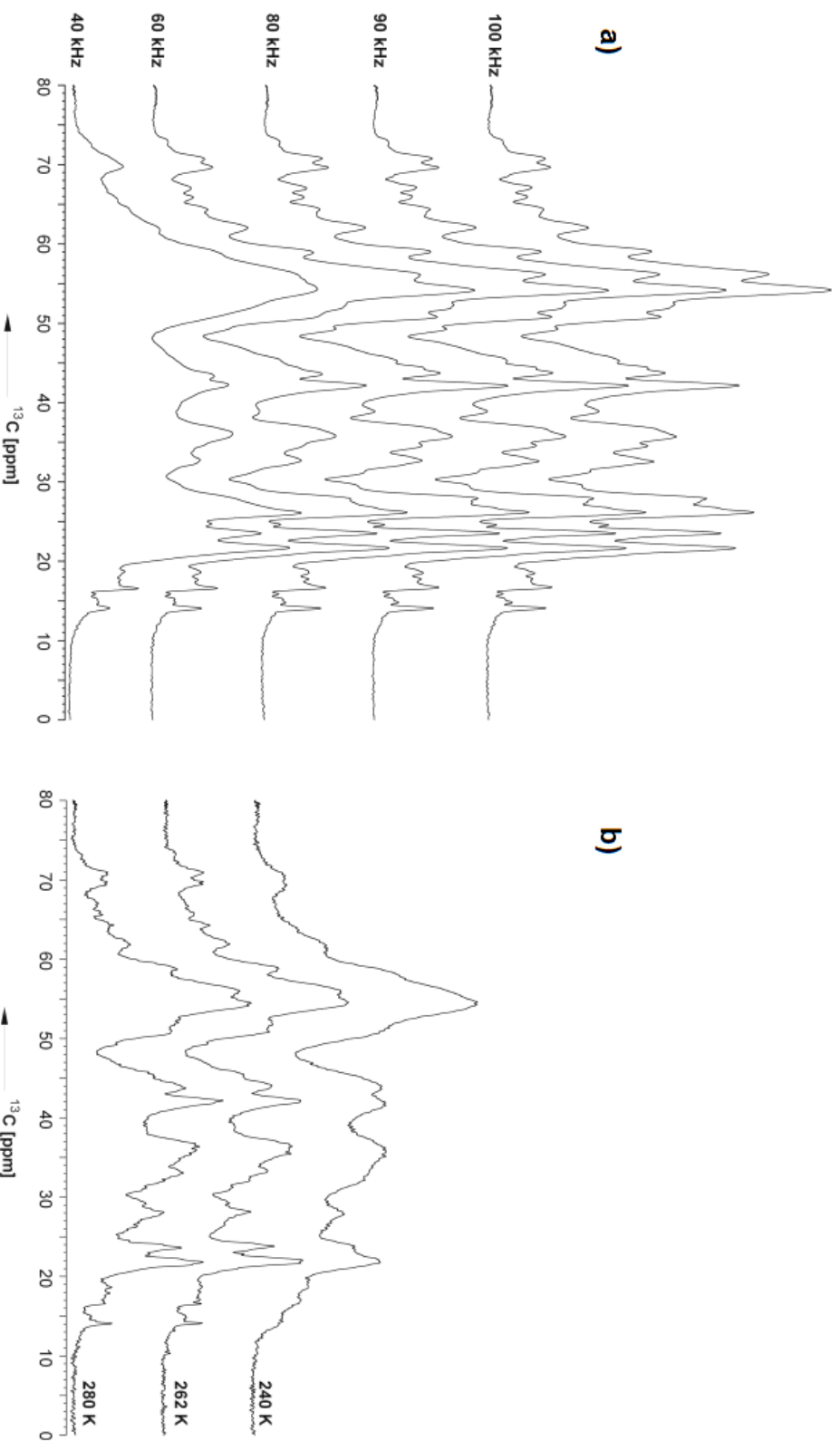
Initial crystallisation screens of OmpG were done in co-operation with K.R. Vinothkumar, from the group of Prof. W. Kühlbrandt (MPI of Biophysics, Frankfurt am Main). OmpG was reconstituted into lipid bilayers and 2D crystals were subsequently formed by dialysis. The highest quality 2D crystals were obtained at 20°C using a total lipid extract from *E.coli* (LPR= 0.5 (w/w)) in the presence of both monovalent and divalent cations with sodium and magnesium preferred (150 mM NaCl and 25 mM  $\text{MgCl}_2$ ). During the first dialysing period lipid vesicles were formed, which mostly clumped together. After the fourth buffer exchange 2D crystals with a tubular shape (1-1.5  $\mu\text{m}$  in length and 130-200 nm in width) were formed (Figure 3.7b insets). However, they were mostly stacked on top of each other, which hampered the effort to collect tilted images for a 3D map construction. The diffraction patterns of these crystals had a

resolution of up to 8 Å. In contrast to that, 2D crystallisation of OmpG solubilised and purified from native membranes led to crystalline sheets (up to about 300 nm x 300 nm) diffracting to 6 Å resolution [117]. However, the unit-cells and the projection-maps were comparable for both types of crystals. The projection maps showed a circular molecule indicative of a  $\beta$ -barrel (diameter of 25 Å x 27 Å) and some strong central density, possibly reflecting the presence of loops covering the barrel opening (Figure 3.7c). These observations additionally confirmed that OmpG refolded from inclusion bodies had attained its native structure.

The 1D  $^{13}\text{C}$  CP-MAS spectrum of U-[ $^{13}\text{C}$ , $^{15}\text{N}$ ]-OmpG exhibits carbonyl signals at around 175 ppm, signals from aromatic carbons between 120 and 140 ppm and signals from aliphatic carbons in the region between 10 and 80 ppm (Figure 3.7a). The overall spectral quality is comparable to that of the  $\alpha$ -spectrin SH3 domain (Figure 3.7b) [60]. This indicates that the local order of low diffracting 2D crystals is sufficient to record high-resolution solid-state MAS NMR spectra. However, it has to be kept in mind that the high molecular weight of OmpG (34 kDa, 281 aa) causes a high degree of spectral overlap and makes it difficult to measure the line width of a single NMR signal in a one dimensional spectrum. Under MAS conditions,  $^1\text{H}$ - $^{13}\text{C}$  dipolar couplings and the degree to which these can be decoupled during signal detection influence the sensitivity and resolution of the spectra. Therefore, the influence of the TPPM and the SPINAL64 proton-decoupling schemes was analysed [83,84]. 1D  $^{13}\text{C}$  CP-MAS spectra recorded with TPPM or SPINAL64 proton decoupling showed no significant differences. The SPINAL64 proton decoupling scheme was used in all other experiments, since it is a broadband decoupling scheme for liquid crystals and solids [84]. Figure 3.8a shows the effect of increased proton-decoupling power. The spectrum with low decoupling power (20 kHz) showed both low sensitivity and low resolution (not shown). When increasing the proton decoupling power level from 40 to 90 kHz, the sensitivity and resolution of the spectrum improved significantly. Higher proton decoupling powers (100 kHz) had no further effect and thus all other spectra were recorded with 90 kHz proton decoupling power (Figure 3.8a). In a third series of 1D  $^{13}\text{C}$  CP-MAS NMR experiments, the influence of temperature on spectral resolution was investigated (Figure 3.8b). Measurements were performed at temperatures ranging from 240 K to 280 K. Similarly narrow lines were observed between 262 and 280 K. Between 265 and 260 K, a change in probe circuit tuning was observed.



**Figure 3.7 Comparison of  $^{13}\text{C}$  CP-MAS NMR spectra of U-[ $^{13}\text{C}$ , $^{15}\text{N}$ ]-OmpG (20 mg) reconstituted 2D crystals (a) and a microcrystalline U-[ $^{13}\text{C}$ , $^{15}\text{N}$ ]-SH3 sample (10 mg) (b).** The spectra were recorded with 1024 scans, with a proton decoupling power of 90 kHz and a spinning frequency of 8 kHz at 280 K on a 400 MHz wide-bore spectrometer. Insets: (top left) Electron micrographs of 2D crystals of u-[ $^{13}\text{C}$ , $^{15}\text{N}$ ]-OmpG, scale bar = 1  $\mu\text{m}$ ; (top right) electron diffraction pattern. (bottom) microscopic picture of a typical microcrystalline  $\alpha$ -spectrin SH3 domain preparation. (c) **Comparison of projection density-maps of 2D crystals from recombinant refolded OmpG (left) and native OmpG (right) at 8  $\text{\AA}$  resolution.**

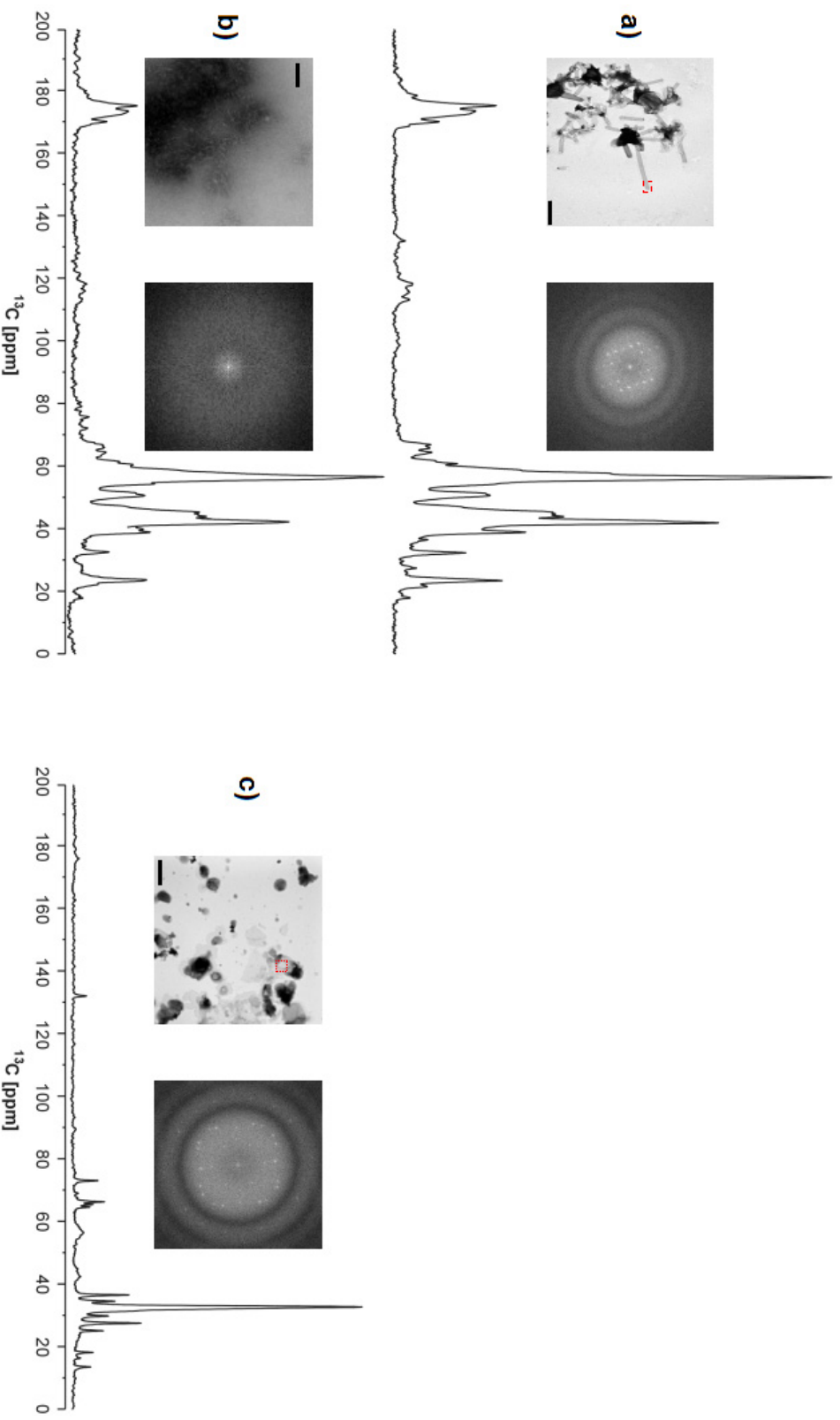


**Figure 3.8 1D  $^{13}\text{C}$  CP-MAS NMR spectra of U-[ $^{13}\text{C}$ , $^{15}\text{N}$ ]-OmpG reconstituted into 2D crystals recorded at 280 K with different SPINAL64 proton decoupling power levels (a) and at different temperatures (b). The spectra were recorded with 1024 scans at a spinning frequency of 8 kHz on a 400 MHz wide-bore spectrometer (a) at different proton decoupling levels (40-100 kHz) at 280 K and (b) at a temperature range from 280-240 K with a proton decoupling power of 90 kHz.**

This can be explained by a phase transition of the lipids from liquid crystalline to gel phase lipids as the sample freezes, which changes the sample conductivity and, therefore, the tuning of the probe circuit. When the sample was frozen the signal intensity of the overall spectrum was increased by a factor of 1.5-2. However, the resolution is lower in comparison to the spectra at 262 and 280 K (Figure 3.8b). These changes were reversible and the higher resolution was fully recovered when the temperature was raised to 280 K again after repeated freeze-thaw cycles. The observed line broadening might be due to the reduced mobility of the side chains in the frozen state, resulting in a reduced self-decoupling. However, increasing the proton decoupling power from 90 kHz to 110 kHz had no effect on the 1D spectrum of a frozen sample. This increase would normally be sufficient to induce heteronuclear decoupling in a rigid body. Therefore, the most likely explanation for the reduced resolution at 240 K is the structural inhomogeneity of side chains immobilised by frozen water in the pore and in the more rigid lipid phase. This is in line with other reports, which show that sample freezing causes incomplete motional averaging and results in inhomogeneous line broadening [73]. Therefore, all further solid-state MAS NMR experiments were recorded at 280 K with a proton decoupling power of 90 kHz.

### **3.2.2 Proteoliposomes of OmpG**

Proteoliposomes were produced in a similar manner to the 2D crystals. The difference between the two methods lies in the chosen LPR. In contrast to 2D crystals, proteoliposomes were prepared with an LPR of 20 (w/w), so that a lipid vesicle surrounds each OmpG molecule. Electron micrographs of the resulting proteoliposome preparation showed that it contained not only vesicles, but also crystalline sheets (Figure 3.9c insets). The 1D  $^{13}\text{C}$  CP-MAS spectrum of the proteoliposome preparation using [GAFY,  $^{15}\text{N}$ ]-OmpG shows very low intensity protein signals and very high intensity signals, which originate from the natural abundance  $^{13}\text{C}$  nuclei in the lipids (Figure 3.9c). One reason for the low protein signal intensity could be the small amount of protein (5mg) in the sample. However, the number of scans recorded for this spectrum was eight times higher than that for the 2D crystalline sample (20 mg protein), which should be enough to obtain the same signal intensities.



**Figure 3.9 1D  $^{13}\text{C}$  CP-MAS NMR spectra of a) [GAFY]-OmpG 2D crystals b) [GAFY]-OmpG PEG precipitate and c) [GAFY]-OmpG proteoliposomes.** The spectra were recorded on a 900 MHz narrow-bore spectrometer with a proton decoupling power of 90 kHz and spinning frequency of 13 kHz at 280 K with 1024 scans. The spectrum of the proteoliposome sample was recorded with 8192. Insets: (left) Electron micrographs of each preparation type, scale bar = 1  $\mu\text{m}$ ; (right) electron diffraction pattern.

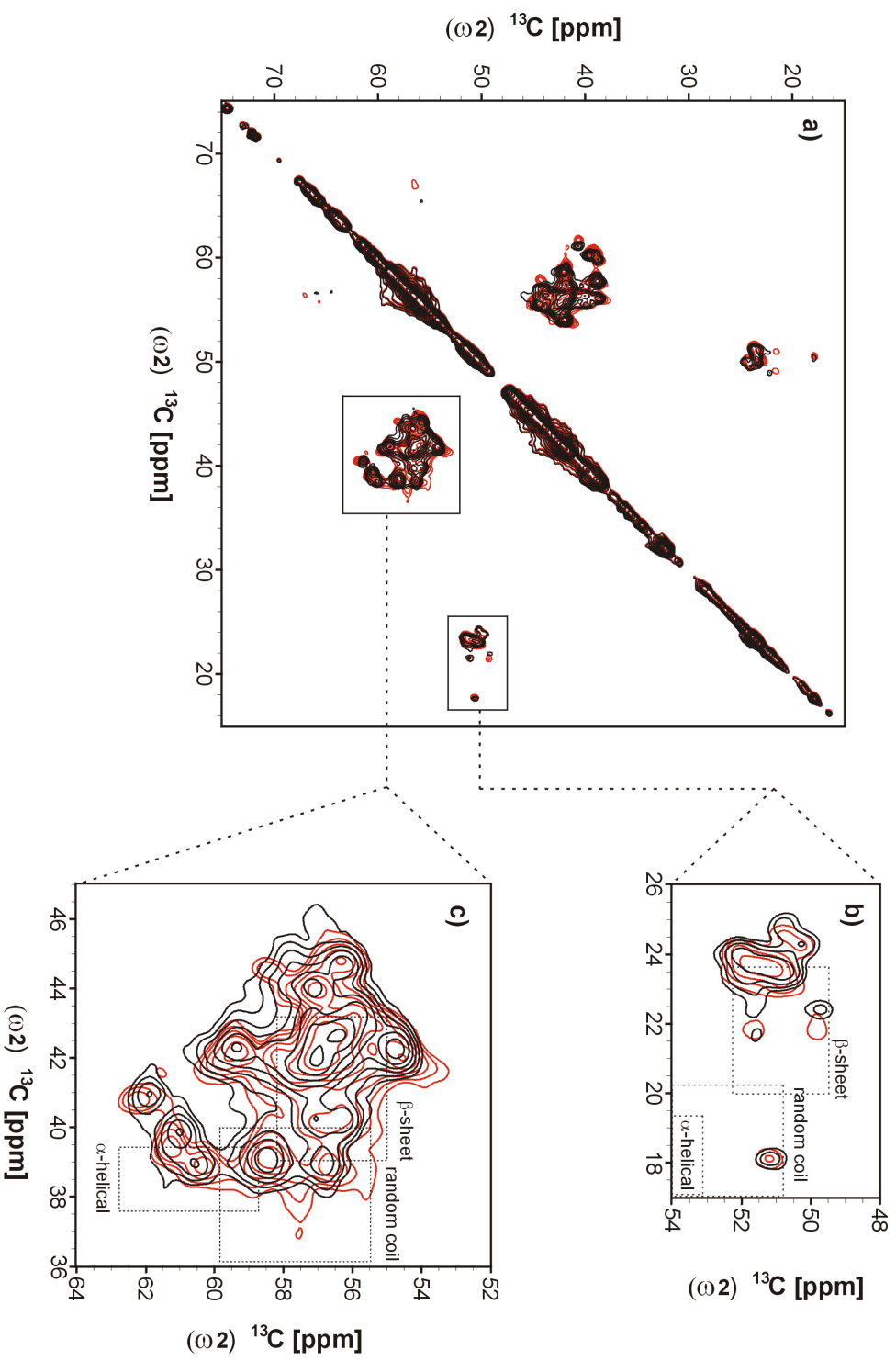


The most likely explanation is that the protein contained in the lipid vesicles is structurally inhomogeneous, which results in a drastically decreased signal intensity due to line broadening. To get a clear view of the spectral resolution of a pure proteoliposome sample, the vesicular part would have to be enriched and isolated from the 2D crystalline part. A further difficulty with proteoliposome samples is already apparent, which is the problem of fitting a sufficient amount of protein into the NMR rotor. In the case of OmpG, it was only possible to fit 5 mg of protein into the rotor due to the large amount of lipids in the sample.

### 3.2.3 PEG precipitate

Martin *et al.* (2003) demonstrated that crystalline material far too small to be useful in X-ray diffraction studies is suitable for structural studies by solid-state MAS NMR using different soluble proteins as examples [73]. Additionally, the structure of OmpG was solved from highly diffracting 3D crystals by Yildiz *et al.* (2006) [120]. Thus, the idea above was followed to produce 10 mg of a crystalline OmpG PEG-precipitate using the known crystallisation conditions (see chapter 3.2.5). The EM analysis showed an amorphous precipitate, which did not diffract (Figure 3.9b insets). The resolution of the 1D  $^{13}\text{C}$  CP-MAS spectrum of [GAFY]-OmpG was similar to that of the 2D crystalline sample (Figure 3.9a). This indicated that although the PEG precipitate did not show any long-range order, which could be detected by X-ray diffraction, the protein was nonetheless highly homogeneous in its structure across the sample.

To compare the spectral quality of the 2D crystalline and the PEG precipitated sample precisely, 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectra of [GAFY]-OmpG were recorded with 20 ms PDS mixing at 280 K (Figure 3.10a). The strong diagonal signals arised from  $^{13}\text{C}$ -labelled sites in OmpG and in the case of 2D crystals from the  $^{13}\text{C}$  natural abundance background of the lipids. The unlabelled lipids did not give rise to detectable cross peaks due to the low abundance of adjacent  $^{13}\text{C}$  nuclei. All observed cross peaks reflect intra-residual correlations in OmpG. In the region of 50-55 ppm ( $\omega_1$ ) and 15-25 ppm ( $\omega_2$ ) cross-peaks involving the  $\text{C}\alpha$  and  $\text{C}\beta$  atoms of alanine residues are visible (Figure 3.10b). The correlations between the  $\text{C}\alpha$  and  $\text{C}\beta$  carbon atoms of the tyrosine and phenylalanine residues are present in the region of 55-60 ppm ( $\omega_1$ ) and 35-45 ppm ( $\omega_2$ ) (Figure 3.10c).



**Figure 3.10** Overlay of 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDSD NMR spectra of [GAFY]-OmpG reconstituted into 2D crystals (black) and of the PEG precipitate (red). (a) alpha-helical region of the spectra. (b and c) alanine and phenylalanine/tyrosine  $\text{C}\alpha$ - $\text{C}\beta$  correlations enlarged. The spectra were recorded at 700 MHz and 280 K with a PDSM mixing time of 20 ms and a MAS frequency of 13 kHz.

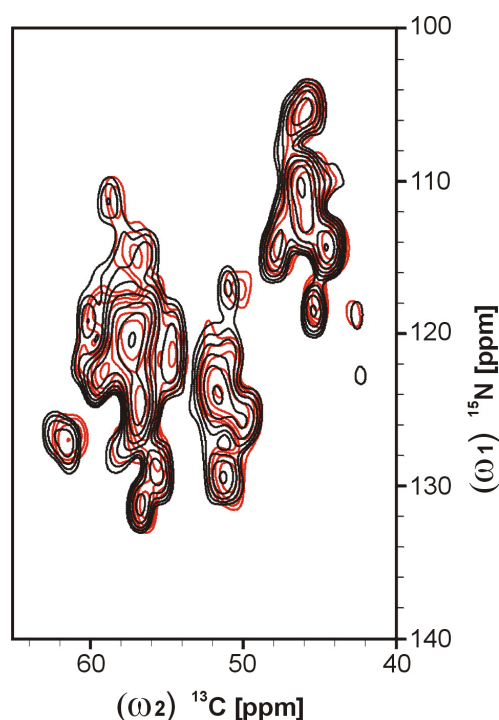
In a similar way, the correlations between the C' and C $\alpha$  or C $\beta$  carbon atoms of the alanine and glycine residues are detectable in the carbonyl region (not shown). Additionally, both spectra show no significant chemical shift differences and the apparent line width of the isolated outlying signals was around  $82 \pm 30$  Hz (including scalar couplings), which was determined by extraction of the relevant spectrum rows. Analysis of the chemical shift ranges of the C $\alpha$  and C $\beta$  signals from alanine and aromatic residues are indicative of their secondary structure (Table 4.2) [136]. These secondary structure trends were compared to those given by the X-ray crystal structure of OmpG [120]. As can be seen in Figure 3.10b, most of the alanine C $\alpha$ -C $\beta$  cross peaks are in the range of a  $\beta$ -sheet conformation and only one signal had a chemical shift indicative for a random coil conformation. Indeed, the distribution of alanine residues in the crystal structure shows that thirteen alanines are located in  $\beta$ -strands and only two alanines are located in loops [120]. In the case of the aromatic amino acids, most of the cross peaks have a chemical shift indicative of  $\beta$ -sheet conformations and only about five to seven cross peaks are indicative of an  $\alpha$ -helical or random coil conformation (Figure 3.10c). The OmpG sequence contains 18 phenylalanines and 22 tyrosines of which 38 residues are located in a  $\beta$ -strand and 7 residues are located in loops [120]. These results indicated that OmpG in 2D crystals and in the PEG precipitate appears to be in the same conformation as the X-ray crystal structure.

*Table 3.2 Averaged chemical shift (in ppm) and standard deviation values (in parentheses) categorized according to secondary structure type (data were taken from [136])*

AS	C $\alpha$ [ppm]			C $\beta$ [ppm]		
	$\beta$ -strand	$\alpha$ -helix	Random coil	$\beta$ -strand	$\alpha$ -helix	Random coil
Ala	$50.86 \pm 1.28$	$54.86 \pm 0.94$	$52.67 \pm 1.76$	$21.72 \pm 1.77$	$18.27 \pm 1.08$	$19.03 \pm 1.27$
Phe	$56.33 \pm 1.31$	$60.74 \pm 1.63$	$56.94 \pm 1.98$	$41.64 \pm 1.65$	$38.91 \pm 1.49$	$39.43 \pm 1.93$
Tyr	$56.56 \pm 1.59$	$61.07 \pm 1.72$	$57.72 \pm 2.14$	$40.79 \pm 1.77$	$38.38 \pm 0.89$	$38.71 \pm 2.00$

The homonuclear  $^{13}\text{C}$ - $^{13}\text{C}$  correlation spectra primarily yielded intra-residue correlations, allowing the identification of amino-acid type specific signals. For the later sequential assignment, however, heteronuclear  $^{15}\text{N}$ - $^{13}\text{C}$  correlation experiments are necessary. Therefore, NCA experiments of both samples were recorded and the spectral resolution was compared (Figure 3.11). The spectra showed resolved cross peaks

involving the backbone nitrogen atoms and the C $\alpha$  carbon atoms. The signals from the glycine residues are at around 45 ppm ( $\omega_2$ ), those from the alanine residues are at around 50-55 ppm ( $\omega_2$ ) and those from the aromatic residues lie between 55 and 65 ppm ( $\omega_2$ ). The line widths of outlying signals were  $110 \pm 20$  Hz for  $^{15}\text{N}$  ( $\omega_1$ ) and  $82 \pm 30$  Hz for  $^{13}\text{C}$  ( $\omega_2$ ), determined by extracting spectrum columns and rows. To summarise, no significant chemical shift differences were observed in either homo- or heteronuclear 2D spectra (Figure 3.10 and 3.11). This indicates that the conformational state of both samples (2D crystals and PEG precipitate) was more or less the same.



**Figure 3.11** Overlay of 2D NCA-spectra of [GAFY]-OmpG reconstituted into 2D crystals (black) and of the PEG precipitate (red). The spectra were recorded at 400 MHz, at 280 K and with a MAS frequency of 8 kHz.

### 3.2.4 Is there a perfect sample-preparation technique?

For the structural analysis of MPs by solid-state MAS NMR, several conditions must be fulfilled: i) milligram amounts of recombinant protein must be available, ii) the incorporation of NMR-active isotopes must be possible and iii) the protein must be structurally homogeneous across the sample, i.e. the local order in the sample is very important. The first two conditions can be addressed with molecular biology techniques, while the third depends on the sample preparation method. The sample preparations can be divided into three groups: 2D crystalline preparations, proteoliposome-embedded

proteins and detergent-embedded proteins. An increasing number of solid-state MAS NMR studies of MPs has been reported in the literature in recent years. By comparing these studies with the data for OmpG it should be possible to draw conclusions about the advantages and disadvantages of the different sample preparation methods and what might be a preferred way of preparing a MP sample when faced with a new protein.

OmpG was reconstituted into lipid bilayers to form tubular 2D crystals. The resulting solid-state MAS NMR spectra of poorly diffracting 2D crystals showed high resolution in both  $^{13}\text{C}$ - and  $^{15}\text{N}$ -detected spectra. In particular, 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS and 2D NCA spectra showed cross peaks with an apparent line width of  $110 \pm 20$  Hz for  $^{15}\text{N}$  and  $82 \pm 30$  Hz for  $^{13}\text{C}$  (including scalar couplings), respectively (Figures 3.10 and 3.11). Additionally, the dispersion of the signals confirmed the structural integrity of OmpG in the 2D crystalline sample. However, the spectral resolution is temperature dependent. Line broadening effects were observed over a temperature range from 262 to 240 K (Figure 3.8b). This is due to the fact that mobile side chains as well as mobile protein segments were frozen in different conformations in different monomers.

The use of poorly diffracting 2D crystals for solid-state MAS NMR was also demonstrated for three other MPs (all  $\alpha$ -helical): the ABC-transporter ArtMP (53 kDa), proteorhodopsin (PR, 25 kDa) and the multidrug transporter EmrE (12 kDa) [137-139]. 2D crystals of PR and ArtMP were found by screening several hundred crystallisation conditions. Although these crystals only yielded low resolution EM data (diffraction to 37 Å for PR and unfortunately unknown for ArtMP) they gave rise to highly resolved solid-state MAS NMR spectra (**recorded at 280 K**), which indicates the structural integrity of both preparations [137,138]. Additionally, to verify the integrity of PR in these 2D crystals, the  $^{15}\text{N}$ -chemical shift of a lysine residue (Lys 231), which is covalently linked to the chromophore, was probed. The chemical shift was almost identical to that of PR in liposomes, which was fully functional [140].

In contrast to that, a 1D  $^{13}\text{C}$  CP-MAS spectrum **recorded at 273 K** of a 2D crystalline sample of EmrE showed low sensitivity [139]. The authors speculate that this might be due to the high mobility of EmrE, since lowering the temperature to 213 K yields a gain in sensitivity by a factor of 2-2.5. Interestingly, 2D crystals used for the structural investigation by EM had a tubular shape. However, this sample also contained triangular vesicles, which are similar to those used for solid-state MAS NMR [139,141]. Additionally, 2D crystals used for solid-state MAS NMR were crystallised without  $\text{MgCl}_2$ . Divalent cations, however, can have important effects on crystal formation [37].

Unfortunately, it was not shown whether the 2D crystals exhibited diffractions during EM recordings. Maybe, these “2D crystals” were triangular vesicles, which had a high mobility at 273 K. Nevertheless, the authors could show sample integrity by a secondary structural analysis of the C $\alpha$  chemical shifts extracted from a 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum [139]. The line widths of the signals were estimated to be around 80 to 140 Hz, which is comparable to those, determined for signals in the ArtMP and OmpG spectra.

Proteoliposome preparations of MPs form the second possible sample type for structural investigations by solid-state MAS NMR. Reconstitution of OmpG into proteoliposomes resulted in an inhomogeneous mixture of vesicles and crystalline sheets (Figure 3.9c insets). The 1D  $^{13}\text{C}$  CP-MAS spectrum showed very low sensitivity in comparison to the 2D crystalline sample (Figure 3.9a). This might be due to the sample heterogeneity. An additional disadvantage of proteoliposome samples was the high LPR (20:1), which made it difficult to fill the NMR rotor with a sufficient amount of protein. However, some examples of proteoliposome-based MP preparations suitable for solid-state MAS NMR experiments could be found in the literature, such as for PR, diacylglycerol kinase (DGK, 13.1 kDa) from *E. coli*, sensory rhodopsin II (NpSRII, 27 kDa) from *Natronomonas pharaonis*, a chimeric potassium ion channel (KcsA-Kv1.3, homotetramer 65 kDa) and for the monomeric form of human phospholamban (AFA-PLN, 6.2 kDa) [140,142-144]. NMR measurements of these proteoliposome samples were conducted in different aggregate states. The KcsA-Kv1.3 channel, for example, was measured at around 285 K (corresponding to the liquid crystalline state) and at 260 K (corresponding to the frozen gel-state) [145]. The spectra recorded at higher temperatures were used to detect mobile protein segments, whereas those recorded at lower temperatures were used to investigate the rigid parts of the protein. The resulting spectra showed an average line width of 100-140 Hz for  $^{13}\text{C}$ -nuclei, which is comparable to that for a 2D crystalline sample.

In the case of the uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labelled AFA-PLN peptide, the full assignment was obtained and a structural model could be calculated from the chemical shift data [144]. In the case of the KcsA-Kv1.3 channel 94 of 160 amino acids could be assigned from a uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$  labelled sample [145]. For NpSRII it was possible to assign several amino-acids of the transmembrane  $\alpha$ -helices and of the loop region, using a selectively labelled sample [143].

The third sample type possible for MAS NMR studies of MPs, consists of MPs surrounded by detergent micelles. Three different ways of preparing such samples have been described in the literature. One possibility is simply to concentrate the protein surrounded by detergent micelles. Solid-state MAS NMR experiments can then be performed using a frozen sample, as described for the light harvesting complex 2 (LH2, homo-nonameric, 261 kDa) from the photosynthetic bacterium *Rhodospseudomonas acidophila* [80]. Alternatively, the excess of detergent could be removed by dialysis. Afterwards, the sample was pelleted using an ultra-centrifugation step and placed in the rotor, as described for the cytochrome  $bo_3$  ubiquinol oxidase (cytochrome  $bo_3$ , 144 kDa) and for the disulfide-bond-forming enzyme (DsbB, 21.3 kDa) [146,147]. The third approach is to perform a PEG precipitation, which is often used in 3D crystallisation screens. This was done for the KcsA channel (homotetramer, 70.4 kDa) and for OmpG [148].

The spectral quality of the amorphous OmpG PEG precipitate was similar to that of the 2D crystalline sample, measured at the same temperature (Figure 3.9b). No significant chemical shift differences were found in 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDSO and 2D NCA spectra (Figure 3.10 and 3.11). However, similar to the 2D crystalline OmpG sample, the spectral resolution was temperature dependent. Comparable effects were also described for the cytochrome  $bo_3$  and the DsbB samples [146,147]. Additionally, in some cases the lines were so broad, that some signals had completely vanished [146]. The line widths of the signals measured in the frozen state were 90 Hz for  $^{13}\text{C}$  nuclei [146,147]. These examples illustrate that MPs in detergent micelles are also suitable for solid-state MAS NMR experiments.

In conclusion, all these examples illustrate that samples of MPs suitable for solid-state MAS NMR experiments can be prepared in two main ways, originating from a MP in detergent micelles (Figure 3.12). In the first method, the protein is reconstituted into liposomes to form proteoliposomes or 2D crystals. The great advantage of using a lipid environment is that the MP of interest can be studied in a quasi-native state. For this reason, this should probably be the method of choice when preparing MP samples for solid-state MAS NMR. While for some proteins proteoliposomes have been successful, the example of OmpG shows that this is not suitable for every protein. Considering, in addition, that maximising the amount of protein in the NMR rotor is of great importance for sensitivity reasons, the formation, if possible, of a 2D crystalline sample should be

the preferred sample preparation method. In this context, an interesting set-up to screen 96 different 2D crystallisation conditions was recently presented by Vink *et al.* [149]. Although this was developed for EM sample preparation, it should be equally useful to solid-state MAS NMR spectroscopists.

The second possibility is to leave the MP in detergent micelles, and to form an amorphous precipitate directly (Figure 3.12). This can be done by removal of excessive detergent or by looking for precipitation conditions using a commercially available crystallisation screen (e.g. MembFac from Hampton research). This is a very straightforward method and may be suitable if 2D crystals or proteoliposomes cannot be obtained. Additionally, similar to the use of 2D crystals, crystalline material (amorphous precipitates) far too small to be useful in X-ray diffraction studies is suitable for solid-state MAS NMR investigations.

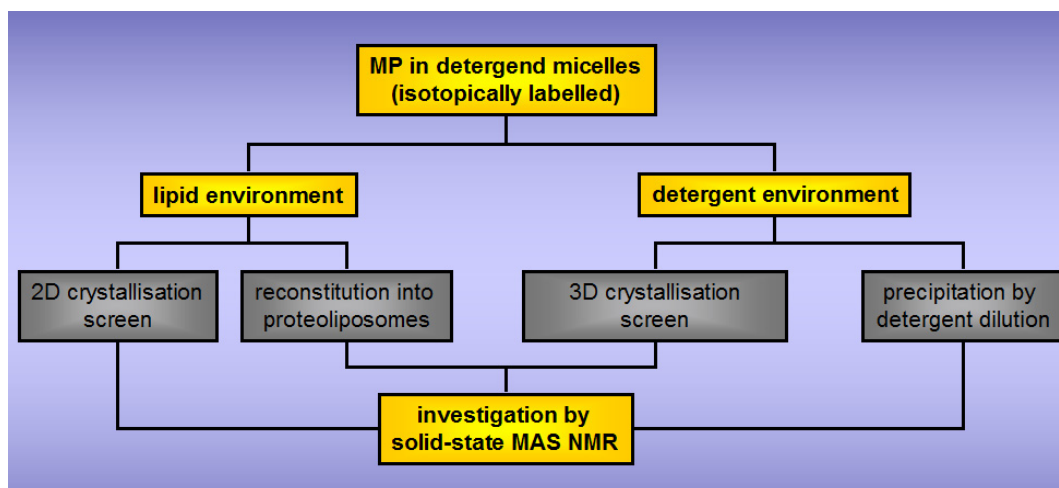


Figure 3.12 Different types of MP sample are suitable for structural investigations by solid-state MAS NMR.

However, the structural and functional integrity of the MP has to be tested at each step to ensure that the protein is correctly folded.

The solid-state MAS NMR studies of MPs conducted to date show that the experimental conditions including the temperature, the proton decoupling power and the spinning frequency have to be optimised individually for each protein and preparation type. In addition, it may be necessary to consider what aspect of the protein is being studied. For example, Andronesi *et al.* and Etzkorn *et al.* found that lower temperatures were useful for studying the rigid parts of the MP using CP transfers, whereas higher temperatures were beneficial for studying the mobile parts of the MP using INEPT transfers [143,144].



### 3.3 Assignment of solid-state MAS NMR spectra of OmpG

As was shown in section 3.2 different types of MP samples are suitable for structural investigation by solid-state MAS NMR. Since, we want to investigate the structure of OmpG in a near native environment, all further experiments were carried out using 2D crystalline samples of OmpG. The major prerequisite for structural studies, and the most time consuming step, is always the assignment of resonances to specific nuclei within the protein. In the beginning of the assignment work, 2D  $^{13}\text{C}$ - $^{13}\text{C}$ , 3D NCACX and NCOCX correlation spectra of a uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled OmpG sample (U-[ $^{13}\text{C}$ , $^{15}\text{N}$ ]-OmpG) were acquired. Using these spectra it was possible to identify well-separated resonances, like those of alanines, glycines, prolines, serines, threonines and valines. However, resonances of all other amino acid types are mostly overlapped and can therefore not be identified unambiguously. Additional loss of resolution results from line broadening due to scalar couplings between  $\text{C}'$ - $\text{C}\alpha$ ,  $\text{C}\alpha$ - $\text{C}\beta$ ,  $\text{C}\beta$ - $\text{C}\gamma$  and the other carbon atoms of the side chain. Additionally, OmpG consists of a large number of amino acids (281) and many of them have similar chemical shift values not only for the backbone nitrogen and carbon atoms but also for the side chain carbon atoms. Side chain signals, however can be very helpful to distinguish between different amino acid types.

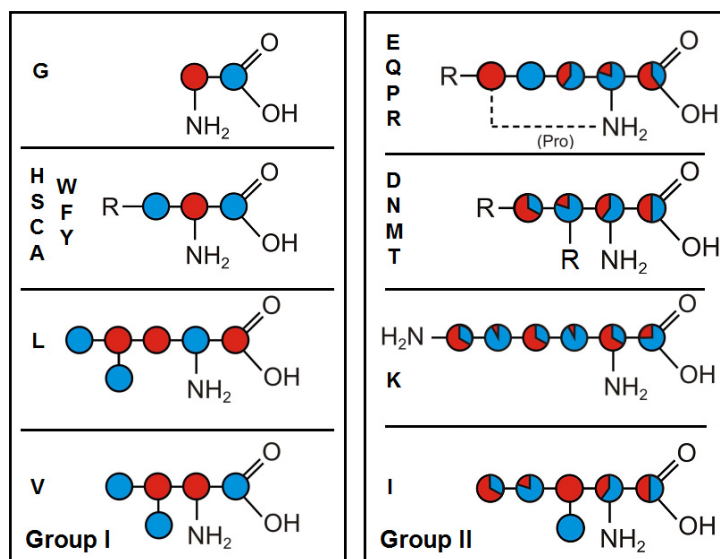
Therefore, a strategy, which allows the sequential assignment of such a large MP, has to be developed. In the first step, a selectively labelled OmpG sample was prepared that allows the ambiguous identification of assignment starting points. In a second step uniformly site-directed  $^{13}\text{C}$ -enriched OmpG samples were used to enhance and to connect these assignment starting points.

#### 3.3.1 [2,3- $^{13}\text{C}$ ]-labelling of aromatic residues – getting a head start in the assignment of OmpG

Castellani *et al.*, describe the use of two biosynthetically site-directed  $^{13}\text{C}$ -enriched samples that can be expressed by using [1,3- $^{13}\text{C}$ ] or [2- $^{13}\text{C}$ ]-labelled glycerol as the sole carbon source in the bacterial growth medium [150,151]. These samples have a defined labelling pattern, which simplifies the spectra (Figure 3.13) [150]. Amino acids in group I are synthesised from precursors made in the glycolytic and the pentose phosphate pathway. This labelling scheme yields either 100% or 0% labelled side chain carbon atoms. Group II contains amino acids produced in the tricarboxylic acid cycle (TCA-

cycle), which exhibit a more complex labelling scheme (Figure 3.13). This is due to the fact, that the TCA-cycle can be followed several times, which results in the production of a mixture of differently labelled amino acids (so called isotopomers).

The reduced labelling partly suppresses or completely removes line-broadening effects due one bond  $^{13}\text{C}$ - $^{13}\text{C}$  scalar couplings. Additionally, such a samples improve the identification of long-range distances [99].



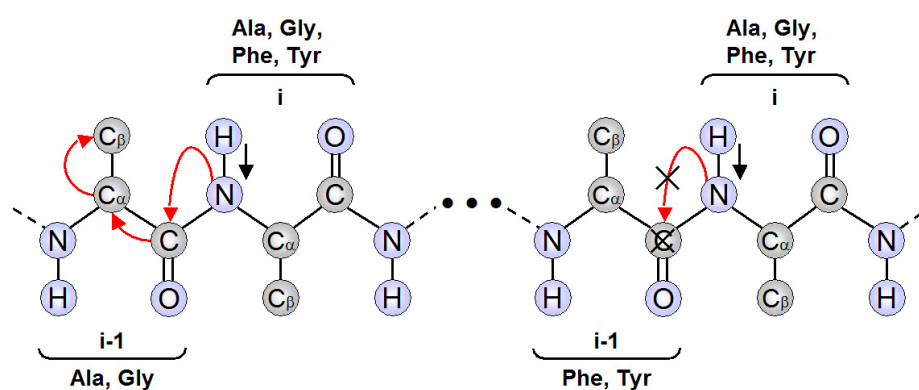
**Fig. 3.13 Schematic representation of the labelling pattern as obtained by protein expression in *E.coli* BL21(DE3).** The red colour corresponds to the degree of  $^{13}\text{C}$  incorporation by growth on [2- $^{13}\text{C}$ ]glycerol. The opposite labelling pattern, obtained by growth on [1,3- $^{13}\text{C}$ ]glycerol, is represented in blue.

For ambiguous sequential assignment of large proteins, however, those samples are unsuitable since the  $\text{C}'$ ,  $\text{C}\alpha$  and  $\text{C}\beta$  carbon atoms are never simultaneously labelled. Additionally, amino acids of the type Arg, Asn, Asp, Gln, Glu, Ile, Lys, Pro and Thr occur as isotopomers with different labelling schemes, which might hamper the identification of side chain signals.

An alternative is to use samples in which labelling is only applied to certain amino acid types. This can be done by growing *E. coli* cultures on M9 minimal media supplemented with labelled and unlabelled amino acids, also termed as forward labelling. However, it would be beneficial to restore the favourable cross relaxation properties of the glycerol samples in order to obtain inter-residue cross peaks which are important for successful sequential assignment. A suitable strategy for this might be to label several amino acids only at the N,  $\text{C}\alpha$  and  $\text{C}\beta$  positions (i.e., [ $^{15}\text{N}$ ]- and [2,3- $^{13}\text{C}$ ]-labelling) while selecting other short amino acid types for full labelling. This will

additionally significantly reduce overlap in NCO-type experiments and aid unambiguous sequential assignment (Figure 3.14). In this context, another important point of consideration is the number of amino acid doublets, which can be assigned from such a specifically labelled sample.

So as to address these issues a specifically labelled sample of OmpG, which contains [ $^{15}\text{N}$ , 2,3- $^{13}\text{C}$ ]-labelled Phe and Tyr and uniformly [ $^{15}\text{N}$ ,  $^{13}\text{C}$ ]-labelled glycine and alanine was prepared. This sample will be hereafter referred as to [GAFY]-OmpG. The OmpG sequence contains numerous sequential pairs of Gly or Ala with Phe and Tyr, which will give rise to fourteen unambiguous cross peaks in an NCO-type experiment (Figure 3.15). Gly and Ala also have the advantage that they can be easily identified by their chemical shifts. The use of the amino acids with short side chains will, additionally, maintain some of the favourable cross relaxation properties of the glycerol samples.



**Figure 3.14** Illustration of the magnetisation transfer during an NCO type experiment using the [GAFY]-OmpG sample. In a first step magnetisation is transferred from protons to the backbone nitrogen atoms (indicated by black arrow). During the specific NCO CP-step, magnetisation is transferred specifically to the carbonyl carbon atom of the previous amino acid. Hence, the magnetisation can only be transferred to alanine or glycine residues since their carbonyl carbon atom is  $^{13}\text{C}$ -labelled (left, indicated by red arrows). In contrast to that, Phe or Tyr are not labelled at this position. Thus, the magnetisation can be transferred (right).

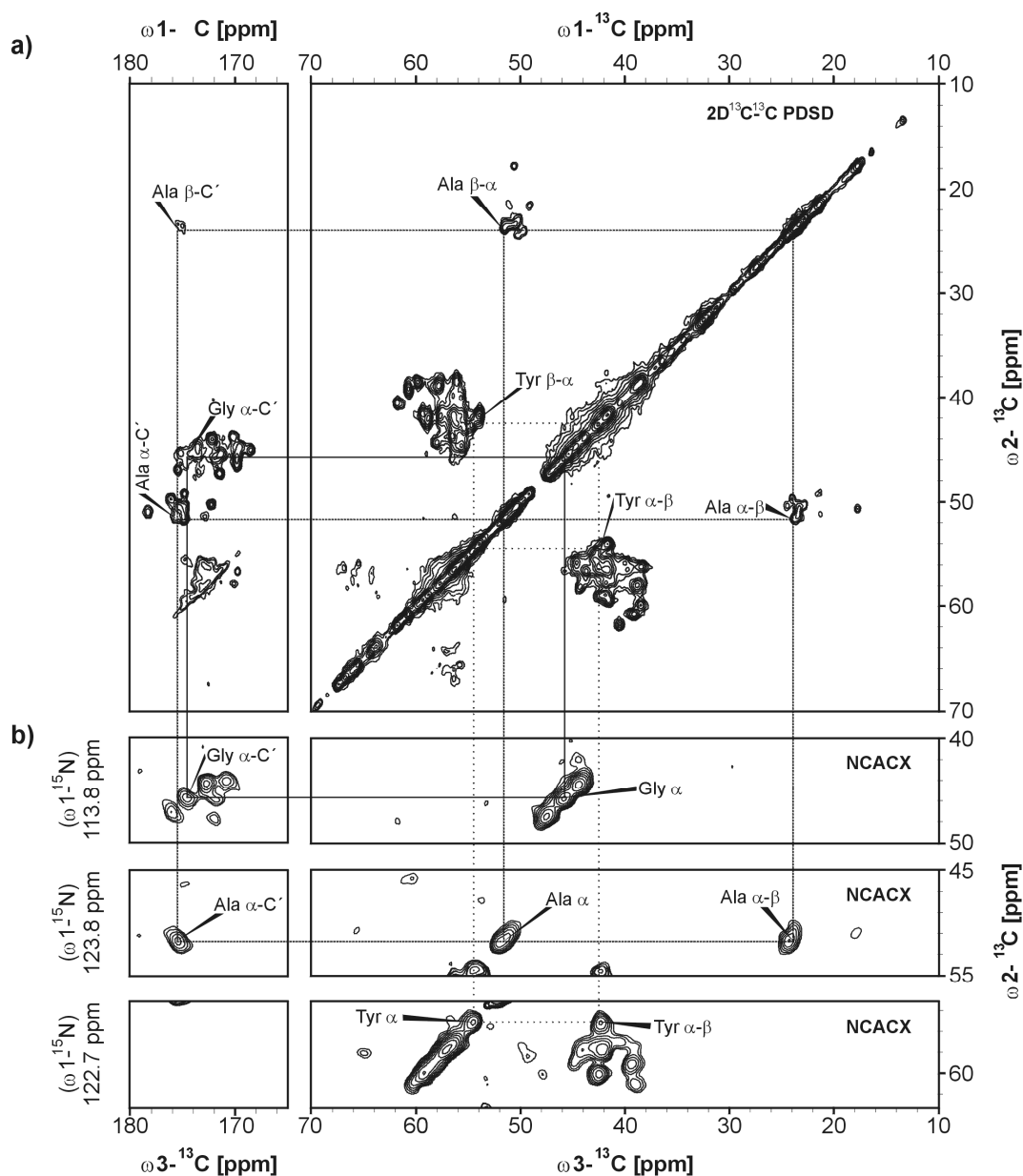
```

1 - MEERNDWHFN IGAMYEIENV EGYGEDMDGL AEPSVYFNAA NGPWRIALAY -50
51- YQEGPVDYSA GKRGTWFDRP ELEVHYQFLE NDDFSFGLTG GFRNYGYHV -100
101-DEPGKDTANM QRWKIAPDWD VKLTDDLRFN GWLSMYKFAN DLNTTGYADT -150
151-RVETETGLQY TFNETVALRV NYYLERGFNM DDSRNNGEFS TQEIRAYLPL -200
201-TLGNHSVTPY TRIGLDRWSN WDWQDDIERE GHDFNRVGLF YGYDFQNGLS -250
251-VSLEYAFEWQ DHDEGDSDKF HYAGVGVNYS F-281

```

**Figure 3.15** OmpG sequence. 2,3  $^{13}\text{C}$  labelled aromatic amino acids are labelled in blue. Uniformly  $^{13}\text{C}$ -labelled amino acids Gly and Ala are labelled in red. Sequential connectivities that will give a cross peak in an NCO-type experiment is underlined.

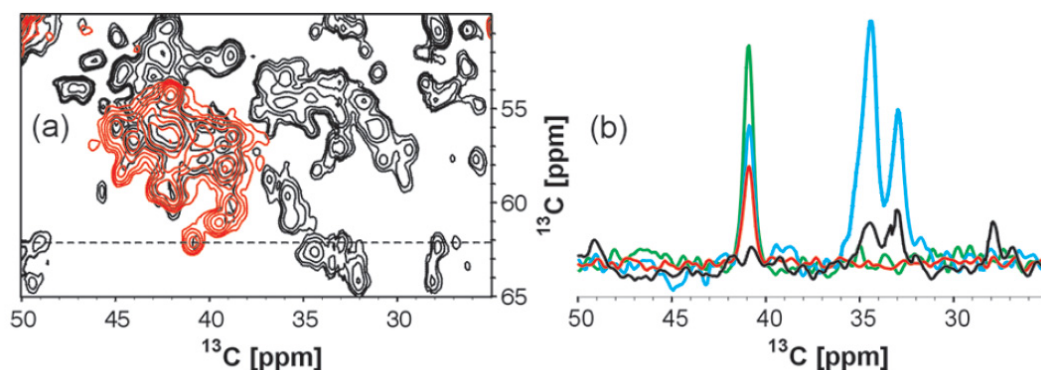
In a first step, the labelling scheme of the [GAFY]-OmpG was proven by a 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS and 3D NCACX correlation experiment. As can be seen in the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum  $\text{C}\alpha$ - $\text{C}\beta$  signals for Phe, Tyr, and Ala and the  $\text{C}\alpha$ - $\text{C}'$  cross-peaks for Gly and Ala are visible (Figure 3.16a). While there are still problems with overlap in the Phe/Tyr and Ala  $\text{C}\alpha$ - $\text{C}\beta$  signal clusters, the spectrum is considerably less crowded than that of the uniformly labelled sample.



**Figure 3.16 Identification of the Ala, Gly, Phe and Tyr spin systems in a 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS (a) and a 3D NCACX spectrum of [GAFY]-OmpG (b).** In the first step, resonances of several spin systems were identified in the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS (a). In the second step, a 3D NCACX spectrum was used to identify the backbone nitrogen atom in the 3D NCACX spectrum ( $^{15}\text{N}$  strips in b). As examples, the identification of Gly, Ala and Tyr spin systems are indicated by solid, dashed and dotted lines, respectively. The  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum was recorded on a 900 MHz spectrometer, at a spinning frequency of 13 kHz and at 280 K with a PDS mixing time of 20 ms. The NCACX spectrum was recorded on an 400 MHz wide-bore spectrometer, at a spinning frequency of 8 kHz and at 280 K with a PDS mixing time of 20 ms.

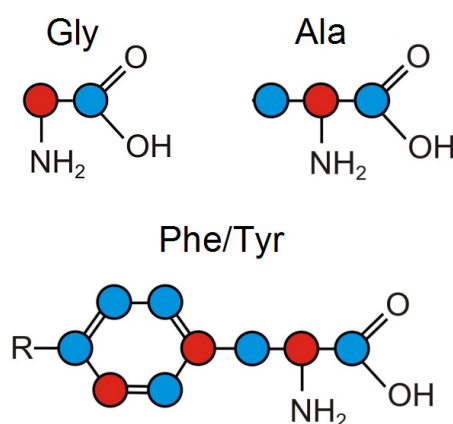
Using  $^{15}\text{N}$ -strips of the 3D NCACX spectrum it was possible to identify the N, C $\alpha$ , C $\beta$ , and C' chemical shifts for each of the spin systems. Some overlap could not be resolved, but all alanines, 96% of the glycines, and 85% of the aromatic amino acids were identified (Figure 3.16b). In addition, weak Ser C $\alpha$ -C $\beta$  ( $\omega_1$ = 56 ppm  $\omega_2$ = 66 ppm) and C $\alpha$ -C' ( $\omega_1$ = 174 ppm  $\omega_2$ = 56 ppm) peaks are also present in the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum (Figure 3.16a). This can be explained by considering the metabolic and catabolic pathways of glycine in *E.coli*, which both proceed via serine [152]. Since all amino acids are provided in excess in the bacterial growth medium, it is not necessary for the cells to synthesise glycine from serine. Thus the isotopic labelling in the glycine is not diluted. However, due to the excess of amino acids, the cells break down some of the glycine, which results in the incorporation of isotopic labels into serine.

Considering the Phe and Tyr C $\alpha$ -C $\beta$  correlations, the cross peaks are more intense and new peaks appear in the [GAFY]-OmpG spectrum compared to the U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -OmpG spectrum (Figure 3.17a). The peak intensities improve more significantly than the line widths, which suggests that the major cause of the poor spectral quality in the U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -OmpG spectrum is the ring system acting as a “magnetization sink”. This is supported by the fact that a  $^{13}\text{C}$ - $^{13}\text{C}$  DREAM experiment optimised for the aromatic C $\alpha$ /C $\beta$  cross-peaks further increases the signal intensity, but to a far greater extent for U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -OmpG than for [GAFY]-OmpG (Figure 3.17b). Line widths are similar in all spectra and are comparable to those of peaks from other spin systems. The [GAFY]-OmpG sample with [2,3- $^{13}\text{C}$ ]-labelled aromatic residues thus has the major advantage that peak intensities are nearly those of a U- $^{13}\text{C}$ ,  $^{15}\text{N}$ -OmpG band-selective experiment, even during broad banded experiments. This is of particular importance for obtaining long-range and certain sequential cross-peaks, which require long mixing times. This would not be possible with DREAM or other band-selective pulse sequences since the intense and continued irradiation over a long mixing time would lead to increased sample heating and thus sample degradation.



**Figure 3.17 Comparison of Phe and Tyr  $C\alpha$ - $C\beta$  cross peak intensities in spectra recorded on U- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG and [GAFY]-OmpG.** (a) Overlay of  $^{13}\text{C}$ - $^{13}\text{C}$  PDSD spectra (20 ms mixing time) recorded at 900 MHz on U- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG (black) and [GAFY]-OmpG (red); (b) 1D traces of the spectra in panel (a) at the dotted line as well as from 2D  $^{13}\text{C}$ - $^{13}\text{C}$  DREAM spectra (1.5 ms, optimized for aromatic  $C\alpha$ / $C\beta$  cross peaks) recorded at 900 MHz on U- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG (blue) and [GAFY]-OmpG (green).

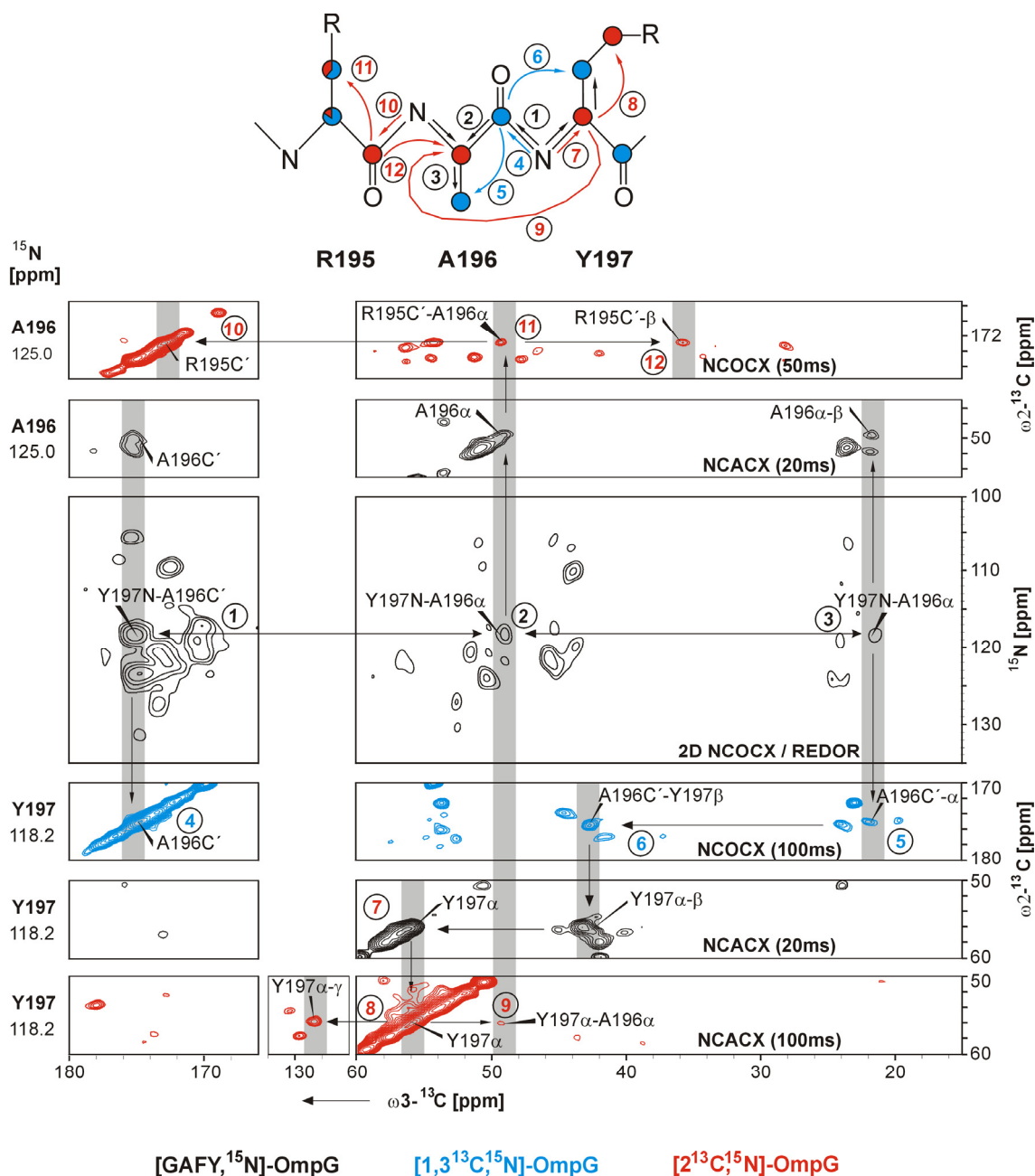
The assignment of [GAFY]-OmpG was achieved by a combinatorial use of 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS, 3D NCACX, 2D NCOCX and 2D REDOR spectra. In the first step, sequential assignments involving Ala and Gly resonances were derived from the 2D NCOCX and 2D REDOR spectra. These spectra provide links from the backbone nitrogen atom (Ala, Gly, Phe or Tyr residues) of the amino acid (i) to Gly ( $C'_{i-1}$  and  $C\alpha_{i-1}$ ) or Ala ( $C'_{i-1}$ ,  $C\alpha_{i-1}$  and  $C\beta_{i-1}$ ) residues. In addition, the assignment is supported by using 3D NCACX and NCOCX spectra of the  $[1,3-^{13}\text{C}, ^{15}\text{N}]$ - and  $[2-^{13}\text{C}, ^{15}\text{N}]$ -OmpG samples (labelling pattern Figure 3.18). The 3D NCACX spectrum of the  $[2-^{13}\text{C}, ^{15}\text{N}]$ -OmpG was particularly helpful to distinguish Phe and Tyr residues, since the chemical shift of their  $C\gamma$  atoms are quite different (Phe  $C\gamma = 138.2 \pm 2,4$  ppm and Tyr  $C\gamma = 129.3 \pm 2,5$  ppm).



**Fig. 3.18 Schematic representation of the labelling pattern of Gly, Ala, Phe and Tyr as it was obtained by protein expression in *E.coli* BL21(DE3).** The red colour corresponds to the degree of  $^{13}\text{C}$  incorporation into the  $[2-^{13}\text{C}, ^{15}\text{N}]$ -labelled sample. The opposite labelling pattern corresponds to the  $[1,3-^{13}\text{C}, ^{15}\text{N}]$ -labelled sample.

The assignment procedure is illustrated for Arg195-Ala196-Tyr197 as an example (Figure 3.19). The 2D NCOCX and 2D REDOR spectra show that residue [i-1] must be an Ala, due to the presence of the C' (174.8 ppm), C $\alpha$  (49.3 ppm) and C $\beta$  (21.6 ppm) resonances at N[i]= 118.2 ppm. However, it is difficult to identify residue [i] in the 3D NCACX spectrum of [GAFY]-OmpG only from the nitrogen resonance.

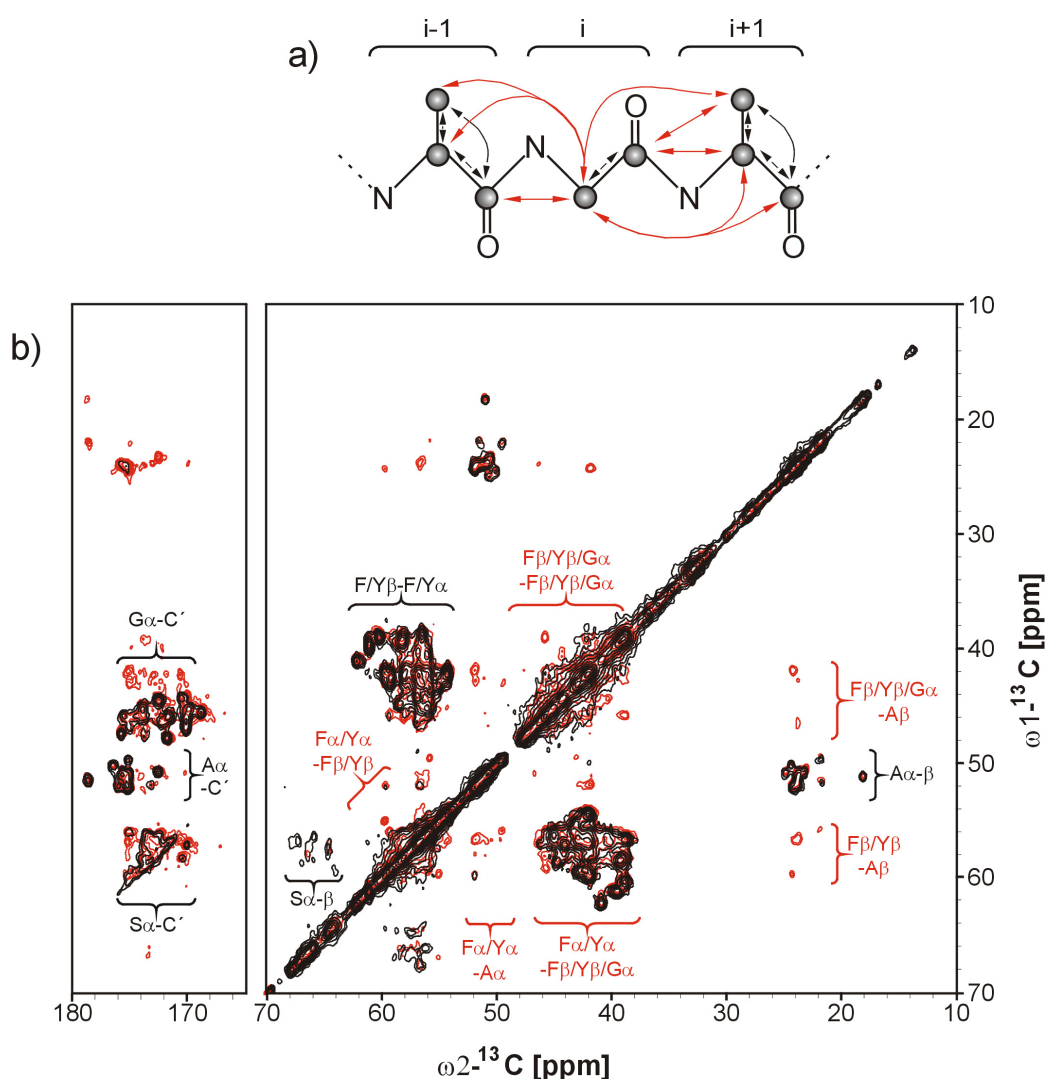
The spectrum of [1,3-<sup>13</sup>C, <sup>15</sup>N]-OmpG, for example, provides the same sequential correlations for Ala197 (except C $\alpha$ -resonances), due to the 100% labelling of the carbonyl and the C $\beta$  carbon atoms. Additionally, the spectrum also provides a link to a C $\beta$  carbon atom of an aromatic residue at 42.8 ppm at the same CO frequency. Thus, the sequential cross peaks belong to an Ala-Phe or Ala-Tyr pair. Using both, the chemical shift of the backbone nitrogen (N[i]=118.2 ppm) and that of the C $\beta$ [i] carbon atom (42.8 ppm), it was possible to identify the resonance of the C $\alpha$ [i] carbon atom at 55.6 ppm in the 3D NCACX spectrum of [GAFY]-OmpG. However, from these data it is unclear which type of aromatic amino acid is present at this nitrogen frequency. In this case, the 3D NCACX spectrum of the [2-<sup>13</sup>C, <sup>15</sup>N]-OmpG sample was helpful. In this spectrum a cross peak at 128.6 ppm, which is typical for a Tyr C $\gamma$  atom, was observed in the nitrogen plane N=118.2 ppm. In addition, a cross peak with the C $\alpha$  frequency of Ala 196 was found in the same nitrogen plane. Thus, it is clear that the sequential correlation in the 2D REDOR and NCOCX spectra arises from an Ala-Tyr pair. However, OmpG contains two Ala-Tyr pairs at positions 49-50 and 196-197 (Figure 3.15). To assign the observed resonances to one of the two pairs, the chemical shift value of the Ala backbone nitrogen atom (N=125 ppm), extracted from the 3D NCACX spectrum of [GAFY]-OmpG was used to find a sequential connectivity to the previous amino acid in the 3D NCOCX spectra of [2-<sup>13</sup>C, <sup>15</sup>N]- or [1,3-<sup>13</sup>C, <sup>15</sup>N]-OmpG samples. In the spectrum of the [2-<sup>13</sup>C, <sup>15</sup>N]-OmpG sample, a peak pattern typical for an Arg residue, including the C $\alpha$  resonance of the alanine, was found. Thus, the Ala-Tyr pair is unambiguously assigned to Arg195-Ala196-Tyr-197 (Figure 3.19). Using this strategy, all sequential pairs involving Ala and Gly as the first amino acid were assigned.



**Figure 3.19** Sequential assignment R195 - Ala196 - Tyr197 using 2D REDOR and 2D NCOCX spectra of the  $[\text{GAFY}]$ -OmpG and 3D NCACX and NCOCX spectra of the  $[\text{GAFY}]$ -OmpG,  $[1,3\text{-}^{13}\text{C}, ^{15}\text{N}]$ - or  $[2\text{-}^{13}\text{C}, ^{15}\text{N}]$ -OmpG samples. Top: Schematic representation of the magnetisation transfer during 2D REDOR, 2D NCOCX, 3D NCACX and 3D NCOCX experiments. Below; the assignment procedure starts from the 2D REDOR/NCOCX spectrum. Circled numbers indicate the magnetisation transfers and the resulting assignment steps (for more details see text). Mixing times in ms are indicated on the right. Spectra are coloured red ( $[2\text{-}^{13}\text{C}, ^{15}\text{N}]$ -OmpG sample), blue ( $[1,3\text{-}^{13}\text{C}, ^{15}\text{N}]$ -OmpG sample) and black ( $[\text{GAFY}]$ -OmpG). 2D and 3D spectra were recorded on a 400 MHz wide-bore spectrometer at 280 K. The spinning frequency was set to 8 kHz.



Castellani *et al.* described a method for assigning long-range distances in  $[2-^{13}\text{C},^{15}\text{N}]$ - or  $[1,3-^{13}\text{C},^{15}\text{N}]$  labelled  $\alpha$ -spectrin SH3 domain samples by using 3D NCACX and NCOCX spectra recorded with longer  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times [99]. Since the [GAFY]-OmpG sample is composed only of short-chain amino acids, it should be possible to assign sequential and long-range connectivities in a similar way (Figure 3.20a). Therefore, 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS and 3D NCACX spectra were recorded with  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times between 300 and 700 ms. As can be seen in the  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum, inter-residual correlations between  $\text{C}\beta$ - $\text{C}\beta$ ,  $\text{C}\alpha$ - $\text{C}\beta$  and  $\text{C}\alpha$ - $\text{C}\alpha$  carbon atoms become visible at longer  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times (700 ms) (Figure 3.20b, red spectrum).

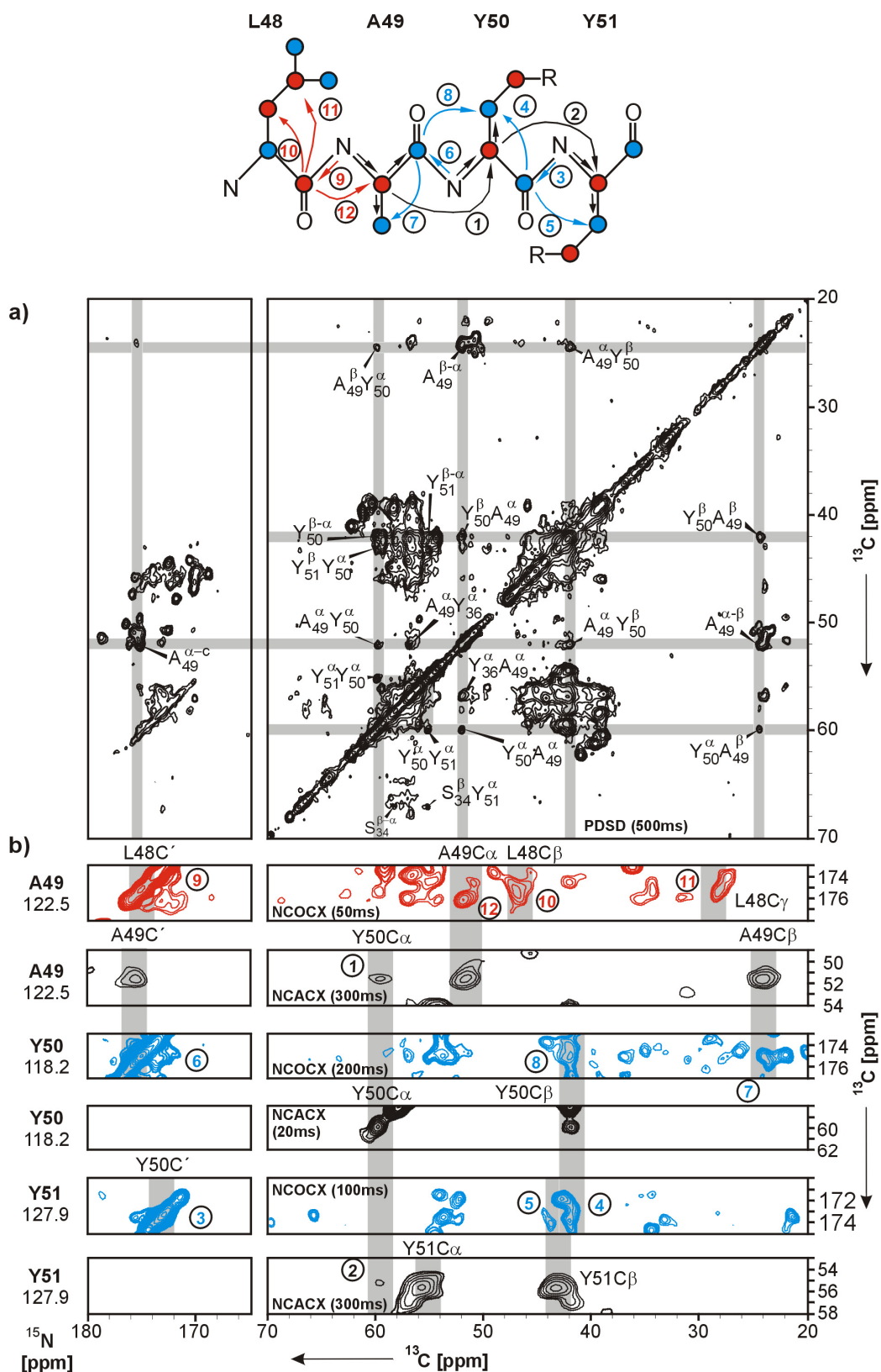


**Figure 3.20** Long-range connectivities can be obtained at longer  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times. a) Schematic representation of the magnetisation transfer with short (black, 20 ms) and long  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times (red, 700 ms). b)  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectra of OmpG-GAFY recorded at 900 MHz with 20 ms (black spectrum) and 700 ms (red spectrum) mixing time. Intra- (black) and inter-residual (red) cross peak regions are indicated.

Since OmpG is an anti-parallel  $\beta$ -sheet protein some of the detected correlations also arise from inter-strand  $C\alpha$ - $C\alpha$  connectivities as the distance between sequential  $C\alpha$ - $C\beta$  carbon atoms are of comparable length (around 5 Å).

In Figure 3.21a, the assignment procedure of the L-A-Y motif (at positions 48-49-50-51) is illustrated as an example. In the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum sequential cross peaks can be traced from  $A49\alpha$ - $Y50\beta$  to  $Y50\alpha$ - $A49\alpha$ ,  $Y50\alpha$ - $Y51\alpha$  and  $Y50\alpha$ - $Y51\beta$ . Additional inter-strand links can be found from  $A49\alpha$  probably to  $Y36\alpha$  and from  $Y51\alpha$  to the  $S34\beta$  carbon atom (Figure 3.21a). Parts of the sequential correlations are also visible in the 3D NCACX spectra recorded with longer  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times. A sequential cross peak from the  $A49\text{NC}\alpha$  group to  $Y50\alpha$  is visible in the 3D NCACX spectrum of the [GAFY]-OmpG sample recorded with 300 ms  $^{13}\text{C}$ - $^{13}\text{C}$  mixing time (Figure 3.21b,  $^{15}\text{N}=122.5$  ppm). The same situation is found in the  $^{15}\text{N}$ -plane corresponding to  $Y51$  (Figure 3.21b,  $^{15}\text{N}=127.9$  ppm). Surprisingly, no inter residual cross peaks can be found in the  $^{15}\text{N}$ -plane corresponding to  $Y50$ . The fact that some sequential cross peaks are present in a 3D NCACX experiment (recorded at longer  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times) and others not, was often observed during the assignment procedure. The reason for this may be the low magnetization transfer during the specific N- $C\alpha$  transfer step in the 3D NCA or NCO type experiments. Additionally, the inclusion of the  $^{13}\text{C}$ - $^{13}\text{C}$  mixing step (e.g. 300ms) results in further loss in sensitivity. In contrast to that, the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum recorded with longer mixing times shows a much higher sensitivity and thus it is a much more robust experiment. Therefore, 2D experiments provide useful starting points for sequential assignments.

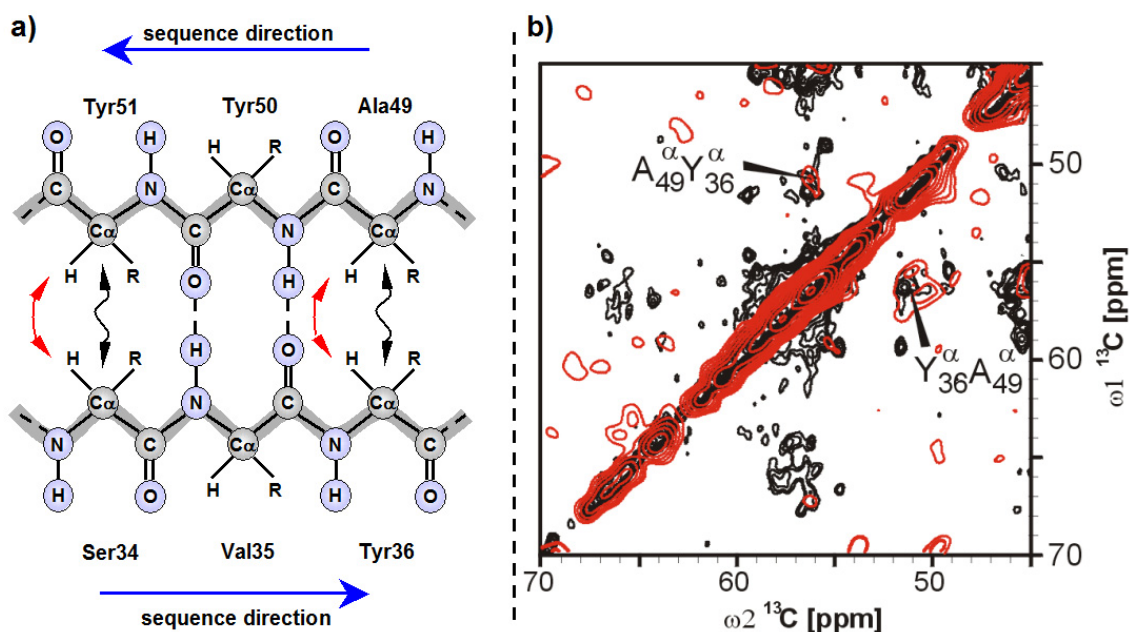
Nevertheless, the chemical shift values in both, the  $^{15}\text{N}$ -plane of  $Y51$  and  $A49$  were used as starting points for the sequential assignment. In the first step, the chemical shift value of the backbone nitrogen atom of  $Y51$  was used to find a sequential connectivity to  $Y50$  in the 3D NCOCX spectrum of the  $[1,3\text{-}^{13}\text{C},^{15}\text{N}]$ -OmpG (Figure 3.21,  $^{15}\text{N}=127.9$  ppm). Indeed, peaks corresponding to the  $C'$  and  $C\beta$  carbon atoms of  $Y50$  were found, due to the 100 % labelling of these carbon atoms. In the same way, a peak pattern corresponding to the  $C'$  and  $C\beta$  of  $A49$  and an additional cross peak to the  $C\beta$  carbon atom of  $Y50$  were found (Figure 3.21,  $^{15}\text{N}=118.2$  ppm). This shows that the sequential cross peaks observed in the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  spectrum belong to the  $A49$ - $Y50$ - $Y51$  triplet in OmpG. To extend the assignment of this triplet to  $L48$ , 3D NCOCX spectra of the  $[2\text{-}^{13}\text{C},^{15}\text{N}]$ - and  $[1,3\text{-}^{13}\text{C},^{15}\text{N}]$ -OmpG samples were used, respectively.



**Figure 3.21** Illustration of the assignment of the OmpG L48-A49-Y50-Y51 motif using spectra of the [GAFY]-, [2- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]- and [1,3- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG samples. (a)  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum (700 ms). Sequential cross-peak connections can be traced from  $A_{49}^{\alpha}$ - $Y_{50}^{\beta}$  to  $Y_{50}^{\alpha}$ - $A_{49}^{\alpha}$ ,  $Y_{50}^{\alpha}$ - $Y_{51}^{\alpha}$  and  $Y_{50}^{\alpha}$ - $Y_{51}^{\beta}$ . Additional cross-strand links can be found from  $A_{49}^{\alpha}$  to a F/Y/S $\alpha$  and from  $Y_{51}^{\alpha}$  to a S $\beta$ . (b) 2D nitrogen planes of several 3D NCACX spectra recorded at different  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times. The scheme on top of the figure indicates the magnetisation transfer and should be used to follow up the sequential assignments (for more details see text).

In the  $^{15}\text{N}$ -plane of A49, a typical Leu cross-peak pattern, including the  $\text{C}\beta$  and  $\text{C}\gamma$  carbon atoms of Leu48 and the  $\text{C}\alpha$  carbon atom of A49 were found (Figure 3.21b,  $^{15}\text{N}=122.5$  ppm).

As described above, inter-strand links between the antiparallel arranged  $\beta$ -sheets (A49 $\alpha$  to Y36 $\alpha$  and from Y51 $\alpha$  to a S34 $\beta$ ) were observed in the  $^{13}\text{C}$ - $^{13}\text{C}$  PDS experiment recorded with 500 ms  $^{13}\text{C}$ - $^{13}\text{C}$  mixing time (Figure 3.21a and 3.22a). However, since the PDS experiment utilises through space  $^{13}\text{C}$ - $^{13}\text{C}$  spin diffusion, these assignments are not unambiguous. In contrast to that, a CHHC experiment employs indirect  $^{13}\text{C}$ - $^{13}\text{C}$  recoupling via the  $\text{C}\alpha$  bonded proton spins (Figure 3.22a). As can be seen in figure 3.22b, using a CHHC spectrum recorded with a short  $^1\text{H}$ - $^1\text{H}$  mixing time of (100  $\mu\text{s}$ ) it was possible to identify the inter-strand correlation between A49 $\alpha$  and Y36 $\alpha$ . The second inter-strand correlation between Y51 $\alpha$  to the S34 $\alpha$  is not visible. This might be due to the poor signal-to-noise ratio in the CHHC spectrum. This experiment has to be optimised in the future.



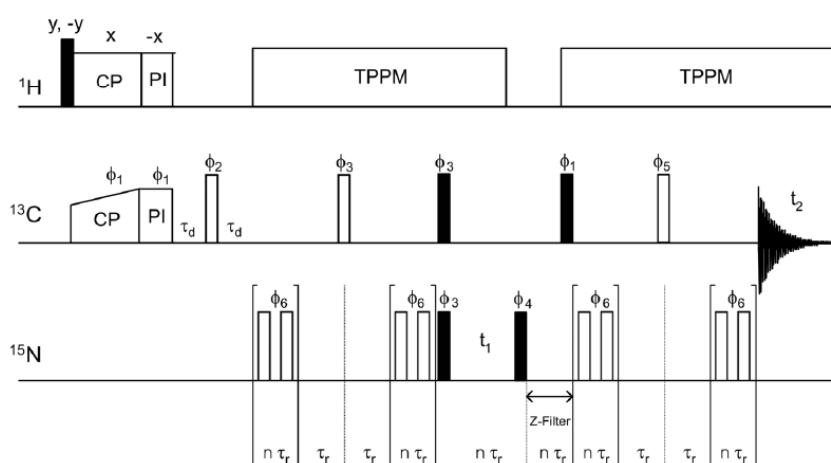
**Figure 3.22** a) Schematic representation of an antiparallel  $\beta$ -sheet as it occurs in the membrane spanning part of OmpG. In a PDS experiment recorded with a long  $^{13}\text{C}$ - $^{13}\text{C}$  mixing time, the magnetisation transfer between  $\text{C}\alpha$  carbon atoms of neighbouring  $\beta$ -strands occurs via through space spin diffusion (black arrow). In contrast to that, in a CHHC experiment recorded with a short  $^1\text{H}$ - $^1\text{H}$  mixing time (e.g. 50  $\mu\text{s}$ ), the magnetisation transfer between  $\text{C}\alpha$  carbon atoms of neighbouring  $\beta$ -strands occurs via the bonded proton spins (distance  $\sim 2.3$  Å, red arrows). The blue arrow indicates the sequence direction of the amino acid chain. b) A superposition of the  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum from figure 3.21 and the CHHC spectrum of the [GAFY]-OmpG sample recorded with an  $^1\text{H}$ - $^1\text{H}$  mixing time of 50  $\mu\text{s}$ . The interstrand cross peak between A49 $\alpha$  and Y36 $\alpha$  is indicated. The CHHC spectrum was recorded on a 900 MHz narrow-bore spectrometer at 280 K with a spinning frequency of 13 kHz.

In summary, on the basis of the GAFY labeling scheme, sequence specific assignments for a total of 45 residues could be made. Thus, the GAFY labelling scheme provided a highly successful starting point for the assignment of the large MP OmpG. Advantages of the labelling scheme lie in the fact that it is composed of a low number of small, isolated spin systems (pairs and triplets of sequentially labelled carbon atoms). Furthermore, the number of inter-residue cross peaks is significantly increased, which is important for the assignment. The use of a mixture of [2,3-<sup>13</sup>C]- and short U-[<sup>13</sup>C]-labelled amino acids has eased the interpretation of NCO- and NCA-type spectra. In addition, it has the advantage compared to the [1,3-<sup>13</sup>C]- and [2-<sup>13</sup>C]-glycerol labelling schemes that there is a direct connection between the C $\alpha$  and C $\beta$  chemical shifts by which the aromatic spin systems can be identified. Furthermore, [2,3-<sup>13</sup>C]-labelling provides the spectral quality of band-selective experiments while allowing broadband experiments to be recorded. Further help and confirmation of assignments was provided by NCACX and NCOCX spectra of OmpG expressed in media containing [2-<sup>13</sup>C]- and [1,3-<sup>13</sup>C]-glycerol as the sole carbon source. In many cases it was possible to distinguish between Phe and Tyr residues on the basis of their C $\gamma$  resonances in the [2-<sup>13</sup>C]-glycerol 100 ms 3D NCACX spectrum. These spectra also allowed several motifs to be extended beyond the GAFY residues. Thus the labelling scheme presented here should be applicable and invaluable to a wide variety of large MPs, enabling assignment of this important class of proteins in a first step toward their structural characterisation.

### 3.3.2 Assignment using spectral editing techniques: selection of methyl groups

As demonstrated in the last section, the use of selectively labelled samples is very helpful in decreasing spectral crowding and providing starting points for the assignment. Another approach is the spectral editing technique in which resonances of specific nuclei can be selected by their chemical and spectral properties. Methyl resonances are good monitors for ligand binding as demonstrated by Zech *et al.* (2004), who investigated changes in the chemical shifts of CH<sub>3</sub>-groups of the anti-apoptotic protein (Bcl-xL) induced by ligand binding [153]. In 1979 it was demonstrated that spectral editing in solid-state NMR could be achieved using dipolar dephasing, in order to distinguish chemically different carbons on the basis of the number of attached protons [154]. Likewise, it was shown that cross polarisation with polarisation inversion (CPPI) could be applied to select between carbons with different multiplicities [155,156].

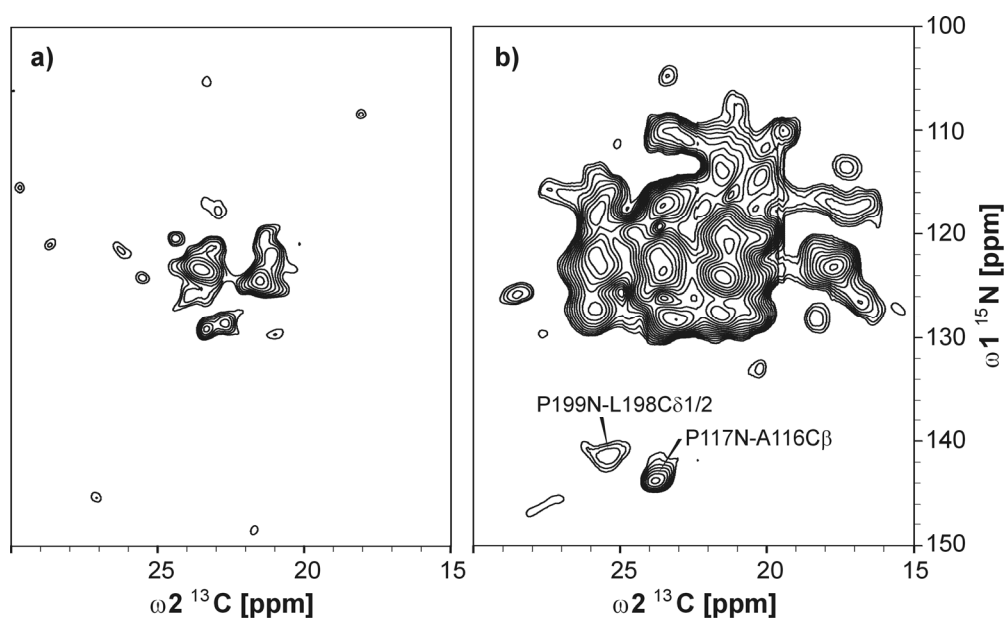
A simple but robust pulse sequence that combines CPPI and dipolar dephasing in a double filtering step to filter methyl resonances was developed in our group by Stefan Jehle [157]. The pulse sequence was designed for a CH<sub>3</sub>-double filtered DARR and a CH<sub>3</sub>-double filtered TEDOR experiment. For the CH<sub>3</sub>-double filtered TEDOR experiment the magnetisation was first transferred from protons to carbons by a ramped CP step (CP-time = 2 ms) (Figure 3.23).



**Figure 3.23** Pulse sequence used for the <sup>15</sup>N-<sup>13</sup>C heteronuclear dipolar correlation experiment. The polarisation was first transferred from protons to carbons using a ramped CP. In the next step, cross polarisation with polarisation inversion (CPPI) combined with dipolar dephasing was used to select methyl groups. The filtering block was followed by a rotational-echo double resonance (REDOR) mixing scheme. During the REDOR mixing and data acquisition a TPPM scheme was applied to decouple protons. The figure was taken from [157].

In the next step, cross polarisation with polarisation inversion (CPPI, with 50  $\mu$ s polarisation inversion) combined with dipolar dephasing was used to select methyl groups. During dipolar dephasing, proton decoupling was not applied. This leads to rapid transverse relaxation of  $\text{CH}_2$  and  $\text{CH}$ -carbon signals due to strong heteronuclear interactions with their directly bonded protons. This efficiently dephased the negative  $\text{CH}_2$  magnetisation and any residual  $\text{CH}$  signals, whilst  $\sim 35\%$  of the  $\text{CH}_3$  signals remained (reference). The CPPI in combination with dipolar dephasing was used in connection with a TEDOR transfer scheme to recouple methyl carbons and nitrogen spins (Figure 3.23).

In a first experiment, the methyl-filtered TEDOR experiment was applied to a  $[1,3^{13}\text{C},^{15}\text{N}]$ -OmpG sample with a short mixing time of 3 ms. As can be seen in figure 3.24, many correlations between the backbone nitrogen atoms and the methyl carbon atoms of alanines are observed (Figure 3.24a). Valine, threonine, leucine and isoleucine methyl groups could also give rise to intraresidual cross peaks.



**Figure 3.24 Methyl-selective TEDOR correlation spectra of the  $[1,3^{13}\text{C},^{15}\text{N}]$ -OmpG sample.** The spectra were recorded using REDOR mixing times of 3 ms (a) and 8 ms (b). The sequential assignments are indicated. Both spectra were recorded on a 400 MHz wide bore spectrometer.

However, they tended to be weaker since the methyl carbon atoms of these amino acids are further away from their backbone nitrogen atoms.

In the spectrum recorded with 8 ms mixing time, many signals were observed (Figure 3.24b). These signals predominantly arise due to transfer from methyl groups of isoleucines ( $\text{C}\gamma 1/2$ ), valines ( $\text{C}\gamma 1/2$ ), leucine ( $\text{C}\delta 1/2$ ) and threonine ( $\text{C}\gamma$ ) to backbone

nitrogens. Furthermore, a strong signal at 23.8 ppm in  $\omega_2$  and 143.7 ppm in  $\omega_1$  was observed. The chemical shift value in  $\omega_1$  (143.7 ppm) indicates that the cross peak can only be attributed to a sequential or interstrand magnetisation transfer to the backbone nitrogen atom of a proline. The protein sequence of OmpG contains only three residue pairs that could account for this transfer: A116-P117, L198-P199 and T208-P209. A Magnetisation transfer from the  $C_\gamma$  of T208 can be excluded since the chemical shifts detected for the threonine  $C_\gamma$ s do not match. The second candidate, L198-P199, can also be excluded, since the leucine  $C\delta_{1/2}$  chemical shifts are slightly higher (around 25 ppm). Thus, this cross peak can most likely be attributed to the A116-P117 pair. A second, weaker, cross peak was observed at 25.4 ppm in  $\omega_2$  and 141.5 ppm in  $\omega_1$ . This peak was initially tentatively assigned to L198, as the chemical shift for the methyl carbon atom is in good agreement with those found for a leucine  $C\delta_1$  or  $C\delta_2$  carbon atom. The weaker intensity of the cross peak is consistent with a larger distance between the methyl group of a leucine and the backbone nitrogen of the following amino acid. However, both assignments need to be verified by other signals, since the cross peaks could also reflect long-range or inter-strand correlations. Further assignments by Dr. Victoria Higman have shown the A116-P117 to be correct, but the other P-L link to arise from a cross-strand correlation between L198 and P209.

The method used here provides a simple way to obtain efficient filtering of methyl groups. The suppression of CH and  $CH_2$  signals can significantly simplify spectral regions containing methyl groups. Additionally, the double filter can be easily implemented as a building block into other multidimensional solid-state MAS NMR experiments. Moreover, it is shown that the methyl selective  $^{15}N$ - $^{13}C$  correlation experiment can be used to find starting points for the sequential assignment.



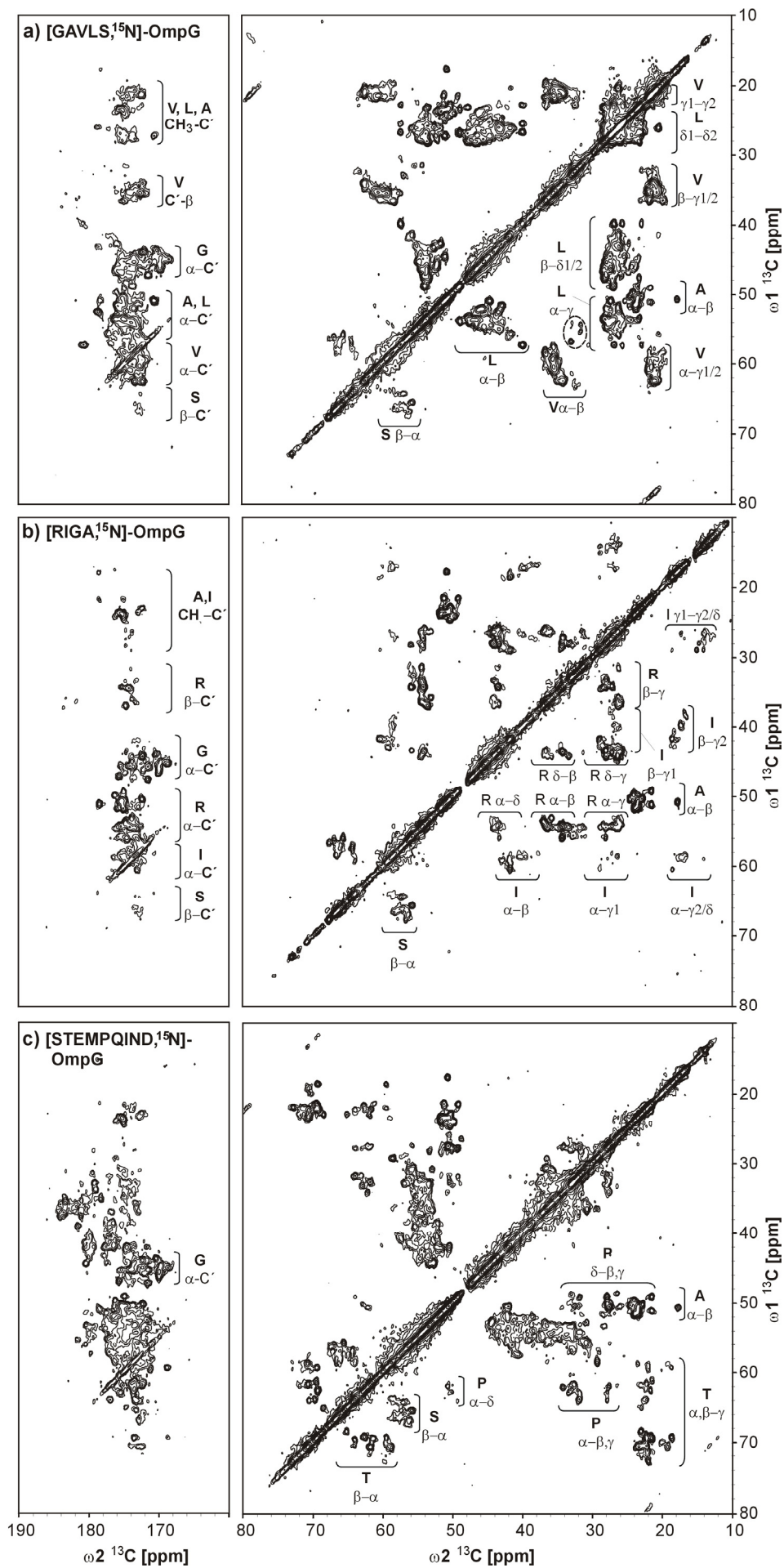
### 3.3.3 Enhancements of the assignment using other specifically labelled samples

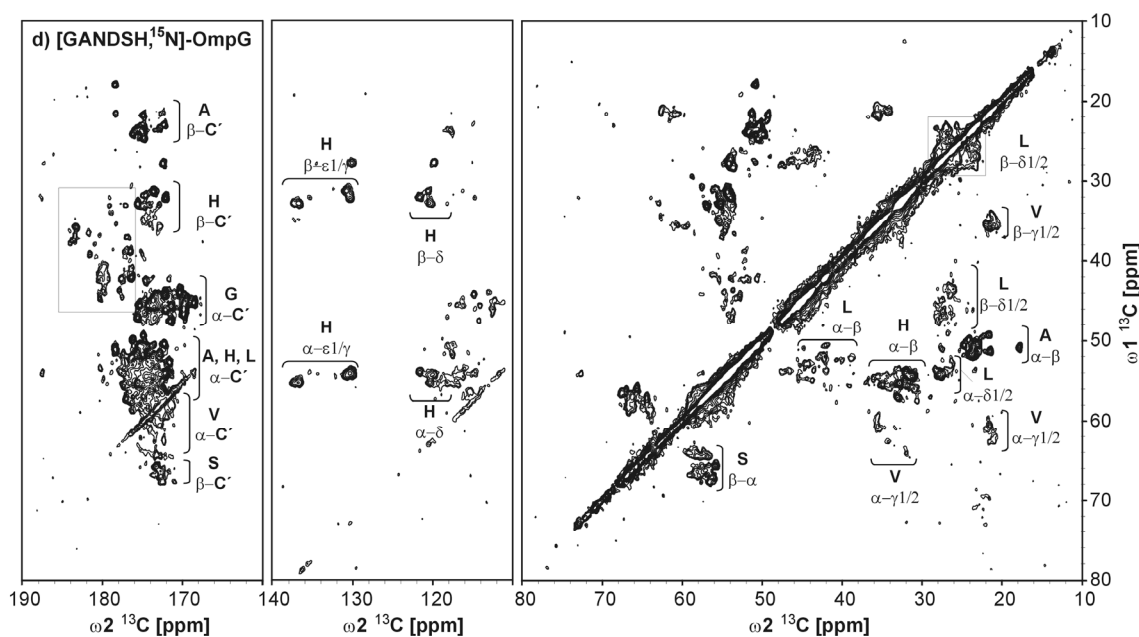
Since the assignment of the GAFY labelled sample was successful, four other samples with different  $^{13}\text{C}$  labelling schemes were prepared (Table 3.3). In these samples, the backbone nitrogen atoms of all 20 amino acids are  $^{15}\text{N}$ -labelled. This was achieved by growing *E. coli* cultures on M9 minimal medium supplemented with  $^{15}\text{NH}_4\text{Cl}$  and labelled and unlabelled amino acids, also termed forward labelling. Another possibility is the reverse labelling strategy, where M9 medium containing  $^{15}\text{NH}_4\text{Cl}$  and uniformly  $^{13}\text{C}$ -labelled glucose or glycerol is supplemented with unlabelled amino acids, whose  $^{13}\text{C}$ -labelling is to be suppressed. However, these samples have to be planned carefully, as several amino acids form the building blocks for the synthesis of other amino acids, and isotope dilution to other amino acids can thus occur. Therefore, the labelling schemes of all four samples were chosen in such a way, as to minimise isotope dilution, to decrease spectral overlap and to provide complementary information for the assignment of OmpG. For each sample OmpG was purified and reconstituted into 2D crystals as described in sections 3.1 and 3.2.

**Table 3.3 Specifically  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled OmpG samples used for assignment**

Sample	$^{13}\text{C}$ labelled residues	labelling type
[GAVLS, $^{15}\text{N}$ ]-OmpG	G, A, V, L, S	forward
[GANDSH, $^{15}\text{N}$ ]-OmpG	G, A, N, D, S, H	forward
[RIGA, $^{15}\text{N}$ ]-OmpG	R, I, G, A	forward
[STEMPQIND, $^{15}\text{N}$ ]-OmpG	S, T, E, M, P, Q, N, D, S	reverse

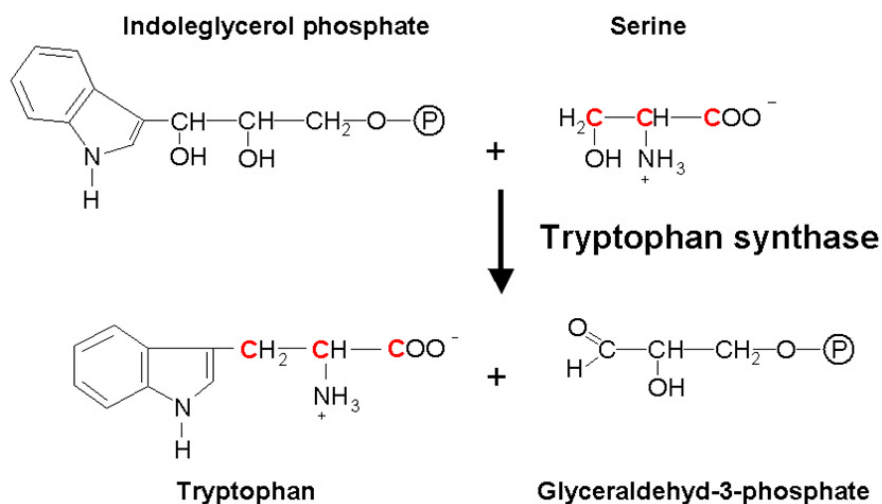
In a first step, the labelling pattern of each sample was analysed by recording 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectra (Figure 3.25). In the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum of the [GAVLS,  $^{15}\text{N}$ ]-OmpG sample the complete side chain signal patterns for Ala, Gly, Leu, Ser and Val are visible (Figure 3.25a). In addition, weak signals in the range of 50-60 ppm ( $\omega_1$ ) and 30-40 ppm ( $\omega_2$ ) are visible along with correlations to the carbonyl region. These signals are in the typical chemical shift range of  $\text{C}\alpha$ - $\text{C}\beta$  correlations of Trp residues. The labelling of Trp  $\text{C}\alpha$  and  $\text{C}\beta$  carbon atoms might be due to the fact that Trp is synthesised through the shikimate pathway in which chorismate is converted to Phe, Tyr and Trp [158].





**Figure 3.25** 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDSM NMR spectra of the [GAVLS,  $^{15}\text{N}$ ]-, [RIGA,  $^{15}\text{N}$ ]-, [STEMPQIND,  $^{15}\text{N}$ ]- and [GANDSH,  $^{15}\text{N}$ ]-OmpG samples. The spectra were recorded at 900 MHz and 280 K with a PDSM mixing time of 20 ms and a MAS frequency of 13 kHz. Brackets indicate the correlation networks. In a), the Trp  $\text{C}\alpha$ - $\text{C}\beta$  correlations are indicated by a circle.

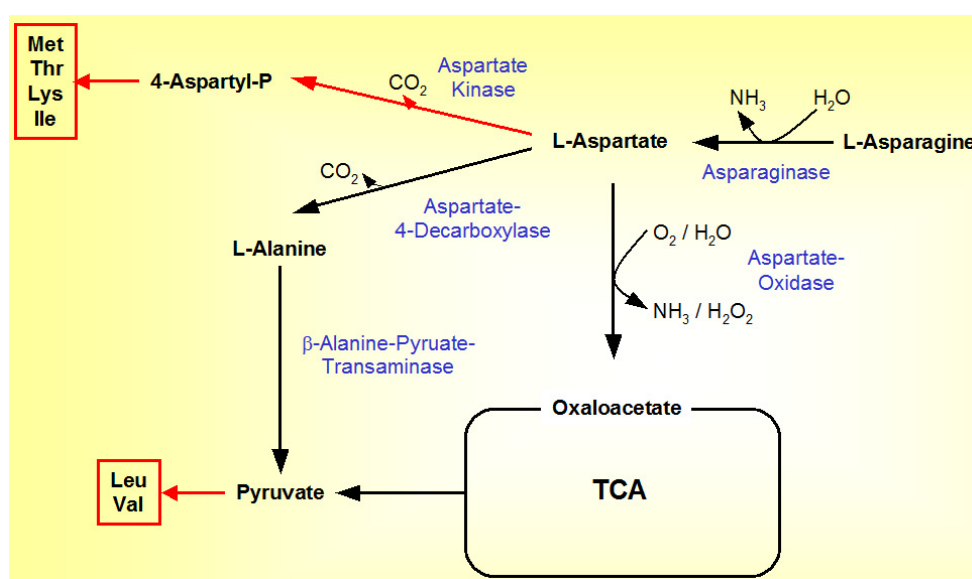
In the last step of the synthesis, the tryptophan synthase converts indoleglycerol phosphate and serine to tryptophan and glyceraldehyde-3-phosphate. Thus, incorporation of  $^{13}\text{C}$  labelled carbon atoms at the  $\text{C}\alpha$ ,  $\text{C}\beta$  and  $\text{C}'$  position of tryptophane originated from the serine catabolism (Figure 3.26).



**Figure 3.26** Last step of the tryptophan synthesis. The tryptophan synthase converts indoleglycerol phosphate and serine to tryptophan and glyceraldehyde 3-phosphate. Red letters indicate the incorporation of  $^{13}\text{C}$ -labelled carbon atoms.

The second sample contains  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled Ala, Arg, Gly and Ile; all other amino acids were added to the growth medium unlabelled (sample referred as [RIGA,  $^{15}\text{N}$ ]-OmpG). The  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum of [RIGA,  $^{15}\text{N}$ ]-OmpG shows the complete side chain signal patterns of Ala, Arg, Gly and Ile (Figure 3.25b). Similar to the GAFY sample, Ser  $\alpha$ - $\beta$  and  $\alpha$ -C' carbon atoms are visible due to the Gly catabolism (section 3.3.1). However, no isotopic spin dilution to Trp occurred, since Ser was added into the growth medium unlabelled and the labelled Ser produced from the labelled glycine was not further metabolised to Trp.

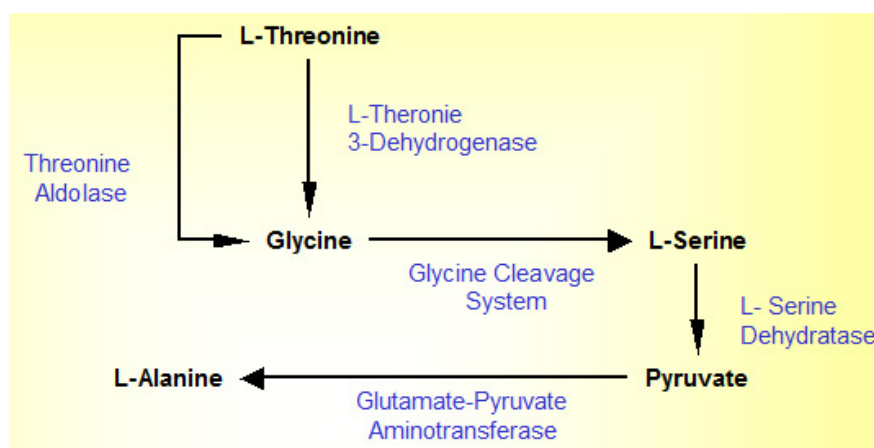
The third sample includes  $^{13}\text{C}$ ,  $^{15}\text{N}$  labelled Ala, Asn, Asp, Gly and His (sample termed [GANDSH,  $^{15}\text{N}$ ]-OmpG), which were prepared using the forward labelling procedure. The  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum of [GANDSH,  $^{15}\text{N}$ ]-OmpG shows the complete side chain signal pattern of Ala, Gly, His and Ser (Figure 3.25d). However, weak signals belonging to the amino acids Leu and Val are also visible. This might be due to the branched catabolism of Asn and Asp (Figure 3.27) [158]. In the first pathway Asn can be converted to Asp, which can be transformed to  $\beta$ -alanine by the catalytic action of the aspartate-4-decarboxylase. Afterwards,  $\beta$ -alanine-pyruvate-transaminase converts alanine to pyruvate. In a second pathway, the catabolism of Asp yields oxaloacetate and pyruvate via aspartate-oxidase and the tricarboxylic acid cycle (TCA cycle). Thus in both directions pyruvate is formed, which is the basis for the Leu and Val metabolism (Figure 3.27).



**Figure 3.27 Pathways of the Aspartate and Asparagine catabolism.** Black arrows indicate catabolic pathways. Red arrows indicate metabolic pathways (for more details see text).

As in the GAVLS sample, C $\alpha$ , C $\beta$  and C' signals from Trp were observed in the spectrum of the GANDSH sample. This sample, illustrates the difficulty of predicting isotopic scrambling. Since neither the GAFY nor the RIGA sample had resulted in the production of Leu or Val from Ala and pyruvate, it had not been anticipated that such severe scrambling would occur in the GANDSH sample.

The fourth sample was initially referred to as [STEMPQIND, <sup>15</sup>N]-OmpG. This sample was produced using the reverse labelling procedure. Therefore, all amino acids, which should be unlabelled, were added to the medium as unlabelled amino acids (Ala, Arg, Gly, His, Leu, Lys, Phe, Tyr, Trp and Val). To label all other amino acids the medium was supplemented with U-<sup>13</sup>C glucose and <sup>15</sup>N-NH<sub>4</sub>Cl. As can be seen in the <sup>13</sup>C-<sup>13</sup>C PDS spectrum the complete side chain pattern of Pro, Ser and Thr are visible (Figure 3.25c). However in the range of 50-60 ppm in  $\omega_1$  and 30-50 ppm in  $\omega_2$  the signal sets of Asn, Asp, Gln, Glu and Met are overlapped, due to the high similarity of the chemical shifts of these amino acids. Thus, it is difficult to exclude isotopically scrambling in this area. Additionally, Ile cross peaks were not visible in this spectrum. This is surprising, since Ile is a product of Thr, which is clearly visible in the spectrum. The reason for this could be the antagonistic inhibition of the Ile synthesis by the presence of unlabelled Leu and Val in the growth medium [152]. Contrary to the expectations, cross peaks arising from Ala and Gly are visible in the spectrum. These must arise from the reversible catabolism of Gly from Ser. The labelling of Ala might be due to both the Thr and the Ser catabolism (Figure 3.28). In the first step of the Thr-catabolism glycine is formed by the catalytic activity of the threonine aldolase and the L-threonine 3-dehydrogenase. In the second step, the Glycine Cleavage System converts glycine to serine, which can be converted to pyruvate by the L-serine dehydratase. Finally, labelled alanine can be formed from pyruvate (Figure 3.28)[152]. Using this information, a further sample with a similar labelling pattern ([TEMPQANDSG, <sup>15</sup>N]-OmpG) was prepared. However, NMR spectra of this sample have still to be recorded.



**Figure 3.28 Pathways of the threonine catabolism.** Black arrows indicate catabolic pathways. Names of enzymes are written in blue (for more details see text).

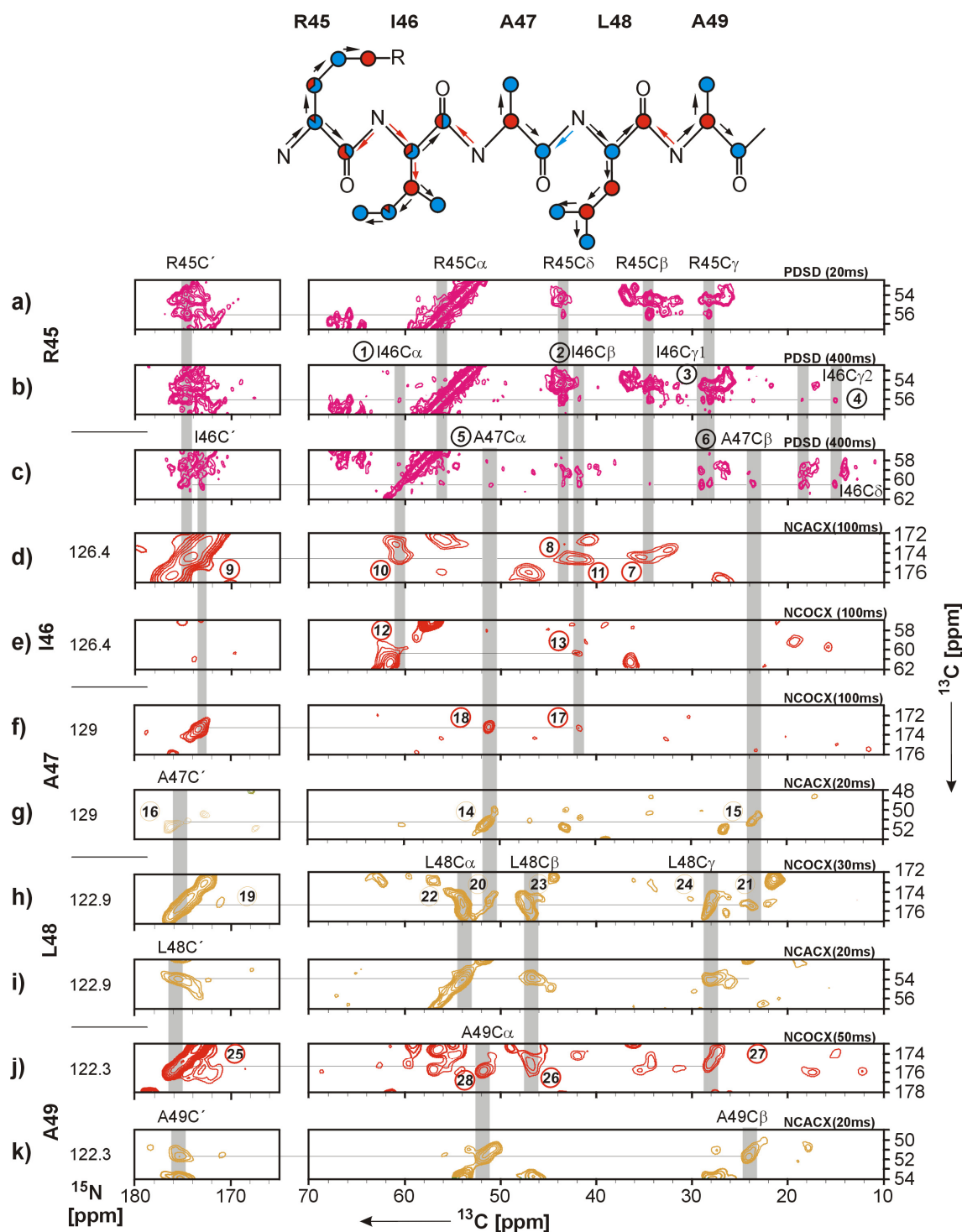
The sequential assignment of these samples was done using the following procedure. In the first step, 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS and 3D NCACX spectra with a short  $^{13}\text{C}$ - $^{13}\text{C}$  mixing time (20 ms) were recorded to identify all spin systems (intra-residue correlations). In the second step, 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS, 3D NCACX experiments recorded with longer mixing times (100-600 ms) and a 3D or 2D NCOCX (recorded with 30 ms  $^{13}\text{C}$ - $^{13}\text{C}$  mixing time) spectra were used for sequential and inter strand assignments. 3D NCACX and NCOCX spectra of  $[2\text{-}^{13}\text{C},^{15}\text{N}]$ - and  $[1,3\text{-}^{13}\text{C},^{15}\text{N}]$ -OmpG samples were used to confirm and extend the assignments. In figure 3.29 the assignment for the R45-I46-A47-L48-A49 motif in OmpG is shown. In the first step we used the 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum of the  $[\text{RIGA},^{15}\text{N}]$ -OmpG sample to identify the complete spin system, including the  $\text{C}\alpha$  (56.01 ppm),  $\text{C}\beta$  (34.5 ppm),  $\text{C}\gamma$  (28.2 ppm),  $\text{C}\delta$  (43.5 ppm) and  $\text{C}'$  (174.6 ppm) of an arginine residue by the typical chemical shift values (Figure 3.29a). In a second step, we use a 2D  $^{13}\text{C}$ - $^{13}\text{C}$  PDS spectrum recorded with 400 ms  $^{13}\text{C}$ - $^{13}\text{C}$  mixing time. This spectrum shows sequential cross peaks to an isoleucine ( $\text{C}\alpha = 60.35$  ppm,  $\text{C}\beta = 41.8$  ppm,  $\text{C}\gamma_1 = 28.9$  ppm) (Figure 3.29b 1-4). By using the  $\text{C}\alpha$  resonance of the isoleucine, sequential cross peaks to the arginine  $\text{C}\alpha$ ,  $\text{C}\beta$ ,  $\text{C}\gamma$ ,  $\text{C}\delta$  and  $\text{C}'$  carbon atoms and to the  $\text{C}\alpha$  (51.03 ppm) and  $\text{C}\beta$  (23.6 ppm) carbon atoms of an alanine were observed (Figure 3.29c 5-6). From this information we can conclude that the observed sequential cross peaks belong to the R-I-A triplet in OmpG at position 45-46-47 (Figure 3.15). However, we have to be careful with this hypothesis, since there is also an I-R-A triplet at position 194-195-196. Therefore, the assignment ideally needs to be confirmed using a NCOCX spectrum, since this is the only spectrum, which provides unambiguous and direction-specific sequential information.

Using the C $\alpha$  and C $\beta$  resonances of the isoleucine residue, it was possible to identify the chemical shift of the backbone nitrogen atom (N = 126.4 ppm) in a 3D NCACX spectrum of the [2-<sup>13</sup>C, <sup>15</sup>N]-OmpG sample (Figure 3.29e 12-13). Since the carbonyl atom of the R45 is 70% labelled in the [2-<sup>13</sup>C, <sup>15</sup>N]-OmpG sample, we used the 3D NCOCX spectrum of this sample to prove the sequential connectivity of R45 and I46. Indeed, there we found the C $\beta$  (34.4 ppm), C $\delta$  (43.4 ppm) and C' (174.4 ppm) cross peaks of R45 (Figure 3.29d 7-9). Additionally, at the same C' resonance two cross peaks at 60.5 ppm and 41.7 ppm, which belong to the C $\alpha$  and C $\beta$  carbon atoms of I46 were found (Figure 3.29d 10-11). Thus, it is clear that the sequential cross peaks observed in the 2D <sup>13</sup>C-<sup>13</sup>C PDSB spectrum (400ms) of the [RIGA, <sup>15</sup>N]-OmpG sample, belong to the R45-I46 pair.

To extend the assignment to A47, the C $\alpha$  (51.03 ppm) and C $\beta$  (23.6 ppm) resonances were used to find the nitrogen plane (N= 129 ppm) in the 3D NCACX spectrum of the [GAVLS, <sup>15</sup>N]-OmpG sample (Figure 3.29g 14-16). Using this nitrogen resonance it was possible to find the sequential cross peak to the C $\beta$  carbon atom of I46 in the 3D NCOCX spectrum of the [2-<sup>13</sup>C, <sup>15</sup>N]-OmpG sample (Figure 3.29e 17-18).

In the next step, we looked for the L48 resonances in the 3D NCOCX of the [GAVLS, <sup>15</sup>N]-OmpG sample. In the nitrogen plane at 122.9 ppm, C', C $\alpha$  and C $\beta$  resonances of A47 were observed (Figure 3.29h 19-21). Additionally, resonances typical for C $\alpha$ , C $\beta$  and C $\gamma$  carbon atoms of leucine residues were found (Figure 3.29h 22-24).

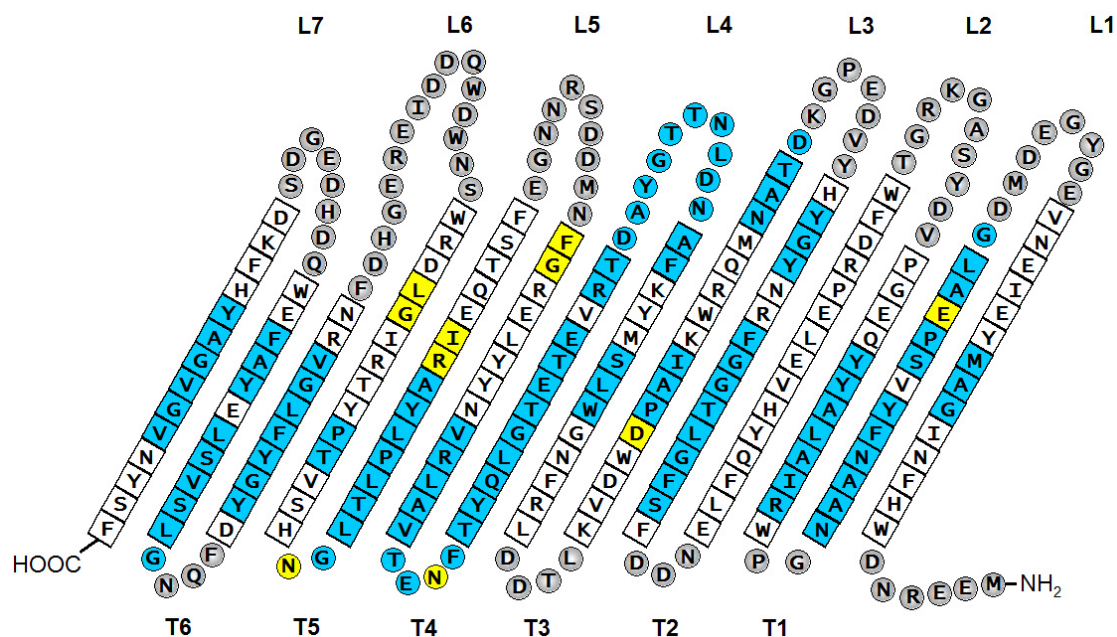
In the last step we looked for a sequential connection to A49 in the 3D NCOCX spectrum of the [2-<sup>13</sup>C, <sup>15</sup>N]-OmpG sample, since leucines are 100 % labelled at the carbonyl carbon atom in this sample. Indeed, we found the C', C $\beta$  and C $\gamma$  resonances of L48 in the nitrogen plane at 122.3 ppm (Figure 3.29i 25-28). In this nitrogen strip, the C $\alpha$  resonance of A49 was found (Figure 3.29i 29). Thus we connect the sequential assignments (Y51-Y50-A49) found by using the GAFY sample (see 3.3.1) with these assignments found here by using the spectra of two differently labelled samples.



**Figure 3.29** Illustration of the assignment of the OmpG R45-I46-A47-L48-A49 motif using spectra of the [RIGA,  $^{15}\text{N}$ ]- (cyan), [GAVLS,  $^{15}\text{N}$ ]- (yellow-brown), [2- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG (red) samples. The assignment is illustrated using strips from 2D  $^{13}\text{C}$ - $^{13}\text{C}$ , 3D NCACX and NCOCX spectra recorded with different  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times. The type of experiment and the  $^{13}\text{C}$ - $^{13}\text{C}$  mixing time is indicated on the top right of each strip. For 3D NCACX and NCOCX spectra the  $^{15}\text{N}$ -frequency is indicated on the left side of the spectra. The vertical dimension is an indirect  $^{13}\text{C}$  dimension containing either  $\text{C}\alpha$  or  $\text{C}'$  frequencies. The horizontal dimension is the direct  $^{13}\text{C}$  dimension containing intra-residue and sequential connectivities, depending on the mixing time and the labelling pattern. Grey lines indicate sequential connectivities (for more details see text above). 2D  $^{13}\text{C}$ - $^{13}\text{C}$  spectra were recorded on a 900 MHz narrow bore spectrometer at 280K and 13 kHz spinning frequency. 3D NCACX and NCOCX spectra were recorded on a 400 MHz wide bore spectrometer at 280 K and with a spinning frequency of 8 kHz.



Using this procedure, and making use of all the specifically labeled OmpG samples, it was possible to assign nearly 50% of the OmpG amino acid sequence (Figure 3.30). The majority of this work was carried out by Dr. Victoria A. Higman. The chemical shifts of the backbone nitrogen and the C $\alpha$ , C $\beta$  and C' atoms are listed in Appendix 7.1. These data were used to calculate the secondary chemical shift, from which the secondary structure elements of a protein can be determined (see section 3.3.3) [159].



**Figure 3.30** Schematic representation of the membrane topology of OmpG extracted from the X-ray structure of OmpG (PDB code 2F1C)[119].  $\beta$ -Strand residues are denoted by squares. Loops (L) and turns (T) are numbered from N to C terminus and indicated by circles. Assigned residues are indicated in blue. Residues whose assignments are still slightly uncertain are indicated in yellow.

### 3.3.4 Secondary structure elements calculated from chemical shift data

In the early 90s it was shown that the C $\alpha$  and C $\beta$  chemical shifts of a given residue correlate closely with its  $\phi$  and  $\psi$  torsion angles [160]. This empirical correlation was found to be useful for the identification of structural elements using only the secondary chemical shifts of C $\alpha$  and C $\beta$  nuclei [161,162]. The secondary chemical of C $\alpha$  or C $\beta$  resonances for each assigned residue was calculated using the following equations:

$$\Delta CA = \delta C_{\alpha} - \delta C_{\alpha(\text{random coil})} \quad [\text{Eq. 5}]$$

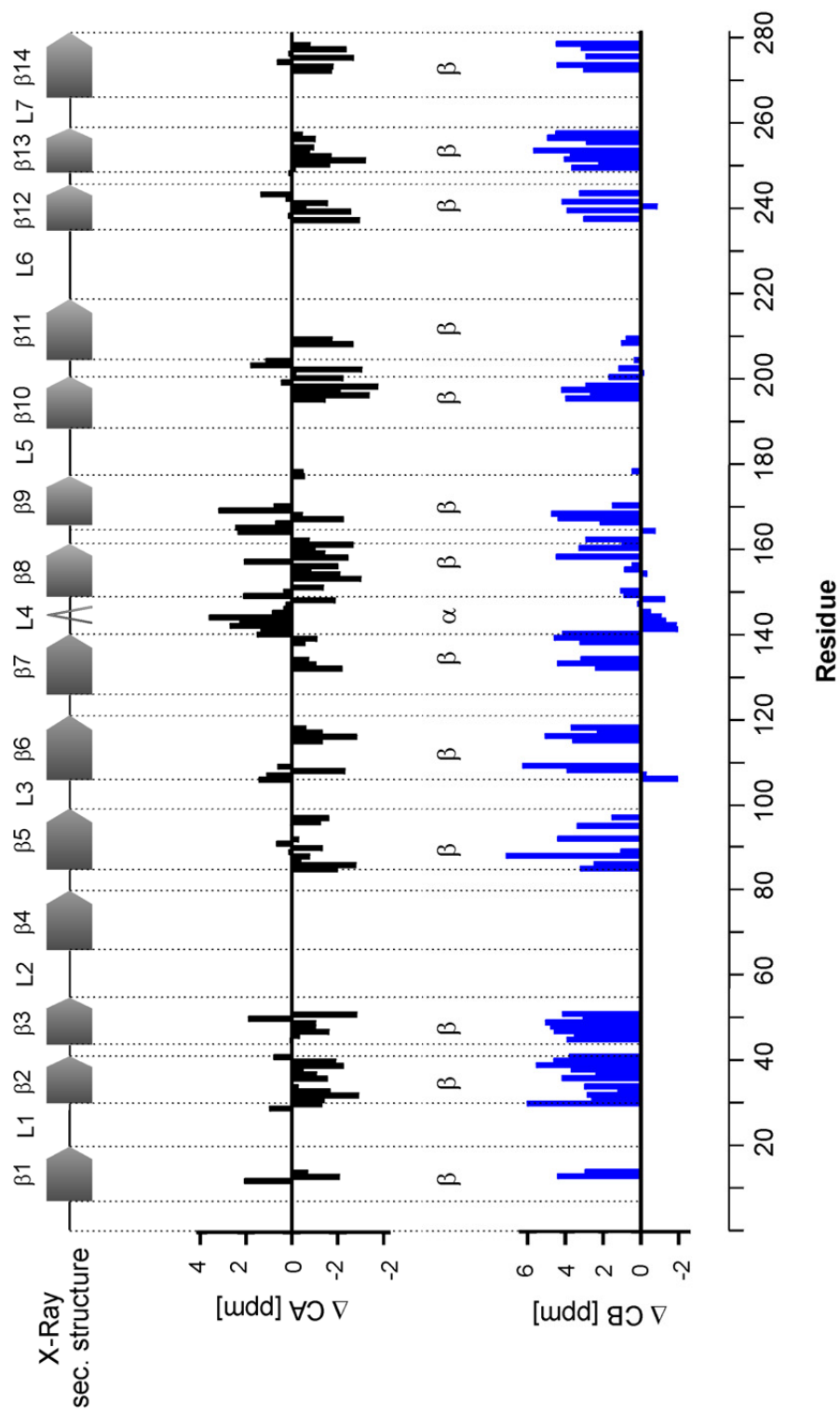
$$\Delta CB = \delta C_{\beta} - \delta C_{\beta(\text{random coil})} \text{ [Eq. 6]}$$

were  $\Delta CA$  and  $\Delta CB$  are the secondary chemical shift,  $\delta(C\alpha)$  and  $\delta(C\beta)$  are the chemical shifts of the protein and  $\delta(C\alpha/\beta \text{ random coil})$  are the averaged chemical shift values of a specific amino acid found in a random coil structure. The  $\delta(C\alpha/\beta \text{ random coil})$  values used were taken from Wang and Jardetzky [136]. Secondary chemical shifts calculated for the R45-I46-A47-L48-A49-Y50-Y51 motif are shown in table 3.4 as an example (Table 3.4). The calculated  $\Delta CA$  values are mostly negative, except for the residues R45 and Y50, whereas the calculated  $\Delta CB$  values are mostly positive. These values are indicative of a  $\beta$ -sheet structure. Indeed, these residues are all located in the  $\beta$ -strand 3 of OmpG (Figure 3.30 and 3.31). In contrast to that,  $\Delta CA$  and  $\Delta CB$  of residues located in an  $\alpha$ -helix would have a positive or negative value, respectively.

*Table 3.4 Secondary chemical shifts calculated for the R45-I46-A47-L48-A49-Y50-Y51 motif*

<b>Residue</b>	<b><math>\delta C\alpha</math> [ppm]</b>	<b><math>\Delta CA</math> [ppm]</b>	<b><math>\delta C\beta</math> [ppm]</b>	<b><math>\Delta CB</math> [ppm]</b>
<b>R45</b>	56,01	0,05	34,4	3,87
<b>I46</b>	60,35	-0,29	41,75	3,49
<b>A47</b>	51,09	-1,58	23,58	4,55
<b>L48</b>	53,85	-1,00	46,62	4,75
<b>A49</b>	51,67	-1,00	24,05	5,02
<b>Y50</b>	59,59	1,87	41,74	3,03
<b>Y51</b>	54,92	-2,8	42,82	4,11

In figure 3.31 all calculated  $\Delta CA$  and  $\Delta CB$  values were plotted as a function of the amino acid number. The calculated structural motifs are mostly  $\beta$ -sheets, typical for a  $\beta$ -barrel protein. Additionally, the positive  $\Delta CA$  and negative  $\Delta CB$  values between strand 7 and 8 indicate a single  $\alpha$ -helix in this region. Most of the calculated motifs are in good agreement with the OmpG structures published so far [49,119,120].



**Figure 3.31** Calculated secondary chemical shifts plotted as a function of the amino acid number in **OmpG**. Top, secondary structure elements of OmpG extracted from the crystal structure (PDB entry 2IWW)[120]. Below,  $\Delta CA$  and  $\Delta CB$  values plotted as a function of the amino acid number. Putative  $\beta$ -strands indicated by a characteristic negative  $\Delta CA$  value and a positive  $\Delta CB$  value. In contrast to that,  $\alpha$ -helical elements are indicated by positive  $\Delta CA$  and negative  $\Delta CB$  values.

## 4 Conclusions and Outlook

This thesis describes the investigation of methods for structure determination of large MPs using solid-state MAS NMR. The potential of the methods employed has been demonstrated by the partial resonance assignment and analysis of the secondary structure of the 281-residue OmpG protein, which was used as a model system.

An essential prerequisite for structural investigations of MPs by NMR in general is the availability of high amounts of a isotopically labelled MP sample, which should be homogenous in terms of purity and conformation. In **chapter 3.1**, the work focused on the preparation of large amounts of isotopically labelled OmpG by using a modified flask expression protocol and a high cell density fermentation (HCDF) procedure. It was shown that the yield of isotopically labelled OmpG could be increased by a factor of up to seven, whereas the costs of isotopically labelled carbon and nitrogen sources could be reduced by a factor of up to five in comparison to a standard flask expression protocol. OmpG was expressed as a non-functional inclusion body precipitate. It was shown that the purification and the subsequent *in vitro* refolding yield high amounts of native OmpG per litre culture. These methods may be helpful for other types of membrane proteins, e.g.  $\alpha$ -helical membrane proteins, which can typically only be expressed with very low yields of correctly folded protein.

Many different preparation types are conceivable for the structural investigation of MPs by solid-state MAS NMR. In **chapter 3.2**, different OmpG preparations ranging from proteoliposomes, 2D crystals and a PEG precipitate were prepared and their quality analysed and compared using EM and 1D and 2D homo- and heteronuclear solid-state MAS NMR experiments. The 2D crystalline and the PEG precipitated samples of OmpG showed highly resolved solid-state MAS NMR spectra with favourable line width, which are indicative for a high degree of conformational homogeneity. However, the cryo-EM investigations showed that a high degree of long-range order, as required for X-ray crystallography, is not present in these samples. A comparison of 2D  $^{13}\text{C}$ - $^{13}\text{C}$  and 2D NCA solid-state MAS NMR spectra of both samples showed no significant chemical shift differences, which indicates that OmpG appears to have the same conformation or folding in both preparations. The proteoliposome

preparation of OmpG, by contrast, did not produce highly resolved NMR spectra. These results illustrate that solid-state MAS NMR investigations of MPs do not require MP preparations with a high degree of long-range order. The most favourable MP preparation types are 2D crystals and PEG-precipitates for the following reasons. Both types can be prepared using well-known crystallisation screens. Additionally, the protein homogeneity and density in such preparation is high, which guarantees high sensitivity in solid-state MAS NMR experiments. Although the suitability of proteoliposome preparations was described for several MPs in the literature, this type of sample also has disadvantages. First, MPs reconstituted into proteoliposomes are often present in an inhomogeneous conformational state, which drastically leads to a line broadening in the spectra. Second, the lipid-to-protein ratio in such preparation is high (50-100 : 1), which decreases the sensitivity of solid-state MAS NMR experiment drastically.

The solid-state MAS NMR studies of OmpG showed that it is necessary to optimise the experimental conditions individually for each preparation type. In particular, temperature is a parameter that can influence the spectral resolution strongly. A high degree of line broadening was observed for OmpG measured in the frozen state (240 K) whereas the spectra recorded at 280 K exhibited very favourable line widths, which enabled further studies to be conducted.

Assignment of OmpG was initially attempted using multi-dimensional correlation spectra of a uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled OmpG sample (U- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-OmpG). However, the NMR spectra are characterised by large amounts of signal overlap. This is due to the fact that OmpG consists of a large number of amino acids with very similar chemical shifts. Furthermore, the resolution is reduced due to scalar couplings and line broadening is observed for many aromatic amino acids due to dynamics processes. This illustrates that uniformly  $^{13}\text{C}$ ,  $^{15}\text{N}$ -labelled samples are not optimal for the assignment of large MPs (> 150 amino acids).

In **chapter 3.3** the development of an assignment strategy using different specifically and biosynthetically site-directed labelled samples is described. Additionally, a heteronuclear methyl-filtered solid-state MAS NMR experiment was used to find further assignment starting points.

An OmpG sample in which only the aromatic amino acids Tyr and Phe are [2,3- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labelled and the amino acids Ala and Gly are U- $^{13}\text{C}$ ,  $^{15}\text{N}$ ]-labelled ([GAFY]-

OmpG) was used in order to find starting points for the assignment of OmpG. Employing 2D  $^{13}\text{C}$ - $^{13}\text{C}$  and 3D NCACX experiments recorded with short  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times it was possible to identify the spin systems of all labelled amino acids. 2D NCOCX and 2D REDOR spectra were used to identify sequential correlations involving Ala and Gly resonances. The [GAFY]-OmpG sample offers two major advantages: firstly it consists only of short isotope labelled sequential sub-sequences, so many inter-residue correlations could be resolved in 3D NCACX and 2D  $^{13}\text{C}$ - $^{13}\text{C}$  spectra recorded with long  $^{13}\text{C}$ - $^{13}\text{C}$  mixing times (300-700 ms). Furthermore, cross peaks of aromatic residues are more intense and new peaks appear in the [GAFY]-OmpG spectra compared to a uniformly labelled sample.

The GAFY labelling scheme has thus provided a highly successful starting point for the assignment of the large membrane protein OmpG. Advantages of the labelling scheme lie in the fact that it is composed of short labelled amino acids (2,3  $^{13}\text{C}$  Tyr and Phe and U- $^{13}\text{C}$  Ala and Gly). Thus the spectral quality is drastically enhanced compared to a uniformly labelled OmpG sample. Additionally, the transfer of magnetisation into the side chain is eliminated which increases the sequential or interstrand magnetisation transfer.

The initial assignments made using GAFY-OmpG were extended using two biosynthetically site-directed labelled OmpG samples. These samples were prepared by using [1,3- $^{13}\text{C}$ ] and [2- $^{13}\text{C}$ ]-labelled glycerol as sole carbon source in the bacterial growth media. The advantage of such samples lies in the fact that each amino acid type has a defined labelling scheme. Additionally, the reduced labelling scheme yields a reduction of scalar couplings, which reduces significantly the line width of the signals [99]. Although such samples alone are not sufficient for a complete unambiguous sequential assignment of large MP, it was shown in this work that in combination with a sample like [GAFY]-OmpG, they were extremely useful for confirming and in some cases extending the sequential assignment of OmpG spectra.

To further extent the sequential assignment, four specifically labelled OmpG samples ([GAVLS, $^{15}\text{N}$ ]-OmpG, [GANDSH, $^{15}\text{N}$ ]-OmpG, [RIGA, $^{15}\text{N}$ ]-OmpG and [STEMPQIND, $^{15}\text{N}$ ]-OmpG) were prepared using both forwards and reverse labelling techniques. Sequential assignments using the spectra of these samples were conducted in a manner analogous to that using spectra of the [GAFY]-OmpG sample. Again, the assignments were confirmed and extended using spectra of the biosynthetically site

directed labelled samples. Using this assignment procedure it was possible to assign nearly 50 % of the amino acids sequence of OmpG from solid-state MAS NMR spectra. Another approach to enhance the resolution and to find starting points for the sequential assignment is the use of methyl filtered pulse programs (section 3.3.2)[157].

The use of the specifically labelled samples, however, demonstrated the need of careful planning to exclude isotope scrambling to unlabelled amino acids. 2D  $^{13}\text{C}$ - $^{13}\text{C}$  spectra show that mainly the catabolic pathways of the amino acids, D, G, N, S and T lead to isotope scrambling.  $^{13}\text{C}$  nuclei from labelled serine were incorporated also into glycine, alanine and tryptophan. Asparagine and aspartic acid lead to labelling of leucine and valine, while isotopically labelled threonine is broken down into glycine and alanine. The knowledge about these isotopic scrambling pathways will be helpful in the future to plan further specifically labelled samples.

In **chapter 3.4**, the chemical shift information extracted from the solid-state MAS NMR data were used to calculate the secondary chemical shift, which provide an insight into the secondary structure of OmpG. The predicted structural motifs are mostly  $\beta$ -sheets, typical for a  $\beta$ -barrel protein. Additionally, between strand 7 and 8 the calculated values indicates a single  $\alpha$ -helical region. Most of the calculated motifs are in good agreement with the published OmpG structures [49,119,120].

As a general summary, the work described in this thesis shows the development of new expression, labelling and sample preparation protocols which enable the extension of solid-state MAS NMR structure determination methods, initially developed using small molecules, to large MPs.

All used assignment procedures presented here rely on correlations between carbons and nitrogens. The detection of protons, which is commonly used in solution NMR, is usually less favourable in solid-state MAS NMR as large dipolar couplings between the abundant  $^1\text{H}$  lead to substantial line broadening. However, it was shown that this disadvantage can be circumvented by a reduction of proton density in the sample [164]. Therefore, an effective proton detected spectrum was achieved by re-crystallisation of a perdeuterated protein in a buffer containing 10%  $\text{H}_2\text{O}$  and 90 %  $\text{D}_2\text{O}$ . In combination with fast MAS frequencies, it was demonstrated that  $^1\text{H}$ - $^{15}\text{N}$  detected spectra shows proton line width of around 20 Hz. Additionally, this labelling

scheme opens up new ways for resonance assignment, as was demonstrated by Linser *et al.*[165]. Modified solution-state pulse sequences like HNCO, HNCA, and HNCACB were successfully applied to microcrystalline SH3 samples in the solid-state. The inclusion of the  $^1\text{H}$  chemical shift in the assignment process yields a better chemical shift dispersion of the amide moiety and decreases the number of ambiguously assigned residues. It would be interesting to see in the future, whether those pulse sequences are applicable for large MP, such as OmpG.



## 5 References

1. Alberts,B., Brey,D. & Lewis,J. Molecular Biology of the Cell. Garland Science, New York (2002).
2. Hubner,C.A. & Jentsch,T.J. Ion channel diseases. *Hum. Mol. Genet.* **11**, 2435-2445 (2002).
3. Schoneberg,T. *et al.* Mutant G-protein-coupled receptors as a cause of human diseases. *Pharmacol. Ther.* **104**, 173-206 (2004).
4. White,S. Membrane Protein Resources. [http://blanco.biomol.uci.edu/Membrane\\_Proteins\\_xtal.html](http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html) (1998).
5. Berman,H.M. *et al.* The Protein Data Bank. *Nucleic Acids Research* **28**, 235-242 (2000).
6. Deisenhofer,J., Epp,O., Miki,K., Huber,R. & Michel,H. Structure of the Protein Subunits in the Photosynthetic Reaction Center of *Rhodospseudomonas viridis* at 3Å Resolution. *Nature* **318**, 618-624 (1985).
7. Palczewski,K. *et al.* Crystal structure of rhodopsin: a G-protein-coupled receptor. *Science* **289**, 739-745 (2000).
8. Toyoshima,C. & Inesi,G. Structural basis of ion pumping by Ca<sup>2+</sup>-ATPase of the sarcoplasmic reticulum. *Annu. Rev. Biochem.* **73**, 269-292 (2004).
9. Kuhlbrandt,W., Wang,D.N. & Fujiyoshi,Y. Atomic Model of Plant Light-Harvesting Complex by Electron Crystallography. *Nature* **367**, 614-621 (1994).
10. Unwin,N. Structure and action of the nicotinic acetylcholine receptor explored by electron microscopy. *FEBS Lett.* **555**, 91-95 (2003).
11. Luirink,J., von Heijne,G., Houben,E. & de Gier,J.W. Biogenesis of inner membrane proteins in Escherichia coli. *Annu. Rev. Microbiol.* **59**, 329-355 (2005).
12. Wickner,W. & Schekman,R. Protein translocation across biological membranes. *Science* **310**, 1452-1456 (2005).
13. Miroux,B. & Walker,J.E. Over-production of proteins in Escherichia coli: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of Molecular Biology* **260**, 289-298 (1996).
14. Reinhart,C. & Krettler,C. Structural Genomics on Membrane Proteins. Lundstrom,K. (ed.), pp. 115-142 (CRC Press, Boca Raton,2006).
15. Bosman,G.J. & deGrip,W.J. Structural Genomics on Membrane Proteins. Lundstrom,K. (ed.), pp. 143-167 (CRC Press, Boca Raton,2006).
16. Lundstrom,K. Structural Genomics on Membrane Proteins., pp. 169-178 (CRS Press, Boca Raton,2006).

17. Klammt,C., Schwarz,D., Dotsch,V. & Bernhard,F. Cell-free production of integral membrane proteins on a preparative scale. *Methods Mol. Biol.* **375**, 57-78 (2007).
18. Grisshammer,R. Understanding recombinant expression of membrane proteins. *Current Opinion in Biotechnology* **17**, 337-340 (2006).
19. Wagner,S., Lerch Bader,M., Drew,D. & de Gier,J.W. Rationalizing membrane protein overexpression. *Trends in Biotechnology* **24**, 364-371 (2006).
20. Lian,L.Y. & Middleton,D.A. Labelling approaches for protein structural studies by solution-state and solid-state NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **39**, 171-190 (2001).
21. Bannwarth,M. & Schulz,G.E. The expression of outer membrane proteins for crystallization. *Biochimica et Biophysica Acta-Biomembranes* **1610**, 37-45 (2003).
22. Kiefer,H. In vitro folding of alpha-helical membrane proteins. *Biochim. Biophys. Acta* **1610**, 57-62 (2003).
23. Diehl,A., Holz,C., Batchelder,L.S. & Oschkinat,H. Hochdurchsatzsystem zum Screenen von rekombinanten Hefeklonen und zur Medienoptimierung . *BIOSpektrum* **06**, 787 (2005).
24. Kenneth,C. Effective uniform isotope labeling of proteins expressed in Baculovirus-infected insect cells using BioExpress® 2000 (Insect cell) medium. [http://www.isotope.com/uploads/File/App\\_Note\\_20.pdf](http://www.isotope.com/uploads/File/App_Note_20.pdf) . 2009.
25. Klammt,C. *et al.* Cell-free production of G protein-coupled receptors for functional and structural studies. *J. Struct. Biol.* **158**, 482-493 (2007).
26. le Maire,M., Champeil,P. & Moller,J.V. Interaction of membrane proteins and lipids with solubilizing detergents. *Biochim. Biophys. Acta* **1508**, 86-111 (2000).
27. Prive,G.G. Detergents for the stabilization and crystallization of membrane proteins. *Methods* **41**, 388-397 (2007).
28. Nelson,N. & Yocum,C.F. Structure and function of photosystems I and II. *Annu. Rev. Plant Biol.* **57**, 521-565 (2006).
29. Tombola,F., Pathak,M.M. & Isacoff,E.Y. How does voltage open an ion channel? *Annu. Rev. Cell Dev. Biol.* **22**, 23-52 (2006).
30. Ostermeier,C. & Michel,H. Crystallization of membrane proteins. *Curr. Opin. Struct. Biol.* **7**, 697-701 (1997).
31. Cherezov,V. *et al.* High-Resolution Crystal Structure of an Engineered Human  $\beta$ 2-Adrenergic G-Protein Coupled Receptor. *Science* **318**, 1258-1265 (2007).
32. Rosenbaum,D.M. *et al.* GPCR Engineering Yields High-Resolution Structural Insights into  $\beta$ 2-Adrenergic Receptor Function. *Science* **318**, 1266-1273 (2007).
33. Zhang,H., Kurisu,G., Smith,J.L. & Cramer,W.A. A defined protein-detergent-lipid complex for crystallization of integral membrane proteins: The cytochrome b6 complex of oxygenic photosynthesis. *Proc. Natl. Acad. Sci. U. S. A* **100**, 5160-5163 (2003).

34. Hunte,C. & Michel,H. Crystallisation of membrane proteins mediated by antibody fragments. *Curr. Opin. Struct. Biol.* **12**, 503-508 (2002).
35. Landau,E.M. & Rosenbusch,J.P. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc. Natl. Acad. Sci. U. S. A* **93**, 14532-14535 (1996).
36. Faham,S. & Bowie,J.U. Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J. Mol. Biol.* **316**, 1-6 (2002).
37. Kuhlbrandt,W. 2-Dimensional Crystallization of Membrane-Proteins. *Quarterly Reviews of Biophysics* **25**, 1-49 (1992).
38. Murata,K. *et al.* Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599-605 (2000).
39. Tate,C.G. Comparison of three structures of the multidrug transporter EmrE. *Current Opinion in Structural Biology* **16**, 457-464 (2006).
40. Chen,Y.J. *et al.* X-ray structure of EmrE supports dual topology model. *Proc. Natl. Acad. Sci. U. S. A* **104**, 18999-19004 (2007).
41. Kunji,E.R.S. & Harding,M. Projection structure of the atractyloside-inhibited mitochondrial ADP/ATP carrier of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **278**, 36985-36988 (2003).
42. Pebay-Peyroula,E. *et al.* Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* **426**, 39-44 (2003).
43. Unwin,N. Acetylcholine-Receptor Channel Imaged in the Open State. *Nature* **373**, 37-43 (1995).
44. Hasler,L., Heymann,J.B., Engel,A., Kistler,J. & Walz,T. 2D crystallization of membrane proteins: Rationales and examples. *Journal of Structural Biology* **121**, 162-171 (1998).
45. Rigaud,J., Chami,M., Lambert,O., Levy,D. & Ranck,J. Use of detergents in two-dimensional crystallization of membrane proteins. *Biochim. Biophys. Acta* **1508**, 112-128 (2000).
46. Wüthrich,K. *NMR of Proteins and Nuclein Acids*. Wiley, New York (1986).
47. Tamm,L.K. & Liang,B.Y. NMR of membrane proteins in solution. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 201-210 (2006).
48. Poget,S.F. & Girvin,M.E. Solution NMR of membrane proteins in bilayer mimics: small is beautiful, but sometimes bigger is better. *Biochim. Biophys. Acta* **1768**, 3098-3106 (2007).
49. Liang,B. & Tamm,L.K. Structure of outer membrane protein G by solution NMR spectroscopy. *Proceedings of the National Academy of Sciences* **104**, 16140-16145 (2007).
50. Hwang,P.M. & Kay,L.E. Solution structure and dynamics of integral membrane proteins by NMR: A case study involving the enzyme PagP. *Nuclear Magnetic Resonance of Biological Macromolecules, Part C* **394**, 335-+ (2005).

51. Hiller, S. *et al.* Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. *Science* **321**, 1206-1210 (2008).
52. Roosild, T.P. *et al.* NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science* **307**, 1317-1321 (2005).
53. Oxenoid, K. & Chou, J.J. The structure of phospholamban pentamer reveals a channel-like architecture in membranes. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10870-10875 (2005).
54. Wu, C.H., Ramamoorthy, A. & Opella, S.J. High resolution heteronuclear dipolar solid-state NMR spectroscopy. *J. Magn. Reson. A* **109**, 270-272 (1994).
55. Opella, S.J. & Marassi, F.M. Structure determination of membrane proteins by NMR spectroscopy. *Chemical Reviews* **104**, 3587-3606 (2004).
56. Nevzorov, A.A., Mesleh, M.F. & Opella, S.J. Structure determination of aligned samples of membrane proteins by NMR spectroscopy. *Magn Reson. Chem.* **42**, 162-171 (2004).
57. Sanders, C.R. & Landis, G.C. Reconstitution of membrane proteins into lipid-rich bilayered mixed micelles for NMR studies. *Biochemistry* **34**, 4030-4040 (1995).
58. Lange, A. & Baldus, M. Structural Biology of Membrane Proteins. Grishammer, R. & Buchanan, S.K. (eds.), pp. 118-130 (RCM, Cambridge UK, 2006).
59. McDermott, A. *et al.* Partial NMR assignments for uniformly (<sup>13</sup>C, <sup>15</sup>N)-enriched BPTI in the solid state. *J. Biomol. NMR* **16**, 209-219 (2000).
60. Pauli, J., van Rossum, B., Forster, H., de Groot, H.J.M. & Oschkinat, H. Sample optimization and identification of signal patterns of amino acid side chains in 2D RFDR spectra of the alpha-spectrin SH3 domain. *J. Magn. Reson.* **143**, 411-416 (2000).
61. Pauli, J., Baldus, M., van Rossum, B., de Groot, H. & Oschkinat, H. Backbone and side-chain C-13 and N-15 signal assignments of the alpha-spectrin SH3 domain by magic angle spinning solid-state NMR at 17.6 tesla. *Chembiochem* **2**, 272-281 (2001).
62. Bockmann, A. *et al.* Solid state NMR sequential resonance assignments and conformational analysis of the 2 x 10.4 kDa dimeric form of the Bacillus subtilis protein Crh. *J. Biomol. NMR* **27**, 323-339 (2003).
63. Igumenova, T.I., Wand, A.J. & McDermott, A.E. Assignment of the backbone resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* **126**, 5323-5331 (2004).
64. Igumenova, T.I. *et al.* Assignments of carbon NMR resonances for microcrystalline ubiquitin. *J. Am. Chem. Soc.* **126**, 6720-6727 (2004).
65. Marulanda, D. *et al.* Magic angle spinning solid-state NMR spectroscopy for structural studies of protein interfaces. resonance assignments of differentially enriched Escherichia coli thioredoxin reassembled by fragment complementation. *J. Am. Chem. Soc.* **126**, 16608-16620 (2004).

66. Lange,A. *et al.* A concept for rapid protein-structure determination by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed Engl.* **44**, 2089-2092 (2005).
67. Franks,W.T. *et al.* Magic-angle spinning solid-state NMR spectroscopy of the beta1 immunoglobulin binding domain of protein G (GB1): <sup>15</sup>N and <sup>13</sup>C chemical shift assignments and conformational analysis. *J. Am. Chem. Soc.* **127**, 12291-12305 (2005).
68. Petkova,A.T., Yau,W.M. & Tycko,R. Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry* **45**, 498-512 (2006).
69. Becker,J. *et al.* A sequential assignment procedure for proteins that have intermediate line widths in MAS NMR spectra: amyloid fibrils of human CA150.WW2. *Chembiochem.* **9**, 1946-1952 (2008).
70. Siemer,A.B. *et al.* <sup>13</sup>C, <sup>15</sup>N resonance assignment of parts of the HET-s prion protein in its amyloid form. *J. Biomol. NMR* **34**, 75-87 (2006).
71. Lange,A. *et al.* Toxin-induced conformational changes in a potassium channel revealed by solid-state NMR. *Nature* **440**, 959-962 (2006).
72. Krabben,L. *et al.* Towards structure determination of neurotoxin II bound to nicotinic acetylcholine receptor: a solid-state NMR approach. *FEBS Lett.* **564**, 319-324 (2004).
73. Martin,R.W. & Zilm,K.W. Preparation of protein nanocrystals and their characterization by solid state NMR. *J. Magn. Reson.* **165**, 162-174 (2003).
74. Petkova,A.T. *et al.* A structural model for Alzheimer's beta -amyloid fibrils based on experimental constraints from solid state NMR. *Proc. Natl. Acad. Sci. U. S. A* **99**, 16742-16747 (2002).
75. Ferguson,N. *et al.* General structural motifs of amyloid protofilaments. *Proc. Natl. Acad. Sci. U. S. A* **103**, 16248-16253 (2006).
76. Heise,H. *et al.* Molecular-level secondary structure, polymorphism, and dynamics of full-length alpha-synuclein fibrils studied by solid-state NMR. *Proc. Natl. Acad. Sci. U. S. A* **102**, 15871-15876 (2005).
77. Tycko,R. Molecular structure of amyloid fibrils: insights from solid-state NMR. *Q. Rev. Biophys.* **39**, 1-55 (2006).
78. deGrip,W.J. *et al.* Photoactivation of rhodopsin: interplay between protein and chromophore. *Novartis. Found. Symp.* **224**, 102-118 (1999).
79. Herzfeld,J. & Lansing,J.C. Magnetic resonance studies of the bacteriorhodopsin pump cycle. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 73-95 (2002).
80. Egorova-Zachernyuk,T.A. *et al.* Heteronuclear 2D-correlations in a uniformly [<sup>13</sup>C, <sup>15</sup>N] labeled membrane-protein complex at ultra-high magnetic fields. *J. Biomol. NMR* **19**, 243-253 (2001).
81. van Gammeren,A.J., Hulsbergen,F.B., Hollander,J.G. & de Groot,H.J.M. Biosynthetic site-specific C-13 labeling of the light-harvesting 2 protein complex: A model for solid state NMR structure determination of transmembrane proteins. *J. Biomol. NMR* **30**, 267-274 (2004).

82. Bloch, F. Theory of Line Narrowing by Double-Frequency Irradiation. *Physical Review* **111**, 841-853 (1958).
83. Bennett, A.E., Rienstra, C.M., Auger, M., Lakshmi, K.V. & Griffin, R.G. Heteronuclear decoupling in rotating solids. *J. Chem. Phys.* **103**, 6951-6958 (1995).
84. Fung, B.M., Khitrin, A.K. & Ermolaev, K. An improved broadband decoupling sequence for liquid crystals and solids. *J. Magn. Reson.* **142**, 97-101 (2000).
85. Laws, D.D., Bitter, H.M. & Jerschow, A. Solid-state NMR spectroscopic methods in chemistry. *Angew. Chem. Int. Ed Engl.* **41**, 3096-3129 (2002).
86. Pines, A., Waugh, J.S. & Gibby, M.G. Proton-Enhanced Nuclear Induction Spectroscopy - Method for High-Resolution Nmr of Dilute Spins in Solids. *J. Chem. Phys.* **56**, 1776-& (1972).
87. Schaefer, J. & Stejskal, E.O. C-13 Nuclear Magnetic-Resonance of Polymers Spinning at Magic Angle. *J. Am. Chem. Soc.* **98**, 1031-1032 (1976).
88. Metz, G., Wu, X.L. & Smith, S.O. Ramped-Amplitude Cross-Polarization in Magic-Angle-Spinning Nmr. *J. Magn. Reson. A* **110**, 219-227 (1994).
89. Szeverenyi, N.M., Sullivan, M.J. & Maciel, G.E. Observation of Spin Exchange by Two-Dimensional Fourier-Transform C-13 Cross Polarization-Magic-Angle Spinning. *J. Magn. Reson.* **47**, 462-475 (1982).
90. Takegoshi, K., Nakamura, S. & Terao, T. C-13-H-1 dipolar-assisted rotational resonance in magic-angle spinning NMR. *Chem. Phys. Lett.* **344**, 631-637 (2001).
91. Baldus, M., Petkova, A.T., Herzfeld, J. & Griffin, R.G. Cross polarization in the tilted frame: assignment and spectral simplification in heteronuclear spin systems. *Mol. Phys.* **95**, 1197-1207 (1998).
92. Baldus, M., Geurts, D.G., Hediger, S. & Meier, B.H. Efficient N-15-C-13 polarization transfer by adiabatic-passage Hartmann-Hahn cross polarization. *J. Magn. Reson. A* **118**, 140-144 (1996).
93. Verel, R., Ernst, M. & Meier, B.H. Adiabatic dipolar recoupling in solid-state NMR: The DREAM scheme. *J. Magn. Reson.* **150**, 81-99 (2001).
94. Franks, W.T., Kloepper, K.D., Wylie, B.J. & Rienstra, C.M. Four-dimensional heteronuclear correlation experiments for chemical shift assignment of solid proteins. *J. Biomol. NMR* **39**, 107-131 (2007).
95. Bennett, A.E., Ok, J.H., Griffin, R.G. & Vega, S. Chemical-Shift Correlation Spectroscopy in Rotating Solids - Radio Frequency-Driven Dipolar Recoupling and Longitudinal Exchange. *J. Chem. Phys.* **96**, 8624-8627 (1992).
96. Gullion, T. & Schaefer, J. Rotational-Echo Double-Resonance Nmr. *J. Magn. Reson.* **81**, 196-200 (1989).
97. Hing, A.W., Vega, S. & Schaefer, J. Transferred-Echo Double-Resonance Nmr. *J. Magn. Reson.* **96**, 205-209 (1992).

98. Ulrich, E.L. *et al.* BioMagResBank. *Nucleic Acids Res.* **36**, D402-D408 (2008).
99. Castellani, F., van Rossum, B.J., Diehl, A., Rehbein, K. & Oschkinat, H. Determination of solid-state NMR structures of proteins by means of three-dimensional N-15-C-13-C-13 dipolar correlation spectroscopy and chemical shift analysis. *Biochemistry* **42**, 11476-11483 (2003).
100. Lange, A., Luca, S. & Baldus, M. Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J. Am. Chem. Soc.* **124**, 9704-9705 (2002).
101. Hong, M. & Griffin, R.G. Resonance assignments for solid peptides by dipolar-mediated C-13/N-15 correlation solid-state NMR. *J. Am. Chem. Soc.* **120**, 7113-7114 (1998).
102. Cornilescu, G., Delaglio, F. & Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **13**, 289-302 (1999).
103. Fossi, M., Castellani, F., Nilges, M., Oschkinat, H. & van Rossum, B.J. SOLARIA: a protocol for automated cross-peak assignment and structure calculation for solid-state magic-angle spinning NMR spectroscopy. *Angew. Chem. Int. Ed Engl.* **44**, 6151-6154 (2005).
104. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiology and Molecular Biology Reviews* **67**, 593+ (2003).
105. Pautsch, A. & Schulz, G.E. High-resolution structure of the OmpA membrane domain. *J. Mol. Biol.* **298**, 273-282 (2000).
106. Vogt, J. & Schulz, G.E. The structure of the outer membrane protein OmpX from *Escherichia coli* reveals possible mechanisms of virulence. *Structure.* **7**, 1301-1309 (1999).
107. Cavard, D. & Lazdunski, C. Colicin cleavage by OmpT protease during both entry into and release from *Escherichia coli* cells. *J. Bacteriol.* **172**, 648-652 (1990).
108. Blaauw, M., Dekker, N., Verheij, H.M., Kalk, K.H. & Dijkstra, B.W. Crystallization and preliminary X-ray analysis of outer membrane phospholipase A from *Escherichia coli*. *FEBS Lett.* **373**, 10-12 (1995).
109. Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K. & Welte, W. Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. *Science* **282**, 2215-2220 (1998).
110. Kurisu, G. *et al.* The structure of BtuB with bound colicin E3 R-domain implies a translocon. *Nat. Struct. Biol.* **10**, 948-954 (2003).
111. Soll, J. & Schleiff, E. Protein import into chloroplasts. *Nat. Rev. Mol. Cell Biol.* **5**, 198-208 (2004).
112. Wiedemann, N., Frazier, A.E. & Pfanner, N. The protein import machinery of mitochondria. *J. Biol. Chem.* **279**, 14473-14476 (2004).

113. Szmelcman,S., Schwartz,M., Silhavy,T.J. & Boos,W. Maltose transport in Escherichia coli K12. A comparison of transport kinetics in wild-type and lambda-resistant mutants as measured by fluorescence quenching. *Eur. J. Biochem.* **65**, 13-19 (1976).
114. Schmid,K., Ebner,R., Jahreis,K., Lengeler,J.W. & Titgemeyer,F. A sugar-specific porin, ScrY, is involved in sucrose uptake in enteric bacteria. *Mol. Microbiol.* **5**, 941-950 (1991).
115. Misra,R. & Benson,S.A. A Novel Mutation, Cog, Which Results in Production of A New Porin Protein (Ompg) of Escherichia-Coli K-12. *Journal of Bacteriology* **171**, 4105-4111 (1989).
116. Fajardo,D.A. *et al.* Biochemistry and regulation of a novel Escherichia coli K-12 porin protein, OmpG, which produces unusually large channels. *Journal of Bacteriology* **180**, 4452-4459 (1998).
117. Behlau,M., Mills,D.J., Quader,H., Kuhlbrandt,W. & Vonck,J. Projection structure of the monomeric porin OmpG at 6 angstrom resolution. *Journal of Molecular Biology* **305**, 71-77 (2001).
118. Conlan,S., Zhang,Y., Cheley,S. & Bayley,H. Biochemical and biophysical characterization of OmpG: A monomeric porin. *Biochemistry* **39**, 11845-11854 (2000).
119. Subbarao,G.V. & van den Berg,B. Crystal structure of the monomeric porin OmpG. *Journal of Molecular Biology* **360**, 750-759 (2006).
120. Yildiz,O., Vinothkumar,K.R., Goswami,P. & Kuhlbrandt,W. Structure of the monomeric outer-membrane porin OmpG in the open and closed conformation. *Embo Journal* **25**, 3702-3713 (2006).
121. Bayley,H. & Cremer,P.S. Stochastic sensors inspired by biology. *Nature* **413**, 226-230 (2001).
122. Rhee,M. & Burns,M.A. Nanopore sequencing technology: nanopore preparations. *Trends Biotechnol.* **25**, 174-181 (2007).
123. Song,L. & Gouaux,E. Crystallization of the alpha-hemolysin heptamer solubilized in decyldimethyl- and decyldiethylphosphine oxide. *Acta Crystallogr. D. Biol. Crystallogr.* **54**, 276-278 (1998).
124. Gu,L.Q., Braha,O., Conlan,S., Cheley,S. & Bayley,H. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature* **398**, 686-690 (1999).
125. Chen,M., Khalid,S., Sansom,M.S. & Bayley,H. Outer membrane protein G: Engineering a quiet pore for biosensing. *Proc. Natl. Acad. Sci. U. S. A* **105**, 6272-6277 (2008).
126. Sambrook,J. & Russell,D. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, U.S, (2001).
127. Marley,J., Lu,M. & Bracken,C. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20**, 71-75 (2001).



128. Fiedler,S., Knocke,C., Vogt,J., Oschkinat,H. & Diehl,A. HCDF as a protein-labeling methodology - Production of H-2-, C-13-, and N-15-labeled OmpG via high cell density fermentation. *Genetic Engineering & Biotechnology News* **27**, 54 (2007).
129. Gill,S.C. & Vonhippel,P.H. Calculation of Protein Extinction Coefficients from Amino-Acid Sequence Data. *Analytical Biochemistry* **182**, 319-326 (1989).
130. Jaroniec,C.P., Filip,C. & Griffin,R.G. 3D TEDOR NMR experiments for the simultaneous measurement of multiple carbon-nitrogen distances in uniformly C-13, N-15-labeled solids. *J. Am. Chem. Soc.* **124**, 10728-10742 (2002).
131. Vranken,W.F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins-Structure Function and Bioinformatics* **59**, 687-696 (2005).
132. Studier,F.W., Rosenberg,A.H., Dunn,J.J. & Dubendorff,J.W. Use of T7 Rna-Polymerase to Direct Expression of Cloned Genes. *Methods in Enzymology* **185**, 60-89 (1990).
133. Kleinschmidt,J.H., Wiener,M.C. & Tamm,L.K. Outer membrane protein A of E-coli folds into detergent micelles, but not in the presence of monomeric detergent. *Protein Science* **8**, 2065-2071 (1999).
134. Beher,M.G., Schnaitman,C.A. & Pugsley,A.P. Major Heat-Modifiable Outer-Membrane Protein in Gram-Negative Bacteria - Comparison with the Ompa Protein of Escherichia-Coli. *Journal of Bacteriology* **143**, 906-913 (1980).
135. Arakawa,T. *et al.* Suppression of protein interactions by arginine: A proposed mechanism of the arginine effects. *Biophysical Chemistry* **127**, 1-8 (2007).
136. Wang,Y. & Jardetzky,O. Probability-based protein secondary structure identification using combined NMR chemical-shift data. *Protein Sci.* **11**, 852-861 (2002).
137. Lange,V. NMR-spektroskopische Untersuchungen an dem ABC-Transporter ArtMP-J aus *G. stearothermophilus*. 22-5-2008. Thesis/Dissertation
138. Shastri,S. *et al.* Proteorhodopsin: characterisation of 2D crystals by electron microscopy and solid state NMR. *Biochim. Biophys. Acta* **1768**, 3012-3019 (2007).
139. Agarwal,V., Fink,U., Schuldiner,S. & Reif,B. MAS solid-state NMR studies on the multidrug transporter EmrE. *Biochim. Biophys. Acta* **1768**, 3036-3043 (2007).
140. Pflieger,N., Lorch,M., Woerner,A.C., Shastri,S. & Glaubitz,C. Characterisation of Schiff base and chromophore in green proteorhodopsin by solid-state NMR. *J. Biomol. NMR* **40**, 15-21 (2008).
141. Ubarretxena-Belandia,I., Baldwin,J.M., Schuldiner,S. & Tate,C.G. Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *EMBO J.* **22**, 6175-6181 (2003).
142. Lorch,M. *et al.* How to prepare membrane proteins for solid-state NMR: A case study on the alpha-helical integral membrane protein diacylglycerol kinase from *E. coli*. *Chembiochem.* **6**, 1693-1700 (2005).

143. Etzkorn, M. *et al.* Secondary structure, dynamics, and topology of a seven-helix receptor in native membranes, studied by solid-state NMR spectroscopy. *Angew. Chem. Int. Ed Engl.* **46**, 459-462 (2007).
144. Andronesi, O.C. *et al.* Probing membrane protein orientation and structure using fast magic-angle-spinning solid-state NMR. *J. Biomol. NMR* **30**, 253-265 (2004).
145. Schneider, R. *et al.* Solid-state NMR spectroscopy applied to a chimeric potassium channel in lipid bilayers. *J. Am. Chem. Soc.* **130**, 7427-7435 (2008).
146. Frericks, H.L., Zhou, D.H., Yap, L.L., Gennis, R.B. & Rienstra, C.M. Magic-angle spinning solid-state NMR of a 144 kDa membrane protein complex: E. coli cytochrome bo<sub>3</sub> oxidase. *J. Biomol. NMR* **36**, 55-71 (2006).
147. Li, Y., Berthold, D.A., Frericks, H.L., Gennis, R.B. & Rienstra, C.M. Partial (13)C and (15)N chemical-shift assignments of the disulfide-bond-forming enzyme DsbB by 3D magic-angle spinning NMR spectroscopy. *ChemBiochem.* **8**, 434-442 (2007).
148. Varga, K., Tian, L. & McDermott, A.E. Solid-state NMR study and assignments of the KcsA potassium ion channel of *S. lividans*. *Biochim. Biophys. Acta* **1774**, 1604-1613 (2007).
149. Vink, M., Derr, K., Love, J., Stokes, D.L. & Ubarretxena-Belandia, I. A high-throughput strategy to screen 2D crystallization trials of membrane proteins. *J. Struct. Biol.* **160**, 295-304 (2007).
150. Castellani, F. *et al.* Structure of a protein determined by solid-state magic-angle-spinning NMR spectroscopy. *Nature* **420**, 98-102 (2002).
151. LeMaster, D.M. & Kushlan, D.M. Dynamical mapping of E-coli thioredoxin via C-13 NMR relaxation analysis. *J. Am. Chem. Soc.* **118**, 9255-9264 (1996).
152. Michal, G. *Biochemical Pathways*. Spektrum Akademischer Verlag, Heidelberg, Berlin (1999).
153. Zech, S.G., Olejniczak, E., Hajduk, P., Mack, J. & McDermott, A.E. Characterization of protein-ligand interactions by high-resolution solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **126**, 13948-13953 (2004).
154. Opella, S.J., Frey, M.H. & Cross, T.A. Detection of individual carbon resonances in solid proteins. *J. Am. Chem. Soc.* **101**, 5856-5857 (1979).
155. Burns, S.T., Wu, X. & Zilm, K.W. Improvement of Spectral Editing in Solids: A Sequence for Obtaining <sup>13</sup>CH + <sup>13</sup>CH<sub>2</sub>-Only <sup>13</sup>C Spectra. *J. Magn. Reson.* **143**, 352-359 (2000).
156. Sangill, R., Rastrupandersen, N., Bildsoe, H., Jakobsen, H.J. & Nielsen, N.C. Optimized Spectral Editing of <sup>13</sup>C MAS NMR Spectra of Rigid Solids Using Cross-Polarization Methods. *J. Magn Reson.* **107**, 67-78 (1994).
157. Jehle, S. *et al.* Spectral editing: selection of methyl groups in multidimensional solid-state magic-angle spinning NMR. *J. Biomol. NMR* **36**, 169-177 (2006).
158. Neidhardt, F. *Escherichia coli* and *Salmonella*. Cellular and Molecular Biology. ASM Press, Washington, D.C. (1996).

159. Wishart,D.S. & Sykes,B.D. Chemical shifts as a tool for structure determination. *Methods Enzymol.* **239**, 363-392 (1994).
160. Spera,S. & Bax,A. Empirical correlation between protein backbone conformation and C.alpha. and C.beta. <sup>13</sup>C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**, 5490-5492 (1991).
161. Wishart,D.S., Sykes,B.D. & Richards,F.M. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222**, 311-333 (1991).
162. Gronenborn,A.M. & Clore,G.M. Identification of N-terminal helix capping boxes by means of <sup>13</sup>C chemical shifts. *J. Biomol. NMR* **4**, 455-458 (1994).
163. Wagner,S. *et al.* Consequences of membrane protein overexpression in Escherichia coli. *Mol. Cell Proteomics.* **6**, 1527-1550 (2007).
164. Chevelkov,V., Rehbein,K., Diehl,A. & Reif,B. Ultrahigh resolution in proton solid-state NMR spectroscopy at high levels of deuteration. *Angew. Chem. Int. Ed Engl.* **45**, 3878-3881 (2006).
165. Linser,R., Fink,U. & Reif,B. Proton-detected scalar coupling based assignment strategies in MAS solid-state NMR spectroscopy applied to perdeuterated proteins. *J. Magn Reson.* **193**, 89-93 (2008).

## 6 Appendix

6.1  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical shifts of OmpG derived from solid-state MAS NMR experiments

Residue		Chemical Shift [ppm]			
		N	CA	CB	C
1	M				
11	I				
12	G	112,18	47,38		173,11
13	A	128,23	50,63	23,4	
14	M		54,47	37,26	
15	Y				
28	D				
29	G	106,04	46,28		174,99
30	L	118,06	53,58	47,86	175,07
31	A	124,94	51,29	21,59	178,29
32	E	121,1	53,51	32,82	
33	P	140,63	61,9	33,05	173,55
34	S	110,93	58,12	66,82	
35	V				
36	Y	125,06	56,21	42,84	
37	F	114,65	55,9	41,77	
38	N	119,37	52,49	41,87	171,26
39	A	120,22	50,46	24,53	174,81
40	A	124,05	50,79	23,61	175,83
41	N	121,06	53,7	41,97	
42	G				
45	R		56,01	34,4	
46	I	126,47	60,35	41,75	173,08
47	A	129,03	51,09	23,58	175,23
48	L	123,11	53,85	46,62	176,02
49	A	122,69	51,67	24,05	175,13
50	Y	119,06	59,59	41,74	173,26
51	Y	129,16	54,92	42,82	
52	Q				
84	F				
85	S	119,89	56,39	67,04	172,75
86	F	122,35	54,18	41,86	174,24
87	G	111,53	44,99		169,66
88	L	117,26	54,11	48,96	173,59
89	T	121,98	61,69	70,77	174,11
90	G	113,86	44,05		172,47
91	G	109,75	45,98		169,86
92	F	121,22	56,68	43,8	
93	R				
94	N				
95	Y	118,97	57,75	42,05	172,74
96	G	114,28	44,13		172,55
97	Y	122,25	56,15	40,21	
98	H				

105	K				
106	D	112,59	55,5	38,87	173,4
107	T	125,72	62,65	69,52	173,37
108	A	128,78	50,39	22,9	172,3
109	N	117,03	53,52	44,43	
110	M				
114	K				
115	I	123,9	59,34	41,83	174,36
116	A	125,63	49,88	24,06	176,34
117	P	142,66	62,24	34,13	
118	D	117,22	53,52	44,41	
119	W				
131	G				
132	W	117,15	55,47	31,64	174,66
133	L	124,13	53,84	46,24	175,6
134	S	120,36	57,65	67,01	172,4
135	M				
136	Y				
137	K				
138	F	121,72	56,4	42,61	172,31
139	A	121,33	51,62	23,59	175,8
140	N	122,04	54,43	42,32	176,29
141	D	112,62	55,42	38,87	177,06
142	L	116,94	57,51	40,03	180,83
143	N	115,82	55,18	36,95	176,53
144	T	116,26	65,17	68,72	174,67
145	T	110,74	62,41	69,28	176,81
146	G	109,37	45,66		173,64
147	Y	121,11	57,95	38,85	175,99
148	A	122,99	50,82	17,81	178,49
149	D	116,23	56,17	41,62	
150	T	118,11	61,91	70,79	172,2
151	R	124,58	54,62		176,36
152	V				
153	E	128,48	53,41		175,14
154	T	122,48	59,53	69,48	168,76
155	E	124,07	55,59	30,85	175,8
156	T	123,26	59,62	70,17	171,07
157	G	114,09	47,39		171,66
158	L	118,71	52,44	46,32	174,77
159	Q	124,03	54,54		174,38
160	Y	126,79	56,75	41,95	171,98
161	T	111,72	58,96	70,7	172,56
162	F	114,74	56,22	42,31	
163	N				
164	E	115,03	58,7	29,31	176,04
165	T	115,93	64,01	69,78	173,38
166	V	120,36	62,48	34,8	173,66
167	A	116,96	50,46	23,38	175,42
168	L	119,4	54,43	46,57	173,84
169	R	124	59,11		175,38
170	V	122,09	62,55	34,15	172,54
171	N				
176	R				
177	G	118,44	44,83		170
178	F	117,75	56,49	39,86	
179	N				

194	I				
195	R		54,54	35,81	172,89
196	A	125,02	49,33	21,66	174,88
197	Y	118,25	55,64	42,88	173,99
198	L	124,1	51,13	44,74	170,58
199	P	134,31	63,96	31,89	176,96
200	L	127,23	52,66	43,53	178,58
201	T	122,09	61,46	69,65	173,61
202	L	128,69	51,83	43	175,96
203	G	112,05	47,1		175,67
204	N	112,39	54,05	38,53	177,55
205	H				
206	S				
207	V				
208	T	122,16	58,96	70,74	171,61
209	P	140,56	61,81	32,62	175,29
210	Y				
236	R				
237	V	112,31	58,88	35,67	
238	G	111,78	45,46		171,72
239	L	126,69	52,32	45,73	173,56
240	F	123,77	56,37	38,62	
241	Y	125,06	56,21	42,84	
242	G	112,53	45,57		171,63
243	Y	122,83	59,041	41,93	
244	D				
247	N				
248	G	113,24	45,44		174,41
249	L	124,44	54,73	45,5	172,97
250	S	119,05	56,73	66,07	172,46
251	V	114,49	58,62	36,69	174,91
252	S	120,25	56,68	67,59	172,22
253	L	121,68	54,12	47,52	174,3
254	E	121,64	55,48		
255	Y	117,7	57,68	41,56	174,89
256	A	123,57	51,69	23,93	174,88
257	F	124,28	56,52	43,9	
258	E				
271	H				
272	Y	131,9	56,02	41,71	173,1
273	A	129,58	50,91	23,44	175,1
274	G	105,97	45,95		170,53
275	V	120,47	59,15	35,55	173,45
276	G	112,39	45,45		171,12
277	V	115,43	59,47	35,8	172,64
278	N	125,09	52,18	42,67	
279	Y				
280	S				
281	F				

## 6.2 Publications

### This thesis partially based on the following publications:

Matthias Hiller, Victoria A. Higman, Stefan Jehle, Barth-Jan van Rossum, Werner Kühlbrandt and Hartmut Oschkinat (2008) **2,3-<sup>13</sup>C labelling of aromatic residues – getting a head start in the MAS-NMR assignment of membrane proteins.**

J Am Chem Soc. **130**, 408-9.

Stefan Jehle, Matthias Hiller, Kristina Rehbein, Anne Diehl, Hartmut Oschkinat and Barth-Jan van Rossum (2006) **Spectral editing: selection of methyl groups in multidimensional solid-state magic-angle NMR.** J Biomol NMR. **36**, 169-77.

Matthias Hiller, Ludwig Krabben, Kutti R. Vinothkumar, Federica Castellani, Barth-Jan van Rossum, Werner Kühlbrandt and Hartmut Oschkinat (2005) **Solid-State Magic-Angle Spinning NMR of Outer-Membrane Protein G from *Escherichia coli*.** Chembiochem. **6**, 1679-84.

### Further Publications:

Victoria A. Higman, Jeremy Flinders, Matthias Hiller, Stefan Jehle, Sebastian Fiedler, Barth-Jan van Rossum and Hartmut Oschkinat (2009) **Assigning Large Proteins in the Solid State: A Straightforward MAS NMR Resonance Assignment Strategy using Selectively and Extensively <sup>13</sup>C-Labeled Proteins.** (Submitted)

Scholz I, Jehle S, Schmieder P, Hiller M, Eisenmenger F, Oschkinat H, van Rossum BJ (2007) **J-deconvolution using maximum entropy reconstruction applied to 13C-13C solid-state cross-polarization magic-angle-spinning NMR of proteins.** J Am Chem Soc. **129**, 6682-3.

**Poster contributions at conferences and workshops:**

M. Hiller, K R. Vinothkumar, B. Heinemann, Prof. H. Oschkinat and Prof. W. Kühlbrandt (2002) **Expression Purification and two-dimensional crystallisation of the outer membrane protein G (OmpG) for electron microscopy and solid-state MAS NMR.** Paris (France), FROM GENE TO FUNCTIONAL PROTEIN, European Workshop

M. Hiller, K R. Vinothkumar, F. Castellani, Bart Jan van Rossum L. Krabben and Hartmut Oschkinat. (2004) **Solid-state MAS NMR investigation of the outer membrane protein G (OmpG) for structure determination.** Monte Verita (Switzerland), THE 2rd INTERNATIONAL CONFERENCE ON THE STRUCTURE, DYNAMICS AND FUNCTION OF BIOLOGICAL MEMBRANES.

M. Hiller, V. Kumar, S. Jehle, B. van Rossum, L. Krabben, W. Kühlbrandt and Hartmut Oschkinat (2006) **Structural investigation of the outer membrane protein G by Solid-State MAS NMR.** Pacific Grove (USA), 2006, 7 th ENC-EXPERIMENTAL NUCLEAR MAGNETIC RESONANCE CONFERENCE

M. Hiller, V. Higman, V. Kumar, S. Jehle, B. van Rossum, W. Kühlbrandt and H. Oschkinat (2007) **Towards Assignment of the outer membrane protein G by Solid-State MAS NMR.** Florence (Italy), ADVANCES AND MANAGEMENT OF NMR IN LIFE SCIENCES.

Victoria A. Higman, Matthias Hiller, Sebastian Fiedler, Stefan Jehle, Barth-Jan van Rossum and Hartmut Oschkinat (2008) **Assigning Outer Membrane Protein G by solid-state MAS NMR.** Berlin (Germany), Joint NMR-Life/Extend-NMR meeting, “New approaches to biomolecular structure determination”



### 6.3 Curriculum vitae

Name	Matthias Hiller
Date of Birth	03 January 1974
Place of Birth	Bautzen, Germany
Nationality	German

#### *Educational*

<b>July 2002 - Present</b>	PhD thesis: “Sample preparation of membrane proteins suitable for solid-state MAS NMR and development of assignment strategies” in the group of Prof. Dr. H. Oschkinat at the "Leibniz-institut für Molekulare Pharmakologie" (FMP), Berlin
<b>June 2001 - July 2002</b>	Technical assistant in the group of Prof. Dr. H. Oschkinat at the "Leibniz Institut für Molekulare Pharmakologie" (FMP), Berlin
<b>April 2000 - June 2001</b>	Diploma thesis: “ <i>Expression of the human endothelin-1 in E. coli</i> ” in the group of Prof. Dr. H. Oschkinat at the "Leibniz Institut für Molekulare Pharmakologie" (FMP), Berlin
<b>September 1998 - February 1999</b>	Work experience: “ <i>Analysis of land ecological characteristics by statistical multivariate methods</i> ”. in the group of Dr. sc. W. Seyfarth at the "Zentrum für Agrar-landschafts- und Landnutzungsforschung e. V" (ZALF) in Müncheberg.
<b>October 1996 - June 2001</b>	Diploma study of chemistry in the field of biotechnology at the "University of applied science" in Senftenberg

Berlin, März 2009                      Matthias Hiller

## Danksagung

Die hier vorgestellte Arbeit entstand am Leibniz Institut für Molekulare Pharmakologie (FMP) auf dem Campus Berlin-Buch. Nachfolgend möchte allen beteiligten Personen für die kooperative Zusammenarbeit während meiner Promotionszeit danken.

Prof. Robert Seckler möchte ich dafür danken, dass er mich unterstützt hat dieses Promotionsvorhaben an der Universität Potsdam einzuleiten und schließlich auch zu beenden.

Prof. Hartmut Oshkinat möchte ich dafür danken, dass er mir die Möglichkeit gab an diesem fesselnden Forschungsprojekt teilzuhaben. Ich möchte mich für seine Unterstützung sowie sein Vertrauen und die zahlreichen Ideen und Vorschläge, wodurch die Durchführung dieser Arbeit erst möglich war, bedanken.

Für die Zusammenarbeit im Labor möchte ich mich bei Dr. Anne Diehl, Kristina Rehbein, Liselotte Handel, Sebastian Fiedler und Markus Pydde bedanken. Anne bei Dir möchte ich mich besonders dafür bedanken, dass Du mich bei allen Tätigkeiten im Labor (und das waren sehr viele) mit viel Geduld unterstützt, gefordert und gefördert hast. Ich habe sehr viel von Dir gelernt!

Dr. Victoria Higman (Davies), möchte ich meinen besonderen Dank für die tolle Zusammenarbeit in diesem Projekt aussprechen. Danke auch dafür, dass Du diese Arbeit so oft gelesen hast und mir mit viel Geduld wissenschaftliches Englisch beigebracht hast. Es macht großen Spaß mit Dir zusammen zu arbeiten!

Für die Zusammenarbeit an den NMR Spektrometern möchte ich mich bei Dr. Johanne Becker, Dr. Federica Castellani, Dr. Barth-Jan van Rossum und besonders bei Dr. Stefan Jehle bedanken. Barth und Federica, ihr habt mir eine sehr gute Einführung in die Festkörper NMR Spektroskopie gegeben. Stefan, du hast oft viele Stunden (manchmal auch vergeblich) mit mir am Spektrometer verbracht um Messungen zu optimieren oder mir geduldig zu erklären, welche theoretischen Grundlagen hinter den einzelnen Experimenten stehen. Vielen Dank!

Bei Frau Dr. Dorothea Lorenz und Martina Ringling möchte ich mich für die Unterstützung bei den elektronenmikroskopischen Untersuchungen meiner Proben bedanken.

Meinen Kooperationspartnern vom MPI für Biophysik in Frankfurt am Main, Prof. Werner Kühlbrandt und Vinothkumar Kutti Rangunath möchte ich für die offene Arbeitsatmosphäre danken. Vinoth, es war gut das Du mir die Präparation von 2D Kristallen ohne viel Hexerei beigebracht hast.

Allen Mitarbeitern der Arbeitsgruppe „NMR-supported structural biology“ gilt mein Dank für die tolle Arbeitsatmosphäre. Besonders bedanken möchte ich mich bei Janina Hahn (die immer ein offenes Ohr für mich gehabt hat), bei Christian Köhler (für die leckere Essen und die entspannte Pfingstpaddeltour), bei Matthias Dorn (der was seinen Rechner angeht sehr depolarisierend wirken kann, aber mit hohem Potential an die Arbeit geht), bei Vivien Lange (die wenn mein ein Versteck für wichtige Dinge sucht, garantiert ein Schließfach findet), bei Dr. Ludwig Krabben (der einen machen und damit reifen lässt) und bei Dr. Peter Schmieder (der einem mit viel Geduld oft auch doppelt Fragen beantwortet).

Ganz besonders bedanken möchte ich mich bei meinen Freunden, Mandy Syring - Edgar Becker (für die schönen Wochenenden und Festivals), Solveig Großmann (für die vielen Kinofilme und daß du mir oft einfach nur zugehört hast), Rudi Ritzer – Suse (für die Übernachtungsmöglichkeiten im Bauwagen und für ganz viel andere Dinge) Michael Soukenik - Irmgard Gran (für die schöne Zeit in den Wein- und etwas höheren Bergen). Dir, Katja, möchte ich dafür danken, dass du mich so nimmst wie ich bin!

Tiefe Dankbarkeit empfinde ich meiner Mutter gegenüber. Du warst immer für mich da und hast mich Zeit meines Lebens unterstützt und an mich geglaubt. Du bist ein besonderer Mensch!

## **Eidesstattliche Erklärung**

Hiermit versichere ich, dass ich die vorliegende Dissertation selbständig und ohne unerlaubte Hilfe verfasst und keine anderen als die angegebene Literatur und Hilfsmittel verwendet habe.

Die dem Verfahren zugrunde liegende Promotionsordnung ist mir bekannt. Ich erkläre, dass ich mich bisher nicht an einer anderen Hochschule um einen Doktorgrad beworben habe und keinen derartigen Titel besitze.

Berlin,

Matthias Hiller