

Forschungsinstitut für Molekulare Pharmakologie Berlin

Development and application of novel triple-resonance
experiments for the assignment of protein NMR spectra

Dissertation

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

vorgelegt am Fachbereich Biologie, Chemie und Pharmazie
der Freien Universität Berlin

von

Mario Schubert

aus Chemnitz

Diese Arbeit wurde von März 1998 bis Dezember 2001 am Forschungsinstitut für Molekulare Pharmakologie Berlin unter Leitung von Dr. Peter Schmieder und Prof. Dr. Hartmut Oschkinat angefertigt.

Dissertation eingereicht am: 20.12.2001
Tag der Disputation: 07.03.2002
Erstgutachter: Prof. Dr. Hartmut Oschkinat
Zweitgutachter: Priv.-Doz. Dr. Gerd Buntkowski
weitere Kommissionsmitglieder: Prof. Dr. Hans-Ulrich Reißig,
Prof. Dr. Ernst-Walter Knapp,
Priv.-Doz. Dr. Burkhard Kirste

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Schubert, Mario:

Development and application of novel triple-resonance experiments for the assignment of protein NMR spectra / vorgelegt von Mario Schubert.

- Berlin : Mensch-und-Buch-Verl., 2002

Zugl.: Berlin, Freie. Univ., Diss., 2001

ISBN 3-89820-336-0

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus zu vervielfältigen.

© **MENSCH & BUCH VERLAG**, Berlin 2002
Nordendstr. 75, 13156 Berlin • ☎ 030 - 45 49 48 66
<http://www.menschundbuch.de> • info@menschundbuch.de

*Meinen Eltern
und Constance*

Acknowledgements

The work presented in this thesis has been carried out at the "Forschungsinstitut für Molekulare Pharmakologie (FMP)", Berlin. I would like to express my highest gratitude to the following people for help and advice throughout this project:

Prof. Dr. Hartmut Oschkinat for the possibility to work in his lab, for the excellent environment and his advises contributing to this work.

Dr. Peter Schmieder for his supervision and continuous support, his excellent introduction into multidimensional NMR spectroscopy of proteins, into the use of NMR machines, the art to write pulse programs and his patience and attention to detail in reviewing the draft thesis.

Prof. Dr. Dieter Oesterhelt, Dr. Michael Kolbe and Brigitte Kessler for the constructive collaboration in the bacteriorhodopsin project, especially for the preparation of the bacteriorhodopsin samples.

Maika Smalla, Martina Leidert and Dr. Thomas Jarchau for the preparation of the labeled samples of the SAM-, SH3- and EVH1-domain.

Dr. Ronald Kühne for his collaboration in the structure calculations with AMBER, fruitful discussions and his introduction into Sybyl.

all colleagues involved in the "Protein Structure Factory" for the weekly meetings and useful discussions, especially

Dr. Rüdiger Winter for the introduction into automated assignment, our uncountable discussions about assignment strategies, his suggestions which selective experiments might be helpful for an automated assignment strategy, for his collaboration especially to write the program SELMA;

Dr. Dietmar Leitner and Dr. Dirk Labudde for their collaboration, their efforts to develop a novel automated assignment software and especially for implementing the amino acid type-selective experiments into the standard set of experiments used for each new protein.

Michele Fossi for his enthusiasm for MUSIC, his critical questions and extremely creative discussions.

I thank all the members of the group for the nice atmosphere and their support, especially Dr. Barth van Rossum and Federica Castellani in my office, Dr. Ricardo Pires, Dr. Jutta Pauli, Maika Smalla, Martin Wahl and Dr. Markus Schade.

Dr. Marc Kelly and Dr. Linda Ball for their help concerning the English language.

Stefan Hübel, Dr. Grant Langdon and Urs Wiedemann, the computer specialists, are acknowledged for their help when the computer did not understand me. Dr. Grant Langdon for introducing me into the program Sparky and other useful software.

Andrea Steuer for her help with bureaucracy.

Dr. Wolfgang Bermel (Bruker AG) for the introduction into the hardware of AVANCE spectrometers, for his help in generating and simulating shaped pulses, his comments on chapter 3, for fruitful discussions and his help concerning hardware problems.

Jürgen M. Schmidt (London) and Frank Löhr (Frankfurt) for helpful discussions and for coupling constant data.

Dr. Peter Schmieder, Dr. Linda Ball, Dr. Wolfgang Bermel, Dr. Dirk Labudde, Dr. Dietmar Leitner, Michele Fossi for carefully reading parts or all of the manuscript.

Financial support by the DFG Graduiertenkolleg 80 "Modellstudien zu Struktur, Eigenschaften und Erkennung Biologischer Makromoleküle auf Atomarer Ebene" is acknowledged.

Dr. Wünsche und Dr. Liesaus (Rudolf-Elle-Krankenhaus, Eisenberg) and the staff of the rehabilitation center Medifit in Jena for healing my damaged knee.

I wish to thank my parents and Constance for their motivation, patience and constant moral support.

Contents

Chapter 1	General introduction	
	1
	References	3
Chapter 2	NMR chemical shift assignment strategies for proteins	
2.1	Introduction	4
2.2	Sequential resonance assignment in unlabeled and ^{15}N labeled proteins	4
2.3	Assignment of uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled proteins using triple-resonance experiments	7
2.4	Frontiers of Biological NMR: Application of TROSY on large uniformly $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled proteins	14
	References	17
Chapter 3	MUSIC with shaped pulses and tuned delays: An introduction to the techniques used to create amino acid type-selective NMR experiments	
3.1	Introduction	19
3.2	Basic building blocks in triple-resonance experiments	19
3.3	Spectral editing techniques - The precursors of MUSIC	21
3.4	MUSIC: A multiplicity selective in-phase coherence transfer	24
3.5	Shaped pulses	30
3.6	Selecting specific coupling patterns by tuning of delays	35
	References	41
Chapter 4	MUSIC in triple-resonance experiments: Amino acid type-selective ^1H-^{15}N correlations	
4.1	Introduction	42
4.2	Previous work	44
4.3	General strategy to design amino acid type-selective NMR experiments	46
4.4	The individual amino acid type-selective ^1H - ^{15}N correlations	48
4.4.1	Selecting Gly: The G-(i+1)- and G-(i,i+1)-HSQC	48
4.4.2	Selecting Ala: The A-(i+1)- and A-(i,i+1)-HSQC	50
4.4.3	Selecting Thr, Val, Ile and Ala: The TA-(i+1)- and TA-(i,i+1)-HSQC, the VIA-(i+1)- and VIA-(i,i+1)-HSQC	51
4.4.4	Selecting Leu and Ala: The LA-(i+1)- and LA-(i,i+1)-HSQC	54
4.4.5	Selecting Ser: The S-(i+1)- and S-(i,i+1)-HSQC	60
4.4.6	Selecting aromatic residues: The W-(i+1)- and W-(i,i+1)-HSQC, the FYH-(i+1)- and FYH-(i,i+1)-HSQC	62
4.4.7	Selecting Asn, Asp and Gly: The DNG-(i+1)- and DNG-(i,i+1)-HSQC	66

4.4.8	Selecting Asp: The D-(i+1)- and D-(i,i+1)-HSQC	68
4.4.9	Selecting Asn: The N-(i+1)- and N-(i,i+1)-HSQC	70
4.4.10	Selecting Gln, Glu and Gly: The EQG-(i+1)- and EQG-(i,i+1)-HSQC.....	72
4.4.11	Selecting Glu: The E-(i+1)-, ED-(i+1)- and E-(i,i+1)-HSQC.....	75
4.4.12	Selecting Gln: The Q-(i+1)-, QN-(i+1)- and Q-(i,i+1)-HSQC	79
4.4.13	Selecting Arg: The R-(i+1)- and R-(i,i+1)-HSQC	81
4.4.14	Selecting Lys: The K-(i+1)-, KR-(i+1)- and K-(i,i+1)-HSQC.....	84
4.4.15	Selecting neighbors of Pro: The P-(i-1)- and P-(i+1)-HSQC.....	87
4.4.16	Selecting Cys and Ser: The CS-(i+1)- and CS-(i,i+1)-HSQC.....	91
4.4.17	Attempts to select Met	93
4.4.18	Selecting the indole NH of Trp: The H ^{ε1} N ^{ε1} -HSQC	94
4.5	Conclusion and outlook	95
	References	97

Chapter 5 Amino acid type-selective experiments open up new strategies in automated assignment

5.1	Introduction	99
5.2	Automated backbone resonance assignment of ¹³ C/ ¹⁵ N labeled proteins - an overview	99
5.3	Problematic steps in automated assignment	103
5.4	Amino acid type identification of sequential pairs using the amino acid type-selective experiments	106
5.5	Potential strategies for implementing the amino acid type-selective experiments into automated assignment procedures	111
5.6	Conclusion and outlook	115
	References	116

Chapter 6 Bacteriorhodopsin in micelles: A model system for exploring the use of NMR on the structural investigation of membrane proteins

6.1	NMR of membrane proteins	117
6.2	Bacteriorhodopsin.....	120
6.3	Samples and NMR experiments	121
6.4	Sequential assignment	123
6.5	Secondary structure and obtained angle and distance restraints of the assigned regions....	128
6.6	Structure calculations of the intracellular loops AB, EF and the C-terminal part of helix G.....	131
6.7	Mobility	133
6.8	Discussion	134
6.9	Conclusion and outlook	137
	References	137

Summary	140
Zusammenfassung	142
Appendix	144

Abbreviations

2D, 3D, 4D	two-dimensional, three-dimensional, four-dimensional
2Q, 3Q	double quantum, triple quantum
1J	one bond spin-spin coupling constant
BR	bacteriorhodopsin
COSY	correlation spectroscopy
CSA	chemical shift anisotropy
DEPT	distortionless enhancement by polarization transfer
GARP	globally optimized, alternating-phase rectangular pulses
GCPRs	G-protein coupled receptors
HMQC	heteronuclear multiple quantum coherence
HSQC	heteronuclear single quantum coherence
INEPT	insensitive nuclei enhanced by polarization transfer
kDa	kilo Dalton
MUSIC	multiplicity selective in-phase coherence transfer
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser enhancement spectroscopy
ns	number of scans
PFG	pulsed field gradients
POMMIE	phase oscillations to maximize editing
ppm	parts per million
rmsd	root mean square deviation
SAR	structure-activity relationship
SW	spectral window
T_1	longitudinal relaxation time
T_2	transversal relaxation time
TOCSY	total correlation spectroscopy
TPPI	time proportional phase incrementation
TROSY	transverse relaxation optimized spectroscopy
WALTZ	wideband, alternating-phase, low-power technique for zero-residual-splitting
WATERGATE	water suppression by gradient tailored excitation

The understanding of physiological processes at a molecular level has revolutionized life science in the last decades. The discovery of the genetic code, the understanding of how this information is translated into proteins, the sequencing of whole genomes of several organisms and the decoding of the human genome have been milestones yielding profound insights into the secrets of life. The explosive growth of knowledge results from an interplay of several disciplines: molecular biology, biochemistry, bioinformatics, structural biology and also many sections of chemistry and physics. The development of new technologies combined with high degrees of automation have allowed crucial breakthroughs such as the deciphering of the human genome.

Following completion of the Human Genome Project, the DNA code has now to be analyzed and many questions arise. What is the function of the encoded proteins and RNA? How do they interact with each other? How does the whole proteome (all proteins of an organism) change during cell differentiation? Which post-translational modifications take place? This is only a small sample of the many questions which will be investigated in the coming decades. This research field has been given the name Proteomics.

One approach to Proteomics is to determine structures of the unknown proteins encoded in the DNA and based on this to propose their possible functions. Several initiatives based on this strategy were founded in the past three years, among them the Protein Structure Factory in Berlin (Heinemann et al., 2000).

Structural biology, the discipline which deals with the three dimensional structure determination of biological macromolecules and their interactions with other molecules, has developed into one of the most important fields for understanding of biological functions. The successful determination of a protein structure at atomic resolution was first performed in 1959 using X-ray crystallography (Kendrew 1960, 1993). Since then X-ray crystallography has been established as the traditional method for solving protein structures. However, it is useful only for proteins that form high quality crystals. Two decades later, nuclear magnetic resonance (NMR) spectroscopy showed the capability to solve structures of proteins and peptides in an aqueous environment, as they exist in biological systems (Wüthrich, 1986).

Additionally, NMR allows one to observe the physical flexibility of proteins and the dynamics of their interaction with other molecules in solution - a huge advantage when studying the biochemical function of a protein (Palmer et al., 2001). Finally, potential binding partners of proteins, even with low affinity, can be screened by NMR spectroscopy and the binding sites can be mapped. Based on this, a novel method was developed in pharmaceutical research to optimize the binding affinity of potential drugs: SAR (structure affinity relationship) by NMR (Shuker et al., 1996).

NMR spectroscopy is a young method and is developing very rapidly. The method allows the determination of protein structures routinely up to a molecular mass of ca. 20 kDa. Several techniques have been recently developed to extend the applications to sizes up to 100 kDa (Pervushin, 2000). There are efforts to fully automate the whole process of protein structure determination (Moseley and Montelione, 1999).

The main project described in the present work is a contribution to the technology which has to be developed for automated assignment of NMR signals, as the necessary first step in automated structure determination. Novel amino acid type-selective NMR experiments were designed which identify the amino acid type corresponding to each NMR signal. Strategies to implement these experiments into an automated assignment procedure are proposed. The implementation itself would go beyond the scope of this work and is tackled by others (Oschkinat group, unpublished results) . The use of this specific information as additional input will greatly facilitate automated assignment, increase the robustness of existing automated procedures and ultimately reduce or even avoid any manual interference.

The second project described here deals with the extension of liquid state NMR spectroscopy to membrane proteins. These can often not be crystallized and therefore X-ray crystallography is not applicable. Recently developed techniques in NMR spectroscopy open the way for obtaining structural information on membrane proteins solubilized in micelles (Arora and Tamm, 2001). Bacteriorhodopsin was chosen as a model system. The micelle and the membrane protein together behave like a very large protein of ca. 100 kDa and thus ranks at the frontiers of NMR accessibility. The aim of the project was to explore how far the spectra of the membrane protein could be assigned with recent methods and which structural information could be extracted. This knowledge allows us to judge whether biologically relevant membrane proteins such as G-protein coupled receptors are already accessible to NMR investigation, or whether the recent techniques have to be further developed.

References

- A. Arora and L. K. Tamm, *Curr. Opin. Struct Biol.* **11**, 540-547 (2001).
- U. Heinemann, J. Frevert, K. Hofmann, G. Illing, C. Maurer, H. Oschkinat, W. Saenger, *Prog. Biophys. Mol. Biol.* **73**, 347-62 (2000).
- J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. R. Davies, D. C. Phillips and V. C. Shore, *Nature* **185**, 422-427 (1960).
- J. C. Kendrew, *Science* **139**, 1259-1266 (1963).
- H. N. Moseley and G. T. Montelione, *Curr. Opin. Struct Biol.* **9**, 635-642 (1999).
- A. G. Palmer III, C.D. Kroenke, J. P. Loria, *Methods Enzymol.* **339**, 204-238 (2001).
- K. Pervushin, *Q. Rev. Biophys.* **33**, 161-197 (2000).
- S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **274**, 1531-1534 (1996).
- K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York (1986).

2.1 Introduction

The development of the Fourier transform NMR in 1966 (Ernst and Anderson, 1966) and the introduction of a second frequency dimension in NMR spectroscopy by Jeener in 1971 (Jeener, 1971) opened the way for structure determination of complex molecules like natural products, sugars and peptides by NMR. A decade later systematic procedures for the analysis of complex protein spectra were introduced by Wüthrich and co-workers to determine three dimensional structures of proteins (Wüthrich et al. 1986). In 1985 the first protein structure was determined by NMR spectroscopy (Williamson et al., 1985). These methods were developed for unlabeled proteins and allowed resonance assignments and measurements of distance restraints for proteins with molecular masses of up to 10-12 kDa. They are briefly described in chapter 2.2 together with methods for the investigation of ^{15}N labeled proteins.

The development of molecular biological methods in the late 1980's allowed ^{15}N -labeling and later on $^{13}\text{C}/^{15}\text{N}$ -labeling of proteins. This resulted in a substantial addition to the assignment strategies by the use of double and triple-resonance NMR experiments. The use of triple-resonance experiments on uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled proteins which is discussed in chapter 2.3 is nowadays a routine method for assigning proteins with molecular masses of up to 15-20 kDa.

The recent introduction of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeling and the TROSY technique has pushed the size limit towards 100 kDa. These techniques are introduced in chapter 2.4.

2.2 Sequential resonance assignment in unlabeled and ^{15}N labeled proteins

For the investigation of structure and dynamics of biomolecules by NMR a complete resonance assignment is a prerequisite (Wüthrich, 1986; Cavanagh et al., 1996). The methods for the sequential assignment of proteins developed by Wüthrich and co-workers (Wüthrich et al., 1982) are based on the identification of spin systems within individual amino acids using through-bond ^1H - ^1H connectivity measured with correlation spectroscopy (COSY) (Jeener, 1971; Aue et al., 1976) and later on by total correlation spectroscopy (TOCSY) (Braunschweiler and Ernst, 1983; Bax and Davis, 1985). In a second step, the

sequential connectivities between neighboring amino acids are established via nuclear Overhauser effect spectroscopy (NOESY) (Jeener et al., 1979). Small peptide segments of different length are thus obtained which are matched to the sequence. In order to complete the resonance assignment these segments are extended and linked.

With increasing molecular weight not only the number of signals but also the linewidth of the signals increases. Additionally, losses of magnetization occur due to relaxation. The development of three-dimensional homonuclear experiments in the late 1980's (Griesinger et al. 1987a; Griesinger et al. 1987b; Oschkinat et al. 1988) offered a better dispersion of signals. However, the number of signals is higher than in 2D spectra, the resolution in the two indirect dimensions is usually inferior and the sensitivity is relatively low due to intensity losses during the additional magnetization transfers. Therefore relatively long measurement times are necessary.

The introduction of ^{15}N labeling into protein NMR offered an additional frequency dimension (^{15}N) which greatly enhanced the dispersion of 2D and 3D NMR spectra as illustrated in Figs. 2.1 and 2.2. The extension of common homonuclear 2D experiments by a third ^{15}N frequency dimension yields 3D ^{15}N edited ^1H - ^1H spectra. The spectra are less complex than homonuclear 3D spectra, only signals of ^1H nuclei which are attached to ^{15}N are present in the edited dimension. While the ^1H chemical shift dispersion is often not sufficient to distinguish all the H^{N} resonances, the ^1H - ^{15}N correlation is used as a "fingerprint" in which every individual amino acid in the chain is represented by one signal (Fig. 2.2). Each residue corresponds to a characteristic $^1\text{H}/^{15}\text{N}$ chemical shift pair. However, the general chemical shift assignment strategy based on homonuclear COSY, TOCSY and NOESY transfers was not revised.

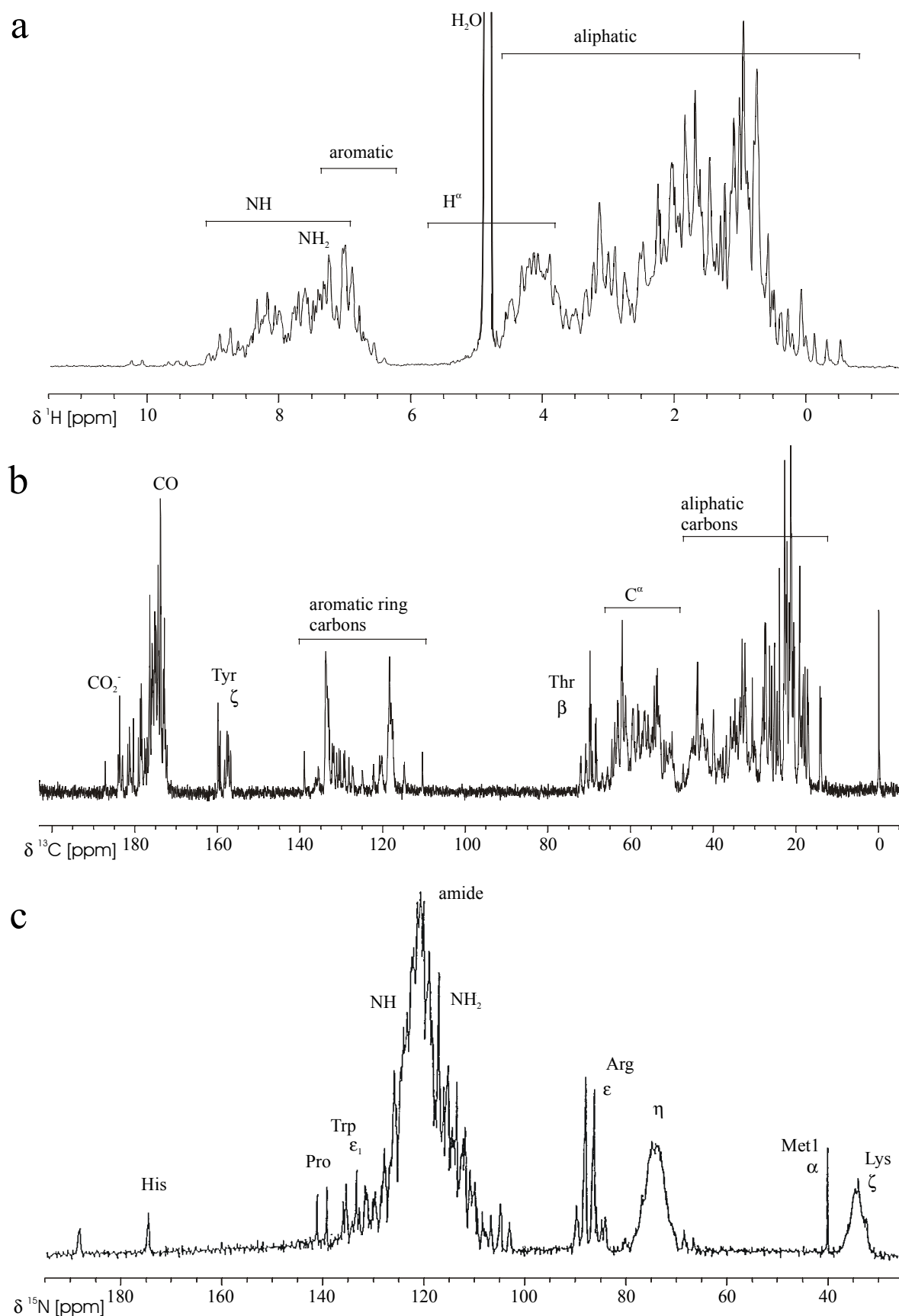


Fig. 2.1: The chemical shift dispersion of one-dimensional protein NMR spectra of the nuclei ^1H , ^{13}C and ^{15}N : (a) ^1H 1D spectrum of the SAM domain (Smalla et al., 1999), (b) ^{13}C 1D spectrum of Tendamistad (Kessler et al., 1990) at natural abundance, (c) ^{15}N 1D spectrum of ^{15}N labeled T_4 Lysozyme (McIntosh & Dahlquist, 1990).

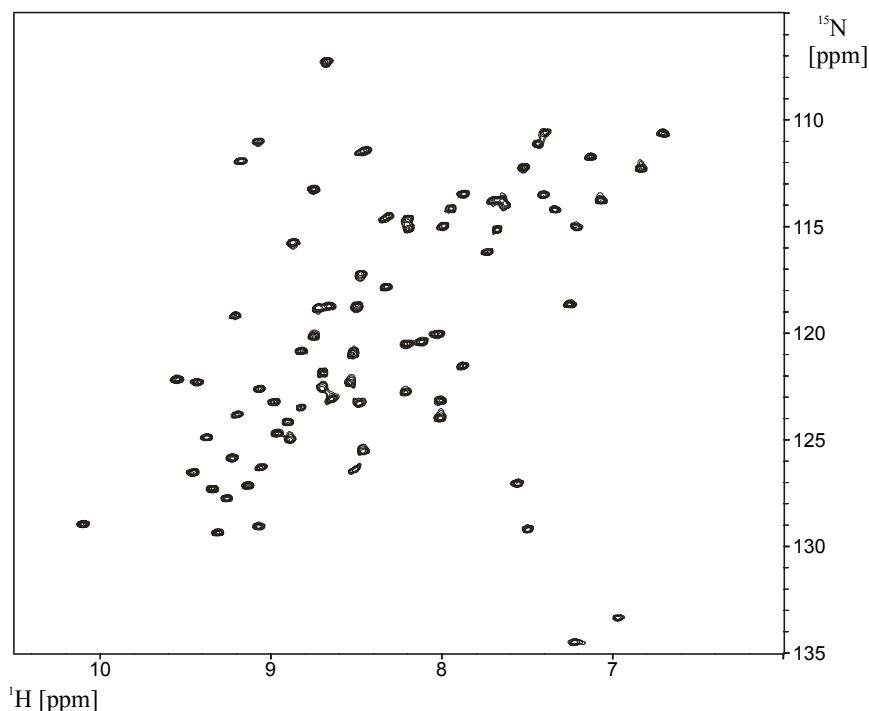


Fig. 2.2: Two-dimensional spectrum of a ^1H - ^{15}N correlation: ^{15}N -HSQC of the α -spectrin SH3 domain (Musacchio et al., 1992; Blanco et al., 1997). Each amide group corresponds to one signal. NH_2 signals of Asn/Gln and NH signals of the side chains of Trp and Arg are present too.

2.3 Assignment of uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled proteins using triple-resonance experiments

The introduction of $^{13}\text{C}/^{15}\text{N}$ labeled proteins initiated the development of a fundamentally new assignment procedure based on triple-resonance NMR experiments (Kay et al., 1990; Montelione and Wagner, 1990). Uniformly $^{13}\text{C}/^{15}\text{N}$ labeled proteins offer very efficient magnetization transfers through a network of coupled heteronuclear spins. Magnetization can be transferred between different ^{13}C nuclei and between ^{15}N and ^{13}C via COSY and INEPT steps using the coupling constants shown in Fig. 2.3.

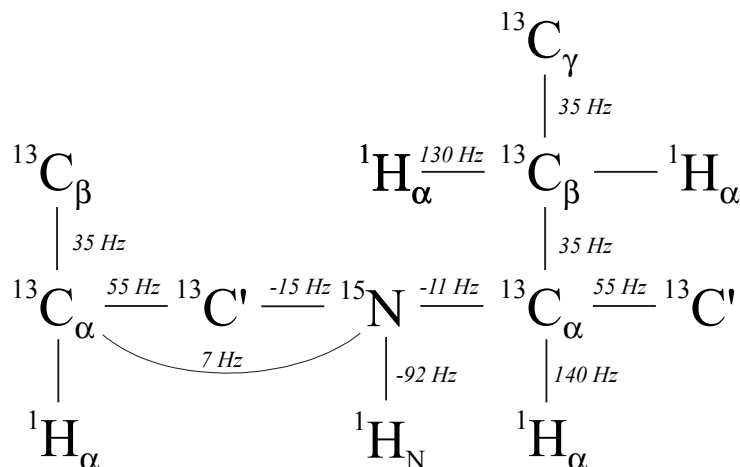


Fig. 2.3: The most important $^1\text{J}/^2\text{J}$ coupling constants in $^{13}\text{C}/^{15}\text{N}$ labeled proteins. They are used in the common triple-resonance experiments.

Due to the values of the 1J coupling constants being large in comparison to the linewidths, the magnetization transfer steps are quite efficient and several of them can be combined. Note that these couplings and transfers are conformation independent. The efficient transfer lead to the development of triple-resonance experiments which make use of radio frequency pulses at the frequencies of the three nucleus types ^1H , ^{13}C and ^{15}N . For reasons of sensitivity, the magnetization transfer usually starts and ends on ^1H .

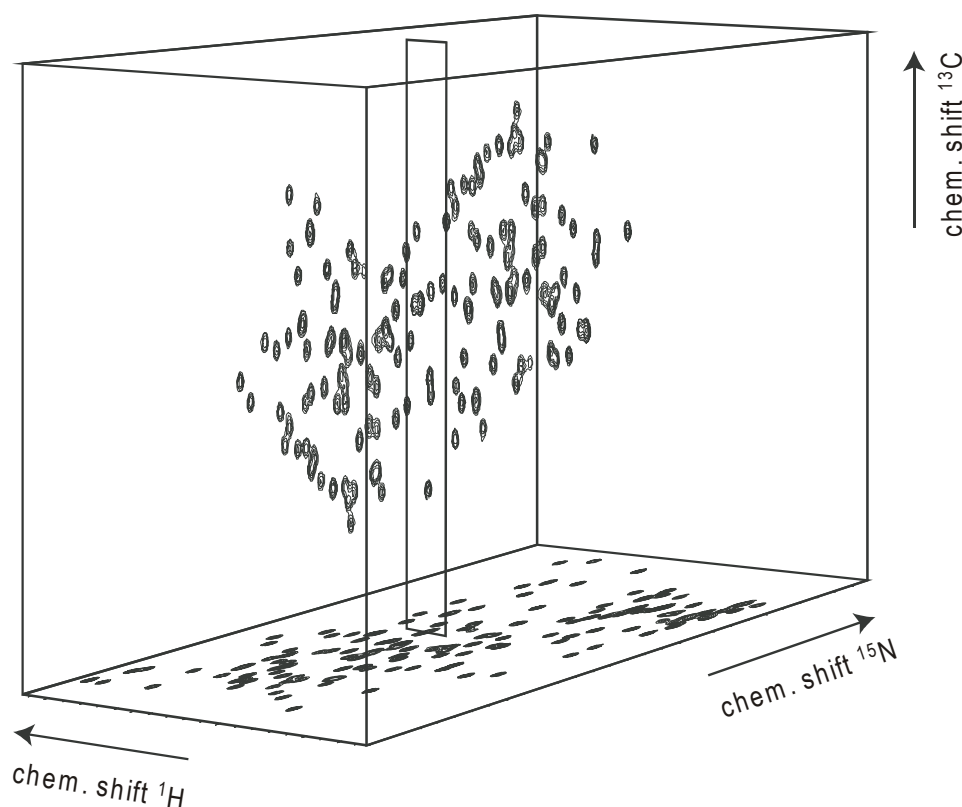


Fig. 2.4: Schematic representation of a three-dimensional triple-resonance spectrum with ^1H , ^{15}N and ^{13}C chemical shift axis. For one (^1H , ^{15}N) frequency pair which corresponds to a residue a “strip” is shown in the center of the 3D cube.

Most 3D triple-resonance experiments designed for proteins record $^1\text{H}^{\text{N}}$ and ^{15}N chemical shifts in addition to a third frequency (Fig. 2.4). 2D projections ($^1\text{H}/^{15}\text{N}$) of these spectra look like a usual 2D ^{15}N -HSQC, which contains a “fingerprint” of the protein, i. e., a pair of ($^1\text{H}^{\text{N}}$, ^{15}N) frequencies for each residue, only some side chain NH signals are missing. It should be emphasized that their use in the sequence specific assignment of ^{15}N , ^{13}C -labeled proteins lead to the most powerful assignment strategy available nowadays for proteins up to ca. 20 kDa (Clare and Gronenborn, 1991; Sattler et al., 1999). The usual strategy is to correlate each ($^1\text{H}^{\text{N}}$, ^{15}N) frequency pair with C^{α} , C^{β} or other frequencies of the same and the sequential amino acid and thus form chains of residue patterns that can be matched to the amino acid sequence of the protein once the type of amino acid corresponding to the pattern has been identified.

To connect ^{13}C chemical shifts of C^α and C^β to each $(^1\text{H}, ^{15}\text{N})$ frequency pair the two 3D experiments CBCA(CO)NH (Grzesiek and Bax, 1992a) and CBCANH (Grzesiek and Bax, 1992b) are usually used. The magnetization transfer pathways are shown schematically in Fig. 2.5 a and b. Note that the CBCA(CO)NH deliver the C^α and C^β chemical shift of the preceding residue (i-1) while the CBCANH delivers the C^α and C^β chemical shift of residue (i-1) and i. Signals of residue (i-1) are recognized by the presence in both spectra, while signals of the amino acid i are only present in the CBCANH (Fig. 2.6). The C^β signals in the CBCANH have opposite sign compared to the signals from C^α .

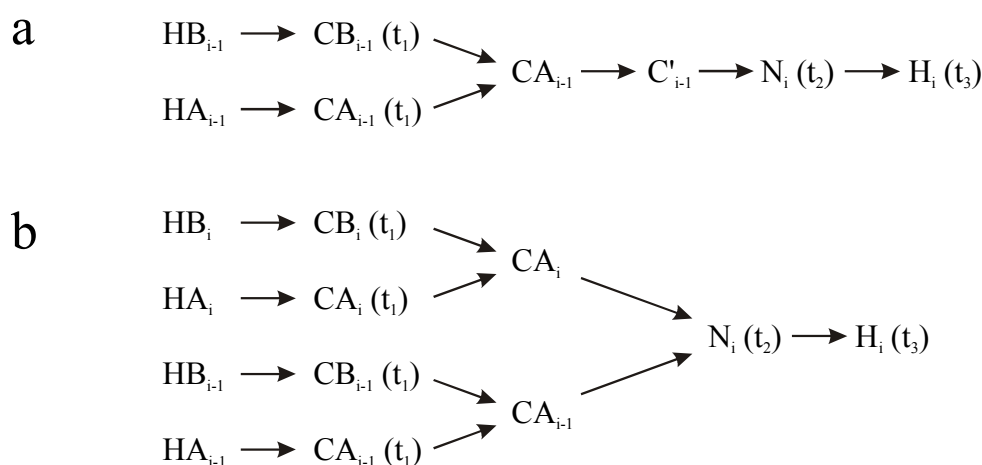


Fig. 2.5: Magnetization transfer pathways of the triple-resonance experiments (a) CBCA(CO)NH, (b) CBCANH used for the sequential backbone assignment of $^{13}\text{C}/^{15}\text{N}$ labeled proteins. Magnetization is transferred using ^1J coupling constants and in the CBCANH additionally $^2\text{J}_{\text{NCA}}$. The ^{13}C and ^{15}N chemical shift evolution periods are indicated with t_1 and t_2 . t_3 is the ^1H acquisition time. The indices i and $i-1$ indicate two neighboring residues in the amino acid sequence.

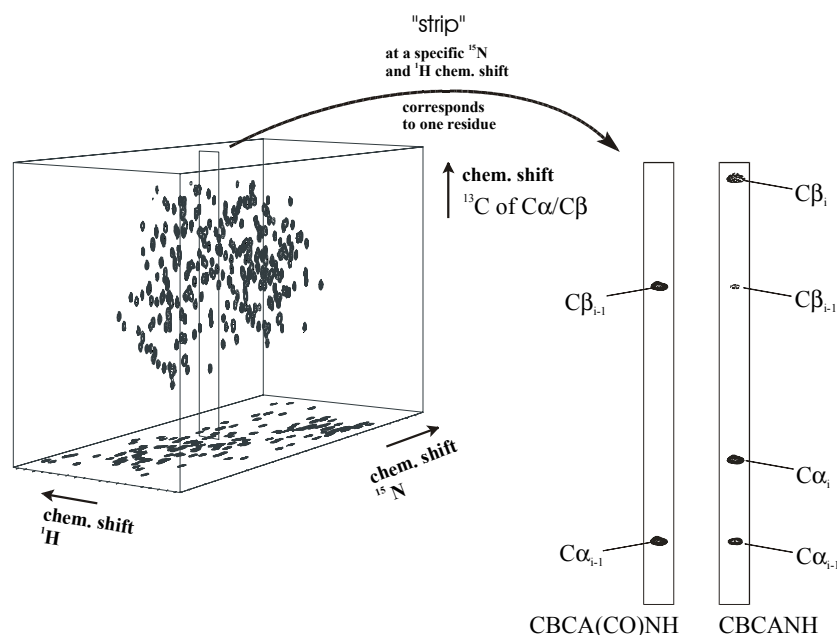


Fig. 2.6: Representation of the signals contained in the triple-resonance experiments CBCA(CO)NH and CBCANH. The signals of the C^β in the CBCANH are negative.

Neighboring residues are found by matching the ^{13}C chemical shifts of residue (i-1) (corresponding to one strip) to the ^{13}C chemical shifts of residue (i) which follows it in the sequence. This is demonstrated in Fig. 2.7.

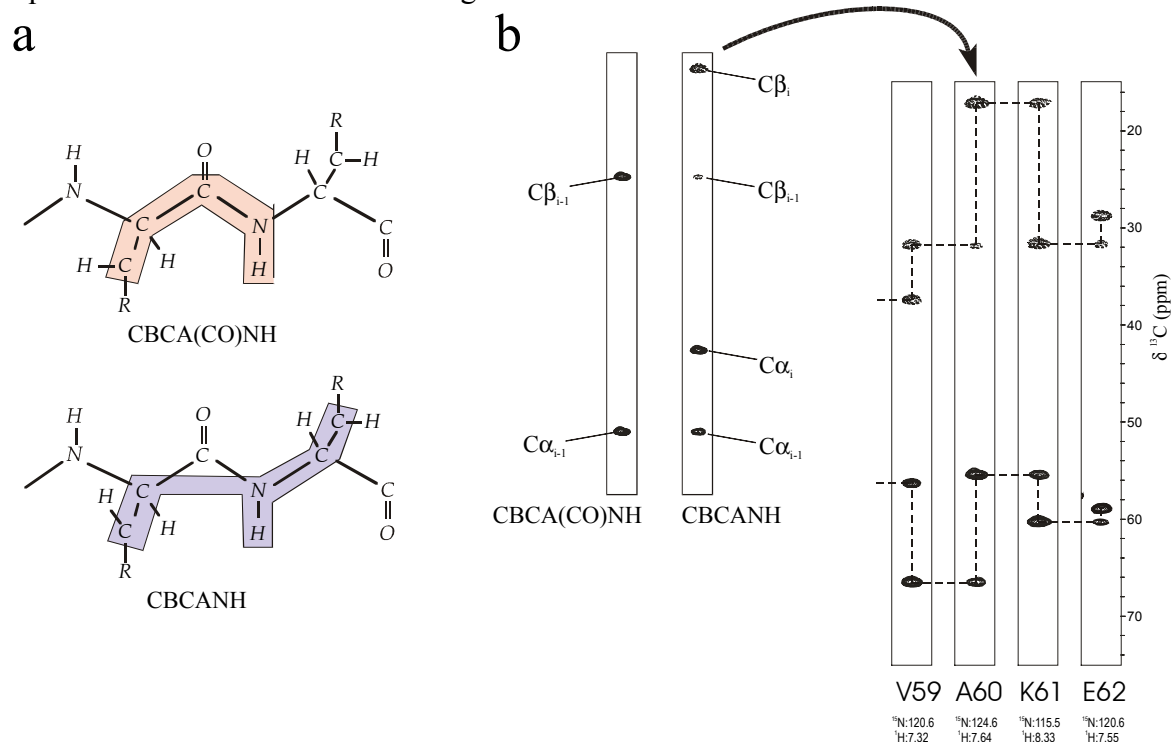


Fig. 2.7: Finding neighboring residues with the experiments CBCA(CO)NH and CBCANH by matching C β and C α chemical shifts of residue i and i-1. (a) Peptide topology with the spins involved in the two experiments highlighted. (b) On the left two "strips" of both experiments at a particular $^1\text{H}/^{15}\text{N}$ frequency pair are shown. On the right the matching of strips of the CBCANH is illustrated. The CBCA(CO)NH is only used to distinguish the ^{13}C frequencies of (i-1) from those of the i residues.

To obtain a robust assignment strategy, however, more 3D experiments should be used, for example a set of the 6 triple-resonance experiments (CBCA(CO)NH, CBCANH, HN(CA)CO, HNC(O), HA(CACO)NH and HA(CA)NH) which deliver the CB, CA, C' and HA frequencies of residues i and of previous residues i-1. All frequencies collected for a $^1\text{H}/^{15}\text{N}$ pair are called in the following a **generic spin system**. The experiments HNC(O) (Ikura et al., 1990; Kay et al., 1990) and HN(CA)CO (Clubb et al., 1992a) are used to obtain the carbonyl ^{13}C chemical shifts of residue i and i-1. Their magnetization transfer pathways are shown schematically in Fig. 2.7 a and b. The experiment pair HA(CACO)NH (Boucher et al., 1992; Montelione et al., 1999) and HA(CA)NH (Kay et al., 1991) or HN(CA)HA (Clubb et al., 1992b) deliver the HA frequencies of residue i and i-1. Most recent versions and a variety of other triple-resonance experiments are reviewed in the literature (Sattler et al., 1999; Montelione et al., 1999). Note that other combinations of triple-resonance experiments are possible. Even a strategy based on 4D experiments is feasible.

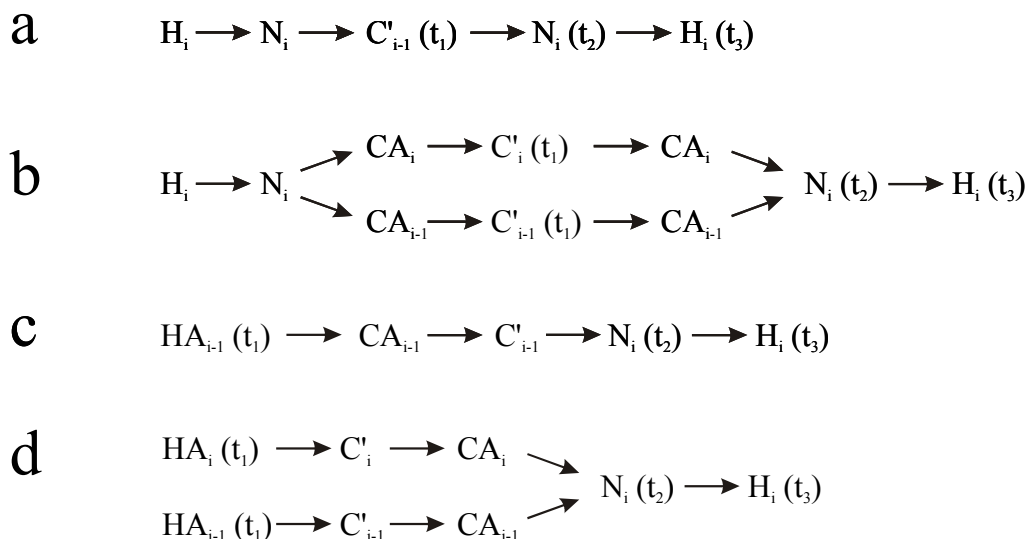


Fig. 2.8: Magnetization transfer pathways of the triple-resonance experiments (a) HNCO, (b) HN(CA)CO, (c) HA(CACO)NH and (d) HA(CA)NH. Magnetization is transferred using 1J coupling constants and in case b and d additionally $^2J_{NCA}$. The ^{13}C and ^{15}N chemical shift evolution periods are indicated with t_1 and t_2 . In (c) and (d) t_1 is the chemical shift evolution of the $^1H^\alpha$. t_3 is the 1H acquisition time. The indices i and $i-1$ indicate two neighboring residues in the amino acid sequence.

The pool of generic spin systems obtained from the set of six 3D experiments are shown schematically in Fig. 2.9 and can be compared to pieces of a jigsaw puzzle. Sequential neighboring residues, are found by matching the CB, CA, C' and HA chemical shifts, and linked up so that chains of residues are formed. Sometimes various pieces would fit to one particular piece and the connection to the neighbor is ambiguous. The chains are then matched to the protein sequence. The CB and CA chemical shifts are used to obtain information about the amino acid type (Grzesiek and Bax, 1993). A correlation of both chemical shifts results in a two-dimensional landscape which is shown in Fig. 2.10. Only in few cases the amino acid type can be identified unambiguously, which is the case for Ala, Ser and Thr. Probabilities for all amino acid types are obtained and the amino acid typing is usually ambiguous. However, there are procedures available which are able to handle probabilities of several amino acid types per residue for the sequence matching and the most probable assignments for a given chain are delivered (Grzesiek and Bax, 1993; Zimmerman et al., 1997).

The segments which have been assigned can be extended by finding matching spin patterns until the protein is completely assigned. Dynamics and exchange can lead to missing signals which prevent complete assignment. A correct backbone assignment is indicated by a full consistency of all data.

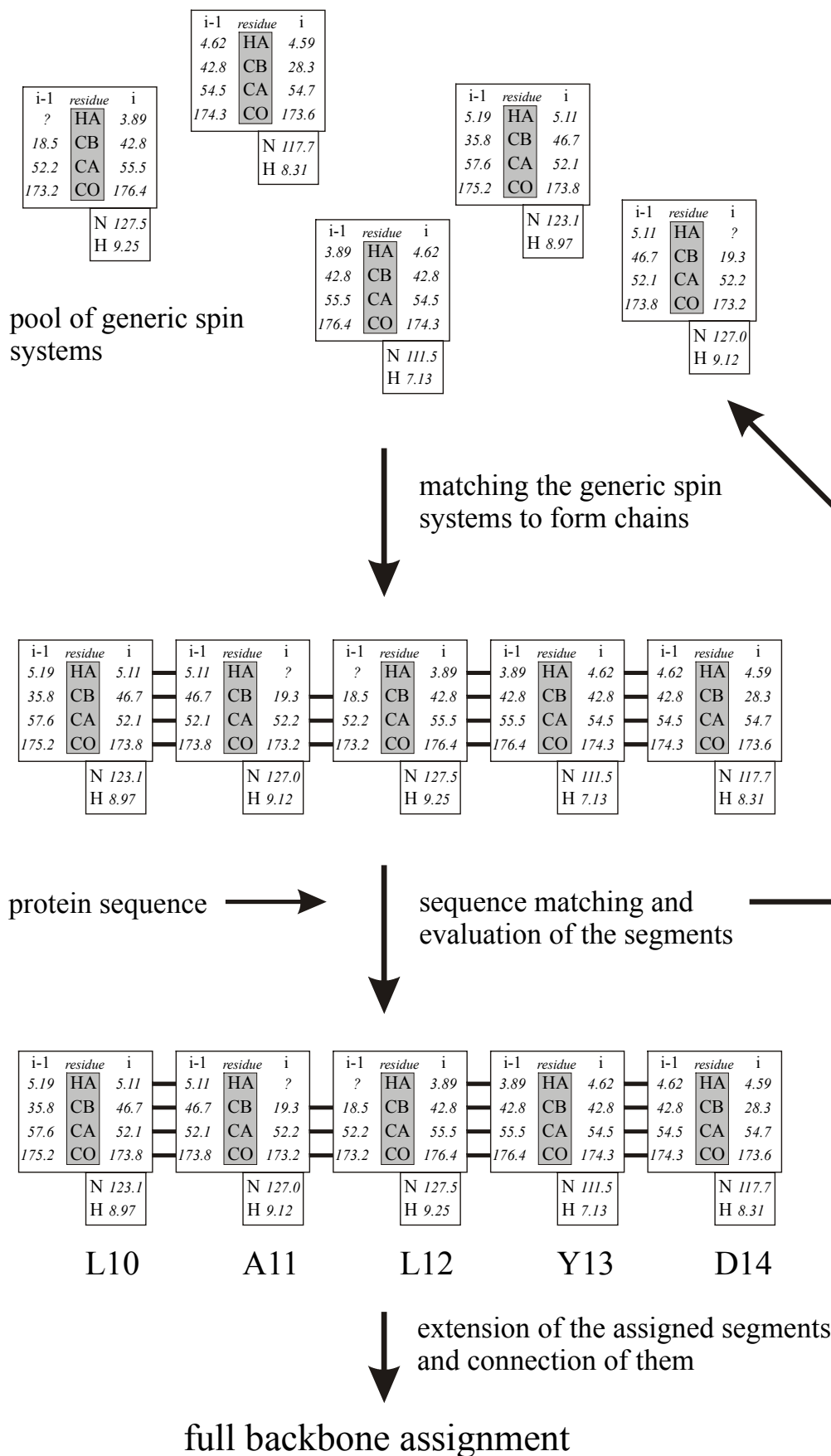


Fig. 2.9: Schematic presentation of the backbone assignment procedure. Missing frequencies which might be caused by missing signals in the spectra or incomplete water suppression are marked with "?".

The identification of the amino acid type is critical. All probable amino acid types for each residue of an unassigned segment have to be considered during the sequence mapping. This can lead to a combinatorial explosion of possible amino acid type combinations which need to be evaluated and scored during sequence mapping. An unambiguous determination of the amino acid type will greatly facilitate the assignment procedure and the development of experiments which deliver this information has been the goal of the present thesis. They are presented in chapter 4.

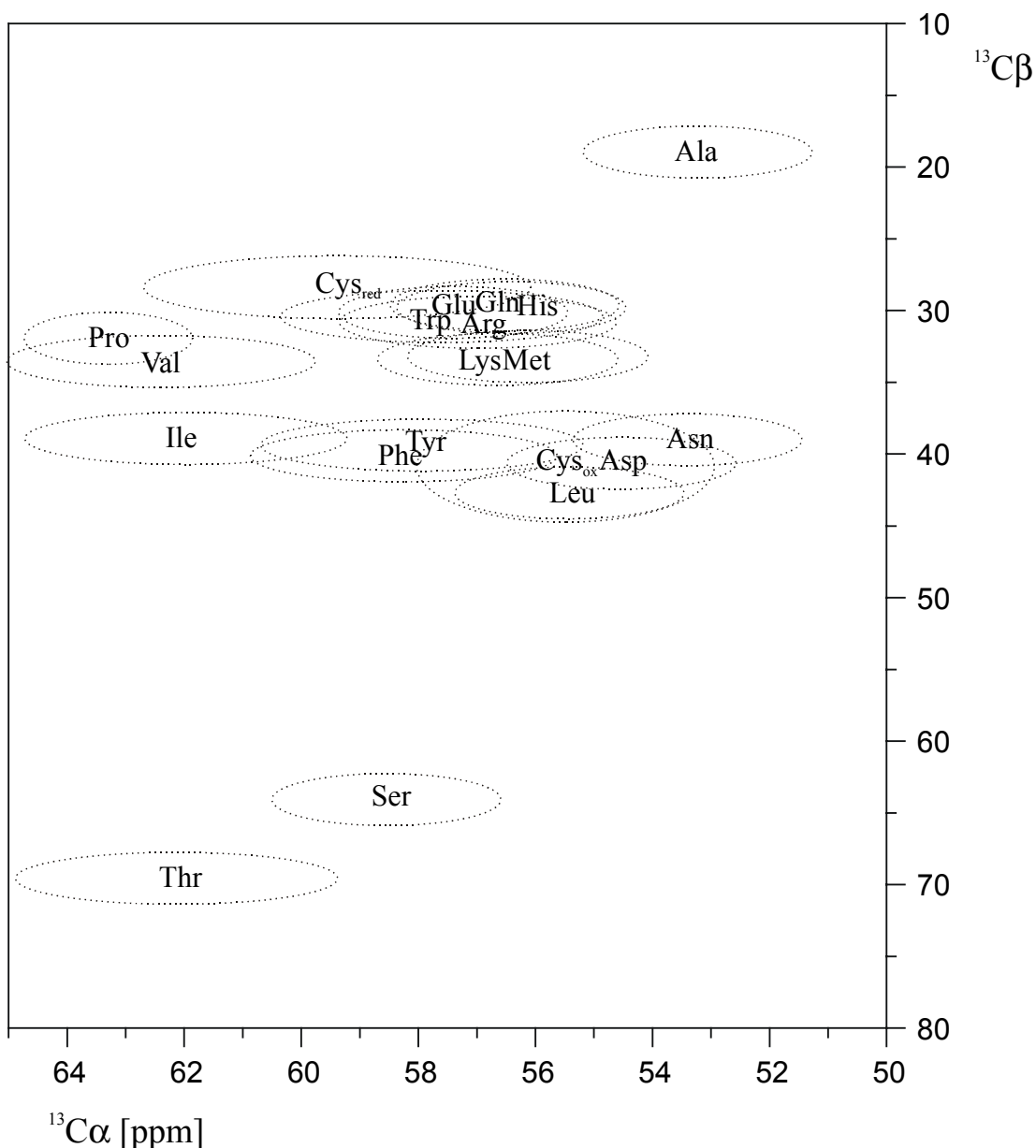


Fig. 2.10: Two-dimensional landscape of the CB and CA chemical shifts with the characteristic areas for the different amino acid types. The ellipses show approximately the standard deviation of the chemical shift values found in the BMRB database (Seavey et al., 1991). Data for Cys in the oxidized and reduced form are taken from Sharma and Rajarathnam (D. Sharma and K. Rajarathnam, 2000).

Proline residues hinder sequential assignment strategies based on a ^1H - ^{15}N correlation because proline lacks a NH group. A sequential assignment then has to rely either on the observation of NOEs from proline to neighboring residues or on triple-resonance experiments which do not involve a NH. A couple of triple-resonance experiments have been designed just for this case, the HACAC(CO)(N) (Olejniczak and Fesik, 1994), the HDHA(NCO)CAHA (Bottomley et al., 1999) and the two 3D experiments HCAN and (HB)CBCA(CO)N(CA)HA (Kanelis et al., 2000). Additional proline selective experiments are presented in chapter 4.

Several approaches to automate the sequential backbone assignment are discussed in chapter 5.

2.4 Frontiers of Biological NMR: Application of TROSY on large uniformly $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled proteins

With increasing protein size not only the number of signals but also the linewidths increase due to increased relaxation which is caused by the slower tumbling times of larger molecules. Standard triple-resonance experiments mentioned in chapter 2.4 experience an exponentially progressive decrease in sensitivity. Proton-proton spin diffusion is one of the largest contributions to the fast relaxation. It can be eliminated by exchanging the ^1H nuclei by ^2H nuclei. Deuteration also decreases ^{13}C relaxation and thus ^{13}C line widths. A deuteron is 42-fold ($\gamma_{\text{H}}^2/\gamma_{\text{D}}^2$) less effective than a proton at causing dipolar relaxation of an attached ^{13}C and the surrounding ^1H .

A variety of deuteration schemes are in use, including random fractional deuteration, uniform deuteration and selective deuteration. Several reviews describe the methods in detail (Gardner & Kay, 1999; Farmer & Venters, 1999). For the backbone assignment of large proteins a uniformly $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled sample in water is usually used. Only exchangeable protons like in NH and NH_2 are present. The assignment strategy is in principle identical to the one for $^{13}\text{C}/^{15}\text{N}$ labeled proteins, however because of missing protons attached to ^{13}C only triple-resonance experiments of the "out and back" type, starting and ending on $^1\text{H}^{\text{N}}$, are applicable. Several triple-resonance experiments were developed for $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled proteins which are reviewed in the literature (Gardner & Kay, 1999; Farmer & Venters, 1999). The magnetization transfer pathways of the experiments used in this work are given in Fig. 2.6.

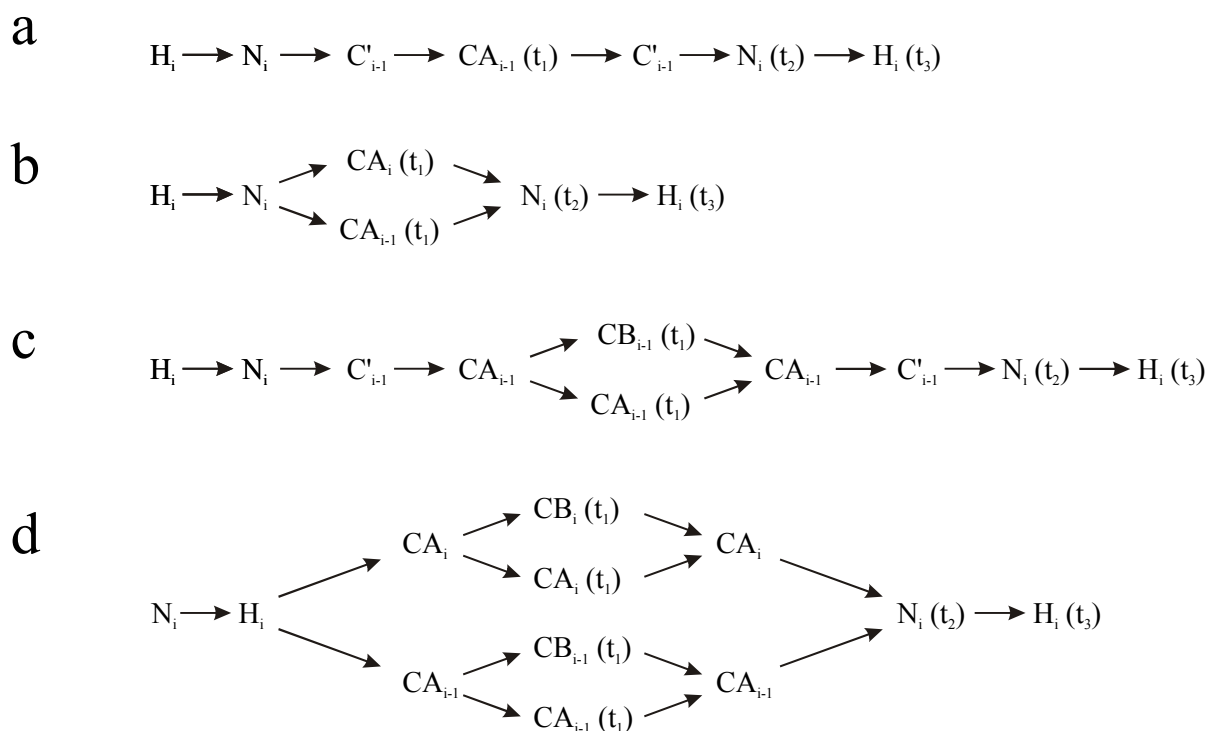


Fig. 2.11: Magnetization transfer pathways of the common triple-resonance experiments used for the sequential backbone assignment of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled proteins, (a) HN(CO)CA, (b) HNCA, (c) HN(CO)CACB, (d) HNCACB. The pathways for the HNCO and HN(CA)CO are given in Fig. 2.5. Magnetization is transferred using ^1J coupling constants and in case b and d additionally $^2\text{J}_{\text{NCA}}$. The ^{13}C and ^{15}N chemical shift evolution periods are indicated with t_1 and t_2 . t_3 is the ^1H acquisition time. The indices i and $i-1$ indicate two neighboring residues in the amino acid sequence.

The TROSY (transverse relaxation-optimized spectroscopy) technique has been developed to reduce relaxation losses during the chemical shift evolution of a heteronucleus X (e. g. ^{15}N), the $X \rightarrow ^1\text{H}$ magnetization transfer and the acquisition time (Pervushin et al., 1997). The longitudinal relaxation rate (R_1) characterizes the return of the z -component of precessing magnetization to thermal equilibrium and the transverse relaxation rate (R_2) characterizes the dephasing of the transverse components. R_2 directly determines the line width of NMR signals and thus indirectly the signal intensity in multidimensional experiments. The transverse relaxation rate is mainly caused by dipole-dipole (DD) coupling and chemical shift anisotropy (CSA). Considering the components of a doublet of I in a weakly coupled two spin system IS the transverse relaxation rates are different due to addition or subtraction of the influence of DD coupling and the CSA. A narrow and a broad component is the result. In a non-decoupled two-dimensional ^{15}N -HSQC spectrum only one component of the quartet has a narrow line width in both dimensions (the TROSY component) as shown in Fig. 2.12. TROSY experiments select solely these narrow components and suppress the broad components of the quartets. Note that the signals in a TROSY experiment are shifted in both dimensions by $1/2J_{\text{NH}}$. To summarize, TROSY

suppresses transverse relaxation, e. g. in ^{15}N - ^1H moieties by constructive use of interference between dipole-dipole (DD) coupling and chemical shift anisotropy (CSA).

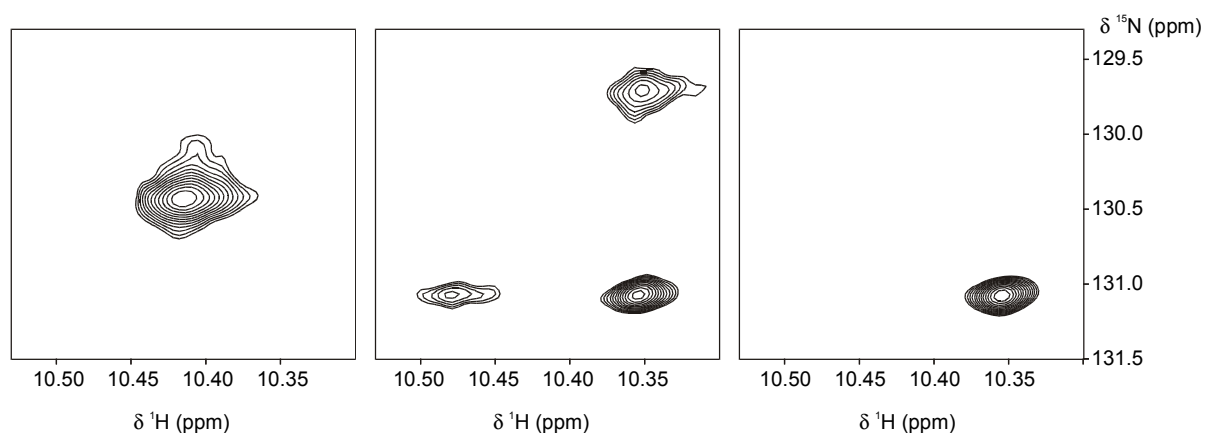


Fig. 2.12: Regions of ^1H - ^{15}N correlations of $^2\text{H}/^{15}\text{N}$ labeled Bacteriorhodopsin solubilized in detergent micelles (chapter 6). Shown is a signal of a NH moiety: (a) conventional ^{15}N -HSQC with decoupling in both dimensions, (b) ^{15}N -HSQC without decoupling in both dimensions and (c) ^{15}N TROSY which selects only the narrow component.

triple-resonance experiments using TROSY	dimensionality	references
HNCA	3D	Salzmann et al., 1998, 1999a, 1999b; Permi and Annala, 2001; Meissner and Sørensen, 2001a.
HNCO	3D	Salzmann et al., 1998; Yang and Kay, 1999a; Loria et al., 1999.
HN(CO)CA	3D	Salzmann et al., 1999c.
HN(CA)CO	3D	Salzmann et al., 1999c.
HNCACB	3D	Salzmann et al., 1999c.; Meissner and Sørensen, 2001b.
HN(CO)CACB	3D	Salzmann et al., 1999c.
HNCOCA	4D	Yang and Kay, 1999b.
HNCACO	4D	Yang and Kay, 1999b.
HNCO _{i-1} CA	4D	Konrat et al., 1999.
HNCAN	3D	Löhr et al., 2000.

Table 2.1: TROSY triple-resonance experiments for sequential resonance assignment of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled proteins.

The TROSY technique can be implemented in every NH detected triple-resonance experiment and TROSY versions of most of the above mentioned triple-resonance

experiments were published. Table 2.1 gives an overview. The advantages are higher sensitivities of the experiments and smaller half height line widths. Therefore the separation of signals in the ^{15}N - ^1H plane is better (Riek et al., 2000).

The application of the TROSY technique to the membrane protein Bacteriorhodopsin solubilized in detergent micelles is presented in chapter 6.

References

- W. P. Aue, E. Bartholdi, and R. R. Ernst, *J. Chem. Phys.* **64**, 2229-2246 (1976).
- A. Bax and D. G. Davis, *J. Magn. Reson.* **65**, 355-360 (1985).
- F. J. Blanco, A. R. Ortiz and L. Serrano, *J. Biomol. NMR* **9**, 347 (1997).
- W. Boucher, E. D. Laue, S. Campbell-Burk and P. J. Dommelle, *J. Am. Chem. Soc.* **114**, 2262-2264 (1992).
- R. Braunschweiler and R. R. Ernst, *J. Magn. Reson.* **53**, 521-528 (1983).
- R. T. Clubb, V. Thanabal and G. Wagner, *J. Magn. Reson.* **97**, 213-217 (1992a).
- R. T. Clubb, V. Thanabal and G. Wagner, *J. Biomol. NMR* **2**, 203-210 (1992b).
- R. R. Ernst and W. A. Anderson, *Rev. Sci. Instr.* **37**, 93-102 (1966).
- B. T. Farmer II and R. A. Venters, *Biol. Magn. Reson.* **16**, 75-120 (1999).
- K. H. Gardner and L. E. Kay, *Biol. Magn. Reson.* **16**, 27-74 (1999).
- C. Griesinger, O. W. Sørensen and R. R. Ernst, *J. Magn. Reson.* **73**, 574-579 (1987a).
- C. Griesinger, O. W. Sørensen and R. R. Ernst, *J. Am. Chem. Soc.* **109**, 7227-7228 (1987b).
- S. Grzesiek and A. Bax, *J. Am. Chem. Soc.* **114**, 6291-6293 (1992a).
- S. Grzesiek and A. Bax, *J. Magn. Reson. B* **99**, 201-207 (1992b).
- S. Grzesiek and A. Bax, *J. Biomol. NMR* **3**, 185-204 (1993).
- M. Ikura, L. E. Kay and A. Bax, *Biochemistry* **29**, 4659-4667 (1990).
- J. Jeener, *Amp. Int. Summerschool II, Basko Polje, Yugoslavia* (1971).
- J. Jeener, B. H. Meier, P. Bachmann and R. R. Ernst, *J. Chem. Phys.* **71**, 4546-4553 (1979).
- V. Kanelis, L. Donaldson, D. R. Muhandiram, D. Rotin, J. D. Forman-Kay and L. E. Kay, *J. Biomol. NMR* **16**, 253-259 (2000).
- H. Kessler, P. Schmieder and W. Bermel, *Biopolymers* **30**, 465-475 (1990).
- L. E. Kay, M. Ikura, R. Tschudin and A. Bax, *J. Magn. Reson.* **89**, 496-514 (1990).
- L. E. Kay, M. Ikura and A. Bax, *J. Magn. Reson.* **91**, 84-92 (1991).
- R. Konrat, D. Yang and L. E. Kay, *J. Biomol. NMR* **15**, 309-313 (1999).
- F. Löhr, S. Pfeiffer, Y.-J. Lin, J. Hartleib, O. Klimmek and H. Rüterjans, *J. Biomol. NMR* **18**, 337-346 (2000).
- J. P. Loria, M. Rance and A. G. Palmer III, *J. Magn. Reson.* **141**, 180-184 (1999).
- L. P. McIntosh and F. W. Dahlquist, *Quart. Rev. Biophys.* **23**, 1-38 (1990).
- A. Meissner and O. W. Sørensen, *J. Magn. Reson.* **150**, 100-104 (2001a).
- A. Meissner and O. W. Sørensen, *J. Magn. Reson.* **151**, 328-331 (2001b).
- G. T. Montelione and G. Wagner, *J. Magn. Reson.* **87**, 183-188 (1990).
- A. Musacchio, M. Noble, R. Paupit, R. Wierenga and M. Serraste, *Nature* **359**, 851-855 (1992).
- H. Oschkinat, C. Griesinger, P. J. Kraulis, O. W. Sørensen, R. R. Ernst, A. Gronenborn and G. M. Glore, *Nature* **332**, 374-376 (1988).
- P. Permi and A. Annala, *J. Biomol. NMR* **20**, 127-133 (2001).
- K. Pervushin, R. Riek, G. Wider and K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **94**, 12366-12371 (1997).

- R. Riek, K. Pervushin and K. Wüthrich, *Trends Biochem. Sci.* **25**, 462-468 (2000).
- M. Salzmann, K. Pervushin, G. Wider, H. Senn and K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **95**, 13585-13590 (1998).
- M. Salzmann, K. Pervushin, G. Wider, H. Senn and K. Wüthrich, *J. Biomol. NMR* **14**, 85-88 (1999a).
- M. Salzmann, G. Wider, K. Pervushin and K. Wüthrich, *J. Biomol. NMR* **15**, 181-184 (1999b).
- M. Salzmann, G. Wider, K. Pervushin, H. Senn and K. Wüthrich, *J. Am. Chem. Soc.* **121**, 844-848 (1999c).
- M. Sattler, J. Schleucher and C. Griesinger, *Prog. NMR Spectrosc.* **34**, 93-158 (1999).
- B. R. Seavey, E. A. Farr, W. M. Westler and J. Markley, *J. Biomol. NMR* **1**, 217-236 (1991);
<http://www.bmrwisc.edu>
- D. Sharma and K. Rajarathnam, *J. Biomol. NMR* **18**, 165-171 (2000).
- M. Smalla, P. Schmieder, M. J. S. Kelly, A. ter Laak, G. Krause, L. J. Ball, M. Wahl, P. Bork and H. Oschkinat, *Protein Sci.* **8**, 1954-1961 (1999).
- M. P. Williamson, T. F. Havel and K. Wüthrich, *J. Mol. Biol.* **182**, 295-315 (1985).
- K. Wüthrich, G. Wider, G. Wagner and W. Braun, *J. Mol. Biol.* **155**, 311-319 (1982).
- K. Wüthrich *NMR of Proteins and Nucleic Acids*, Wiley, New York (1986).
- D. Yang and L. E. Kay, *J. Biomol. NMR* **13**, 3-10 (1999a).
- D. Yang and L. E. Kay, *J. Am. Chem. Soc.* **121**, 2571-2575 (1999b).
- D. E. Zimmerman, C. A. Kulikowski, Y. Huang, W. Feng, M. Tashiro, S. Shimotakahara, C. Chien, R. Powers and G. Montelione, *J. Mol. Biol.* **269**, 592-610 (1997).

Chapter 3

MUSIC with shaped pulses and tuned delays: An introduction to the techniques used to create amino acid type-selective NMR experiments

Parts of this chapter were published in *J. Magn. Reson.* **141**, 34-43 (1999).

3.1 Introduction

This chapter gives an introduction into the basic techniques of triple-resonance experiments and into selection mechanisms which are used in chapter 4 to create amino acid type-selective NMR experiments. Basic building blocks to achieve the desired magnetization transfers are discussed in chapter 3.2. More sophisticated techniques which are able to edit multiplicities are introduced in chapter 3.3. MUSIC, a novel powerful building block which transfers magnetization and selects multiplicities is discussed in chapter 3.4. A short introduction into selective pulses is given in chapter 3.5. It includes an analysis of the evolution of spin-spin coupling during long selective pulses. How carefully tuned delays can select specific coupling topologies is described in chapter 3.6.

3.2 Basic building blocks in triple-resonance experiments

The common triple-resonance experiments consist of a combination of simple building blocks. The most common ones are introduced in Fig. 3.1.

INEPT (insensitive nuclei enhanced by polarization transfer) (Morris and Freeman, 1979) is the standard magnetization transfer between heteronuclear spins (Fig 3.1 a). A delay 2Δ is set between $1/4J_{IS}$ and $1/2J_{IS}$ optimized to minimize relaxation losses. A simple chemical shift evolution in the indirect dimension is shown in Fig. 3.1b. A 180° pulse in the center of the chemical shift evolution period refocuses J coupling to a coupled spin I. A constant time chemical shift evolution combined with a magnetization transfer between two different types of nuclei is shown in Fig. 3.1c. The optimal value for the delay 2Δ lies in between $1/4J_{IS}$ and $1/2J_{IS}$ depending on the relaxation losses. A constant time chemical shift evolution during which J-coupling between S and I is refocused is presented in Fig. 3.1d. A homonuclear through-bond magnetization transfer is the COSY (correlation spectroscopy) step (Jeener, 1971; Aue et al., 1976). A refocused COSY transfer without chemical shift evolution is shown in Fig. 3.1e. The two pairs of delays Δ are separated by 180° pulses to refocus chemical shift evolution. The tuning of the delays Δ_1 and Δ_2 is discussed in detail in chapter

3.6. Total correlation spectroscopy (TOCSY) (Fig. 3.1f) is a technique which transfers in-phase magnetization to all J-coupled homonuclear nuclei within one spin system (Braunschweiler and Ernst, 1983; Bax and Davis, 1985). Different pulse trains were designed for that purpose which is reviewed in the literature (Glaser and Quant, 1996). Pulse trains designed for broad band decoupling of a nucleus type are for example WALTZ-16 (Shaka et al., 1983) and GARP-1 (Shaka et al., 1985) (Fig. 3.1g).

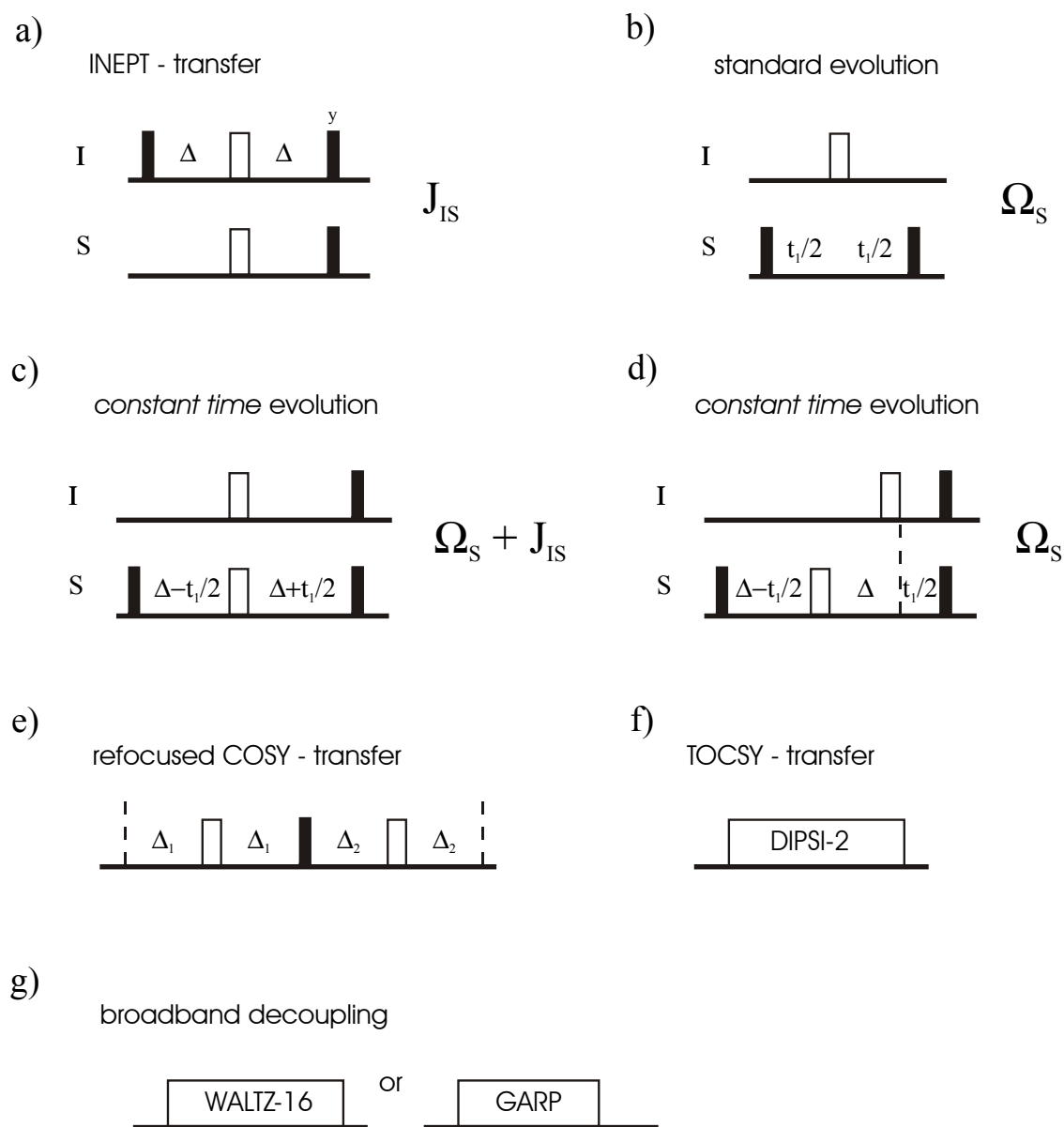


Fig. 3.1: The most common building blocks used in triple resonance experiments. 90° and 180° pulses are represented by thin black filled and thick unfilled bars, respectively. (a) The INEPT transfer, (b) a standard chemical shift evolution time during which J-coupling between spin I and S is refocused, (c) a constant time chemical shift evolution time during which J-coupling between spin I and S is active, (d) a constant time evolution time during which J-coupling between spin I and S is refocused, (e) a refocused COSY transfer (without chemical shift evolution), (f) a TOCSY transfer and (g) two broadband decoupling schemes. WALTZ-16 is widely used to decouple ^1H and GARP to decouple heteronuclei.

3.3 Spectral editing techniques - The precursors of MUSIC

Several techniques were developed to obtain multiplicity information of a certain heteronucleus (SI , SI_2 or SI_3). The different multiplicities can be distinguished by different signs of the signals which is called editing. The most common editing techniques are DEPT (Doddrell et al., 1982) and refocused INEPT (Morris and Freeman, 1979) whose pulse sequences are shown in Fig. 3.2. POMMIE is a rarely used technique (Bulsing et al., 1984) but it offers some advantages.

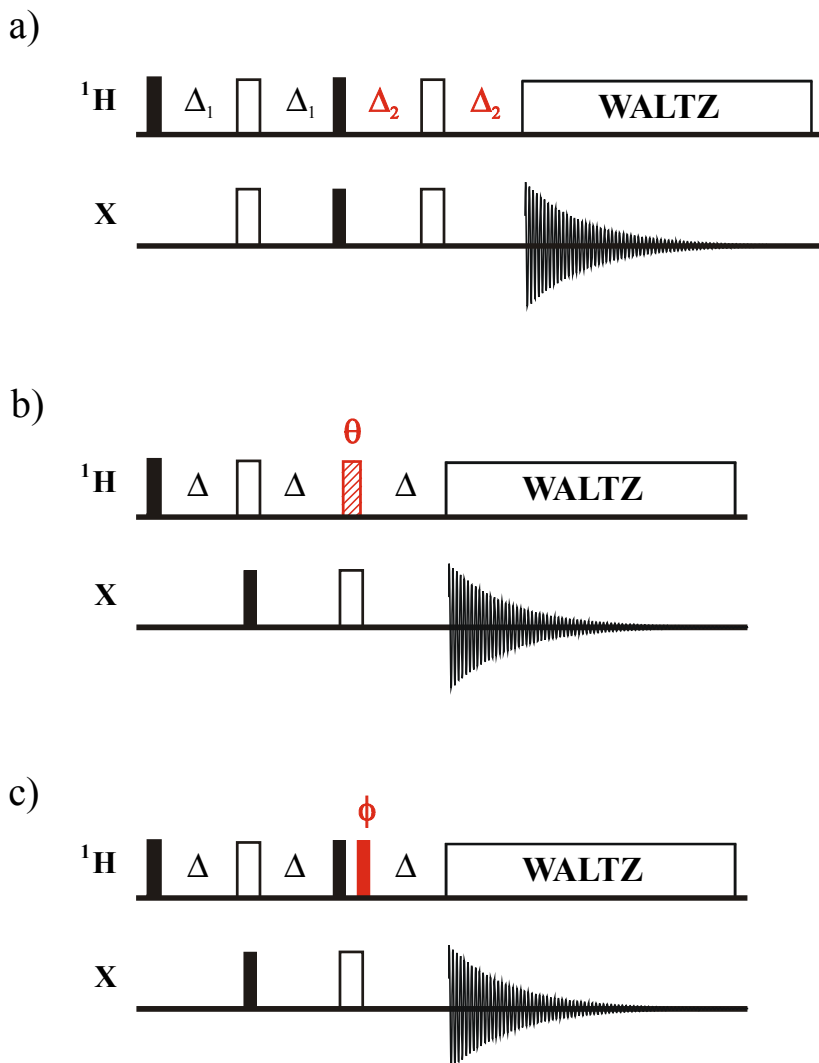


Fig. 3.2: Editing techniques: (a) refocused INEPT, (b) DEPT and (c) POMMIE.

The INEPT element in Fig. 3.1a leads to antiphase magnetization $-2I_zS_y$. A refocusing step has to follow to convert the antiphase magnetization into in-phase magnetization S_x . The different behavior of the multiplicities can be described by the product operator formalism (Sørensen et al., 1982; Sørensen and Ernst, 1983). Irrelevant terms are omitted.

$$\mathbf{SI}_1: \quad -2I_z S_y \xrightarrow{2\Delta_2} S_x \sin \pi J_{IS} 2\Delta_2$$

$$\mathbf{SI}_2: \quad -2I_z S_y \xrightarrow{2\Delta_2} S_x \sin \pi J_{IS} 2\Delta_2 \cos \pi J_{IS} 2\Delta_2$$

$$\mathbf{SI}_3: \quad -2I_z S_y \xrightarrow{2\Delta_2} S_x \sin \pi J_{IS} 2\Delta_2 \cos^2 \pi J_{IS} 2\Delta_2$$

The signal intensities of the different multiplicities are a function of $\pi J_{IS} 2\Delta_2$ and thus of the delay Δ_2 (Fig. 3.3). By appropriate adjustment of the parameter Δ_2 different signs of the signals depending on the multiplicities are obtained.

In the DEPT experiment the pulse length θ determines the sign of a particular multiplicity. The delay Δ is set to $1/2J_{IS}$. A detailed description of the experiment in the product operator formalism is published elsewhere (Sørensen and Ernst, 1983; Chandrakumar and Subramanian, 1987). The intensities as a function of θ are given in the following. Irrelevant terms are omitted.

$$\mathbf{SI}_1: \quad S_x \sin \theta$$

$$\mathbf{SI}_2: \quad 2 S_x \sin \theta \cos \theta$$

$$\mathbf{SI}_3: \quad 3 S_x \sin \theta \cos^2 \theta$$

The functions are shown in Fig. 3.3.

The POMMIE experiment is able to edit multiplicities by changing the phase ϕ of a 90° pulse. The intensity dependence on ϕ is almost identical to the DEPT sequence. The only difference is that the editing parameter is the pulse angle in the DEPT and a pulse phase in the POMMIE sequence (Fig. 3.3):

$$\mathbf{SI}_1: \quad S_x \sin \phi$$

$$\mathbf{SI}_2: \quad 2 S_x \sin \phi \cos \phi$$

$$\mathbf{SI}_3: \quad 3 S_x \sin \phi \cos^2 \phi$$

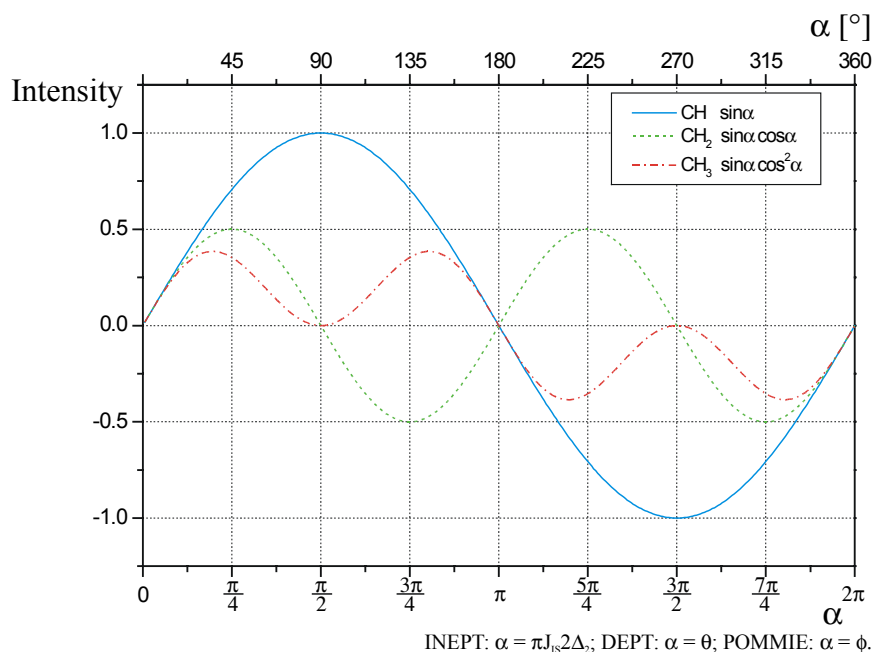


Fig. 3.3: Dependence of the signal intensities of the particular multiplicities as a function of the editing parameter α which is identical to the pulse angle θ in the DEPT experiment and to the phase ϕ in the POMMIE experiment. In the refocused INEPT α is correlated to the refocusing delay Δ_2 : $\alpha = \pi J_{IS} 2\Delta_2$. Note that the curves show the intensities resulting from one excited proton. Since there are two or three of them in a CH_2 , CH_3 group, respectively, these intensities have to be multiplied by 2 or 3, respectively.

Several attempts have been made not only to edit, but also to select a particular multiplicity. The difference between editing and selecting multiplicities is that editing just gives signals of certain multiplicities different signs while selecting results in a suppression of undesired multiplicities. A particular multiplicity selection requires a linear combination of several experiments with a different parameter α . An example with the DEPT experiment using three different pulse angles is given by Doddrell et al. (Doddrell et al., 1982). The individual experiments had to be scaled before they were added or subtracted. The selection is not completely clean, residual signals of undesired multiplicities are present. The shortcomings mainly result from different J_{IS} coupling constants in the sample and an incorrect setting of the flip angle. To ensure a correct flip angle is very difficult since it depends on the offset. Even if the flip angle is exactly 90° SI_2 and SI_3 multiplicities will not be totally suppressed due to different J_{IS} couplings. This problem is known as “J-cross-talk” (Sørensen and Ernst, 1983).

Deviations of J_{IS} couplings are even more severe in the refocused INEPT, since it leads to phase distortions and to different values of α for a given delay. Linear combinations of spectra with different delays Δ_2 for the selection of a particular multiplicity do therefore not lead to satisfying results.

Only the POMMIE sequence allows exact adjustment of the parameter α since in this case it is the phase of a 90° pulse. Adding FID’s with different α values is accomplished simply

by phase cycling. POMMIE extended by a proper phase cycle is the only sequence among those three which allows the selection of proton double- and triple-quantum coherence.

3.4 MUSIC: A multiplicity selective in-phase coherence transfer

MUSIC (multiplicity selective in-phase coherence transfer) is based on the POMMIE sequence and selects ^1H double and triple quantum coherence by a proper phase cycle. ^1H triple quantum coherence can only be formed by a XH_3 group ($\text{X} = \text{spin } \frac{1}{2} \text{ nucleus}$) whereas double quantum coherence can be formed by XH_2 but potentially also by XH_3 . MUSIC selects higher multiplicities and suppresses lower ones in contrast to editing techniques which just give opposite signs for different multiplicities.

In the following, multiplicity selection with ^{13}C as the heteronucleus is discussed by the product operator formalism. At the beginning n quantum coherence has to be excited efficiently. In order to select ^1H n quantum coherence a phase cycle with π/n phase shifts is necessary.

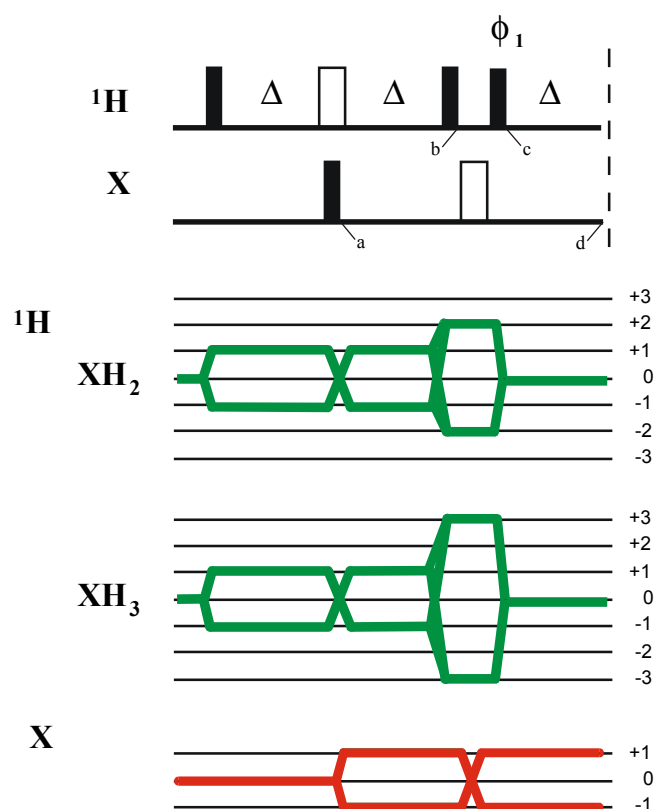


Fig. 3.4: The MUSIC pulse sequence is shown on the top, the coherence pathway diagrams for ^1H and X at the bottom. Δ is set to $1/2J_{\text{HX}}$. The phase cycle for the selection of XH_2 is: $\phi_1 = 45^\circ, 135^\circ, 225^\circ, 315^\circ, \phi_{\text{rec}} = 0^\circ, 180^\circ$; for the selection of XH_3 it is: $\phi_1 = 30^\circ, 90^\circ, 150^\circ, 210^\circ, 270^\circ, 330^\circ, \phi_{\text{rec}} = 0^\circ, 180^\circ$. The phase cycle can be combined with conventional phase cycling used in standard triple-resonance experiments. To further improve the performance of the sequence, three gradients of equal sign and strength can be implemented in each of the three delays. These gradients reject imperfections of the two 180° pulses.

The transfer function up to point a is:

$$\mathbf{H}_{1z} \xrightarrow{90^\circ_x (^1\text{H})} -\mathbf{H}_{1y} \xrightarrow{\Delta} 2 \mathbf{H}_{1x} \mathbf{C}_z \sin(\pi J_{\text{CH}} \Delta) \xrightarrow[90^\circ_x (^{13}\text{C})]{180^\circ_y (^1\text{H})} 2 \mathbf{H}_{1x} \mathbf{C}_y \sin(\pi J_{\text{CH}} \Delta)$$

Note that till this point the transfer functions of CH, CH₂ and CH₃ are identical except the fact that in CH₂ and CH₃ the second and third proton H₂ and H₃ can also be excited leading to the same transfer function where just H₁ is replaced by H₂ or H₃, respectively. The delay Δ is set to $1/2J_{\text{CH}}$ and the sine term is equal to 1. For the sake of clarity $\sin(\pi J_{\text{CH}} \Delta)$ terms are omitted in the following. The different multiplicities are discussed separately up to point b. Chemical shift evolution of ¹³C is not taken into account since it is refocused at the end of MUSIC. Coupling to other ¹³C nuclei is omitted too.

$$\text{CH: } 2 \mathbf{H}_{1x} \mathbf{C}_y \xrightarrow{\Delta} 2 \mathbf{H}_{1x} \mathbf{C}_y \xrightarrow{90^\circ_x (^1\text{H})} 2 \mathbf{H}_{1x} \mathbf{C}_y = (\mathbf{H}_1^+ + \mathbf{H}_1^-) \mathbf{C}_y$$

$$\begin{aligned} \text{CH}_2: \quad 2 \mathbf{H}_{1x} \mathbf{C}_y + 2 \mathbf{H}_{2x} \mathbf{C}_y &\xrightarrow{\Delta} -4 \mathbf{H}_{1x} \mathbf{H}_{2z} \mathbf{C}_x - 4 \mathbf{H}_{1z} \mathbf{H}_{2x} \mathbf{C}_x \xrightarrow{90^\circ_x (^1\text{H})} 4 \mathbf{H}_{1x} \mathbf{H}_{2y} \mathbf{C}_x + 4 \mathbf{H}_{1y} \mathbf{H}_{2x} \mathbf{C}_x \\ &= 2 (2 \mathbf{H}_{1x} \mathbf{H}_{2y} + 2 \mathbf{H}_{1y} \mathbf{H}_{2x}) \mathbf{C}_x = \frac{2}{i} (\mathbf{H}_1^+ \mathbf{H}_2^+ - \mathbf{H}_1^- \mathbf{H}_2^-) \mathbf{C}_x \end{aligned}$$

$$\begin{aligned} \text{CH}_3: \quad 2 \mathbf{H}_{1x} \mathbf{C}_y + 2 \mathbf{H}_{2x} \mathbf{C}_y + 2 \mathbf{H}_{3x} \mathbf{C}_y &\xrightarrow{\Delta} -8 \mathbf{H}_{1x} \mathbf{H}_{2z} \mathbf{H}_{3z} \mathbf{C}_y - 8 \mathbf{H}_{1z} \mathbf{H}_{2x} \mathbf{H}_{3z} \mathbf{C}_y - 8 \mathbf{H}_{1z} \mathbf{H}_{2z} \mathbf{H}_{3x} \mathbf{C}_y \\ &\xrightarrow{90^\circ_x (^1\text{H})} -8 \mathbf{H}_{1x} \mathbf{H}_{2y} \mathbf{H}_{3y} \mathbf{C}_y - 8 \mathbf{H}_{1y} \mathbf{H}_{2x} \mathbf{H}_{3y} \mathbf{C}_y - 8 \mathbf{H}_{1y} \mathbf{H}_{2y} \mathbf{H}_{3x} \mathbf{C}_y \\ &= 3 (\mathbf{H}_1^+ \mathbf{H}_2^+ \mathbf{H}_3^+ + \mathbf{H}_1^- \mathbf{H}_2^- \mathbf{H}_3^-) \mathbf{C}_y - (\mathbf{H}_1^- \mathbf{H}_2^+ \mathbf{H}_3^+ + \mathbf{H}_1^+ \mathbf{H}_2^- \mathbf{H}_3^-) \mathbf{C}_y \\ &\quad - (\mathbf{H}_1^+ \mathbf{H}_2^- \mathbf{H}_3^+ + \mathbf{H}_1^- \mathbf{H}_2^+ \mathbf{H}_3^-) \mathbf{C}_y - (\mathbf{H}_1^+ \mathbf{H}_2^+ \mathbf{H}_3^- + \mathbf{H}_1^- \mathbf{H}_2^- \mathbf{H}_3^+) \mathbf{C}_y \end{aligned}$$

CH groups stay as ¹H single quantum coherence till point b, methylene groups form 100% ¹H double quantum coherence and methyl groups form a mixture of ¹H triple quantum coherence and three spin single quantum coherences.

The function of the consecutive 90° pulse is to convert the multiple quantum coherences of a particular multiplicity to antiphase single quantum coherence $2 \mathbf{H}_{1z} \mathbf{C}_y$, $4 \mathbf{H}_{1z} \mathbf{H}_{2z} \mathbf{C}_x$, $8 \mathbf{H}_{1z} \mathbf{H}_{2z} \mathbf{H}_{3z} \mathbf{C}_y$, respectively and to select ¹H double or triple quantum coherence by a proper phase cycle. The selection of n-quantum coherence requires a 2n step phase cycle with relative phases of $\Delta\phi = 2\pi k/2n$ [$k = 0, 1, \dots, 2n-1$] where n is the order of the multiple quantum coherence (Braunschweiler et al., 1983; Wokaun and Ernst, 1977). To select (± 2) quantum coherence a 4 step phase cycle with 90° phase shifts has to be applied. To select

(±3) quantum coherence a 6 step phase cycle with 60° phase shifts has to be used. To determine the optimal absolute phases in the MUSIC experiment the transfer function is derived with unknown phase ϕ .

$$\text{CH: } 2 \mathbf{H}_{1x} \mathbf{C}_y \xrightarrow{90^\circ_\phi (^1\text{H})} 2 (\mathbf{H}_{1x} \cos^2\phi + \mathbf{H}_{1y} \cos\phi \sin\phi - \mathbf{H}_{1z} \sin\phi) \mathbf{C}_y$$

$$\begin{aligned} \text{CH}_2: \quad 4 \mathbf{H}_{1x} \mathbf{H}_{2y} \mathbf{C}_x + 4 \mathbf{H}_{1y} \mathbf{H}_{2x} \mathbf{C}_x &\xrightarrow{90^\circ_\phi (^1\text{H})} 4 (\mathbf{H}_{1x} \cos^2\phi + \mathbf{H}_{1y} \cos\phi \sin\phi - \mathbf{H}_{1z} \sin\phi) \times \\ &\quad (\mathbf{H}_{2x} \cos\phi \sin\phi + \mathbf{H}_{2y} \sin^2\phi + \mathbf{H}_{2z} \cos\phi) \mathbf{C}_x \\ &\quad + 4 (\mathbf{H}_{1x} \cos\phi \sin\phi + \mathbf{H}_{1y} \sin^2\phi + \mathbf{H}_{1z} \cos\phi) \times \\ &\quad (\mathbf{H}_{2x} \cos^2\phi + \mathbf{H}_{2y} \cos\phi \sin\phi - \mathbf{H}_{2z} \sin\phi) \mathbf{C}_x \end{aligned}$$

$$\begin{aligned} \text{CH}_3: \quad -8 \mathbf{H}_{1x} \mathbf{H}_{2y} \mathbf{H}_{3y} \mathbf{C}_y - 8 \mathbf{H}_{1y} \mathbf{H}_{2x} \mathbf{H}_{3y} \mathbf{C}_y - 8 \mathbf{H}_{1y} \mathbf{H}_{2y} \mathbf{H}_{3x} \mathbf{C}_y &\xrightarrow{90^\circ_\phi (^1\text{H})} \\ -8 (\mathbf{H}_{1x} \cos^2\phi + \mathbf{H}_{1y} \cos\phi \sin\phi - \mathbf{H}_{1z} \sin\phi) (\mathbf{H}_{2x} \cos\phi \sin\phi + \mathbf{H}_{2y} \sin^2\phi + \mathbf{H}_{2z} \cos\phi) \times \\ &\quad (\mathbf{H}_{3x} \cos\phi \sin\phi + \mathbf{H}_{3y} \sin^2\phi + \mathbf{H}_{3z} \cos\phi) \mathbf{C}_y \\ -8 (\mathbf{H}_{1x} \cos\phi \sin\phi + \mathbf{H}_{1y} \sin^2\phi + \mathbf{H}_{1z} \cos\phi) (\mathbf{H}_{2x} \cos^2\phi + \mathbf{H}_{2y} \cos\phi \sin\phi - \mathbf{H}_{2z} \sin\phi) \times \\ &\quad (\mathbf{H}_{3x} \cos\phi \sin\phi + \mathbf{H}_{3y} \sin^2\phi + \mathbf{H}_{3z} \cos\phi) \mathbf{C}_y \\ -8 (\mathbf{H}_{1x} \cos\phi \sin\phi + \mathbf{H}_{1y} \sin^2\phi + \mathbf{H}_{1z} \cos\phi) (\mathbf{H}_{2x} \cos\phi \sin\phi + \mathbf{H}_{2y} \sin^2\phi + \mathbf{H}_{2z} \cos\phi) \times \\ &\quad (\mathbf{H}_{3x} \cos^2\phi + \mathbf{H}_{3y} \cos\phi \sin\phi - \mathbf{H}_{3z} \sin\phi) \mathbf{C}_y \end{aligned}$$

Only antiphase single quantum coherence $4 \mathbf{H}_{1z} \mathbf{H}_{2z} \mathbf{C}_x$ and $8 \mathbf{H}_{1z} \mathbf{H}_{2z} \mathbf{H}_{3z} \mathbf{C}_y$ in the case of CH_2 and CH_3 , respectively, leads to detectable magnetization at the end of the sequence and is relevant for the optimization of the phase cycle. The formation of $4 \mathbf{H}_{1z} \mathbf{H}_{2z} \mathbf{C}_x$ depends on ϕ with the factor $\sin\phi \times \cos\phi$, $8 \mathbf{H}_{1z} \mathbf{H}_{2z} \mathbf{H}_{3z} \mathbf{C}_y$ with the factor $\sin\phi \times \cos^2\phi$.

The amplitude of the resulting signal after a 4 step phase cycle to select CH_2 with a phase incrementation of 90° and an alternating sign of the receiver phase can be described as follows:

$$\begin{aligned} &\sin\phi \cos\phi - \sin(\phi + 90^\circ) \cos(\phi + 90^\circ) + \sin(\phi + 180^\circ) \cos(\phi + 180^\circ) - \sin(\phi + 270^\circ) \cos(\phi + 270^\circ) \\ &= 2 \sin\phi \cos\phi - 2 \sin(\phi + 90^\circ) \cos(\phi + 90^\circ) \\ &= 4 \sin\phi \cos\phi \end{aligned}$$

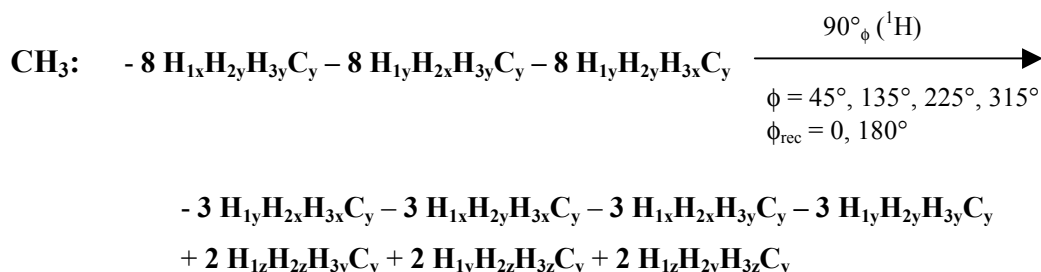
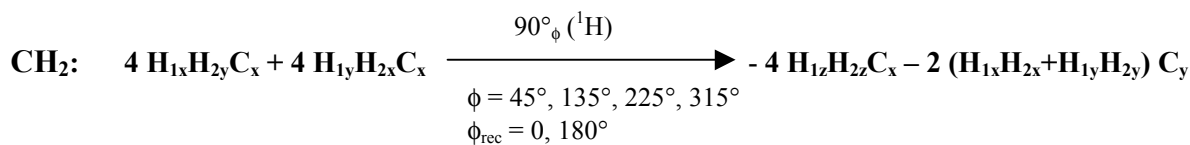
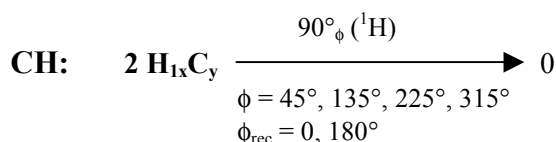
Maximum intensity will be obtained with $\phi = 45^\circ$ and thus with a phase cycle of $\phi = 45^\circ, 135^\circ, 225^\circ, 315^\circ$ with $\phi_{\text{rec}} = 0, 180^\circ$. The normalized intensity factor is 0.5 (per scan, per excited proton) due to the phase cycle.

The signal intensity after a 6 step phase cycle to select CH₃ with a phase incrementation of 60° and an alternating sign of the receiver phase can be described as follows:

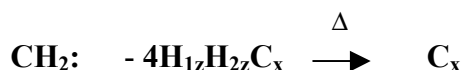
$$\begin{aligned} & \sin\phi \cos^2\phi - \sin(\phi + 60^\circ) \cos^2(\phi + 60^\circ) + \sin(\phi + 120^\circ) \cos^2(\phi + 120^\circ) \\ & - \sin(\phi + 180^\circ) \cos^2(\phi + 180^\circ) + \sin(\phi + 240^\circ) \cos^2(\phi + 240^\circ) - \sin(\phi + 300^\circ) \cos^2(\phi + 300^\circ) \\ & = 2 \sin\phi \cos^2\phi - 2 \sin(\phi + 60^\circ) \cos^2(\phi + 60^\circ) + 2 \sin(\phi + 120^\circ) \cos^2(\phi + 120^\circ) \end{aligned}$$

The function has a maximum at $\phi = 30^\circ$. The optimal phase cycle is thus $\phi = 30^\circ, 90^\circ, 150^\circ, 210^\circ, 270^\circ, 330^\circ$ with $\phi_{\text{rec}} = 0, 180^\circ$. The normalized intensity factor is 0.25 (per scan, per excited proton) due to the phase cycle.

The effect of the 4 step phase cycle (CH₂ selection) on all three multiplicities till point c is discussed in the following (factors are normalized to one scan):



Signals resulting from CH are totally suppressed. CH₂ groups pass the phase cycle to give the desired antiphase term $4 \text{ H}_{1z}\text{H}_{2z}\text{C}_x$ which is refocused during the last delay Δ to give in-phase C_x magnetization (till point d):



Additionally ¹H two spin zero quantum coherence is formed at point c which does not lead to detectable magnetization. Therefore 50% intensity is lost. However, the sum of the two terms resulting from the two individual protons is 100% (compared to a signal of a CH group in an unselective refocused INEPT).

In the case of CH₃ groups several multiple quantum terms are able to pass the phase cycling, but they do not lead to detectable magnetization. The term $8 \text{ H}_{1z}\text{H}_{2z}\text{H}_{3z}\text{C}_y$ is

suppressed. However, if Δ is not exactly $1/2J$ which is usually the case due to small differences of $^1J_{CH}$ coupling constants in a protein, methyl groups can also form double quantum coherence like $4 \mathbf{H}_{1x}\mathbf{H}_{2y}\mathbf{C}_x$ with a probability of $\sin^2(\pi J_{CH}\Delta) \cos(\pi J_{CH}\Delta)$. As shown above, these terms can pass the filter and therefore signals of methyl groups are not completely suppressed.

The effect of the 6 step phase cycle of the MUSIC-CH₃ selection on all three multiplicities till point c is shown in the following (factors are normalized to one scan):

$$\text{CH: } 2 \mathbf{H}_{1x}\mathbf{C}_y \xrightarrow[\substack{\phi = 30^\circ, 90^\circ, 150^\circ, 210^\circ, 270^\circ, 330^\circ \\ \phi_{\text{rec}} = 0, 180^\circ}]{90^\circ_\phi (^1\text{H})} 0$$

$$\text{CH}_2: 4 \mathbf{H}_{1x}\mathbf{H}_{2y}\mathbf{C}_x + 4 \mathbf{H}_{1y}\mathbf{H}_{2x}\mathbf{C}_x \xrightarrow[\substack{\phi = 30^\circ, 90^\circ, 150^\circ, 210^\circ, 270^\circ, 330^\circ \\ \phi_{\text{rec}} = 0, 180^\circ}]{90^\circ_\phi (^1\text{H})} 0$$

$$\begin{aligned} \text{CH}_3: & -8 \mathbf{H}_{1x}\mathbf{H}_{2y}\mathbf{H}_{3y}\mathbf{C}_y - 8 \mathbf{H}_{1y}\mathbf{H}_{2x}\mathbf{H}_{3y}\mathbf{C}_y - 8 \mathbf{H}_{1y}\mathbf{H}_{2y}\mathbf{H}_{3x}\mathbf{C}_y \xrightarrow[\substack{\phi = 30^\circ, 90^\circ, 150^\circ, 210^\circ, 270^\circ, 330^\circ \\ \phi_{\text{rec}} = 0, 180^\circ}]{90^\circ_\phi (^1\text{H})} \\ & -6 \mathbf{H}_{1z}\mathbf{H}_{2z}\mathbf{H}_{3z}\mathbf{C}_y \\ & + 2.5 \mathbf{H}_{1x}\mathbf{H}_{2x}\mathbf{H}_{3z}\mathbf{C}_y + 2.5 \mathbf{H}_{1x}\mathbf{H}_{2z}\mathbf{H}_{3x}\mathbf{C}_y + 2.5 \mathbf{H}_{1z}\mathbf{H}_{2x}\mathbf{H}_{3x}\mathbf{C}_y \\ & + 3.5 \mathbf{H}_{1y}\mathbf{H}_{2y}\mathbf{H}_{3z}\mathbf{C}_y + 3.5 \mathbf{H}_{1y}\mathbf{H}_{2z}\mathbf{H}_{3y}\mathbf{C}_y + 3.5 \mathbf{H}_{1z}\mathbf{H}_{2y}\mathbf{H}_{3y}\mathbf{C}_y \\ & = -6 \mathbf{H}_{1z}\mathbf{H}_{2z}\mathbf{H}_{3z}\mathbf{C}_y \\ & + 3 (\mathbf{H}_{1x}\mathbf{H}_{2x} + \mathbf{H}_{1y}\mathbf{H}_{2y})\mathbf{H}_{3z}\mathbf{C}_y + 3 (\mathbf{H}_{1x}\mathbf{H}_{3x} + \mathbf{H}_{1y}\mathbf{H}_{3y})\mathbf{H}_{2z}\mathbf{C}_y + 3 (\mathbf{H}_{2x}\mathbf{H}_{3x} + \mathbf{H}_{2y}\mathbf{H}_{3y})\mathbf{H}_{1z}\mathbf{C}_y \\ & - \frac{1}{2} (\mathbf{H}_{1x}\mathbf{H}_{2x} - \mathbf{H}_{1y}\mathbf{H}_{2y})\mathbf{H}_{3z}\mathbf{C}_y - \frac{1}{2} (\mathbf{H}_{1x}\mathbf{H}_{3x} - \mathbf{H}_{1y}\mathbf{H}_{3y})\mathbf{H}_{2z}\mathbf{C}_y - \frac{1}{2} (\mathbf{H}_{2x}\mathbf{H}_{3x} - \mathbf{H}_{2y}\mathbf{H}_{3y})\mathbf{H}_{1z}\mathbf{C}_y \end{aligned}$$

All signals resulting from CH and CH₂ are suppressed by the ¹H triple quantum filter. In the case of CH₃ the desired antiphase magnetization $6 \mathbf{H}_{1z}\mathbf{H}_{2z}\mathbf{H}_{3z}\mathbf{C}_y$ passes the filter together with a variety of ¹H two spin zero quantum terms. The latter will not lead to detectable magnetization and are responsible for intensity losses. 75% of the magnetization is lost due to losses during ¹H triple quantum formation and due to the phase cycling. However, the sum of the three terms resulting from each individual proton gives 75% intensity compared to a signal of a CH group in an unselective refocused INEPT. The magnetization transfer between points c and d can be described by:

$$\text{CH}_3: -6 \mathbf{H}_{1z}\mathbf{H}_{2z}\mathbf{H}_{3z}\mathbf{C}_y \xrightarrow{\Delta} \frac{3}{4} \mathbf{C}_x$$

The MUSIC sequence was tested on a sample of menthol (natural abundance), spectra and the spectra are shown in Fig. 3.5. The MUSIC-CH₃ selection based on the ¹H triple quantum filter, gives a clean spectrum with only CH₃ signals remaining. Lower multiplicities are completely suppressed. The same is true in the spectrum of the MUSIC-CH₂ selection based on the ¹H double quantum filter, no CH groups are visible. However small signals of CH₃ are present, because higher multiplicities are not completely suppressed due to the aforementioned reasons. However, this is not a problem in the applications in chapter 4 since MUSIC is used to suppress lower multiplicities together with other selections. The selection of CH₂ groups adjacent to a carbonyl for example does not suffer from incomplete CH₃ suppression since the topology CH₃-CO is not present in proteins.

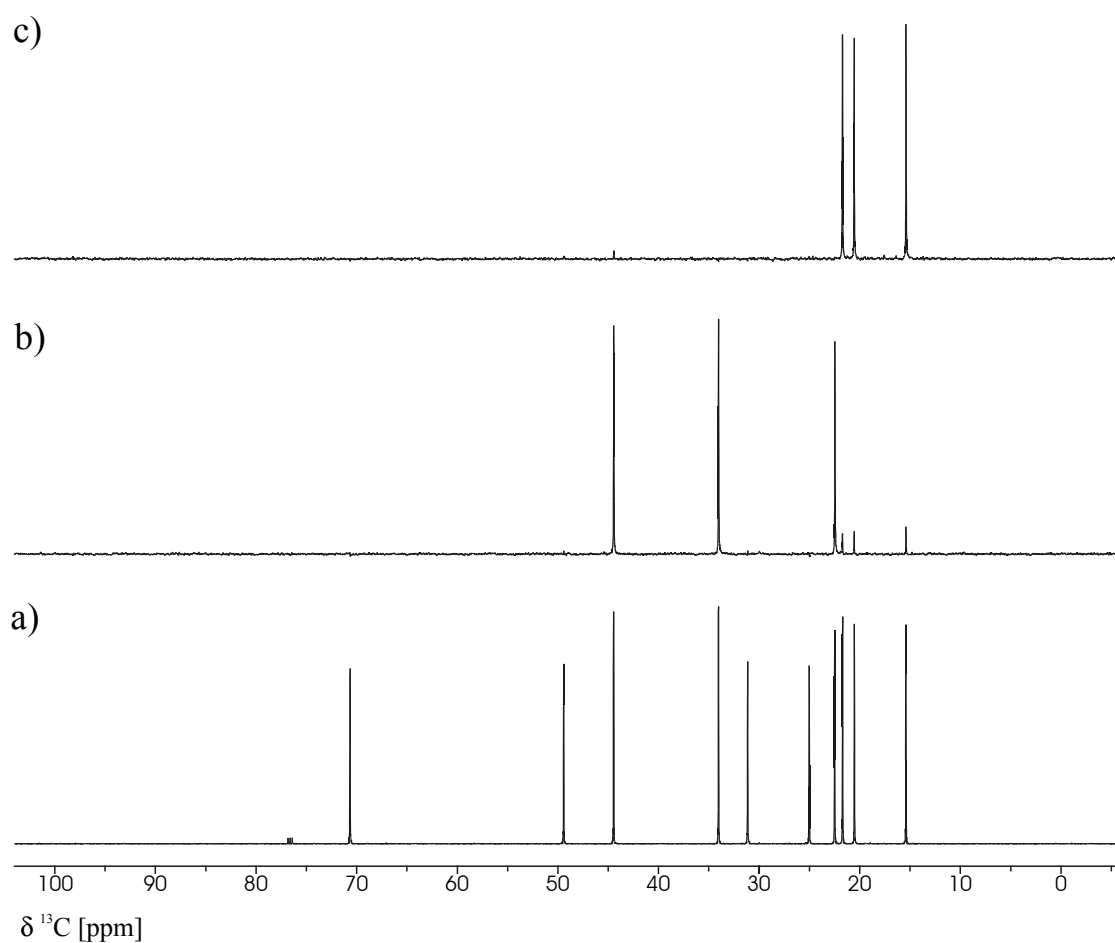


Fig. 3.5: Broadband decoupled ¹³C spectra of menthol. (a) conventional ¹³C spectrum, (b) MUSIC-CH₂ selection with the following phase cycle: $\phi_1 = 45^\circ, 135^\circ, 225^\circ, 315^\circ, \phi_{\text{rec}} = 0, 180^\circ$, (c) MUSIC-CH₃ selection with the following phase cycle: $\phi_1 = 30^\circ, 90^\circ, 150^\circ, 210^\circ, 270^\circ, 330^\circ, \phi_{\text{rec}} = 0, 180^\circ$.

3.5 Shaped pulses

Shaped pulses are pulses which consist of sequences of short rectangular pulses with different amplitudes and/or phases (Fig. 3.6). Modern spectrometers allow the implementation of these as one pulse. The amplitude and phase information of each individual rectangular pulse, the “shape”, is given in a separate file. Selective pulses have been created to span certain bandwidths, others are designed for decoupling or for homonuclear mixing, etc. The basic principles of shaped pulses are described in the literature (Freeman, 1998) and will not be discussed here in detail.

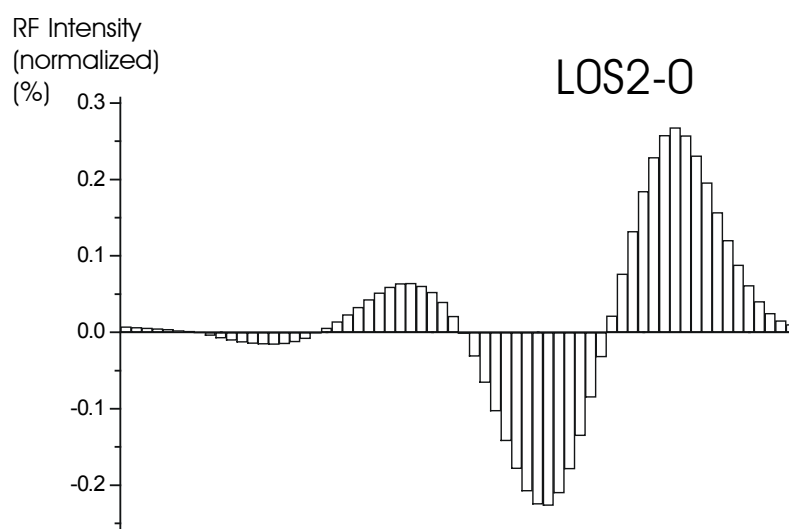


Fig. 3.6: A shaped pulse consists of a number of rectangular pulses with different amplitudes and/or different phases. Here the LOS2-0 pulse, a selective excitation pulse, is shown. In this example, 64 points (rectangular pulses) were used. Negative amplitudes are obtained combining positive values with a phase shift of 180° .

There are three categories of pulses: excitation pulses ($M_z \rightarrow M_y$), inversion pulses ($M_z \rightarrow -M_z$) and refocusing pulses ($M_x \rightarrow -M_x$ and $M_y \rightarrow -M_y$). The desirable characteristics of a selective pulse in general are:

1. uniform excitation, inversion or refocusing profile over the desired frequency range (Fig. 3.7)
2. no effect on resonances outside the desired frequency range
3. the phase of the transverse magnetization should be unaffected or should be at least a linear function of the offset
4. small transition area in comparison to the bandwidth
5. invariance to small miscalibrations of the pulse flip angle and spatial inhomogeneity

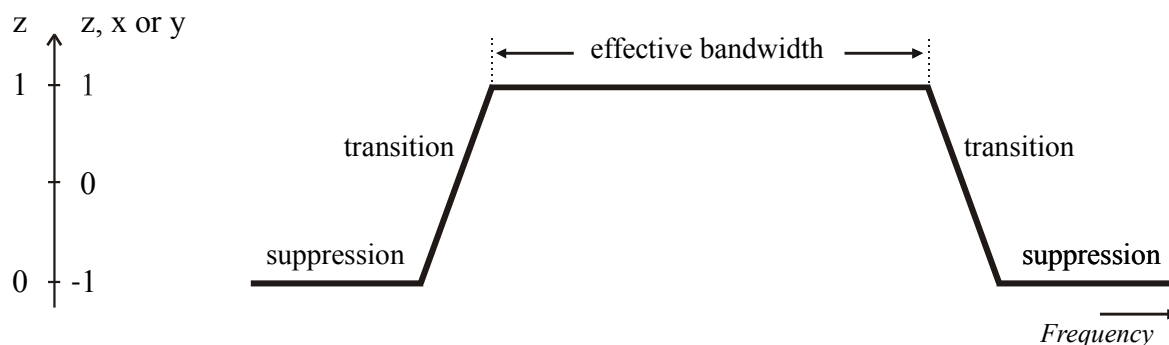


Fig. 3.7: Schematic excitation, inversion or refocusing profile of a selective pulse. In the case of an excitation pulse the ordinate shows transverse magnetization ranging from 0 to 1. In the profile of an inversion pulse the ordinate shows z magnetization ranging from -1 to 1. In the case of an refocusing pulse the transverse magnetization ranging from -1 to 1 is shown. For the sake of clarity the chemical shift evolution, which occurs outside the affected frequency range, if transverse magnetization is present, is ignored.

A large variety of shaped pulses is available and criteria to classify them have been proposed. The selectivity σ of a pulse is given by (Geen and Freeman, 1991; Lunati et al., 1998):

$$\sigma = \frac{1}{BW \cdot T_p} \quad [3.1]$$

where BW is the effective bandwidth of the pulse in Hertz and T_p the pulse length in seconds. The effective bandwidth is the frequency range of the excitation/inversion plateau shown in Fig. 3.7. Each pulse shape has its characteristic selectivity, some examples are given in table 3.1. The pulse with a low selectivity σ effects a broad bandwidth and vice versa for a given pulse length.

Pulse	type	σ	k	reference
G4	excitation	0.166 *	0	Emsley and Bodenhausen, 1990
SNEEZE	excitation	0.23 *	0	Nuzillard and Freeman, 1994
L0S2-0	excitation	0.25	0	Lunati et al., 1998
G3	inversion	0.333 *	-	Emsley and Bodenhausen, 1990
IBURP2	inversion	0.25	-	Geen and Freeman, 1991
REBURP	refocusing/inversion	0.25	-	Geen and Freeman, 1991

Table 3.1: Common shaped pulses with their selectivity σ and the proportionality constant of the phase shift k . (Lunati et al JMR 1998). (*) value estimated from simulation.

Some excitation pulses introduce an offset dependent phase shift, which is a linear function of the offset (Lunati et al., 1998):

$$\phi = k\Delta\Omega. \quad [3.2]$$

φ is the phase shift, $\Delta\Omega$ the offset. The constant of proportionality k may be positive, as for ordinary pulses, negative, as for the “prefocused” pulses (Ngo and Morris, 1987), or zero, as for self-refocusing pulses (Geen and Freeman, 1991; Nuzillard and Freeman, 1994; Kupce and Freeman, 1995). The phase shift due to a positive k can be refocused with a hard 180° pulse. When k is negative maximum signal is achieved at a time after the end of the pulse. When $k = 0$, all signal components are in-phase at the end of the pulse. Unless the pulse shape is symmetrical the time-reversed pulse shape is used to flip transverse magnetization back to z .

The excitation profiles of the G4, SNEEZE and the LOS2-0 pulses are given in Fig. 3.8. G4 is a widely used pulse but the excitation profile is far from the ideal trapezoidal one. The SNEEZE and the LOS2-0 pulse are improved pure phase pulses of the BURP family (band-selective, uniform response, pure phase) developed from the EBURP-2 shape. The excitation profiles of the SNEEZE and LOS2-0 pulses are close to ideal, the LOS2-0 pulse shows a narrower bandwidth and transition band for a given pulse length, has a better out-of-band performance and was therefore used in the experiments introduced in chapter 4.

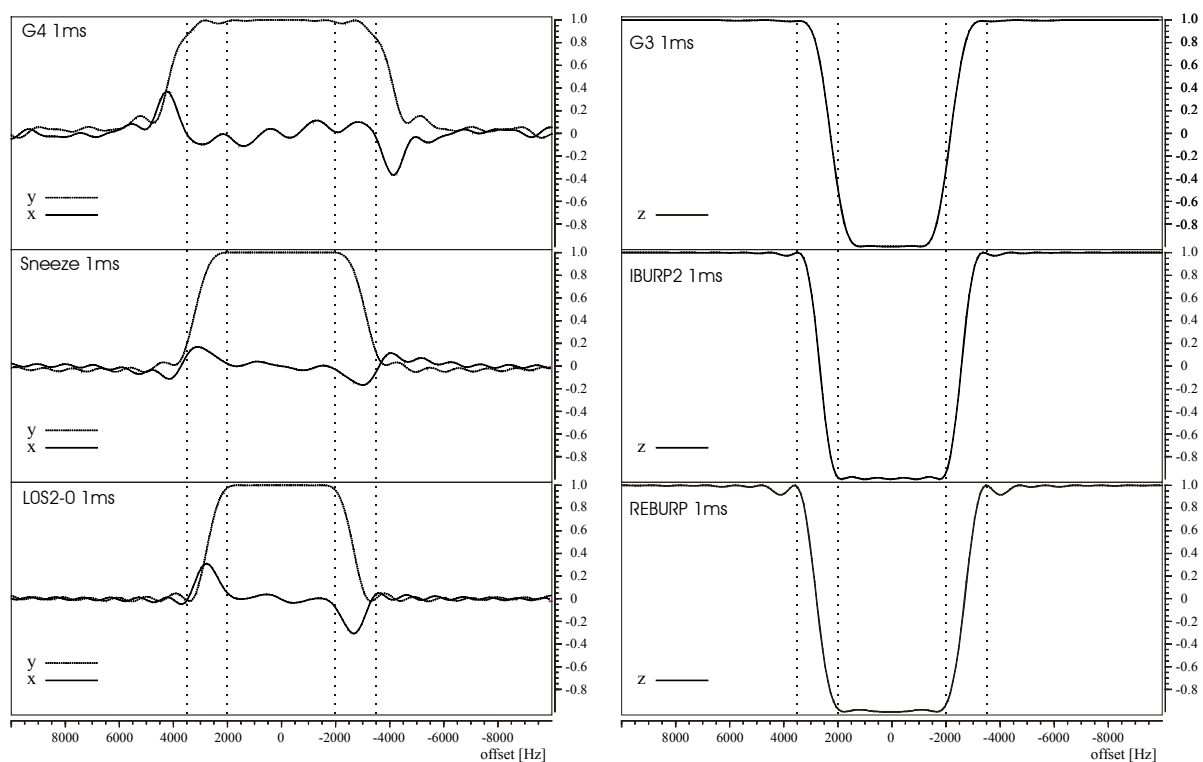


Fig. 3.8: Excitation profiles of G4, Sneeze and LOS2-0 pulses on the left, inversion profiles of G3, IBURP2 and REBURP pulses on the right. All excitation profiles were simulated for 1ms long pulses using the “Shape Tool” of XWINNMR 2.6 (Bruker AG).

The inversion profiles of the G3 and the IBURP2 pulses are also shown in Fig. 3.8. Both pulses show an almost ideal inversion profile. G3 has the advantage over IBURP2 that shorter pulses can be applied for a certain bandwidth due to its higher selectivity. G3 is used to invert carbonyl carbon magnetization in most of the experiments in chapter 4.

In some cases it is desirable to refocus transverse magnetization and to invert z magnetization at the same time. For such a purpose the REBURP pulse was created. The inversion profile is shown in Fig. 3.8 as well. The REBURP pulse can also be used solely as an inversion pulse and, because of its symmetric shape, it has some advantages over more common inversion pulses like IBURP or G3.

Scalar coupling is usually only considered during free precession periods and ignored during pulses, which are normally too short for J coupling to have an effect. However, if pulses get longer (ms range) the question arises how much J coupling is active during the pulses. Two cases are of interest for the applications in chapter 4: J coupling during a shaped 180° inversion pulse in an INEPT transfer (see also Zwaalen et al., 1998) and during a shaped 90° pulse in an HMQC experiment. To analyze this experimentally two modified HSQC experiments were used. In the first INEPT step either the 180° pulse or the 90° pulse on the heteronucleus are performed by a selective pulse (Fig. 3.9).

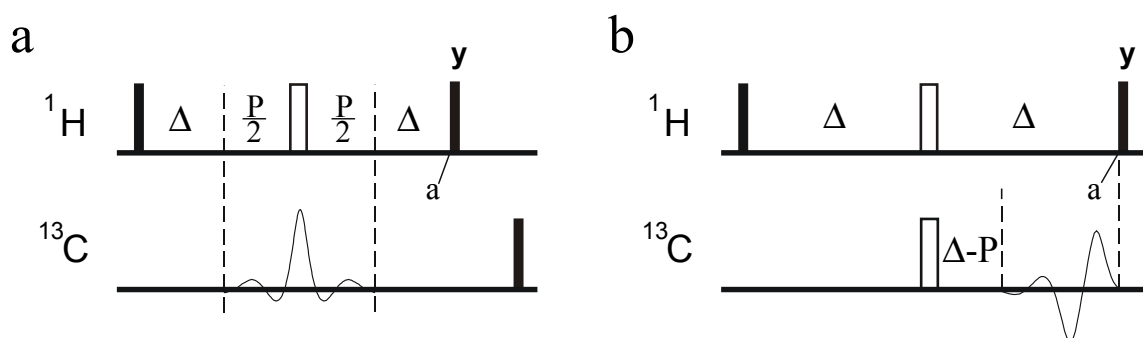


Fig. 3.9: Modified INEPT for the analysis of effective J coupling between ^1H and ^{13}C during (a) a centered selective 180° and (b) a 90° selective pulse. As part of a ^{13}C -HSQC experiment the ^{13}C frequency of the selective pulse is set on resonance. The signal of interest should be isolated in the ^1H dimension. The isolated signal is analyzed in a series of 1D experiments as a function of the delay Δ . The same is done in a reference experiment with a hard pulse instead of the selective one. A ^{13}C labeled sample of methanol was used for the analysis.

The signal intensity of a series of 1D spectra is analyzed as a function of the delay Δ with the modified and unmodified HSQC. The transfer function in an INEPT just with hard pulses (Fig. 3.1a) can be described by:

$$\mathbf{I}_z \xrightarrow{90^\circ_x (^1\text{H})} -\mathbf{I}_y \xrightarrow[\begin{smallmatrix} \Delta 180^\circ_x (^1\text{H}) \Delta \\ 180^\circ_x (^{13}\text{C}) \end{smallmatrix}]{\Delta 180^\circ_x (^1\text{H}) \Delta} \mathbf{I}_y \cos(\pi J_{\text{IS}} 2\Delta) - 2 \mathbf{I}_x \mathbf{S}_z \sin(\pi J_{\text{IS}} 2\Delta)$$

J_{IS} is the coupling constant between the I and S spin. The pulse length of the hard pulses is neglected as is potential coupling during this time. The implementation of a selective 180° pulse (Fig. 3.9 a) allows an effective J coupling during the pulse, leading to the following transfer function up to point a:

$$\mathbf{I}_z \xrightarrow{90^\circ_x (^1\text{H})} -\mathbf{I}_y \xrightarrow[\substack{\Delta 180^\circ_x (^1\text{H}) \Delta \\ 180^\circ_x (^{13}\text{C})}]{\phantom{\Delta 180^\circ_x (^1\text{H}) \Delta}} \mathbf{I}_y \cos(\pi(J_{IS}2\Delta + J_{IS}^{\text{eff}}P)) - 2 \mathbf{I}_x \mathbf{S}_z \sin(\pi(J_{IS}2\Delta + J_{IS}^{\text{eff}}P))$$

J_{IS}^{eff} is the effective coupling constant during the pulse length P. Only the $2 \mathbf{I}_x \mathbf{S}_z$ term is relevant since this is converted into $2 \mathbf{I}_z \mathbf{S}_y$ magnetization which is selected by phase cycling and frequency labeled during t_1 . The $2 \mathbf{I}_x \mathbf{S}_z$ term vanishes if $J_{IS}2\Delta + J_{IS}^{\text{eff}}P = n$ ($n=1,2,3\dots$) and the detected signal intensity is zero. The effective coupling during the pulse can be determined by using two series of 1D spectra as a function of the delay Δ one with a hard pulse, one with a shaped pulse:

$$J_{IS}^{\text{eff}}/J_{IS} = (2\Delta_{0(\text{H})} - 2\Delta_{0(\text{S})})/P \quad [3.3]$$

$J_{IS}^{\text{eff}}/J_{IS}$ is the ratio of effective coupling, $\Delta_{0(\text{H})}$ the zero point in the series with a hard pulse, $2\Delta_{0(\text{S})}$ the zero point with a selective pulse.

The effective coupling during centered REBURP and IBURP2 pulses was determined to be 90 ± 2 % and 28 ± 2 %, respectively. In the case of an IBURP2 pulse the delay of free precession ($2\Delta + P$) has to be increased considerably (Zwahlen et al., 1998). Therefore the REBURP pulse should be preferred if J_{IS} coupling is desired during the pulse in order to minimize Δ and thus relaxation losses. Note that the effective coupling J_{IS}^{eff} in this experiment depends on the position of the shaped pulse, a centered pulse is optimal for a symmetrical pulse shape like REBURP, because the trajectory of the magnetization passes the transversal plane in the middle of the pulse. Optimal effective coupling J_{IS}^{eff} during asymmetrical pulses is achieved by positioning the pulse in a way that the moment of inversion occurs simultaneously with the 180° proton pulse. However, the narrowest bandwidth (best selectivity) in an INEPT step is achieved if the shaped pulse fills the whole time between the 90° pulses and this is only possible if it is centered and a symmetrical shape is used.

The implementation of a selective 90° pulse during the second delay Δ (Fig. 3.9 b) allows the determination of J coupling during long 90° pulses which occur for example in an HMQC with shaped heteronuclear pulses. The experiment can be described up to point a by the following transfer function (irrelevant terms are omitted):

$$\mathbf{I}_z \xrightarrow{90^\circ_x (^1\text{H})} -\mathbf{I}_y \xrightarrow[\substack{\Delta 180^\circ_x (^1\text{H}) \Delta \\ 180^\circ_x (^{13}\text{C}) \quad 90^\circ_x (^{13}\text{C})}]{\phantom{\Delta 180^\circ_x (^1\text{H}) \Delta}} 2 \mathbf{I}_x \mathbf{S}_y \sin(\pi(J_{IS}(2\Delta - P) + J_{IS}^{\text{eff}}P))$$

The determination of the effective coupling during the pulse is achieved by using two series of 1D spectra as a function of the delay Δ , one with a hard pulse, one with a shaped pulse:

$$J_{IS}^{\text{eff}}/J_{IS} = (2\Delta_{0(H)} - 2\Delta_{0(S)} + P)/P \quad [3.4]$$

$\Delta_{0(H)}$ is the zero point in the series with a hard pulse, $2\Delta_{0(S)}$ the zero point with a selective pulse, P the pulse length of the shaped 90° pulse. The effective coupling during a LOS2-0 pulse was determined to be $63 \pm 2\%$.

3.6 Selecting specific coupling patterns by tuning of delays

The use of carefully tuned delays allows the selection or suppression of particular side chain topologies. The principle is similar to the spin topology filter described by Levitt and Ernst (Levitt and Ernst, 1985) which was proposed for homonuclear ^1H spin systems. A disadvantage of ^1H spin systems is the large variation in proton-proton coupling constants and therefore the method was only rarely used.

^{13}C labeled amino acid side chains offer a network of ^{13}C spins with almost uniform $^1J_{CC}$ coupling constants. This allows the exploitation of the number of coupling partners of a particular ^{13}C nucleus in order to select a certain topology. For a magnetization transfer between aliphatic ^{13}C nuclei, the refocused COSY-step (Fig. 3.10a) is used. Several topologies are shown schematically in Fig. 3.11. The ^{13}C nuclei are represented by ①, ②, ③ and ④.

a refocused COSY step



b two refocused COSY steps



c three refocused COSY steps



Fig. 3.10: Pulse sequences to transfer magnetization in a linear chain of spin $\frac{1}{2}$ nuclei coupled with uniform 1J coupling constants. Particular spin systems can be suppressed by appropriate tuning of delays.

In the applications in chapter 4 the magnetization transfers $C^\beta \rightarrow C^\alpha$, $C^\gamma \rightarrow C^\alpha$ and $C^\delta \rightarrow C^\alpha$ are of interest. During these transfers undesired topologies can be suppressed by tuned delays.

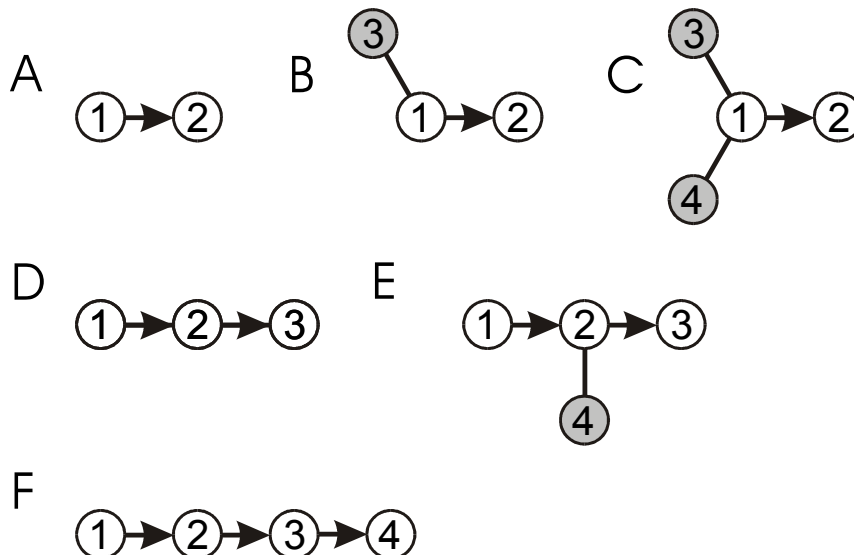


Fig. 3.11: Different spin topologies with up to four spins of the same type. Lines and arrows represent scalar couplings with an uniform 1J coupling constant. White spins are actively involved in magnetization transfer while passive spins have a gray background. In case A-C magnetization should be transferred from ① to ②, in case D and E from ① to ③ and in case F from ① to ④ as indicated by the arrows.

It is important whether the magnetization transfer starts from in-phase or antiphase magnetization. First the magnetization transfer from nucleus ① to ② in the three topologies A, B and C (Fig.3.11) is discussed, starting with in-phase ^{13}C magnetization:

$$\text{A: } \mathbf{I}_{1x} \xrightarrow[J_{12}]{2\Delta_1} 2 \mathbf{I}_{1y}\mathbf{I}_{2z} \sin \pi J_{CC}2\Delta_1 \xrightarrow{90^\circ_x} -2 \mathbf{I}_{1z}\mathbf{I}_{2y} \sin \pi J_{CC}2\Delta_1$$

$$\text{B: } \mathbf{I}_{1x} \xrightarrow[J_{12}, J_{13}]{2\Delta_1} 2 \mathbf{I}_{1y}\mathbf{I}_{2z} \sin \pi J_{CC}2\Delta_1 \cos \pi J_{CC}2\Delta_1 \xrightarrow{90^\circ_x} -2 \mathbf{I}_{1z}\mathbf{I}_{2y} \sin \pi J_{CC}2\Delta_1 \cos \pi J_{CC}2\Delta_1$$

$$\text{C: } \mathbf{I}_{1x} \xrightarrow[J_{12}, J_{13}, J_{14}]{2\Delta_1} 2 \mathbf{I}_{1y}\mathbf{I}_{2z} \sin \pi J_{CC}2\Delta_1 \cos^2 \pi J_{CC}2\Delta_1 \xrightarrow{90^\circ_x} -2 \mathbf{I}_{1z}\mathbf{I}_{2y} \sin \pi J_{CC}2\Delta_1 \cos^2 \pi J_{CC}2\Delta_1$$

Coupling to one, two or three ^{13}C nuclei is evolving during the delay $2 \Delta_1$ (Fig. 3.10a). The 1J coupling constant between spin ① and ③ leads to a cosine term in case B, 1J coupling between ①, ③ and ④ causes the squared cosine term in case C, leading to similar dependencies on the delay $2 \Delta_1$ as in the INEPT experiment for the different multiplicities (Fig. 3.3). Beside the desired magnetization transfer from ① to ② undetectable multiple quantum coherences can be formed with spins ③ and ④ or magnetization can be transferred with equal probability to ③ or ④ which lead all to transfer losses. Optimal magnetization

transfer takes place using $2\Delta_1$ values of $1/2J$, $1/4J$ or $1/6J$ for case A, B or C, respectively. Signals of topology B and C can be suppressed by setting $2\Delta_1 = 1/2J$: all cosine terms get zero while the sine term and thus the transfer function of topology A has a maximum. An application is given in chapter 4.4.16.

The magnetization transfer from ① to ③ starting with in-phase magnetization is discussed with the topologies D and E. To achieve this transfer two COSY steps (relayed COSY) are necessary (Fig. 3.10b). At the beginning, a transfer from ① to ② takes place, which results in antiphase magnetization $2 I_{1z}I_{2y}$. During the delay $2\Delta_2$, coupling between ① and ② has to evolve to refocus the antiphase term $2 I_{1z}I_{2y}$, and coupling between ② and ③ evolves to create the antiphase term $2 I_{2y}I_{3z}$:

$$\begin{aligned}
 \text{D:} \quad & I_{1x} \xrightarrow[J_{12}]{2\Delta_1} 2 I_{1y}I_{2z} \sin \pi J_{CC}2\Delta_1 \xrightarrow{90^\circ_x} -2 I_{1z}I_{2y} \sin \pi J_{CC}2\Delta_1 \xrightarrow[J_{12}, J_{23}]{2\Delta_2} \\
 & 2 I_{2y}I_{3z} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \xrightarrow{90^\circ_x} -2 I_{2z}I_{3y} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \\
 \\
 \text{E:} \quad & I_{1x} \xrightarrow[J_{12}]{2\Delta_1} 2 I_{1y}I_{2z} \sin \pi J_{CC}2\Delta_1 \xrightarrow{90^\circ_x} -2 I_{1z}I_{2y} \sin \pi J_{CC}2\Delta_1 \xrightarrow[J_{12}, J_{23}, J_{24}]{2\Delta_2} \\
 & 2 I_{2y}I_{3z} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_2 \xrightarrow{90^\circ_x} \\
 & -2 I_{2z}I_{3y} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_2
 \end{aligned}$$

Signals resulting from topology E can be suppressed by setting $2\Delta_2 = 1/2J_{CC}$ while optimal transfer will occur in case D. An application with a similar topology is provided in chapter 4.4.4.

Two refocused COSY steps (Fig. 3.10b) can also be used to enable the distinction between the topologies A and D. Magnetization will be transferred from ① to ③ in case D while A is suppressed. In spin system A only magnetization on spin ② would lead to detectable magnetization. Magnetization transfer starts with in-phase magnetization on ①:

$$\begin{aligned}
 \text{A:} \quad & I_{1x} \xrightarrow[J_{12}]{2\Delta_1} I_{1x} \cos \pi J_{CC}2\Delta_1 + 2 I_{1y}I_{2z} \sin \pi J_{CC}2\Delta_1 \\
 & \xrightarrow{90^\circ_x} I_{1x} \cos \pi J_{CC}2\Delta_1 - 2 I_{1z}I_{2y} \sin \pi J_{CC}2\Delta_1
 \end{aligned}$$

$$\begin{aligned}
& \xrightarrow[J_{12}]{2\Delta_2} \mathbf{2 I_{1y}I_{2z} \cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 + I_{2x} \sin \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2} \\
& \xrightarrow{90^\circ_x} \mathbf{2 I_{1z}I_{2y} \cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 + I_{2x} \sin \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2} \\
& \xrightarrow[J_{12}]{2\Delta_3} \mathbf{I_{2x} \cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \sin \pi J_{CC}2\Delta_3} \\
& \quad + \mathbf{I_{2x} \sin \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_3} \\
& = \mathbf{I_{2x} \sin \pi J_{CC}2\Delta_2 (\cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_3 + \sin \pi J_{CC}2\Delta_1 \cos \pi J_{CC}2\Delta_3)} \\
& = \mathbf{I_{2x} \sin \pi J_{CC}2\Delta_2 \sin \pi J_{CC} (2\Delta_1 + 2\Delta_3)}
\end{aligned}$$

$$\begin{aligned}
\mathbf{D:} \quad \mathbf{I_{1x}} & \xrightarrow[J_{12}]{2\Delta_1} \mathbf{2 I_{1y}I_{2z} \sin \pi J_{CC}2\Delta_1} \xrightarrow{90^\circ_x} \mathbf{-2 I_{1z}I_{2y} \sin \pi J_{CC}2\Delta_1} \\
& \xrightarrow[J_{12}, J_{23}]{2\Delta_2} \mathbf{2 I_{2y}I_{3z} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2} \xrightarrow{90^\circ_x} \mathbf{-2 I_{2z}I_{3y} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2} \\
& \xrightarrow[J_{23}]{2\Delta_3} \mathbf{I_{3x} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \sin \pi J_{CC}2\Delta_3}
\end{aligned}$$

Signals of spin system A are suppressed if $2\Delta_1 + 2\Delta_3 = 1/J$. Using this condition the transfer amplitude for spin system D can be optimized. An application is given in chapter 4.4.10.

A very similar distinction between spin system D and F can be applied using a combination of three refocused COSY steps (Fig. 3.10c). Magnetization is transferred from ① to ④ in spin system F while signals from spin system D can be suppressed. In an application in chapter 4 only magnetization on ③ in case D and on ④ in case F are further transferred to other nuclei and are therefore considered. The transfer starts with in-phase magnetization:

$$\begin{aligned}
\mathbf{D:} \quad \mathbf{I_{1x}} & \xrightarrow[J_{12}]{2\Delta_1} \mathbf{I_{1x} \cos \pi J_{CC}2\Delta_1 + 2 I_{1y}I_{2z} \sin \pi J_{CC}2\Delta_1} \\
& \xrightarrow{90^\circ_x} \mathbf{I_{1x} \cos \pi J_{CC}2\Delta_1 - 2 I_{1z}I_{2y} \sin \pi J_{CC}2\Delta_1}
\end{aligned}$$

$$\begin{aligned} \xrightarrow[\substack{J_{12}, J_{23}}]{2\Delta_2} & \quad 2 \mathbf{I}_{1y}\mathbf{I}_{2z} \cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \\ & \quad + \mathbf{I}_{2x} \sin \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_2 \\ & \quad + 2 \mathbf{I}_{2y}\mathbf{I}_{3z} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \end{aligned}$$

$$\begin{aligned} \xrightarrow{90^\circ_x} & \quad -2 \mathbf{I}_{1z}\mathbf{I}_{2y} \cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \\ & \quad + \mathbf{I}_{2x} \sin \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_2 \\ & \quad - 2 \mathbf{I}_{2z}\mathbf{I}_{3y} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \end{aligned}$$

$$\begin{aligned} \xrightarrow[\substack{J_{12}, J_{23}}]{2\Delta_2} & \quad 2 \mathbf{I}_{2y}\mathbf{I}_{3z} \cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \sin^2 \pi J_{CC}2\Delta_3 \\ & \quad + 2 \mathbf{I}_{2y}\mathbf{I}_{3z} \sin \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_3 \sin \pi J_{CC}2\Delta_3 \\ & \quad + \mathbf{I}_{3x} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \sin \pi J_{CC}2\Delta_3 \end{aligned}$$

$$\begin{aligned} \xrightarrow{90^\circ_x} & \quad -2 \mathbf{I}_{2z}\mathbf{I}_{3y} \cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \sin^2 \pi J_{CC}2\Delta_3 \\ & \quad - 2 \mathbf{I}_{2z}\mathbf{I}_{3y} \sin \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_3 \sin \pi J_{CC}2\Delta_3 \\ & \quad + \mathbf{I}_{3x} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \sin \pi J_{CC}2\Delta_3 \end{aligned}$$

$$\begin{aligned} \xrightarrow[\substack{J_{23}}]{2\Delta_2} & \quad \mathbf{I}_{3x} \cos \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \sin^2 \pi J_{CC}2\Delta_3 \sin \pi J_{CC}2\Delta_4 \\ & \quad + \mathbf{I}_{3x} \sin \pi J_{CC}2\Delta_1 \sin \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_2 \cos \pi J_{CC}2\Delta_3 \sin \pi J_{CC}2\Delta_3 \times \\ & \quad \quad \sin \pi J_{CC}2\Delta_4 \\ & \quad + \mathbf{I}_{3x} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \sin \pi J_{CC}2\Delta_3 \cos \pi J_{CC}2\Delta_4 \end{aligned}$$

$$\mathbf{F}: \quad \mathbf{I}_{1x} \xrightarrow[\substack{J_{12}}]{2\Delta_1} 2 \mathbf{I}_{1y}\mathbf{I}_{2z} \sin \pi J_{CC}2\Delta_1 \xrightarrow{90^\circ_x} -2 \mathbf{I}_{1z}\mathbf{I}_{2y} \sin \pi J_{CC}2\Delta_1$$

$$\xrightarrow[\substack{J_{12}, J_{23}}]{2\Delta_2} 2 \mathbf{I}_{2y}\mathbf{I}_{3z} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \xrightarrow{90^\circ_x} -2 \mathbf{I}_{2z}\mathbf{I}_{3y} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2$$

$$\xrightarrow[\substack{J_{23}, J_{34}}]{2\Delta_3} 2 \mathbf{I}_{3y}\mathbf{I}_{4z} \sin \pi J_{CC}2\Delta_1 \sin^2 \pi J_{CC}2\Delta_2 \sin^2 \pi J_{CC}2\Delta_3$$

$$\xrightarrow[J_{34}]{2\Delta_4} \mathbf{I}_{4x} \sin \pi J_{CC} 2\Delta_1 \sin^2 \pi J_{CC} 2\Delta_2 \sin^2 \pi J_{CC} 2\Delta_3 \sin \pi J_{CC} 2\Delta_4$$

Appropriate tuning of the delays Δ_1 , Δ_2 , Δ_3 and Δ_4 leads to a cancellation of the three transfer pathways in spin system D, while the transfer amplitude in spin system F (from ① to ④) is high. A numerical approach can be used to find a solution. An application is given in chapter 4.4.14. However, an experimental optimization will be necessary to take relaxation effects and the individual $^1J_{CC}$ couplings into account.

The principles of selection of certain spin systems and suppression of others using tuned delays in COSY and relayed COSY sequences were demonstrated for networks of ^{13}C spins. However, the $^1J_{CC}$ coupling constants in protein side chains are not absolutely uniform. It has been shown that the $^1J_{C\alpha C\beta}$ coupling constants depend on the secondary structure and on the amino acid type (Cornilescu et al., 2000). The $^1J_{C\alpha C\beta}$ coupling constants measured in ubiquitin ranged from 31.7 Hz to 41.2 Hz. It has to be assumed that other $^1J_{CC}$ coupling constants are similarly dependent on the conformation and the type of the side chain.

The suppression of undesired signals based on the tuning of delays is therefore often incomplete. The delays were optimized experimentally for each application in chapter 4 to account for relaxation effects and individual 1J coupling constants. Small positive or negative signals of undesired side chains can often not be avoided. The clean selection of MUSIC or the use of shaped pulses should be preferred for suppressing unwanted signals. The tuning of delays should only be used at a second stage, for further simplification of the spectra.

References

- W. P. Aue, E. Bartholdi and R. R. Ernst, *J. Chem. Phys.* **64**, 2229-2246 (1976).
 A. Bax and D. G. Davis, *J. Magn. Reson.* **65**, 355-360 (1985).
 L. Braunschweiler, G. Bodenhausen and R. R. Ernst, *Mol. Phys.* **48**, 535-560 (1983).
 R. Braunschweiler and R. R. Ernst, *J. Magn. Reson.* **53**, 521-528 (1983).
 J. M. Bultsing, W. M. Brooks, J. Field and D. M. Doddrell, *J. Magn. Reson.* **56**, 167-173 (1984).
 J. Cavanagh, W. J. Fairbrother, A. G. Palmer III and N. J. Skelton, "Protein NMR Spectroscopy", Academic Press, San Diego (1996).
 N. Chandrakumar and S. Subramanian, *Modern Techniques in High-Resolution FT-NMR*, Springer-Verlag, New York (1987).
 G. Cornilescu, A. Bax and D. A. Case *J. Am. Chem. Soc.* **122**, 2168-2171 (2000).
 D. M. Doddrell, D. T. Pegg and M. R. Bendall, *J. Magn. Reson.* **48**, 323-327 (1982).
 L. Emsley and G. Bodenhausen, *Chem. Phys. Lett.* **165**, 469-476 (1990).
 R. Freeman, *Progr. NMR Spectrosc.* **32**, 59-106 (1998).
 H. Geen and R. Freeman, *J. Magn. Reson.* **93**, 93-141 (1991).

- S. J. Glaser and J. J. Quant, *Adv. Magn. Opt. Reson.* **19**, 59-252 (1996).
- J. Jeener, *Amp. Int. Summerschool II, Basko Polje, Yugoslavia* (1971).
- E. Kupče and R. Freeman, *J. Magn. Reson. A* **112**, 134-137 (1995).
- M. H. Levitt and R. R. Ernst, *J. Chem. Phys.* **83**, 3297-3310 (1985).
- E. Lunati, P. Cofrancesco, M. Villa, P. Marzola and F. Osculati, *J. Magn. Reson.* **134**, 223-235 (1998).
- G. A. Morris and R. Freeman, *J. Am. Chem. Soc.* **101**, 760-762 (1979).
- J. T. Ngo and P. J. Morris, *Magn. Reson. Med.* **5**, 217-237 (1987).
- J.-M. Nuzillard and R. Freeman, *J. Magn. Reson. A* **110**, 252-256 (1994).
- M. Sattler, J. Schleucher and C. Griesinger, *Prog. NMR Spectrosc.* **34**, 93-158 (1999).
- A. J. Shaka, J. Keeler, T. Frenkiel and R. Freeman, *J. Magn. Reson.* **52**, 335-338 (1983).
- A. J. Shaka, P. B. Barker and R. Freeman, *J. Magn. Reson.* **64**, 547-552 (1985).
- O. W. Sørensen, G. W. Eich, M. H. Levitt, G. Bodenhausen and R. R. Ernst, *Progr. NMR Spectrosc.* **16**, 163-192 (1982).
- O. W. Sørensen and R. R. Ernst, *J. Magn. Reson.* **51**, 477-489 (1983).
- A. Wokaun and R. R. Ernst, *Chem. Phys. Lett.* **52**, 407-412 (1977).
- C. Zwahlen, S. J. F. Vincent, K. H. Gardner and L. E. Kay, *J. Am. Chem. Soc.* **120**, 4825-4831 (1998).

Parts of this chapter were published in

J. Magn. Reson. **141**, 34-43 (1999),

J. Biomol. NMR **17**, 331-335 (2000),

J. Magn. Reson. **148**, 61-72 (2001),

J. Biomol. NMR **20**, 379-384 (2001) and

J. Magn. Reson. **153**, 186-192 (2001).

4.1 Introduction

In this section the design of each of the amino acid type-selective experiments for uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled proteins is described and the results are discussed.

As was explained in chapter 2 the ^1H - ^{15}N correlation serves as a “fingerprint” of a protein. Each individual residue is represented by its NH signal. A $^1\text{H}/^{15}\text{N}$ frequency pair corresponds to one amino acid in the chain (Fig. 4.1). The task of amino acid type-selective ^1H - ^{15}N correlations is the **identification of the amino acid type** corresponding to each $^1\text{H}/^{15}\text{N}$ frequency pair. This is accomplished by the fact that only signals of a certain amino acid type appear in each spectrum, e.g. just signals originating from Ala in an Ala selective spectrum. The advantage is the dramatic simplification of spectra: 2D spectra are sufficient to deliver the information about the amino acid type. A further goal is to use the spectra of these amino acid type-selective experiments as additional input in an automated sequential assignment procedure which deals with 3D triple-resonance spectra.

During the design of the pulse sequences emphasis was put on the clean selection of an amino acid type or a small group of amino acid types because for using the experiments as specific information for an automated assignment procedure the information should be 100% correct. Missing signals mean less additional information, but wrong signals would give wrong additional information which would then mislead the automated assignment program.

MUSIC guarantees a clean selection, while tuning of delays and the use of shaped pulses give good but not completely clean selections. Thus MUSIC was applied whenever it was possible. Further selections are accomplished by tuned delays and shaped pulses, but these selections might not be completely clean which has to be considered in an automated assignment procedure.

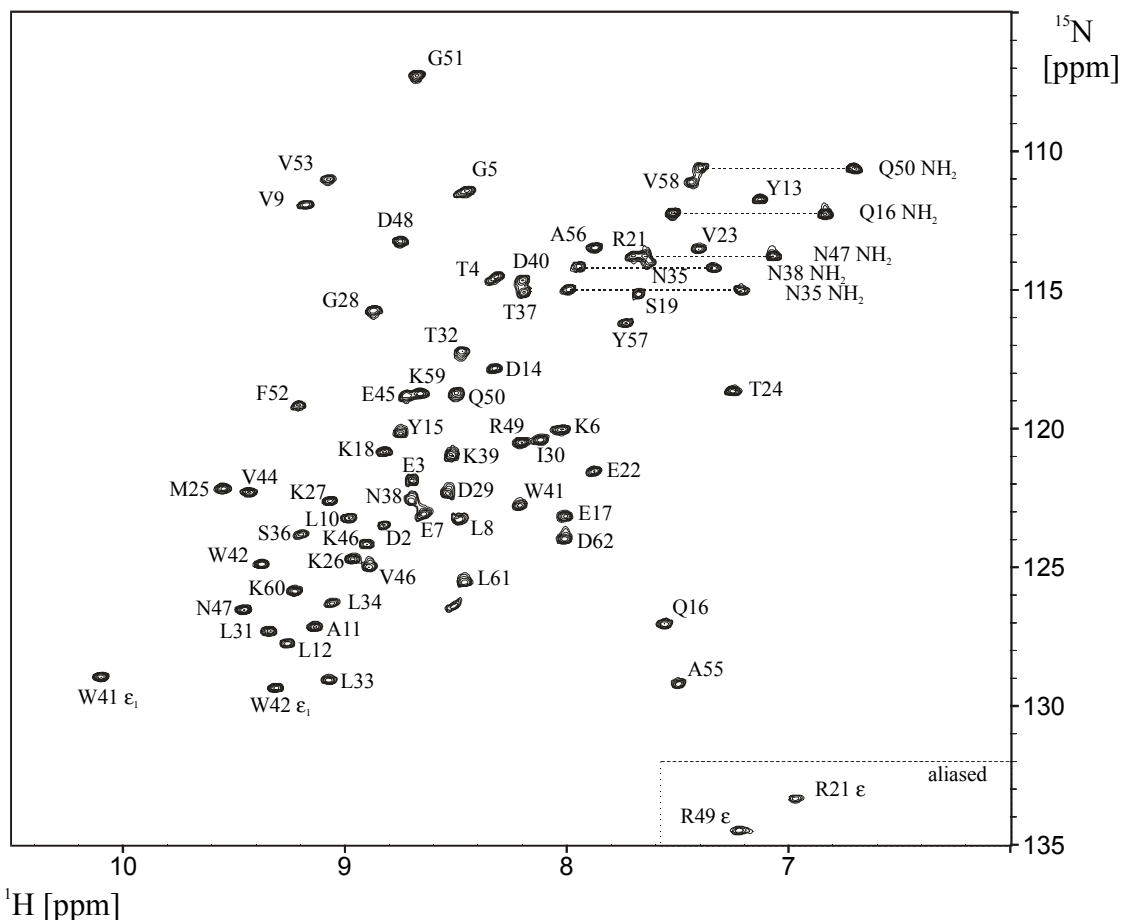


Fig. 4.1: ^{15}N -HSQC of the α -spectrin SH3 domain (Musacchio et al., 1992; Blanco et al., 1997) at 300 K with assignment. Each amide group corresponds to one signal. Additionally side chain signals of Trp, Asn, Gln and Arg are present.

The general strategy for designing selective pulse sequences is given in chapter 4.3 after a brief review of selective/semi-selective experiments developed by others in chapter 4.2. Each amino acid type-selective ^1H - ^{15}N correlation is described in chapter 4.4.

Because Asn, Gln and Trp side chains are also present in a ^1H - ^{15}N correlation and they appear in the same region as backbone NH groups, two additional experiments are presented which select only these signals: the NH_2 of Asn/Gln and the NH of the indole ring of Trp.

The novel experiments are either applied to the chicken EphB2 SAM domain (Smalla et al., 1999; PDB: 1SGG), a mostly helical protein of 83 amino acids (τ_c ca. 6.5 ns) or to the EVH1 domain from human VASP (Ball et al., 2000; PDB: 1EGX) a mostly β -sheet protein of 115 amino acids (τ_c ca. 7.7 ns).

4.2 Previous work

Earlier attempts to develop amino acid type-selective experiments were based mainly on the application of tuned delays or selective pulses. Because of their different backbone topology two of the first selective experiments were created for Gly and Pro.

An out-and-back HN(CA)HA selective for Gly uses the CH₂ multiplicity as selection criteria (Wittekind et al., 1993). The key element of the sequence is a cascade of two 45° pulses. The filter is based on a proper phase cycling selecting ¹H single quantum coherence of methylene groups while rejecting zero and double quantum coherence of CH₂ and single quantum coherence of CH. Note that no proton double quantum filter was used. The experiment was further developed to record zero and double quantum coherence of CH₂ in the indirect dimension (Bazzo et al. 1999). A Gly-edited HN(CO)(CA) was proposed which selects signals of residues following Gly in a ¹H-¹⁵N correlation (Gehring and Guittet, 1995). The filter is based on the tuning of a delay. The suppression of resonances originating from non-glycine is not perfect, since the aim was not zeroing this undesired signals but just to get an edited ¹H-¹⁵N correlation with different signs of Gly neighbors.

Proline residues hinder sequential assignment strategies based on a ¹H-¹⁵N correlation because proline lacks a NH group. A sequential assignment then has to rely either on the observation of NOEs from proline to neighboring residues or on triple-resonance experiments which do not involve an NH. For the identification of residues preceding prolines, (i-1) residues, a proline selective HACA(CO)(N) triple-resonance experiment was proposed earlier (Olejniczak and Fesik, 1994). The selection is based on the absence of an amide proton in proline. Transverse ¹⁵N magnetization is allowed to couple to the H^N during a delay 1/2J_{NH} which lead to a dephasing and suppression of all non-proline residues. The (i-1) residues are characterized by their ¹H^α and ¹³C^α chemical shifts. Recently, a proline selective 3D HDHA(NCO)CAHA (Bottomley et al., 1999) was proposed, which uses the same selection mechanism. The chemical shifts of the proline ¹³C^α and ¹³C^δ are correlated with the ¹H^α and ¹³C^α chemical shifts of the preceding residue via scalar coupling. A potential drawback of these experiments is that they rely on the ¹H^α and ¹³C^α chemical shifts. The dispersion of the ¹H^α and ¹³C^α chemical shifts is bad compared with that of the ¹H^N and ¹⁵N chemical shifts. In addition, the signals appear in the same region of the spectrum as the water signal and the experiments therefore require an excellent water suppression.

Several amino acid type selective ¹H/¹³C correlations of the side chains based on selective pulses were proposed, including experiments for assigning aromatic side chains (Yamazaki et al., 1993a; Grzesiek and Bax, 1993; Löhr and Rüterjans, 1996; Sudmeier et al. 1996; Prompers et al., 1998). Special TOCSY pulse trains were designed for the mixing between aromatic ring carbons and side chain C^β, called AMNESIA and PLUSH TACSYS (Grzesiek

and Bax, 1995; Carlomagno et al., 1996). A 3D NOESY experiment was designed which selects Val and Ile C^γ (Zwahlen et al., 1998).

Several selective experiments for the assignment of Arg side chains are based on ^{15}N selective pulses affecting only the N^ϵ or N^η of Arg (Yamazaki et al., 1993b, Yamazaki et al., 1995; Farmer and Venters, 1996; Rao et al. 1996, Vis et al., 1994, Pellecchia et al., 1997a). A NH_2 -filter based on the tuning of delays and proper phase settings was used to create a variety of selective Asn and Gln side chain experiments for deuterated proteins (Farmer II and Venters, 1996).

A selective experiment for Glu and Asp was proposed selecting CO_2 groups by proper tuning of delays (Pellecchia et al., 1997b). Dötsch et al. proposed several amino acid type-selective ^1H - ^{15}N correlations based on the tuning of delays. A CBCA(CO)NH based 2D experiment selects sequential neighbors of a group of amino acids which lacks a $^{13}\text{C}^\beta$ - $^{13}\text{C}^\gamma$ coupling i. e., Gly, Ala, Cys, Ser (Dötsch et al., 1996a and b). Signals resulting from Asn, Asp and aromatic residues can be selected additionally by refocusing their $^{13}\text{C}^\beta$ - $^{13}\text{C}^\gamma$ coupling using a shaped inversion pulse affecting just the C^γ . A HN(CO)CACB based experiment which use the same mechanism was proposed for deuterated proteins (Dötsch et al., 1996d). Another CBCA(CO)NH based experiment uses the presence of a CH group in the C^β position and selects sequential neighbors of Thr/Val/Ile or with an additional selective pulse solely of Thr (Dötsch et al., 1996c). Since $^1\text{J}_{\text{CH}}$ and $^1\text{J}_{\text{CC}}$ can vary, the suppression of undesired signals is not completely clean.

Kay and co-workers created a HC(CO)NH-TOCSY experiment for fractionally deuterated proteins which selects CH_2D groups initially, to yield a ^1H - ^{15}N correlation of residues immediately following a methyl-containing amino acid (Muhandiram et al., 1997).

Exchangeable OH protons are used to transfer magnetization to aromatic protons in order to obtain a selective experiment for Tyr ring resonances (Whitehead et al., 1997; Dalvit, 1996). The experiment proposed by Dalvit makes use of a double quantum filter.

Only few pulse schemes are based on ^1H double or triple quantum filters. MUSIC was used to create selective side chain experiments of Asn and Gln (Schmieder et al., 1998) and (H)CCH₃-TOCSY and H(C)CH₃-TOCSY experiments for methyl containing amino acids (Uhrin et al., 2000).

In summary, several selective ^1H - ^{15}N correlations have been proposed to date. The suppression of undesired signals is, however, often incomplete and the selection not very specific, i. e. a group of several amino acids are selected. Some of the above mentioned selection mechanisms proposed for ^1H - ^{13}C correlations were used in addition to MUSIC in the present work.

4.3 General strategy to design amino acid type-selective NMR experiments

To create amino acid type-selective NMR experiments, the selection methods outlined in chapter 3 such as MUSIC, selective pulses or tuned delays are introduced into common triple-resonance experiments of the type discussed in chapter 2. The different amino acids are selected based on their side chain topology and characteristic chemical shifts. MUSIC is the most powerful selection filter because it guarantees a clean selection and is therefore preferred.

The topologies of the 20 natural amino acids are shown schematically in Fig. 4.2, including the groups which can in principle be selected by MUSIC.

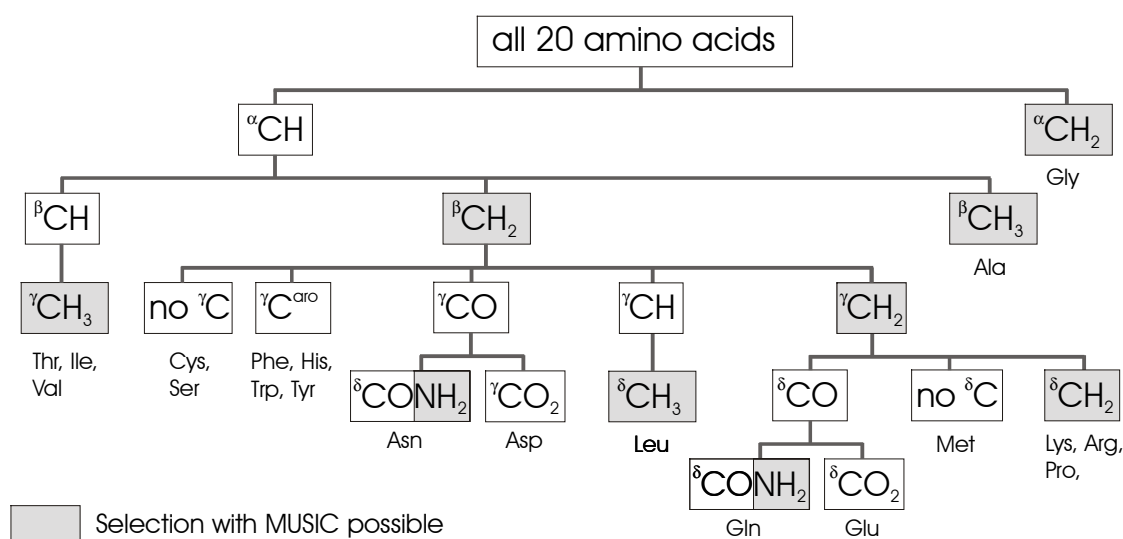


Fig. 4.2: Characteristic side chain topologies of the 20 common amino acids. The groups which can potentially be selected by MUSIC in order to design amino acid type-selective experiments are highlighted in gray. Note that only one MUSIC selection per experiment is practical, either CH₃, CH₂ or NH₂ can be selected. MUSIC is able to suppress lower multiplicities completely, while higher multiplicities are not completely suppressed. Additional selection can be achieved by the use of selective pulses affecting nuclei with characteristic chemical shifts and by carefully tuned delays, both methods are not highlighted in the scheme for the sake of clarity. In the case of Lys, Arg and Pro the selection is solely based on the application of selective pulses and tuned delays, a selection with MUSIC is not necessary.

The proposed amino acid type-selective ¹H-¹⁵N correlations are based on the triple-resonance experiments CBCA(CO)NH, CBCANH (both introduced in chapter 2.3), CA(CO)NH (Boucher et al., 1992a) and CANH (Boucher et al., 1992b). In most of the experiments the initial refocused INEPT step is replaced by MUSIC. For example a MUSIC-CH₃ selection implemented in a CBCA(CO)NH lead to a selection of Ala side chains. Further extensions like additional COSY or INEPT steps allow the magnetization transfer from/to other side chain nuclei like C^γ and C^δ. This is discussed in detail together with the

use of shaped pulses and tuned delays in the section of each particular experiment in chapter 4.4.

After selection of a specific side chain the magnetization is transferred via $^1J_{\text{CC}}$ couplings to the C^α and further to the NH. There are two possibilities for the latter transfer. The transverse C^α magnetization can be transferred via $^1J_{\text{C}\alpha\text{C}'}$ coupling to the carbonyl carbon, further via $^1J_{\text{C}'\text{N}}$ coupling to the sequential $\text{N}_{(i+1)}$ and $\text{NH}_{(i+1)}$ is finally detected as in the CA(CO)NH and CBCA(CO)NH. Because these ^1H - ^{15}N correlations contain only signals of the sequential neighbors (i+1) of the selected amino acid side chains, they are called the (i+1) experiments. The $^1J_{\text{C}\alpha\text{C}'}$ and $^1J_{\text{C}'\text{N}}$ coupling constants are larger than the line widths of the involved nuclei which makes the transfer very efficient.

In the other version the C^α magnetization is transferred via $^1J_{\text{C}\alpha\text{N}}$ and $^2J_{\text{C}\alpha\text{N}}$ coupling to the own and to the sequential nitrogens $\text{N}_{(i)}$ and $\text{N}_{(i+1)}$ as in the CANH and CBCANH experiments. The resulting spectra contain two signals per selected side chain, a signal of their own $\text{N}_{(i)}\text{H}_{(i)}$ and one of the sequential $\text{N}_{(i+1)}\text{H}_{(i+1)}$. These experiments are called the (i,i+1) experiments. Because of the relatively small coupling constants $^1J_{\text{C}\alpha\text{N}}$ and $^2J_{\text{C}\alpha\text{N}}$ compared to the C^α linewidth, magnetization transfer is less efficient and the (i,i+1) experiment is therefore less sensitive than the (i+1) experiment. Signals of sequential neighbors (i+1) are usually weaker than signals of the selected amino acid types themselves due to the smaller coupling constant $^2J_{\text{C}\alpha\text{N}}$.

4.4 The individual amino acid type-selective ^1H - ^{15}N correlations

4.4.1 Selecting Gly: The G-(i+1)- and G-(i,i+1)-HSQC

Glycine is the amino acid with the simplest topology, it lacks a C^β and contains two H^α . The methylene group in the C^α position is thus unique among the naturally occurring amino acids and implementation of a MUSIC- CH_2 selection into the CA(CO)NH (Boucher et al., 1992a) and the CANH (Boucher et al., 1992b) results in the selection of the Gly-side chain. The pulse sequences of the G-(i+1)-HSQC and the G-(i,i+1)-HSQC are given in Fig. 4.3 a and 4.3 b, respectively. The corresponding two-dimensional spectra are shown in Fig. 4.4 a and 4.4 b.

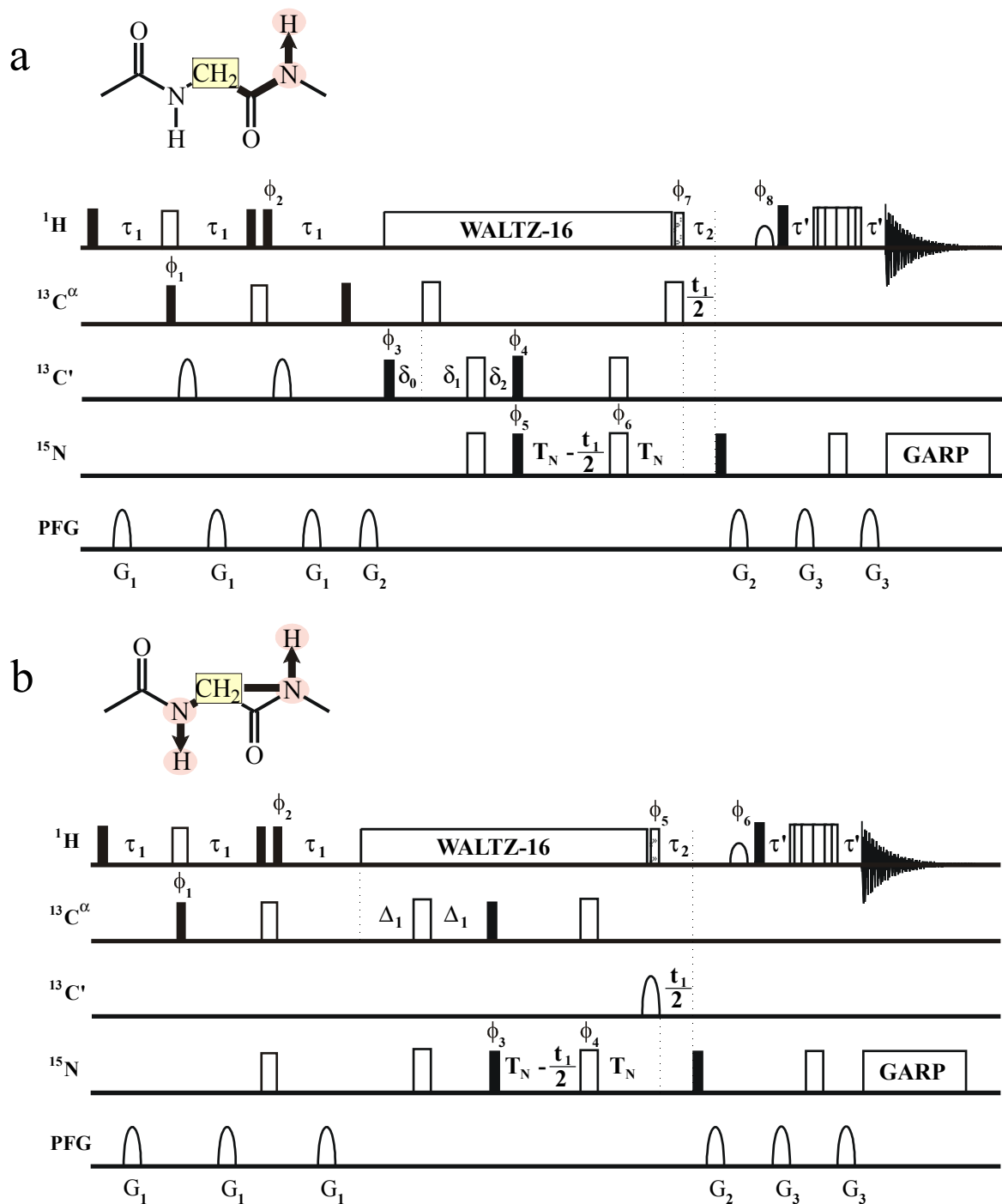


Fig. 4.3: Pulse sequences of the two Gly selective experiments: (a) G-(i+1)-HSQC and (b) G-(i,i+1)-HSQC. 90° and 180° pulses are represented by thin black filled and thick unfilled bars, respectively. The water-selective 90° ¹H flip-back pulse is represented by a sine shape, a hatched thin bar stands for a 90° flip-back ¹H pulse at the end of the ¹H-decoupling sequence. Pulsed field gradients (PFG) as well as shaped 180° ¹³CO pulses are represented by sine shapes. Unless indicated otherwise, pulses are applied with phase x. The following delays were used: τ₁ = 3.5 ms, τ₂ = 5.5 ms, τ' = 2.25 ms, Δ₁ = 7.0 ms, δ₀ = 4.5 ms, δ₁ = 6.9 ms, δ₂ = 11.4 ms, T_N = 11 ms. More experimental details are given in the appendix.

A similar geometry to that of Gly may also be found in the side chains of Asn and Gln and are selected by the CA(CO)NH experiment. Similarly, Arg side chains are selected by the CANH experiment. Since the delay for magnetization transfer from the nitrogen to the amino proton is chosen to be $1/2J_{\text{NH}}$, the signals involving NH_2 groups of Asn and Gln should not show up. However, the spectra are recorded in 90% H_2O / 10% D_2O and some signals from the NHD groups in these side chains are visible. Since side chain amide groups are easily distinguished or identified by a separate experiment, this does not complicate the information in the Gly-selective spectra. The same is true for the Arg side chain signals that show up in the (i,i+1) experiment. They can easily be identified in a conventional HSQC spectrum based on their chemical shifts.

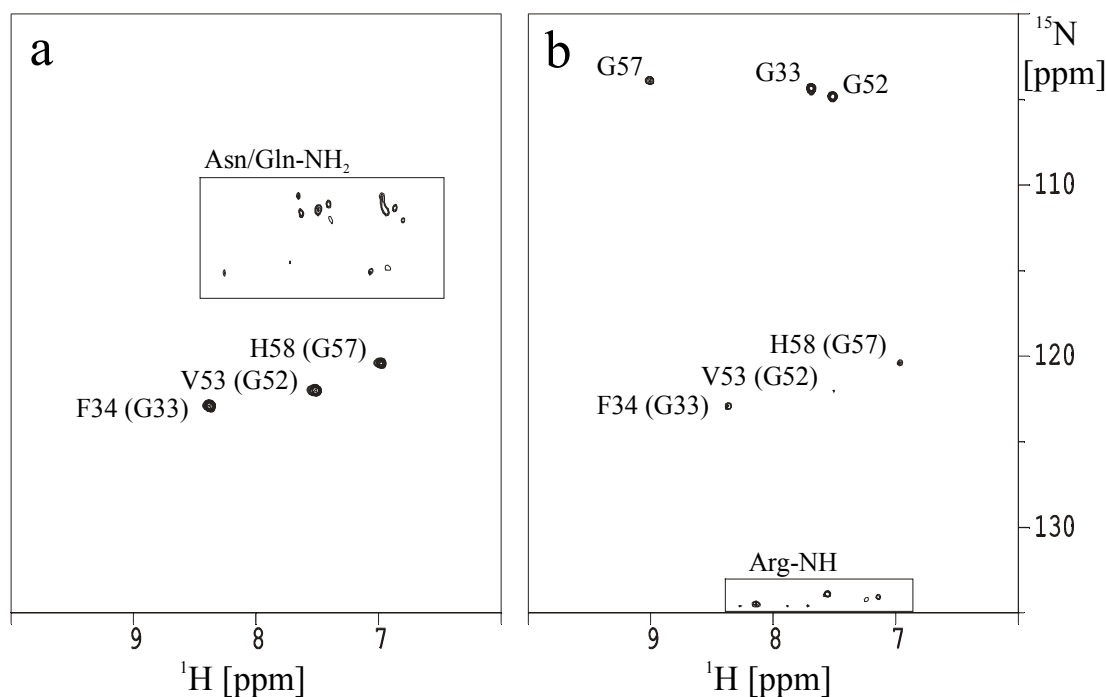


Fig. 4.4: Glycine selective ^1H - ^{15}N -correlations of the SAM domain from EphB2. (a) The G-(i+1)-HSQC was acquired in 25 min using 8 scans. The SAM domain contains three glycines which results in three strong signals from sequential neighbors of Gly. Weak signals from the Asn and Gln are present, which are due to NDH-groups (see text). (b) The G-(i,i+1)-HSQC was acquired in 50 min using 16 scans. In addition to the three signals of the Gly residues the sequential neighbors are also present with weak intensity. Arg $\text{N}^{\epsilon}\text{H}^{\epsilon}$ peaks also appear in the spectrum (see text). These peaks are weaker and distorted since Arg C^{δ} is not in the chemical shift range covered by the semi-selective pulses.

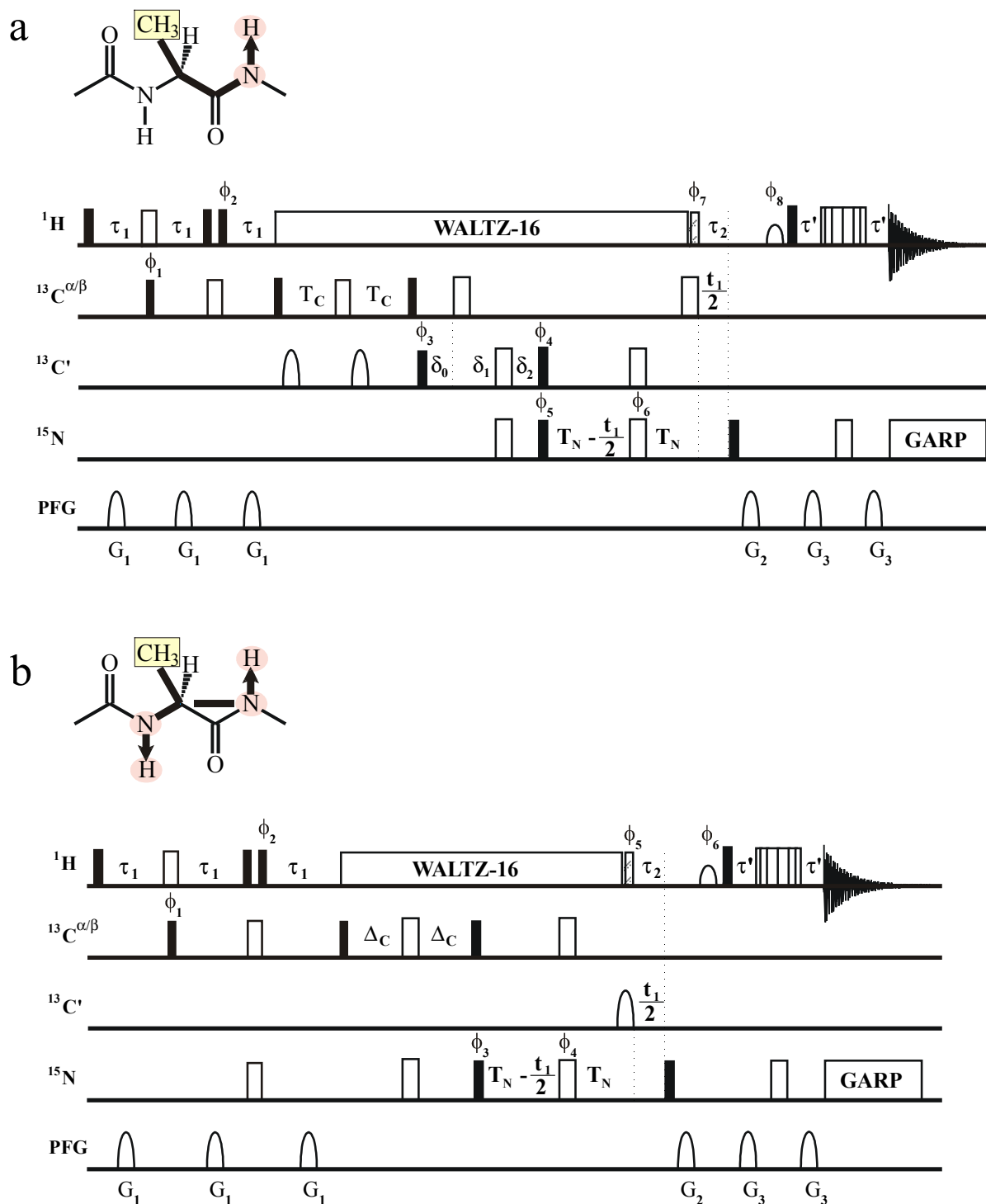


Fig. 4.5: Pulse sequences of the two Ala selective experiments: (a) A-(i+1)-HSQC and (b) A-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_C = 9.3$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $T_C = 4.5$ ms. More experimental details are given in the appendix.

4.4.2 Selecting Ala: The A-(i+1)- and A-(i,i+1)-HSQC

Alanine is the only amino acid with a methyl group in the C^β position. By implementing a MUSIC- CH_3 selection into the CBCA(CO)NH and the CBCANH (both introduced in

chapter 2.3) alanine is the only amino acid whose coherences can pass through the experiments. The pulse sequences of the A-(i+1)- and A-(i,i+1)-HSQC are shown in Figs. 4.5 a and b and the spectra in Figs. 4.6 a and b.

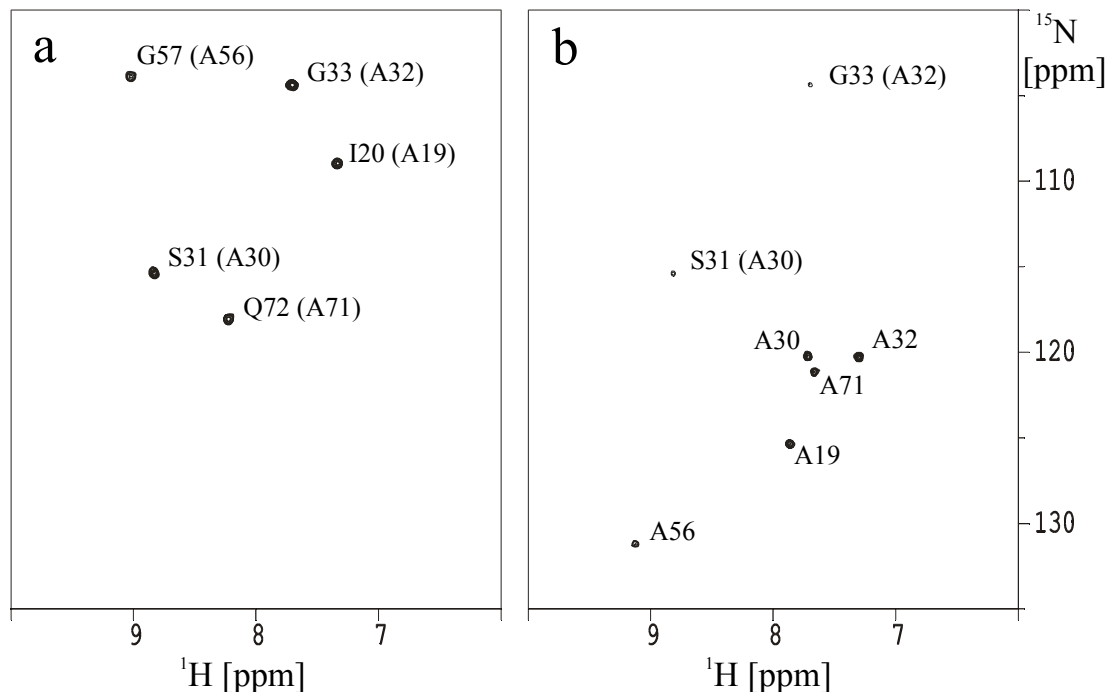
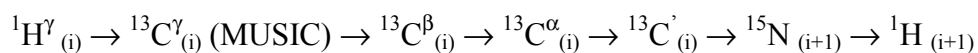


Fig. 4.6: Alanine selective ^1H - ^{15}N -correlations of the SAM domain from EphB2. The A-(i+1)-HSQC and the A(i,i+1)-HSQC were acquired in 2 h 15 min each using 48 scans. The SAM domain contains five alanines. In the (i+1) experiment (a) the expected five signals from sequential neighbors of Ala are present. In the (i,i+1) experiment (b) all five Ala residues show strong signals while two weak signals result from amino acids in the (i+1) position of the alanine residues.

4.4.3 Selecting Thr, Val, Ile and Ala: The TA-(i+1)- and TA-(i,i+1)-HSQC, the VIA-(i+1)- and VIA-(i,i+1)-HSQC

Threonine, valine and isoleucine are the only amino acids with a methyl group in the C^γ position. The implementation of a MUSIC- CH_3 selection into CBCA(CO)NH and CBCANH experiments extended by a relay step in which magnetization from C^γ , C^β and C^α is transferred to the NH groups, results in the selection of Ala, Ile, Val and Thr side chains. Note that magnetization from $^\beta\text{CH}_3$ as well as from $^\gamma\text{CH}_3$ results in signals. The magnetization transfer in the TAVI-(i+1)-HSQC can be described by:



Since the chemical shifts of the $^{13}\text{C}^\beta$ nuclei of Ile/Val are sufficiently different from those of Thr, selective pulses can be applied to suppress either Val/Ile or Thr.

Ile and Val can be suppressed by a band selective 180° REBURP pulse (Geen and Freeman, 1991) with a band width of 26 ppm at 68.5 ppm (Fig. 4.7) during refocusing $2C_z^\beta C_y^\alpha$ magnetization (delay $2T_C$ and $2\Delta_C$): only the C^α but not the C^β nuclei of Ile and Val are affected. Therefore antiphase magnetization can not be refocused and this prevents the creation of detectable magnetization. In the case of Thr, both nuclei, the C^β and C^α are affected as in the experiment without the band selective pulse and $2C_z^\beta C_y^\alpha$ is refocused to C_x^α . In Alanine the magnetization is already transferred to the C^α and the coupling already refocused at the beginning of that delay. The selective pulse will only prevent loss of intensity due to the evolution of coupling to the C^β . The pulse sequences of the TA-(i+1)-HSQC and the TA-(i,i+1)-HSQC are shown in Fig. 4.8, the resulting spectra in Fig. 4.9.

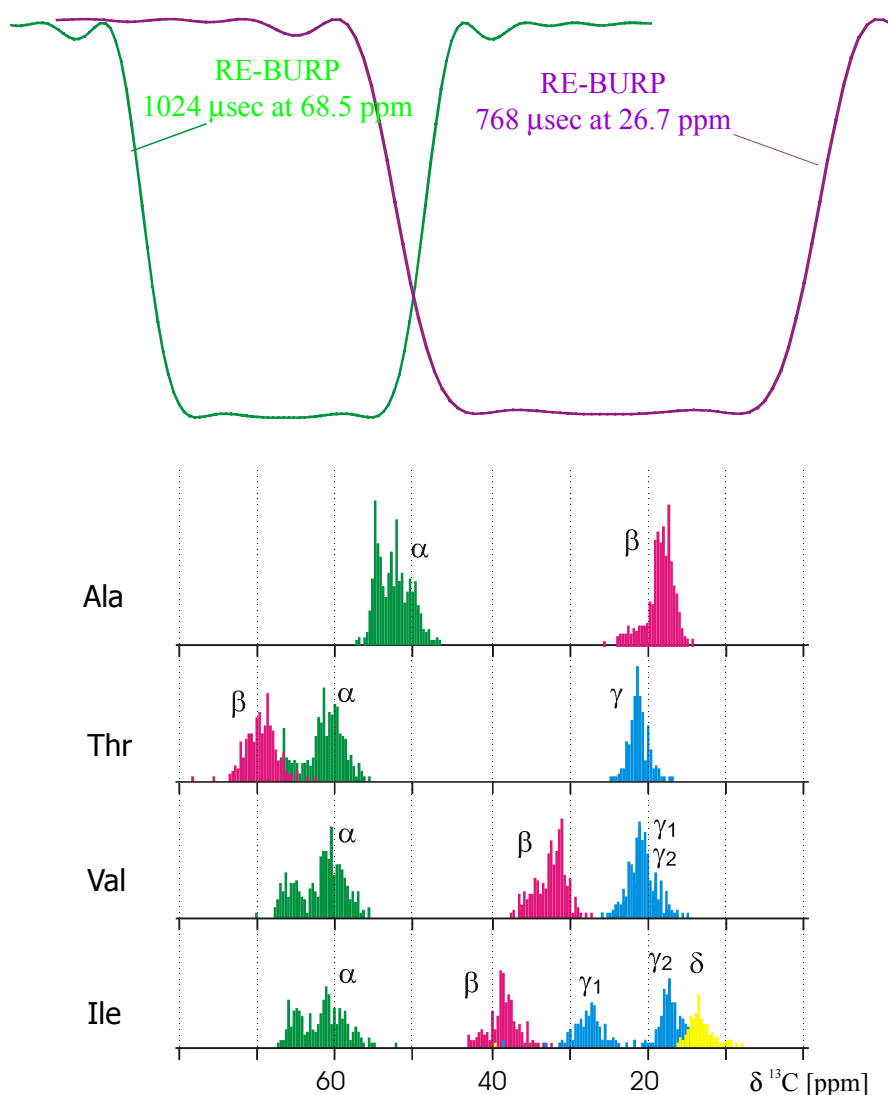


Fig. 4.7: Statistical ^{13}C chemical shift data from the BMRB database (Seavey et al., 1991) are given at the bottom for Leu, Ala, Thr, Val, and Ile. At the top the inversion profiles of the three applied REBURP pulses used in the TA- and VIA-HSQC experiments S_1 and S_2 (pulse lengths are given for 600 MHz spectrometer).

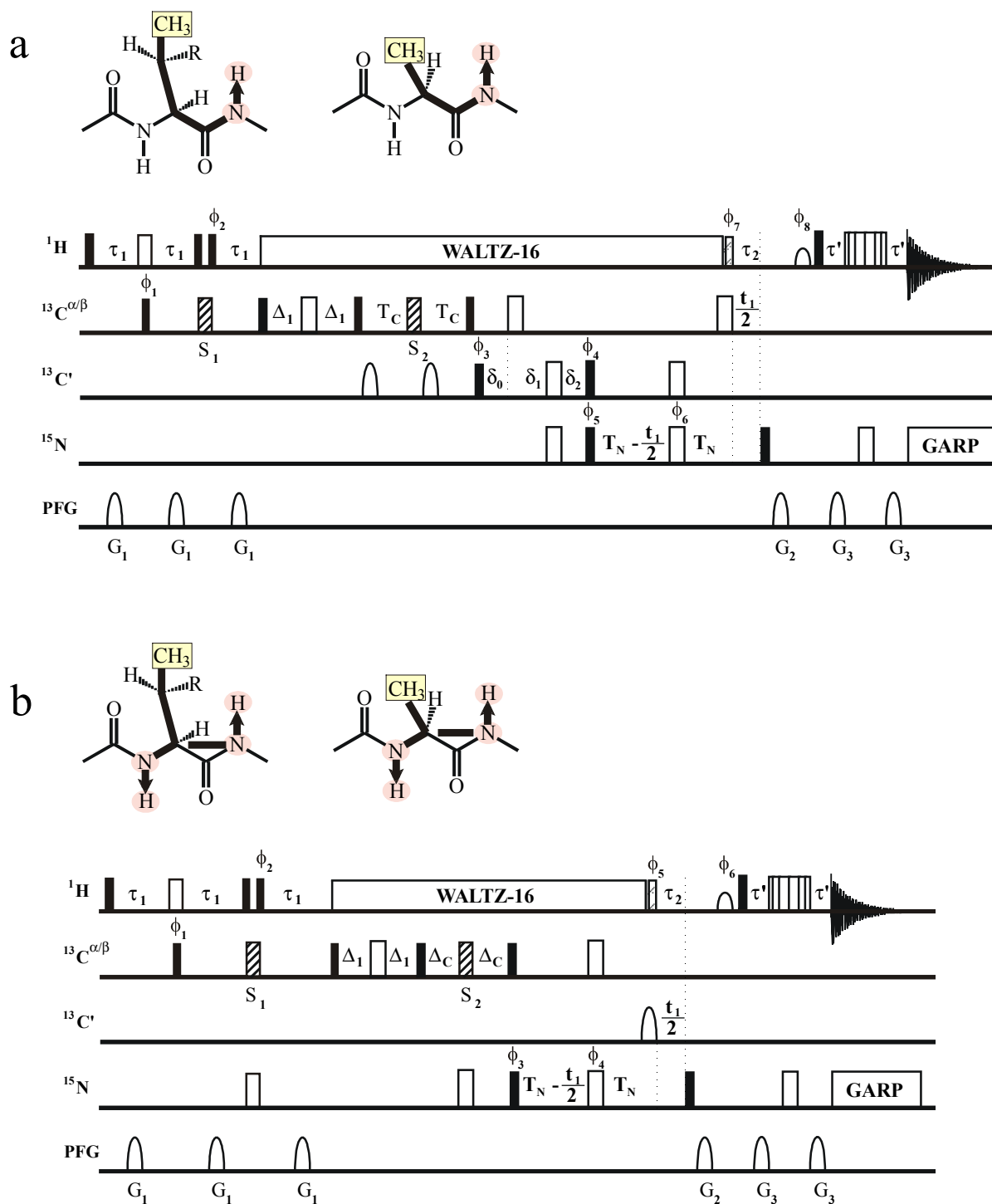


Fig. 4.8: Pulse sequences of the Thr/Ala and Val/Ile/Ala selective experiments: (a) TA-(i+1)-HSQC with a normal rectangular pulse S_1 and the selective pulse S_2 at 68.5 ppm or VIA-(i+1)-HSQC with the selective pulse S_1 at 26.7 ppm and a normal rectangular pulse S_2 , (b) TA-(i,i+1)-HSQC with the selective pulse S_2 at 68.5 ppm or VIA-(i,i+1)-HSQC with the selective pulse S_1 at 26.7 ppm. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_1 = 4$ ms, $\Delta_C = 9.3$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $T_C = 4.5$ ms. The selective pulses S_1 and S_2 are carried out as REBURP pulses of 768 μs and 1024 μs , respectively (pulse lengths for 600 MHz spectrometer). More experimental details are given in the appendix.

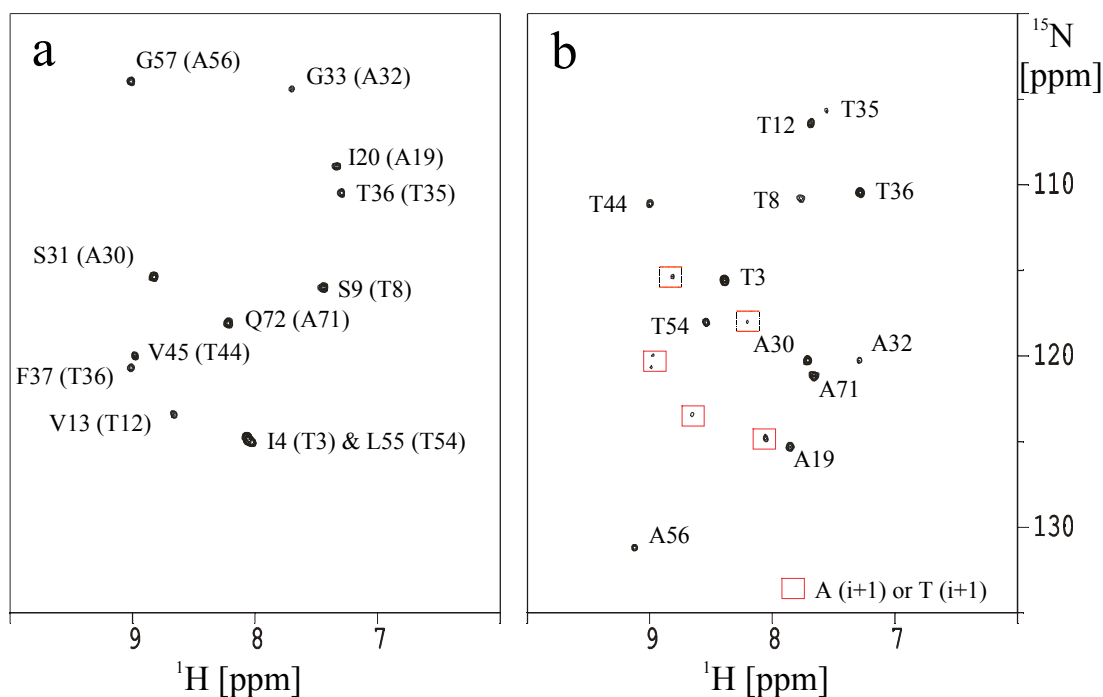


Fig. 4.9: Threonine/alanine selective ^1H - ^{15}N correlations of the SAM domain from EphB2. (a) The TA-(i+1)-HSQC was acquired in 2 h 15 min using 48 scans. (b) The TA-(i,i+1)-HSQC was acquired in 4 h 30 min using 96 scans. The SAM domain contains 7 threonins. In the TA-(i+1)-HSQC (a) signals from sequential neighbors of all 7 Thr and 5 Ala are present. In the TA-(i,i+1)-HSQC all expected signals of the 7 Thr and 5 Ala are visible. Additionally small signals of sequential neighbors of Thr and Ala are present which can be identified by comparison with the TA-(i+1)-HSQC. The (i+1) signals are marked with thin rectangles.

A Val/Ile/Ala (VIA) selective experiment can be obtained by replacing the 180° -pulse S_1 (Fig. 4.7) by a REBURP pulse of 768 μsec (covering a range of 48 ppm) which is centered at 26.7 ppm. It will affect the magnetization of the C^γ and C^β of both Val and Ile but not the C^β magnetization of Thr (Fig. 4.9). The transfer of magnetization from C^γ to C^β in Thr and thus the transfer to the C^α in the next COSY-step will not be effective anymore, hence Thr is suppressed. The C^α of Ala is not affected by the REBURP pulse at 26.7 ppm and the C^β magnetization stays in-phase during the first COSY-step. As in the Ala selective experiment magnetization is then transferred to the C^α by the consecutive COSY-step and subsequently to the NH. The pulse sequences of the VIA-(i+1)-HSQC and the VIA-(i,i+1)-HSQC are shown in Fig. 4.8, spectra in Fig. 4.10.

4.4.4 Selecting Leu and Ala: The LA-(i+1)- and LA-(i,i+1)-HSQC

Leucine is one of the six amino acids which contain methyl groups. It has two methyl groups in the C^δ position. Leu has characteristic carbon chemical shifts and does not branch in the C^β position compared to Val, Ile and Thr. The implementation of a MUSIC- CH_3 selection into the CBCA(CO)NH and the CBCANH that have been extended by two relay

steps, results in a selection of Ala, Ile, Val, Thr and Leu side chains. In order to transfer the magnetization of a selected Leu δCH_3 group to the C^α , three COSY steps are necessary. To suppress the signals of Ile, Val and Thr, a combination of several selective pulses and tuned delays was used. The pulse sequences of the two Leu/Ala-selective experiments, the LA-(i+1)- and the LA-(i,i+1)-HSQC, are shown in Fig 4.11.

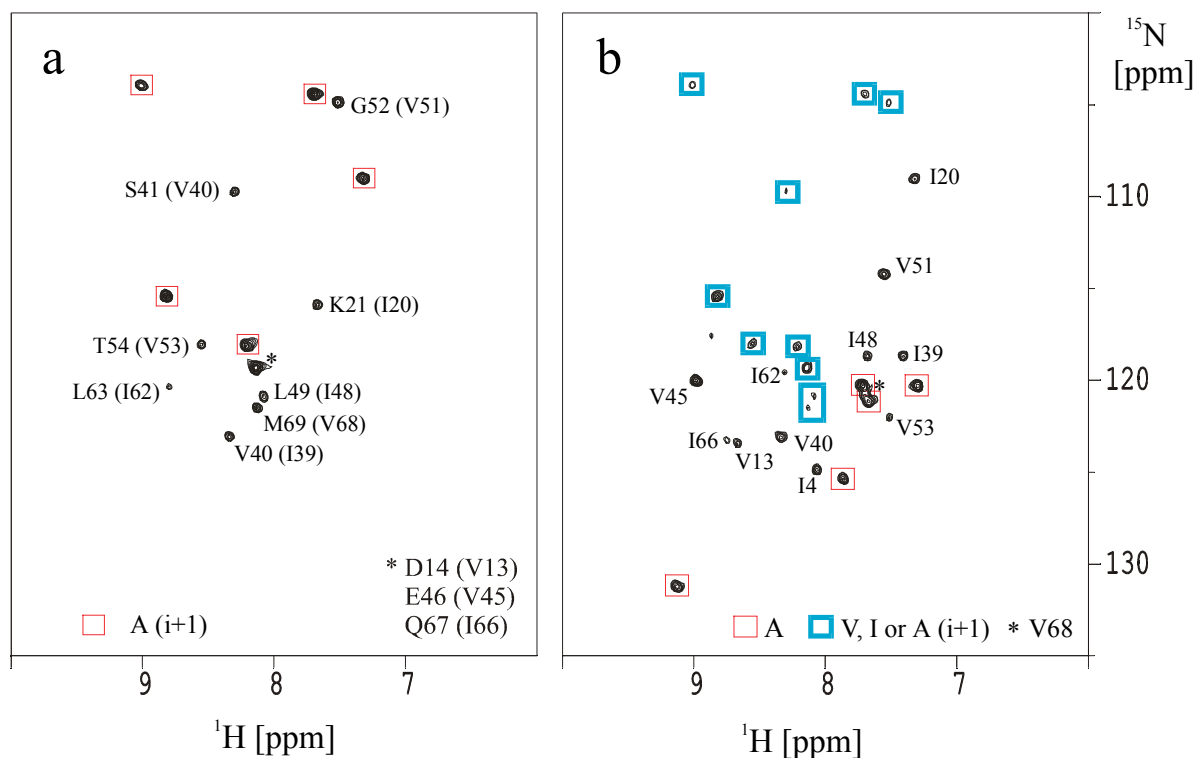
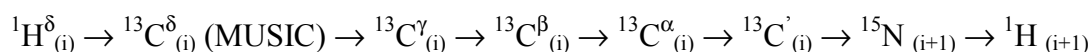


Fig. 4.10: Threonine/alanine selective ^1H - ^{15}N correlations of the SAM domain from EphB2. (a) The VIA-(i+1)-HSQC was acquired in 2 h 15 min using 48 scans. (b) The VIA-(i,i+1)-HSQC was acquired in 4 h 30 min using 96 scans. The SAM domain contains 6 Ile and 6 Val. In the VIA-(i+1)-HSQC (a) all expected signals from sequential neighbors of Ile, Val and Ala are present (only five Ile (i+1) signals can appear since I4 has a Pro as sequential neighbor). Ala neighbors are marked with thin rectangles. In the VIA-(i,i+1)-HSQC all expected signals of the 6 Ile, 6 Val and 5 Ala are visible. Additionally small signals of sequential neighbors of Ile, Val and Ala are present which can be identified by comparison with the VIA-(i+1)-HSQC. The (i+1) signals are marked with thick rectangles, Ala signals with thin rectangles. Thr signals are well suppressed in both experiments.

The magnetization transfer of Leu in the LA-(i+1)-HSQC can be described by:



Three selective REBURP pulses were implemented in the pulse sequence (S_1 , S_2 and S_3 in Fig. 4.11), their excitation profile is shown in Fig. 4.12 together with the statistical chemical shifts of the amino acids involved.

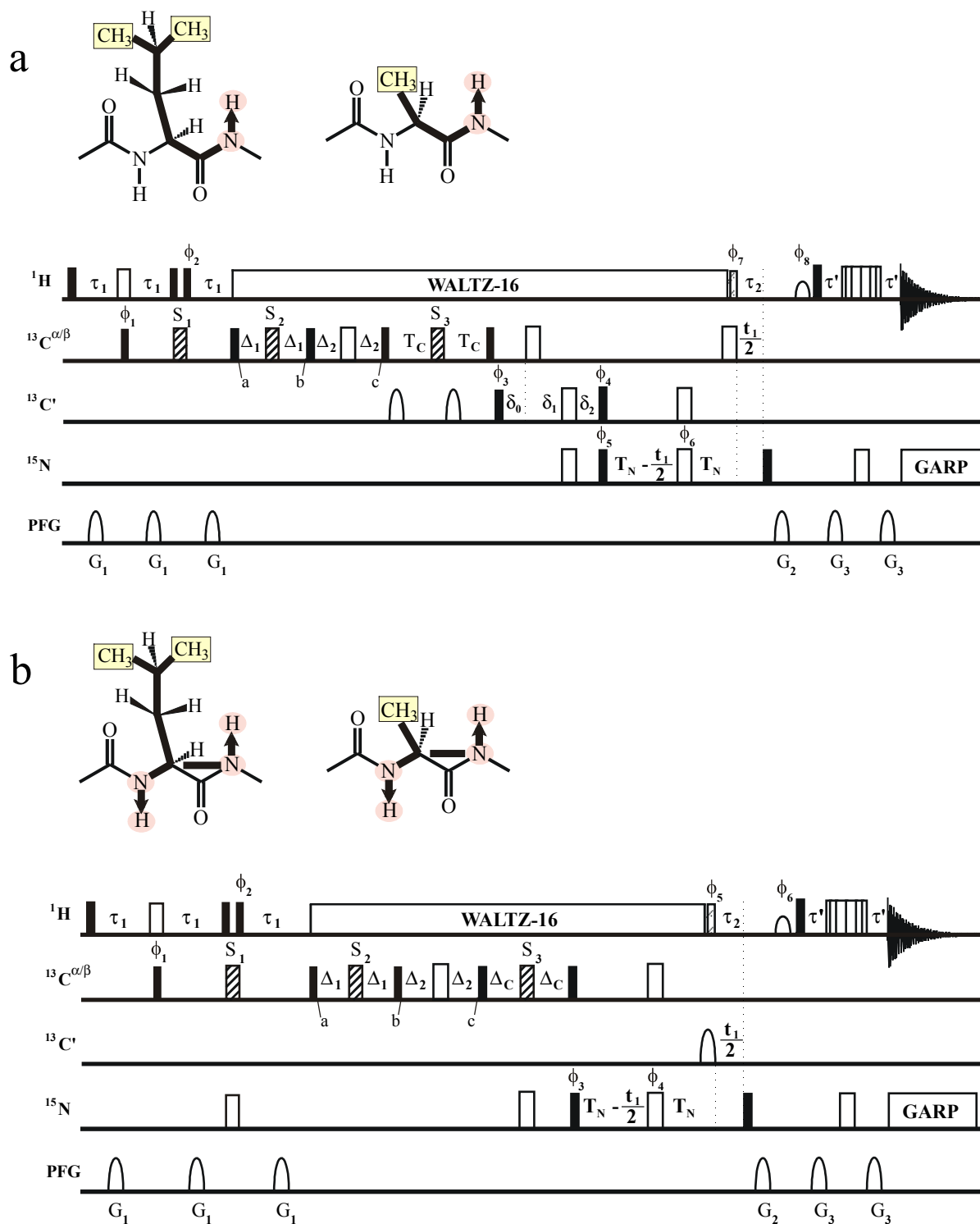


Fig. 4.11: Pulse sequences of the Leu/Ala selective experiments: (a) LA-(i+1)-HSQC and (b) LA-(i,i+1)-HSQC with three selective REBURP pulses (striped thick bars): $S_1 = 1024 \mu\text{s}$ at 18.5 ppm, $S_2 = 768 \mu\text{s}$ at 35 ppm and $S_3 = 1024 \mu\text{s}$ at 48 ppm (pulse lengths are given for a 600 MHz spectrometer). The following delays were used: $\tau_1 = 4 \text{ ms}$, $\tau_2 = 5.5 \text{ ms}$, $\tau' = 2.25 \text{ ms}$, $\Delta_1 = 3.6 \text{ ms}$, $\Delta_2 = 7.1 \text{ ms}$, $\Delta_C = 9 \text{ ms}$, $\delta_0 = 4.5 \text{ ms}$, $\delta_1 = 6.9 \text{ ms}$, $\delta_2 = 11.4 \text{ ms}$, $T_N = 11 \text{ ms}$, $T_C = 4.5 \text{ ms}$. More experimental details are given in the appendix.

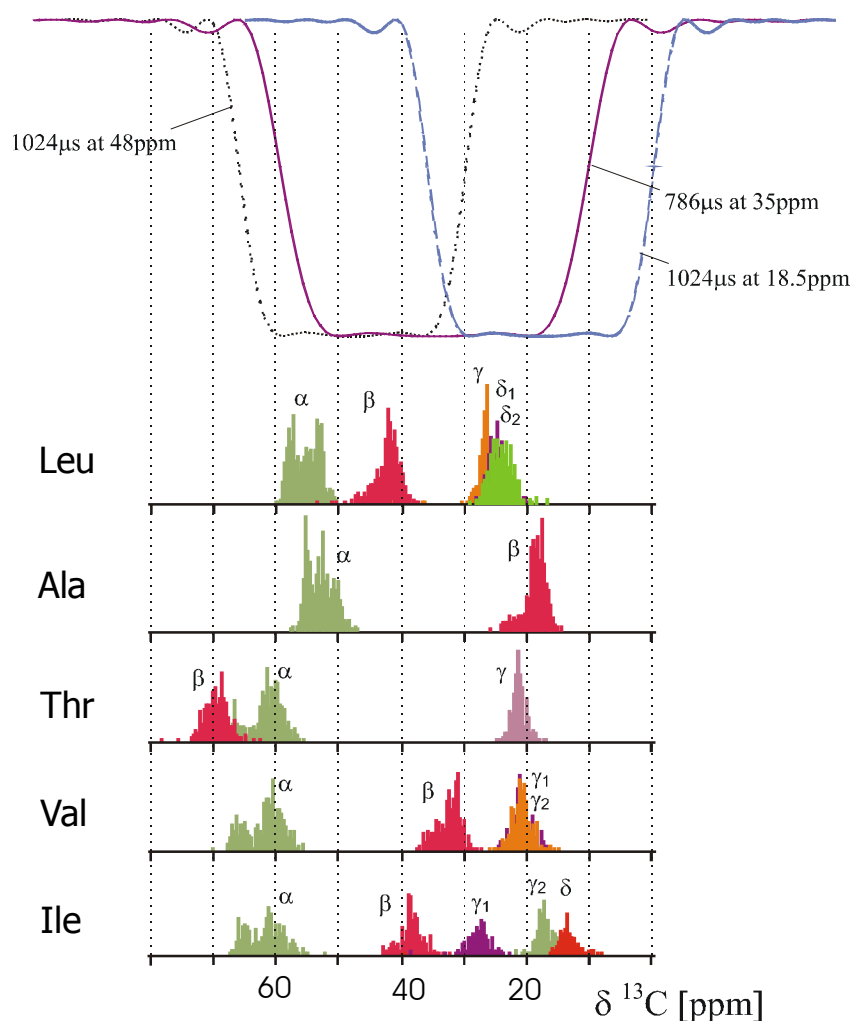


Fig. 4.12: Statistical ^{13}C chemical shift data from the BMRB database (Seavey et al., 1991) are given at the bottom for Leu, Ala, Thr, Val, and Ile. At the top the inversion profiles of the three applied REBURP pulses used in the LA-HSQC experiments S_1 , S_2 and S_3 (pulse lengths are given for 600 MHz spectrometer).

The REBURP pulse S_1 had a length of 1024 μs and was centered at 18.5 ppm. The pulse affects both the C^δ and C^γ of Leu during $2\tau_1$. Antiphase magnetization $2\text{C}_y^\delta\text{C}_z^\gamma$ is created and converted by a 90° pulse to $2\text{C}_z^\delta\text{C}_y^\gamma$ (point a). The second REBURP pulse S_2 at 35 ppm (with a length of 768 μs) affects the C^δ , C^γ and C^β of Leu. Antiphase magnetization $2\text{C}_z^\delta\text{C}_y^\gamma$ is refocused with respect to the C^δ and evolves into antiphase magnetization $2\text{C}_y^\gamma\text{C}_z^\beta$ during $2\Delta_1$ (point b). During the same time the magnetization of interest stays in-phase with respect to the other C^δ . This antiphase magnetization $2\text{C}_y^\gamma\text{C}_z^\beta$ is converted to $2\text{C}_y^\gamma\text{C}_z^\beta$ and refocused during $2\Delta_2$ while coupling to the α -carbon leads to $2\text{C}_y^\beta\text{C}_z^\alpha$ (point c). This is converted into $2\text{C}_z^\beta\text{C}_y^\alpha$ and refocused during $2T_C$ (or $2\Delta_C$) because the third REBURP pulse S_3 at 48 ppm (with a length of 1024 μs) allows coupling between Leu C^α and C^β . Taken together, this results in the following transfer functions for Leu:

$$\sin(\pi J_{CC}2\tau_1) \sin^2(\pi J_{CC}2\Delta_1) \cos(\pi J_{CC}2\Delta_1) \sin^2(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2T_C) \quad [4.1]$$

$$\sin(\pi J_{CC}2\tau_1) \sin^2(\pi J_{CC}2\Delta_1) \cos(\pi J_{CC}2\Delta_1) \sin^2(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_C) \quad [4.2]$$

The Thr- C^β is not affected by any of the three selective pulses. The magnetization C^γ_x can therefore not evolve into antiphase magnetization with respect to the C^β during $2\tau_1$ and $2\Delta_1$. During $2T_R$ the antiphase magnetization $2C^\gamma_y C^\beta_z$ is created and converted into $2C^\gamma_z C^\beta_y$ by the next 90° pulse. This antiphase magnetization is not refocused during T_C (or $2\Delta_C$). No magnetization is therefore transferred to C^α and signals of Thr are suppressed. From similar consideration it is clear that signals from Ala will not be suppressed.

Starting from the γ -CH₃ of Val and Ile there are three possible magnetization pathways. The pulse sequence contains three COSY steps in the beginning while it takes only two steps to transfer the magnetization from the C^γ to the C^α . Consequently the magnetization may stay in-phase during any of the steps and still result in a detectable signal. Depending on the delay during which no antiphase magnetization is created, this results in three possible transfer functions for the (i+1) sequence:

$$\cos(\pi J_{CC}2\tau_1) \sin(\pi J_{CC}2\Delta_1) \sin^2(\pi J_{CC}2\Delta_2) \cos(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2T_C) \quad [4.3]$$

$$\sin(\pi J_{CC}2\tau_1) \sin(\pi J_{CC}2\Delta_1) \cos^2(\pi J_{CC}2\Delta_1) \sin(\pi J_{CC}2\Delta_2) \cos^2(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2T_C) \quad [4.4]$$

$$\sin(\pi J_{CC}2\tau_1) \sin^2(\pi J_{CC}2\Delta_1) \cos(\pi J_{CC}2\Delta_1) \sin(\pi J_{CC}2\Delta_2) \cos(\pi J_{CC}2T_C) \quad [4.5]$$

The transfer functions for the (i,i+1) experiment are identical, except that T_C is replaced by Δ_C . If no selective pulses are used, all three pathways lead to detectable magnetization. The selective pulse S_1 refocuses the coupling between Ile- C^β and $C^{\gamma 2}$ and partially also between Val- C^β and C^γ during $2\tau_1$ so that the last two pathways [4.4] and [4.5] are suppressed. Pathways [4.3] and [4.4] for Ile/Val are suppressed in analogy to case E and D described in chapter 3.6, the delay $2\Delta_2$ is set to $1/2J_{CC}$ (14.2 ms) and $\cos(\pi J_{CC}2\Delta_2)$ becomes zero. C^β magnetization of Val and Ile evolves into antiphase magnetization with respect to the second C^γ during $2\Delta_2$ and the next 90° pulse converts it into multiple quantum coherence that is not converted back into detectable magnetization during the rest of the pulse sequence.

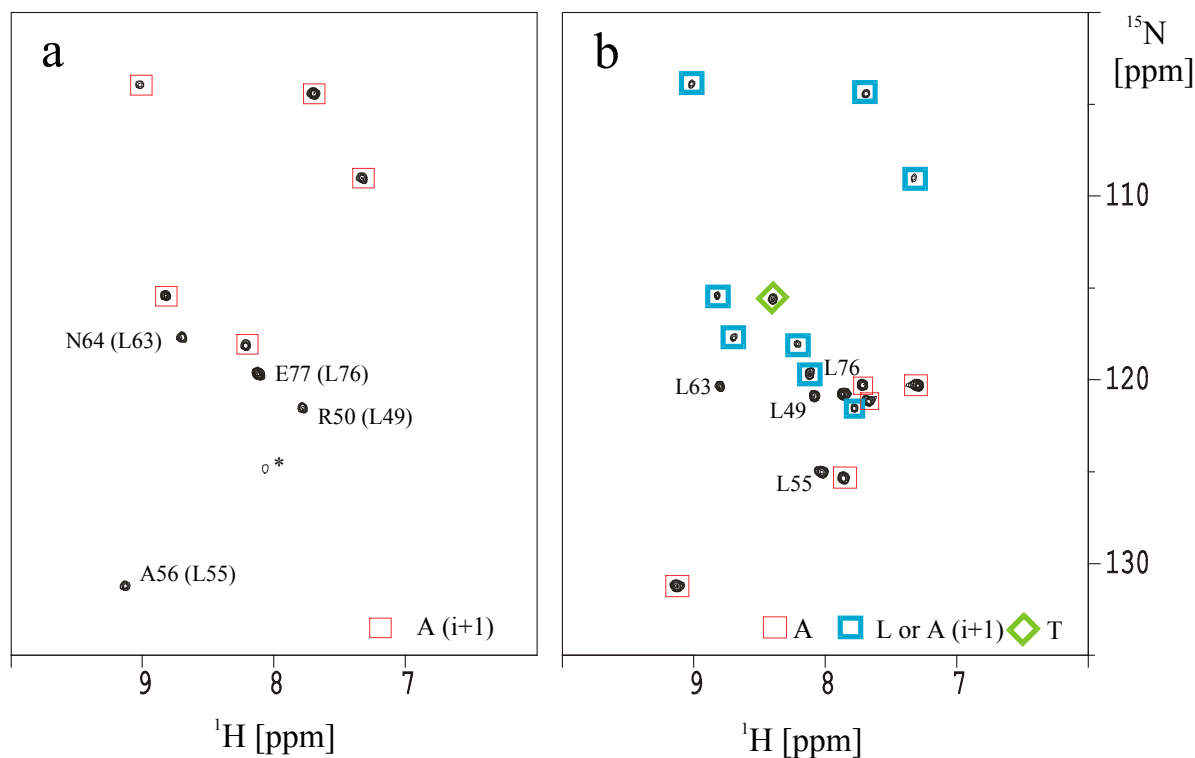


Fig. 4.13: Leu/Ala selective ^1H - ^{15}N correlations of the SAM domain from EphB2. (a) The LA-(i+1)-HSQC was acquired in 4 h 30 min using 96 scans. (b) The LA-(i,i+1)-HSQC was acquired in 6 h 30 min using 144 scans. The SAM domain contains five Leu and five Ala. In the LA-(i+1)-HSQC (a) signals from sequential neighbors of Leu and Ala are present. Ala neighbors are marked with thin rectangles. D18, the sequential neighbor of L17 is absent. Since the suppression of Val/Ile/Thr is solely based on appropriate tuning the delay Δ_2 together with selective pulses, the suppression is not perfect. Weak negative or positive signals of Val/Ile/Thr could appear (marked with *). In the LA-(i,i+1)-HSQC (b) four Leu signals are visible, L17 is absent. All signals Ala-residues sequential (marked with thin rectangles) and their sequential neighbors are present. The (i+1) peaks are marked with thick rectangles, they can be identified by comparison with the LA-(i+1)-HSQC. Again the suppression of Thr/Ala/Val is not perfect and the signal of T3 is visible.

Starting from the δ - CH_3 of Ile magnetization can also be transferred to the C^α . This pathway is suppressed in a similar manner described in chapter 3.6 by setting $2\Delta_2 = 1/2J_{\text{CC}}$. During this delay magnetization originating from the C^δ of Ile has reached the C^β . Antiphase magnetization with respect to the C^α and the second C^γ of Ile is created and subsequently converted into undetectable multiple quantum coherence.

The suppression of Val, Ile and Thr signals is good but not perfect, because of small variations of the coupling constants $^1J_{\text{CC}}$ between the side chain carbons. The corresponding spectra are shown in Fig. 4.13.

4.4.5 Selecting Ser: The S-(i+1)- and S-(i,i+1)-HSQC

Serine, one of the 15 amino acids with a methylene group in the C^β position, has a characteristic C^β chemical shift between 60 and 70 ppm which makes it unique among these 15 amino acid types. The implementation of a MUSIC- CH_2 selection into the CBCA(CO)NH and the CBCANH experiments, together with selective pulses exploiting the characteristic C^β chemical shift, results in two serine-selective experiments: the S-(i+1)- and the S-(i,i+1)-HSQC (Fig. 4.15).

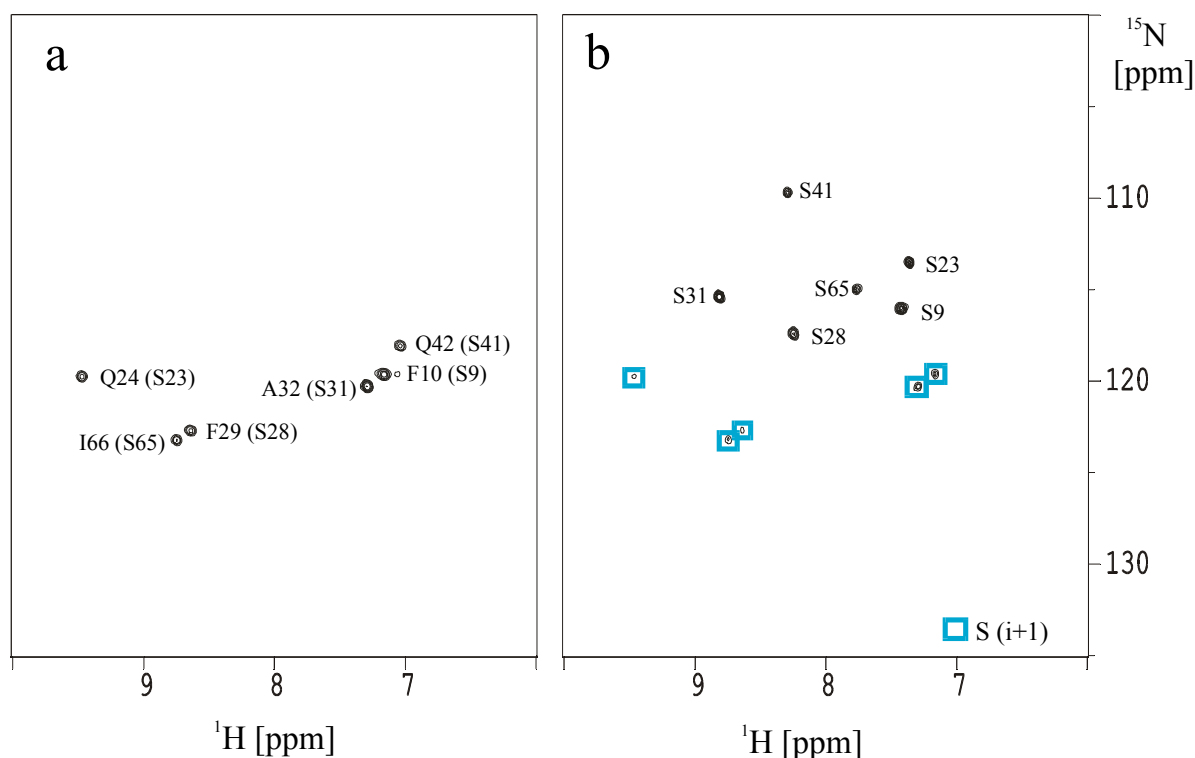


Fig. 4.14: Ser selective 1H - ^{15}N -correlations of the SAM domain from EphB2. (a) The S-(i+1)-HSQC was acquired in 1 h 30 min using 32 scans. The SAM domain contains six serines which results in six strong signals from sequential neighbors of Ser. (b) The S-(i,i+1)-HSQC was acquired in 1h 30 min using 32 scans. In addition to the six signals of the Ser residues five of the sequential neighbors are also present with weak intensity. They are marked with small rectangles.

Both carbon 180° pulses at the beginning of the sequence are replaced by REBURP pulses centered at 60.5 ppm with a bandwidth of 18 ppm (pulse length of 2048 μ sec at 600 MHz) which affect both the serine C^α and C^β nuclei. The C^β magnetization of the other amino acids selected by MUSIC remains, however, unperturbed by these REBURP-pulses while their C^α magnetization is inverted. Since no coupling between C^β and C^α will thus evolve, no magnetization can be transferred to the C^α and hence to the NH in non-serine amino acids. The corresponding spectra are shown in Fig. 4.14.

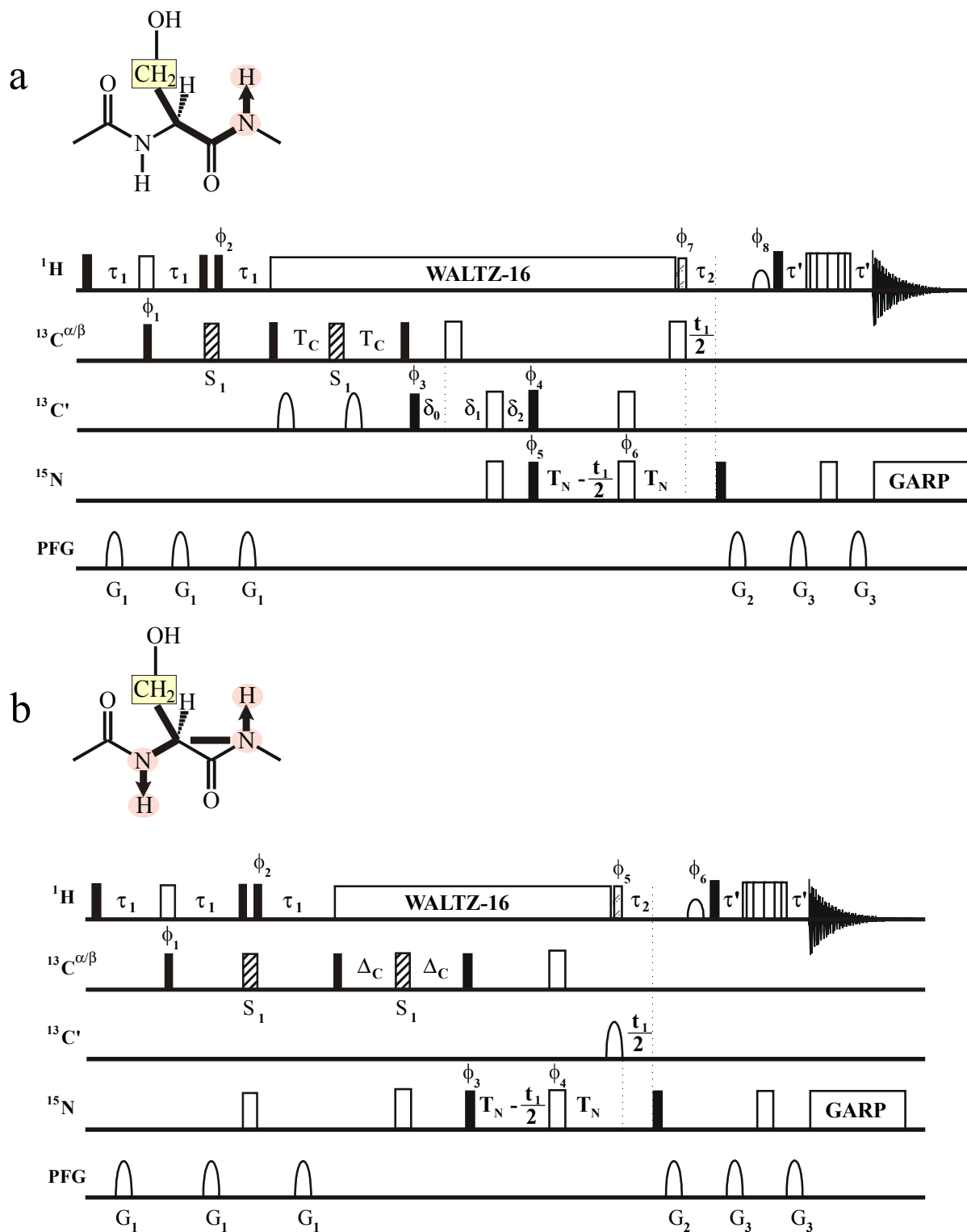


Fig. 4.15: Pulse sequences of the Ser selective experiments: (a) S-(i+1)-HSQC and (b) S-(i,i+1)-HSQC with two selective REBURP pulses (S_1 , indicated by the striped thick bars) with a pulse length of 2048 μs at 60.5 ppm (the pulse length is given for a 600 MHz spectrometer). The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_C = 9$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $T_C = 4.5$ ms. More experimental details are given in the appendix.

4.4.6 Selecting aromatic residues: The W-(i+1)- and W-(i,i+1)-HSQC, the FYH-(i+1)- and FYH-(i,i+1)-HSQC

The four amino acids containing an aromatic ring: His, Phe, Trp and Tyr have a methylene group in the C^β position adjacent to a C^γ which is part of the aromatic system. Aromatic carbons have characteristic chemical shifts which are clearly separated from the regions of aliphatic or carbonyl carbons. This criterion is exploited by using selective pulses which affect only aromatic carbons.

The experiments are based on the CBCA(CO)NH and the CBCANH experiments in which the initial refocused INEPT element is replaced by MUSIC selective for CH₂. To select only the desired methylene carbons, the MUSIC element is combined with an HMQC-like sequence (Müller, 1979; Bax et al., 1983) with shaped 90° pulses. During the HMQC-like sequence 2C^β_yC^γ_y double-quantum coherence is selected by phase cycling. Subsequently, the magnetization is transferred from the C^β to the C^α as in the original experiments. The two pulse sequences are shown in Fig. 4.16.

Self-refocusing LOS2-0 pulses as described in chapter 3.5 are used which affect only aromatic C^γ carbons. This selection requires relatively long selective pulses. Since the J_{C^βC^α} and J_{C^βC^γ} coupling is active during the selective pulse, the delay Δ₁ has to be adjusted to compensate for this. The effective coupling between the C^β and C^γ carbon during the pulse is reduced and was determined experimentally to be 60% of the coupling during a free precession period as outlined in chapter 3.5. The resulting transfer functions for the selected amino acids in the experiments in Figs. 4.16a and b are therefore:

$$\sin^2(\pi J_{C^{\beta}C^{\gamma}}(\Delta_1+0.6P)) \sin(\pi J_{C^{\alpha}C^{\beta}}(2\Delta_1+2P)) \sin(\pi J_{C^{\alpha}C^{\beta}}2\Delta_2) \sin(\pi J_{C^{\alpha}C^{\gamma}}2\Delta_2) \quad [4.6]$$

$$\sin^2(\pi J_{C^{\beta}C^{\gamma}}(\Delta_1+0.6P)) \sin(\pi J_{C^{\alpha}C^{\beta}}(2\Delta_1+2P)) \sin(\pi J_{C^{\alpha}C^{\beta}}2\Delta_3) \sin(\pi J_{C^{\alpha}N}2\Delta_3) \quad [4.7]$$

P is the pulse length of the LOS2-0 pulses. Note that the value for the J_{C^βC^γ} coupling constants depends on the type of aromatic amino acid (Löhr et al., 2000).

The statistical chemical shift regions of the C^γ of the aromatic amino acids are shown in Fig. 4.17. A self-refocusing LOS2-0 pulse with a duration of 640 μs centered at 120 ppm covers the full range of all aromatic C^γ carbons. A more detailed view shows that the area of the ¹³C chemical shift of the C^γ of Tyr, Phe and His is well separated from the C^γ of Trp. Therefore, selective pulses with a narrower bandwidth were chosen to select either Phe, Tyr and His or Trp.

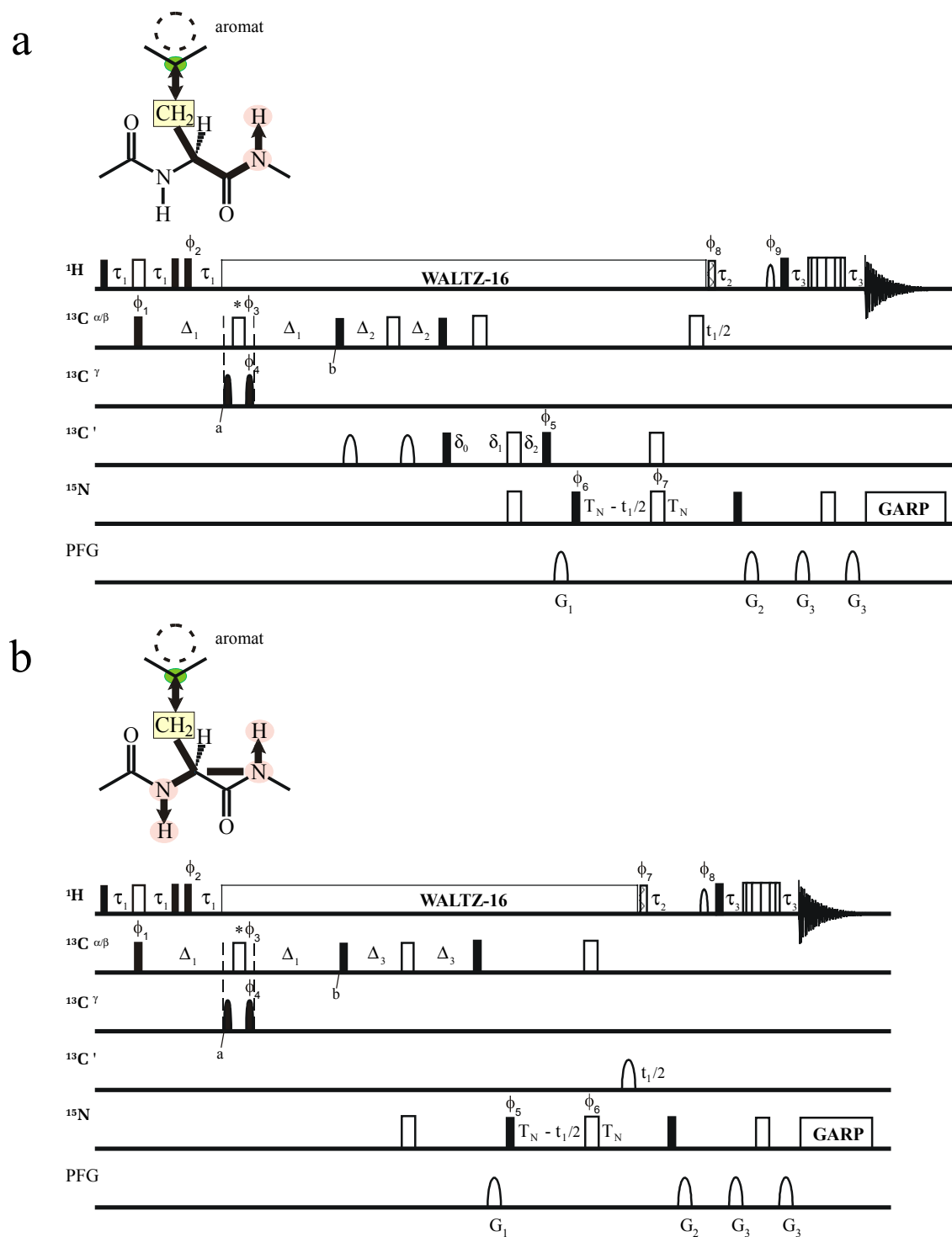


Fig. 4.16: Pulse sequences for the aromatic amino acid-selective ^1H - ^{15}N correlations. The pulse sequences yield spectra selective for Phe/Tyr/His or solely Trp, depending on the choice of the offset for the two selective L0S2-0 pulses effecting the C^γ . (a) FYH-($i+1$)- or W-($i+1$)-HSQC and (b) FYH-($i,i+1$)- or FYH-($i,i+1$)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 6.7$ ms for FYH and 6.5 ms for W, $\Delta_2 = 4.5$ ms, $\Delta_3 = 9$ ms. The pulse lengths of the two L0S2-0 pulses (filled sine shapes) were set to 1024 μs , offsets to 135 ppm for the FYH-($i+1$)- and FYH-($i,i+1$)-HSQC, but to 105 ppm for the W-($i+1$)- and W-($i,i+1$)-HSQC. The rectangular 180° pulse (marked with *) had a pulse length of 72 μs and was applied at 45 ppm in the FYH-($i+1$)- and FYH-($i,i+1$)-HSQC, but at 30 ppm in the W-($i+1$)- and W-($i,i+1$)-HSQC. All pulse lengths are given for a 600 MHz spectrometer. More experimental details are given in the appendix.

To excite and select only the C^γ of His, Phe and Tyr but not of Trp, a selective pulse covering a range of 33 ppm (5000 Hz at 600 MHz) centered 135 ppm was necessary. Self-refocusing LOS2-0 pulses with a duration of 1024 μs were applied. The rectangular 180° pulse (72 μs) in between the LOS2-0 pulses was executed at 45 ppm and was adjusted to have a null in the excitation profile at 125 ppm. Assuming a $J_{C\beta C\gamma} = 44$ Hz (20) and neglecting relaxation effects, nominal values for Δ_1 , Δ_2 and Δ_3 are 8.15 ms, 5.3 ms and 9 ms, respectively. However, after experimental optimization, values for Δ_1 , Δ_2 and Δ_3 were chosen to be 6.7 ms, 4.5 ms and 9 ms, respectively. The resulting spectra of the FYH-(i+1)- and FYH-(i,i+1)-HSQC are shown in Fig. 4.18. They contain only signals originating from His, Phe and Tyr.

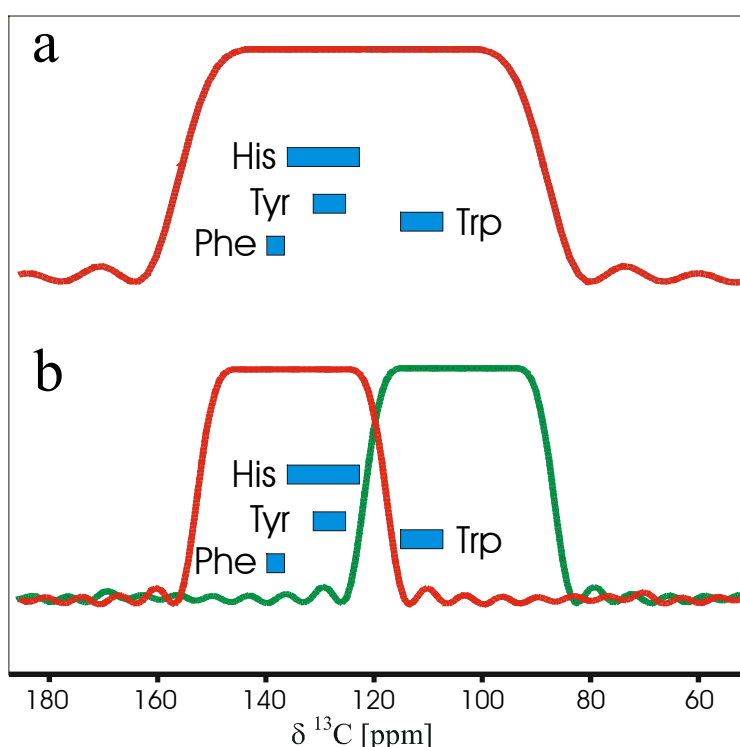


Fig. 4.17: ^{13}C chemical shift distribution for the C^γ of the aromatic amino acids taken from the BioMagRes database (Seavey et al., 1991) and the excitation profiles of the LOS2-0 pulses: (a) with a length of 640 μs to select all aromatic amino acids at 120 ppm and (b) with a length of 1024 μs at either 135 ppm or 105 ppm to select Phe, Tyr and His or solely Trp.

In order to select Trp residues, but not Tyr, His and Phe, the same LOS2-0 pulses as in the FYH experiments but centered at 105 ppm were used (Fig. 4.17). The rectangular 180° pulse in between the LOS2-0 pulses has the same length as in the FYH experiments, but the offset is now set to 30 ppm and adjusted to exhibit a null in the excitation profile at 110 ppm. Using a $J_{C\beta C\gamma} = 46.6$ Hz (Löhr et al., 2000) and neglecting relaxation effects, the nominal value for Δ_1 is 7.95 ms. The best signal to noise ratio was found with a value of 6.5 ms for Δ_1 . The values for the delays Δ_2 and Δ_3 are identical to those in the FYH experiments. The

spectra of the W-(i+1)- and W-(i,i+1)-HSQC contain only signals originating from Trp and are shown in Fig. 4.18.

The quality of the suppression in both types of experiments is determined by the excitation profile of the two selective 90° pulses. Small mis-calibrations of the pulse power level, can lead to breakthrough peaks from Asn/Asp and Ser.

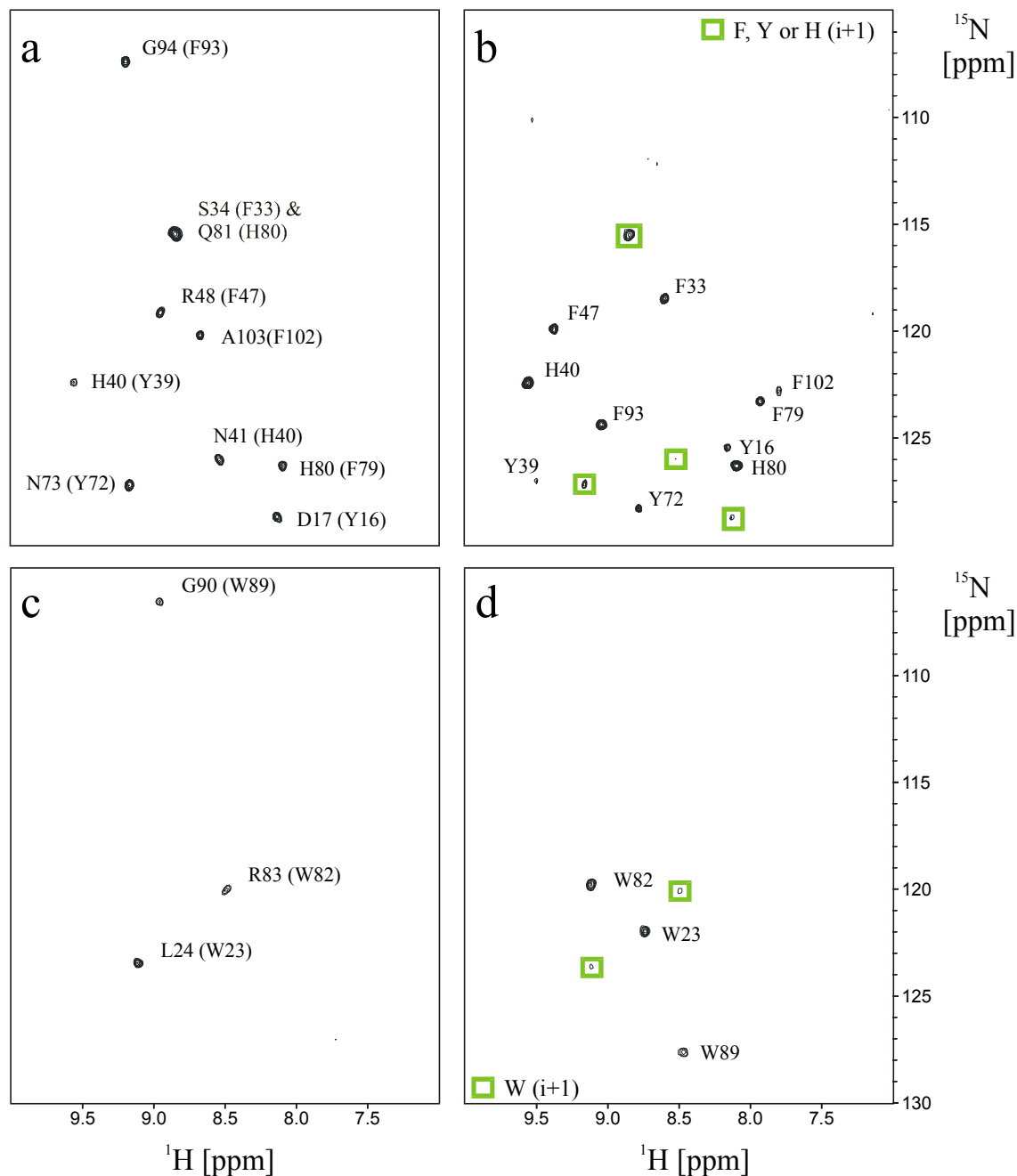


Fig. 4.18: Aromatic amino acid-selective ^1H - ^{15}N -correlations of the EVH1 domain from VASP. The EVH1 domain contains 5 Phe, 3 Tyr, 2 His and 3 Trp. (a) The FYH-(i+1)-HSQC was acquired in 3 h using 64 scans. All expected signals are present. (b) The FYH-(i,i+1)-HSQC was acquired in 6 h using 128 scans. Signals from all His, Phe and Tyr residues occur and of most of their sequential neighbors (marked with rectangles). (c) The W-(i+1)-HSQC was acquired in 3 h using 64 scans. All three expected signals are visible. (d) The W-(i,i+1)-HSQC was acquired in 6 h using 128 scans. The signals of all three Trp residues and two of their sequential neighbors (marked with rectangles) are present.

4.4.7 Selecting Asn, Asp and Gly: The DNG-(i+1)- and DNG-(i,i+1)-HSQC

Asparagine and aspartate, two of the 15 amino acids with a methylene group in the C^β position, offer a carbonyl adjacent to the C^β as additional selection criteria. The chemical shifts of carbonyls are well separated from all other carbons in a protein. Similar to the selection of aromatic residues, selective pulses are used to affect the C^γ nuclei independently. Note that the C^β of Asn and Asp are not the only CH₂ groups that show a coupling to an adjacent carbonyl, the same motive is found in Gly (C^α) and in Gln/Glu (C^γ).

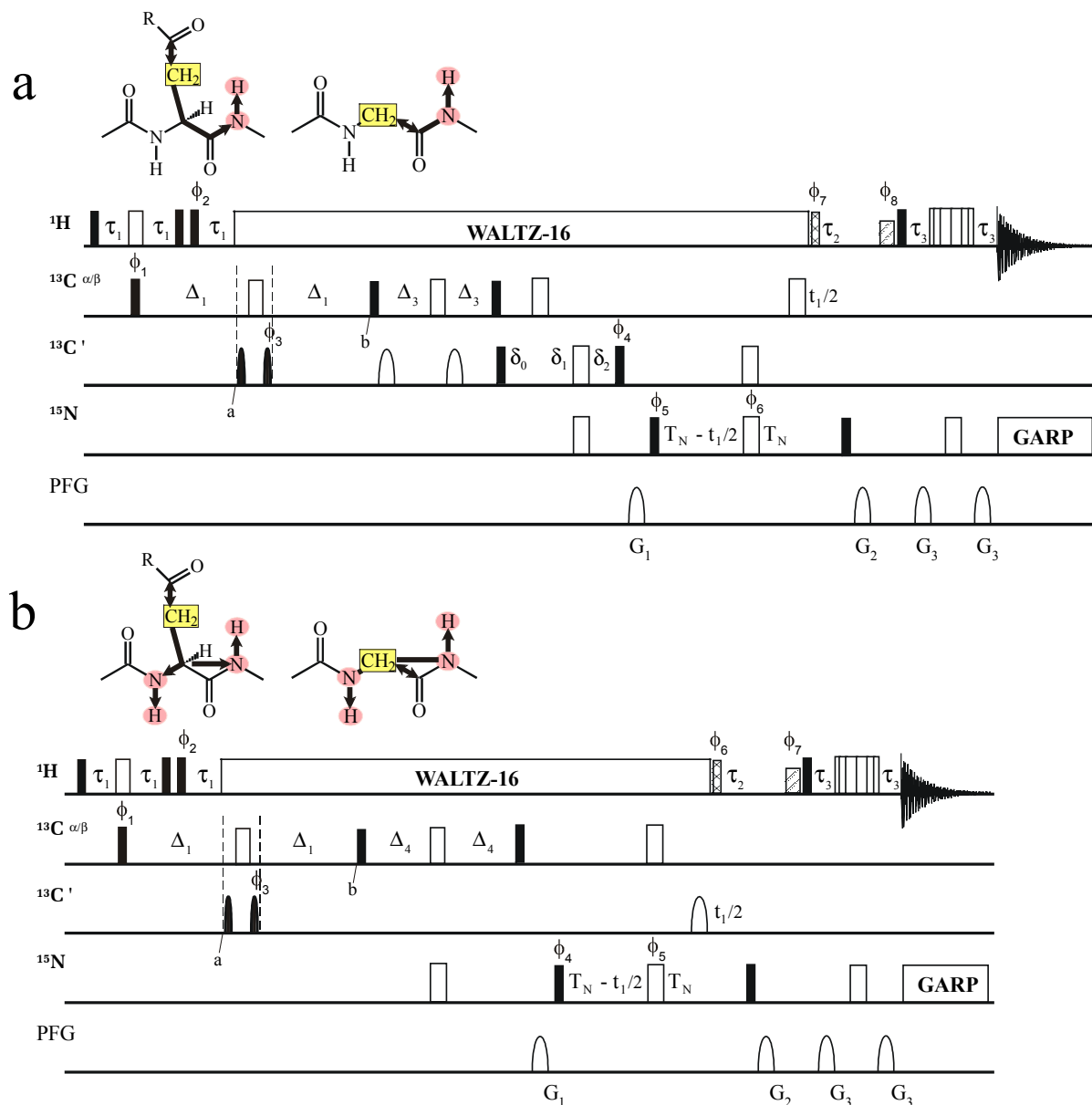


Fig. 4.19: Pulse sequences of the DNG-(i+1)- and DNG-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ ms (not including the pulse length of the LOS2-0 pulse), $\Delta_3 = 4.5$ ms, $\Delta_4 = 9$ ms. Two LOS2-0 pulses (filled sine shapes) with a duration of 512 μ s (at 600 MHz) centered at 185 ppm are used. More experimental details are given in the appendix.

The MUSIC- CH_2 selection combined with an HMQC like sequence (Müller, 1979; Bax et al., 1983) with shaped pulses is implemented into the CBCA(CO)NH and the CBCANH experiments. The resulting pulse sequences are shown in Figure 4.19. During Δ_1 antiphase magnetization $2\text{C}_y^\beta\text{C}_z'$ is created (point a). A selective 90° pulse affecting only the carbonyl region converts it into multiple quantum coherence $2\text{C}_y^\beta\text{C}_y'$ which is selected by alternating the phase of the consecutive, selective 90° carbonyl pulse. This pulse converts the multiple quantum coherence back to $2\text{C}_y^\beta\text{C}_z'$ magnetization which is refocused during the second Δ_1 delay. Self-refocusing LOS2-0 pulses with a bandwidth of 72 ppm, centered at 185 ppm were used (duration of 512 μs at 600 MHz). During $2\Delta_1$ the coupling between C^β and C^α can evolve and antiphase magnetization $2\text{C}_y^\beta\text{C}_z^\alpha$ is created (point b). The next 90° carbon pulse converts it into $2\text{C}_z^\beta\text{C}_y^\alpha$ which is refocused during $2\Delta_3$ (or $2\Delta_4$). The magnetization is then transferred from the C^α to the amide proton.

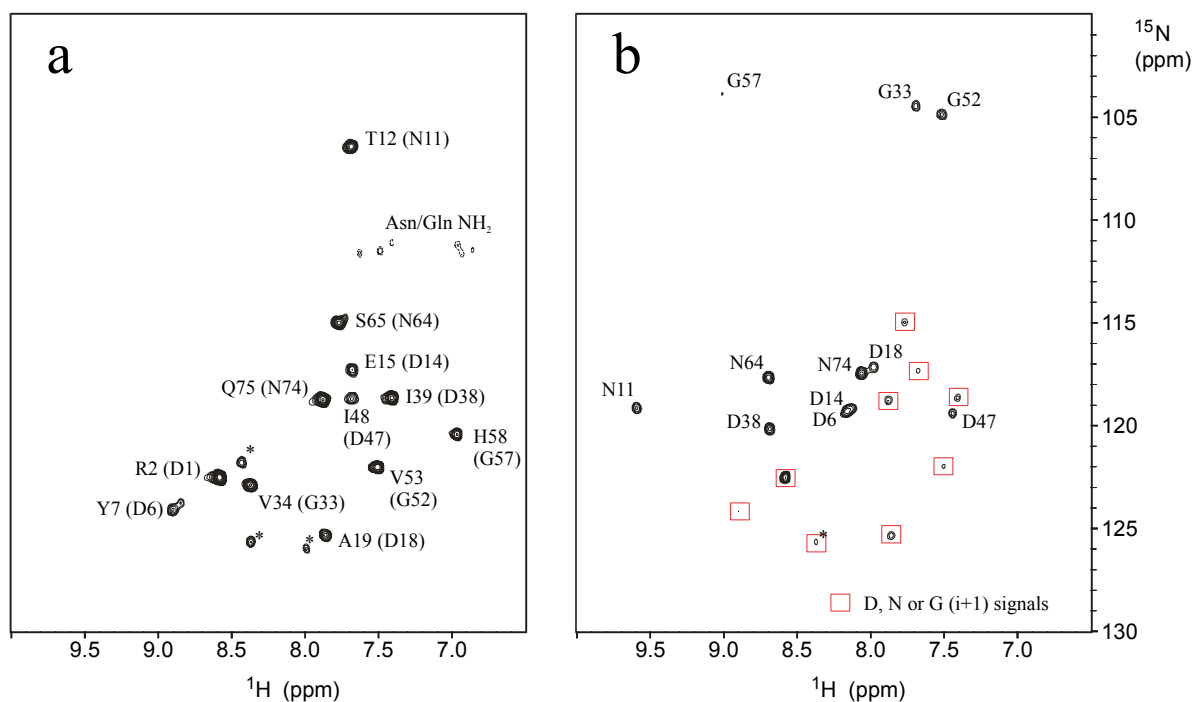


Fig. 4.20: Asp/Asn/Gly selective ^1H - ^{15}N -correlations of the SAM domain from EphB2. The SAM domain contains 6 Asp, 3 Asn and 3 Gly. (a) The DNG-($i+1$)-HSQC was acquired in 1 h 30 min using 32 scans. All expected signals (three N ($i+1$), six D ($i+1$)) are visible. The glycine neighbors (marked with thin rectangles) can be identified using the G-($i+1$)-HSQC, the asparagine neighbors using the N-($i+1$)-HSQC. Some weak and negative peaks from side chain NDH-groups show up. (b) The DNG-($i,i+1$)-HSQC was acquired in 3 h using 64 scans and contains all expected Asp, Asn and Gly signals. D1 is not visible as in the normal ^{15}N -HSQC. The ($i+1$) peaks are marked with rectangles and can be verified by the DNG-($i+1$)-HSQC.

Since the $J_{\text{C}^\beta\text{C}^\alpha}$ and $J_{\text{C}^\beta\text{C}^\gamma}$ coupling is active during the selective pulses, the delay Δ_1 has to be adjusted to compensate for this. The effective coupling between the C^β and C^γ carbon during the pulse is reduced and was determined to be 60% of the coupling during a free precession period (see chapter 3.5). The resulting transfer functions for the experiments in

Figure 4.19 a and b are:

$$\sin^2(\pi J_{CCO}(\Delta_1+0.6P)) \sin(\pi J_{CC}(2\Delta_1+2P)) \sin(\pi J_{CC}2\Delta_3) \sin(\pi J_{CC}2\Delta_3) \quad [4.8]$$

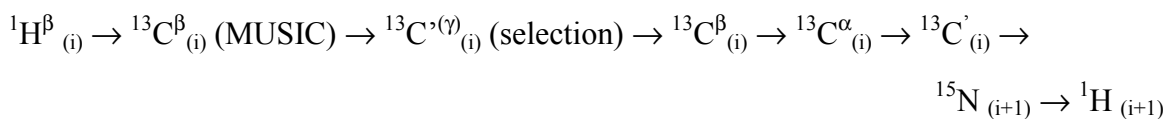
$$\sin^2(\pi J_{CCO}(\Delta_1+0.6P)) \sin(\pi J_{CC}(2\Delta_1+2P)) \sin(\pi J_{CC}2\Delta_4) \sin(\pi J_{CN}2\Delta_4) \quad [4.9]$$

P is the pulse length of the LOS2-0 pulses. Neglecting relaxation effects, the nominal values for Δ_1 , Δ_3 and Δ_4 are 7.8 ms, 5.3 ms and 9 ms, respectively. An experimental optimization resulted in values of 8.0 ms, 4.5 ms and 9 ms respectively. Gln and Glu are suppressed because there is only one COSY transfer step and magnetization from the C^γ can not be transferred to the C^α and thus to the amino proton. Magnetization from the C^α of Gly is not affected by the 90° pulse after point b. Coupling between the C^α and C' (or N) will evolve during $2\Delta_3$ (or $2\Delta_4$) and magnetization is transferred to the amide proton. The DNG-(i+1)- and DNG-(i,i+1)-HSQC experiments are quite sensitive, the resulting spectra contain signals originating from Asp, Asn and Gly and are shown in Figure 4.20.

4.4.8 Selecting Asp: The D-(i+1)- and D-(i,i+1)-HSQC

To select solely aspartate but not asparagine and glycine, the DNG experiments were modified using the fact that a CO_2 carbon lacks a $^1J_{C'N}$ coupling. The selection mechanism was described earlier by Pellecchia et al. to create the H(C)CO₂ experiment, selective for Asp and Glu side chains (Pellecchia et al., 1997b).

The MUSIC-CH₂ selection is merged with an HSQC like magnetization transfer to the carbonyl carbon instead of an HMQC like transfer in order to obtain single quantum coherence on C' . The pulse sequences are shown in Fig. 4.21. The magnetization transfer in the D-(i+1)-HSQC can be described by:



After the magnetization transfer from the selected CH₂ group, single quantum coherence $2C'_yC^\beta_z$ is created (point a). During this transfer a selective 180° pulse effecting only C^β refocuses coupling to C^α . A REBURP pulse with a bandwidth of 13 ppm (duration of 2048 μ s at 600 MHz) at 42 ppm is used. During the delay $2\Delta_2$ coupling to a coupled ^{15}N is allowed to evolve. Setting $2\Delta_2$ to $1/2J_{C'N}$ lead to the formation of antiphase magnetization $4C'_xC^\beta_zN_z$ in the case of Asn and Gly (point b). The next 90° ^{15}N pulse converts it into multiple quantum coherence $4C'_xC^\beta_zN_y$ which is dephased by the consecutive gradient and will not be refocused during the rest of the sequence. Since Asp lacks a coupled ^{15}N nucleus the $2C'_yC^\beta_z$ magnetization stays as it is during $2\Delta_2$ and is converted after point b into $2C'_zC^\beta_z$. Only

magnetization generated in $\text{CH}_2\text{-CO}_2$ motives can pass the filter. Note that a C-terminal Gly would also lead to a signal in the D-(i,i+1)-HSQC. Spectra of the D-(i+1)- and D-(i,i+1)-HSQC are shown in Fig. 4.22. The two experiments are less sensitive than the DNG-(i+1)- and DNG-(i,i+1)-HSQC due to relaxation during the additional delay $2\Delta_2$ of 32 ms and the two in-phase magnetization transfers to and from the C' .

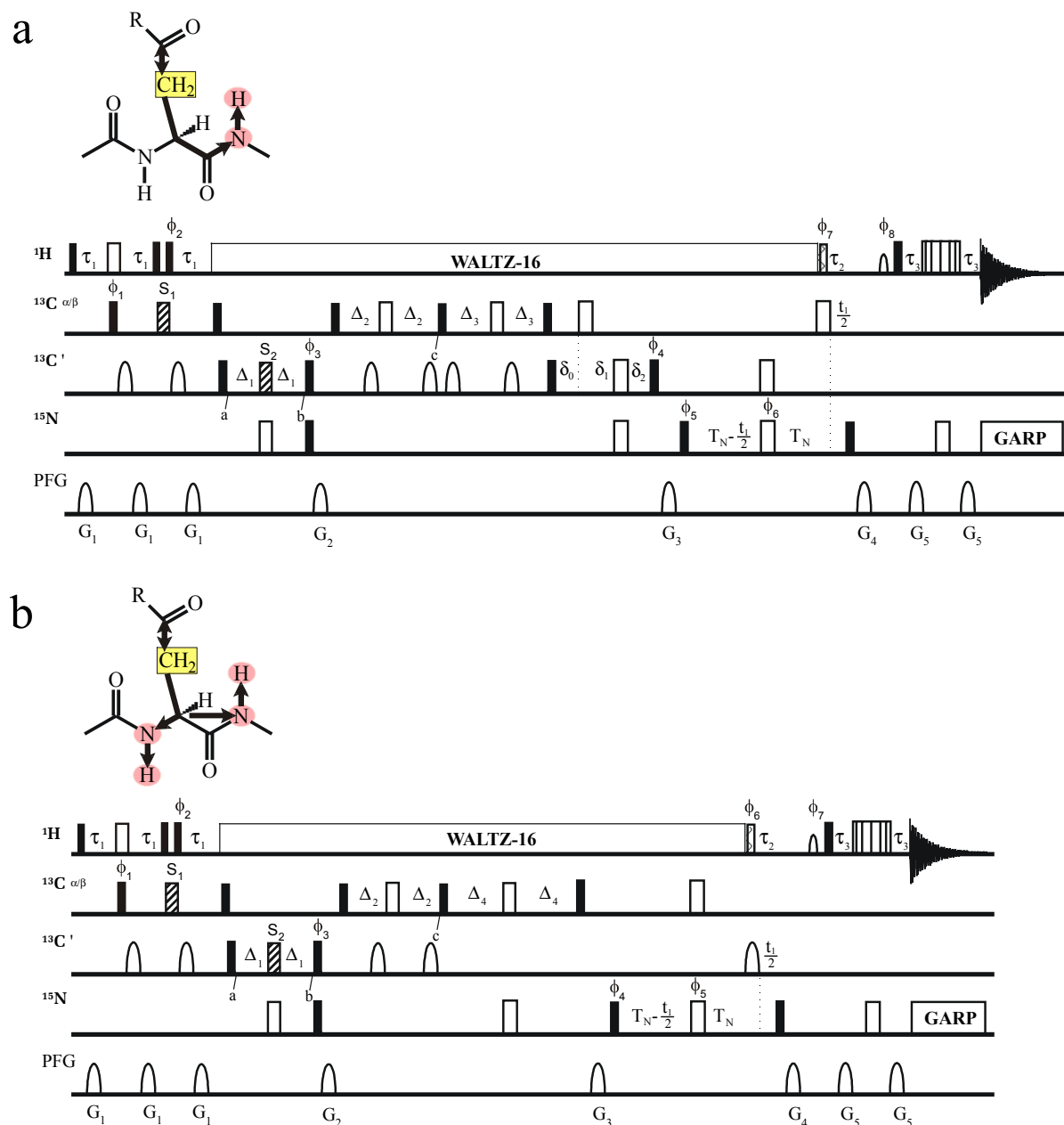


Fig. 4.21: Pulse sequences of the aspartate selective experiments D-(i+1)- and D-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 16$ ms, $\Delta_2 = 4.5$ ms, $\Delta_3 = 4.5$ ms, $\Delta_4 = 9$ ms. The REBURP pulse S_1 with a duration of 2048 μs is applied at 42.5 ppm, the REBURP pulse S_2 with a duration of 409.6 μs is applied at 175 ppm. Pulse lengths are given for a 600 MHz spectrometer. More experimental details are given in the appendix.

To aid sequential assignment is not the only application of the D-(i+1)-HSQC. Degradation of $^{13}\text{C}/^{15}\text{C}$ labeled proteins due to deamination can also be easily monitored. If

an Asn residue is deaminated to Asp, an additional signal appears in the spectrum. An introduction of a chemical shift evolution period on the side chain carbonyl carbon would allow to measure pK_a values of the CO_2 groups using a pH titration (Pellecchia et al., 1997b).

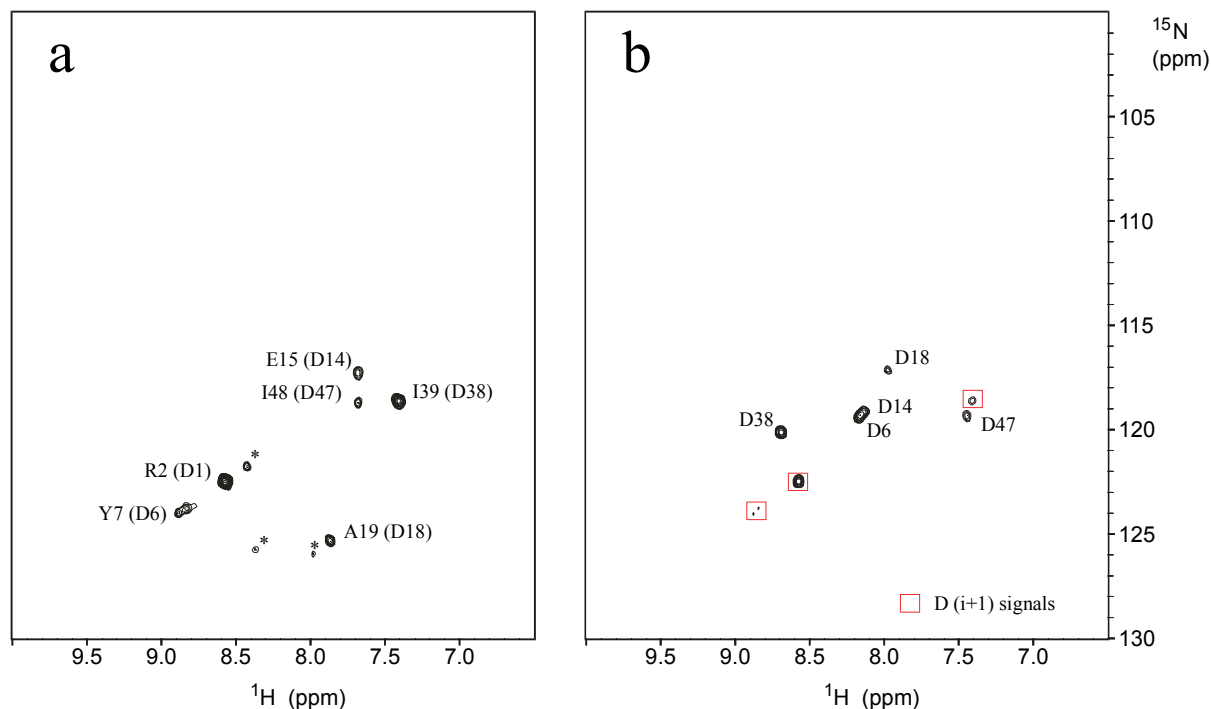
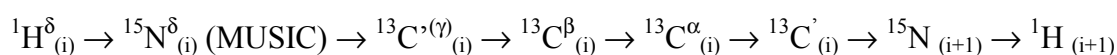


Fig. 4.22: Aspartate selective 1H - ^{15}N -correlations of the SAM domain from EphB2. The SAM domain contains 5 Asp. (a) The D-(i+1)-HSQC was acquired in 1 ½ h using 32 scans. All expected signals are visible. Three small signals originating from a different conformation of the protein are marked with an asterisk (in the conventional ^{15}N -HSQC only very weak signals are present at these positions). (b) The D-(i,i+1)-HSQC was acquired in 3 h using 64 scans and contains signals of all five Asp residues. Peaks marked with rectangles are signals of sequential neighbors and can be verified by the D-(i+1)-HSQC.

4.4.9 Selecting Asn: The N-(i+1)- and N-(i,i+1)-HSQC

Asparagine and glutamine are the only amino acids containing a $CONH_2$ group. Schmieder et al. showed previously that a MUSIC-NH₂ selection in combination with a magnetization transfer to the carbonyl carbon yields a selection of Asn and Gln (Schmieder et al., 1998). In order to achieve a 1H - ^{15}N correlation of the backbone the selected magnetization has to be transferred through the whole side chain from the NH_2 to the C^α and then further either via the carbonyl carbon or directly to the NH resulting in the (i+1) and (i,i+1) spectra, respectively. The magnetization transfer in the N-(i+1)-HSQC can be described by:



The pulse sequences for the Asn selective experiments, the N-(i+1)- and N-(i,i+1)-HSQC are shown in Fig. 4.23. The selected magnetization is transferred by a refocused INEPT from the N^δ to the carbonyl $\text{C}^{\gamma(\prime)}$ and by another refocused INEPT to the C^β . From the C^β magnetization is transferred to the NH of the sequential neighbor of Asn or to the NH of Asn. Since the magnetization has to be transferred via the whole side chain, loss of magnetization due to the transfers and due to relaxation lead to relatively low sensitivity of the experiments. Again the (i+1) version is more sensitive than the (i,i+1) version. The spectra are shown in Fig. 4.24.

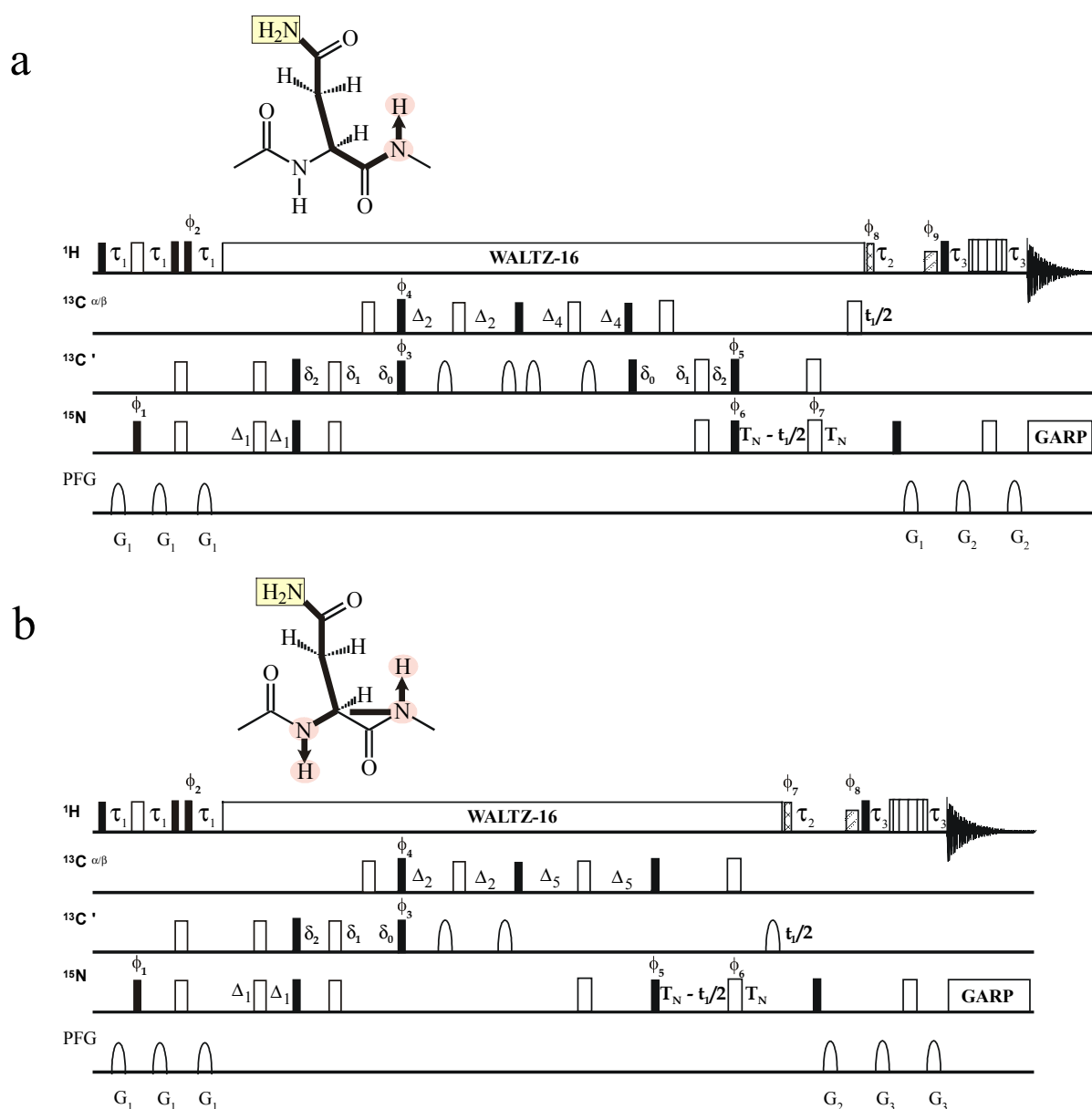


Fig. 4.23: Pulse sequences of the asparagine selective experiments N-(i+1)- and N-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 5.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ ms, $\Delta_2 = 4.5$ ms, $\Delta_4 = 4.5$ ms, $\Delta_5 = 9.3$ ms. More experimental details are given in the appendix.

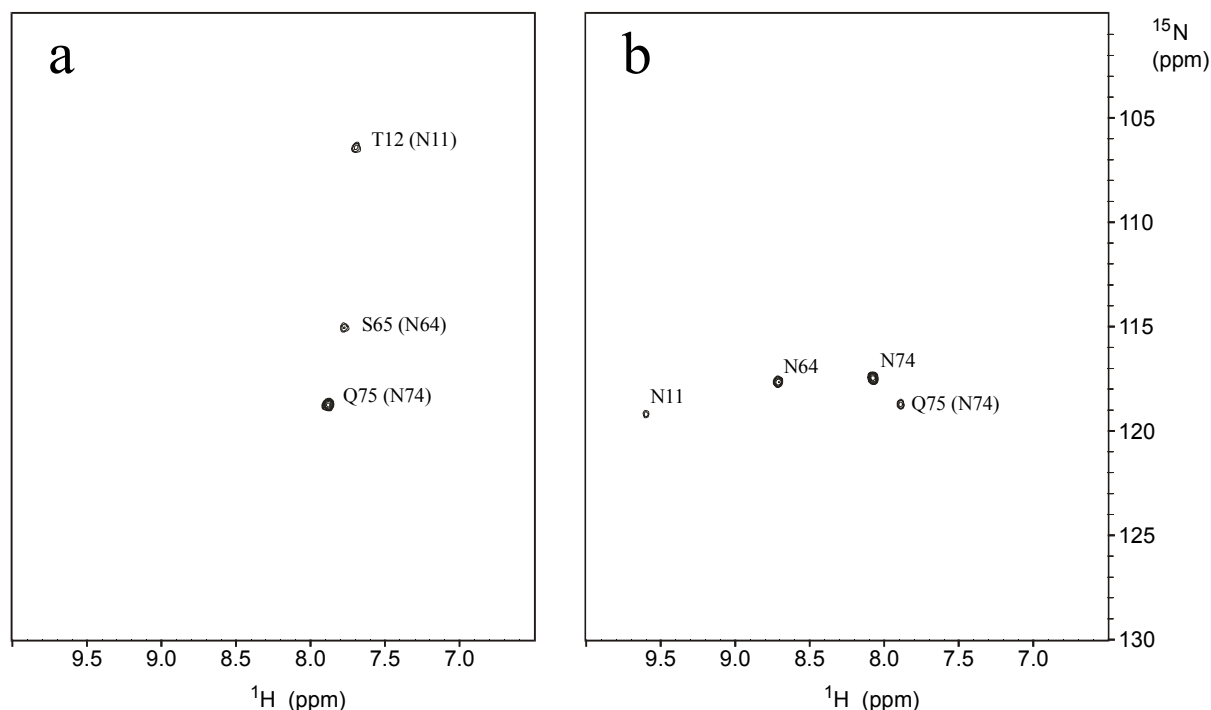


Fig. 4.24: Asparagine selective ^1H - ^{15}N -correlations of the SAM domain from EphB2. The SAM domain contains 3 Asn. (a) The N-(i+1)-HSQC was acquired in 3 h using 64 scans. All three expected signals are visible. (b) The N-(i,i+1)-HSQC was acquired in 6 h using 128 scans and contains all three Asn residues and a sequential peak. Q75 is a the sequential neighbor of N74.

4.4.10 Selecting Gln, Glu and Gly: The EQG-(i+1)- and EQG-(i,i+1)-HSQC

Gln, Glu, Asn, Asp and Gly have a $\text{CH}_2\text{-CO}$ element in common. As already mentioned a combination of the MUSIC- CH_2 selection and an out and back transfer to the carbonyl carbon selects the C^α of Gly, the C^β of Asn/Asp and the C^γ of Gln/Glu. An implementation of these techniques into CBCA(CO)NH and CBCANH experiments extended by a relay step, achieves a magnetization transfer of the selected C^γ of Gln/Glu to their C^α and further to NH. The resulting experiments are almost identical to the DNG-(i+1)- and DNG-(i,i+1)-HSQC but have been extended by an additional COSY step. They select Gln and Glu in addition to Asp, Asn and Gly. As was outlined in chapter 3.6 (case A and D) the signals of Asp and Asn can be suppressed by carefully tuning of delays to yield the EQG-(i+1)- and EQG-(i,i+1)-HSQC which are just selective for Gln, Glu and Gly.

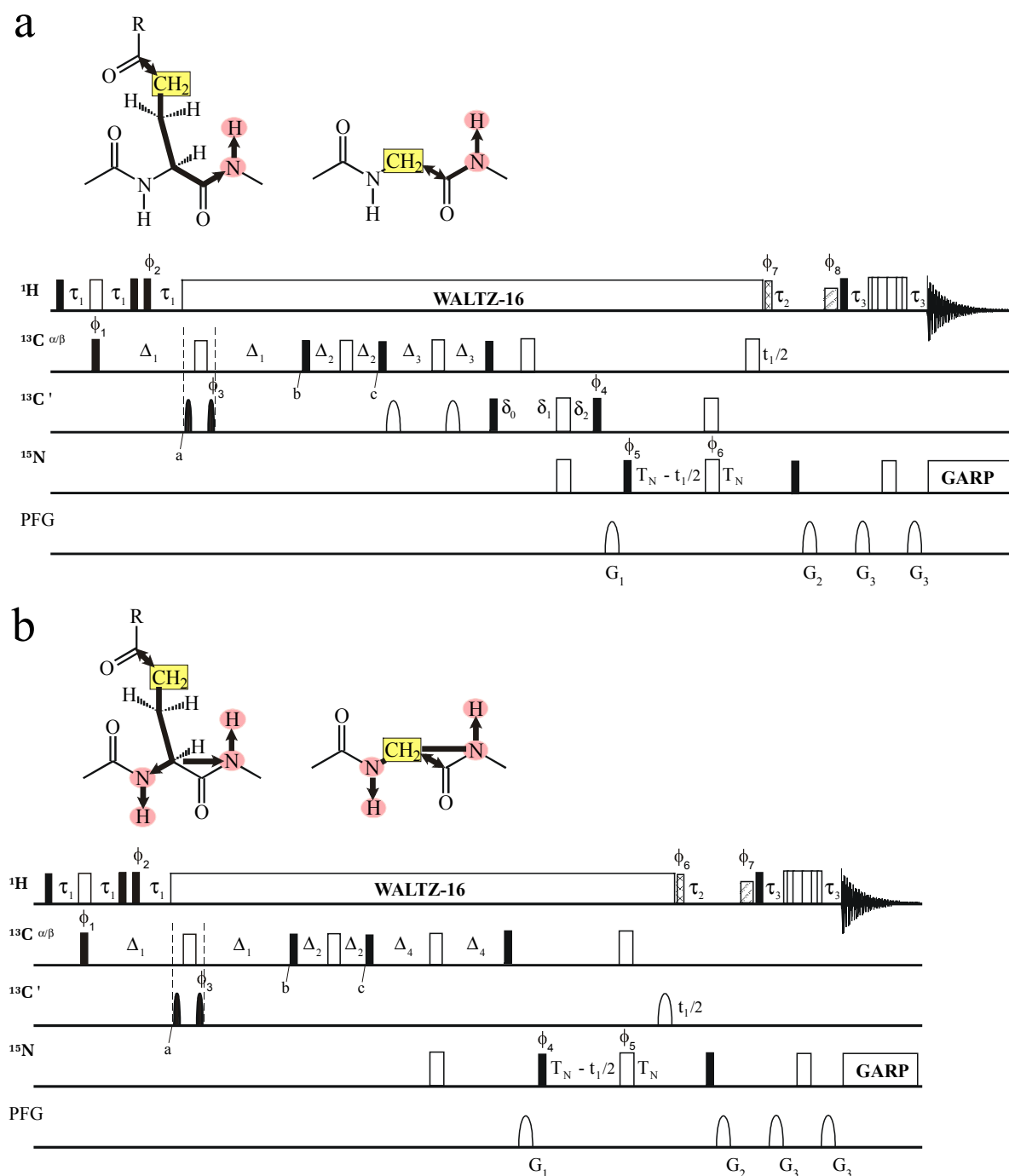


Fig. 4.25: Pulse sequences of the amino acid-type selective experiments EQG-(i+1)- and EQG-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 8.5$ ms in the EQG-(i+1)-HSQC and 7.2 ms in the EQG-(i,i+1)-HSQC, $\Delta_2 = 4.5$ ms, $\Delta_3 = 5.5$ ms, $\Delta_4 = 7.1$ ms. Two LOS2-0 pulses (filled sine shapes) with a duration of 512 μs (at 600 MHz) centered at 185 ppm are used. More experimental details are given in the appendix.

In analogy to the DNG experiments antiphase magnetization $2C_y^y C_z^z$ is created, selected and refocused during $2\Delta_1$, while coupling between the C^γ and C^β is active and antiphase magnetization $2C_y^\gamma C_z^\beta$ is created (point b). It is converted into $2C_z^\gamma C_y^\beta$ by the next 90° carbon pulse. During $2\Delta_2$ antiphase magnetization $2C_z^\gamma C_y^\beta$ is refocused and coupling to C^α leads to antiphase magnetization $2C_y^\beta C_z^\alpha$ (point c). The next 90° carbon pulse creates

antiphase magnetization $2C_z^{\beta}C_y^{\alpha}$ which is refocused during $2\Delta_3$ (or $2\Delta_4$). From the C^{α} the magnetization is transferred to the amino proton in the usual manner. The resulting transfer functions for the experiment EQG-(i+1) in Fig. 4.25 a is:

$$\sin^2(\pi J_{CCO}(\Delta_1+0.6P)) \sin(\pi J_{CC}(2\Delta_1+2P)) \sin^2(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_3) \sin(\pi J_{CC}2\Delta_3) \quad [4.10]$$

P is the pulse length of the selective LOS2-0 pulses. In addition, there are two possible magnetization pathways for Asn and Asp in this experiment: C_x^{β} remains in-phase with respect to C^{α} during $2\Delta_1$ and is converted into C_x^{α} magnetization during $2\Delta_2$ and $2\Delta_3$ or C_x^{α} is already formed at the end of $2\Delta_2$ and remains in-phase with respect to the C^{β} during $2\Delta_3$. The resulting transfer functions are:

$$\sin^2(\pi J_{CCO}(\Delta_1+0.6P)) \cos(\pi J_{CC}(2\Delta_1+2P)) \sin(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_3) \sin(\pi J_{CC}2\Delta_3) \quad [4.11]$$

$$\sin^2(\pi J_{CCO}(\Delta_1+0.6P)) \sin(\pi J_{CC}(2\Delta_1+2P)) \sin(\pi J_{CC}2\Delta_2) \cos(\pi J_{CC}2\Delta_3) \sin(\pi J_{CC}2\Delta_3) \quad [4.12]$$

Analyzing transfer function [4.10], using a pulse length P of 512 μ s and assuming J_{CCO} to be 55 Hz (Bystrov, 1976) nominal delays for Δ_1 , Δ_2 , and Δ_3 are 7.8 ms, 7.1 ms and 5.3 ms, respectively, to obtain optimal signal to noise for Glu and Gln. With these delays the signals of Gly, Glu and Gln have opposite sign to those of Asn and Asp. This can result in cancellation of signals if peaks from both types of amino acids overlap. To avoid this, the delays Δ_1 and Δ_3 can be tuned such that signals from Asp and Asn are suppressed. The two pathways [4.11] and [4.12] cancel if $\Delta_1 + \Delta_3 + P = 1/2J_{CC} = 14.3$ ms (see also chapter 3.6). Neglecting relaxation effects, optimal delays for Glu and Gln are then $\Delta_1 = 8.2$ ms, $\Delta_3 = 5.6$ ms. Experimental optimization with the EVH1 domain showed that the suppression of signals from Asp and Asn worked best with $\Delta_1 = 8.5$ ms, $\Delta_2 = 4.5$ ms and $\Delta_3 = 5.5$ ms.

The transfer functions for the EQG-(i,i+1) pulse sequence in Fig. 4.25 b are similar to [4.10], [4.11] and [4.12], only Δ_3 and J_{CC} are replaced by Δ_4 and J_{CN} , leading to the nominal values for Δ_1 , Δ_2 and Δ_4 of 7.8 ms, 7.1 ms and 9 ms, respectively. As in the (i+1) sequence these delays will lead to opposite sign for signals from Gly, Glu and Gln relative to those from Asn and Asp. Cancellation will occur not only if peaks from both types of amino acids overlap but also if amino acids of both types are adjacent in the sequence. Using the condition $\Delta_1 + \Delta_4 + P = 1/2J_{CC} = 14.3$ ms for the suppression of signals from Asn and Asp, optimal delays are $\Delta_1 = 6.7$ ms, $\Delta_4 = 7.1$ ms. The result of an experimental optimization was $\Delta_1 = 7.2$ ms, $\Delta_2 = 4.5$ ms and $\Delta_4 = 7.1$ ms.

Asp and Asn can not always be totally suppressed, however, since J_{CC} can vary (Cornilescu et al., 2000) and some small negative signals appear in the spectra. These can be seen in Fig. 4.26 a and b.

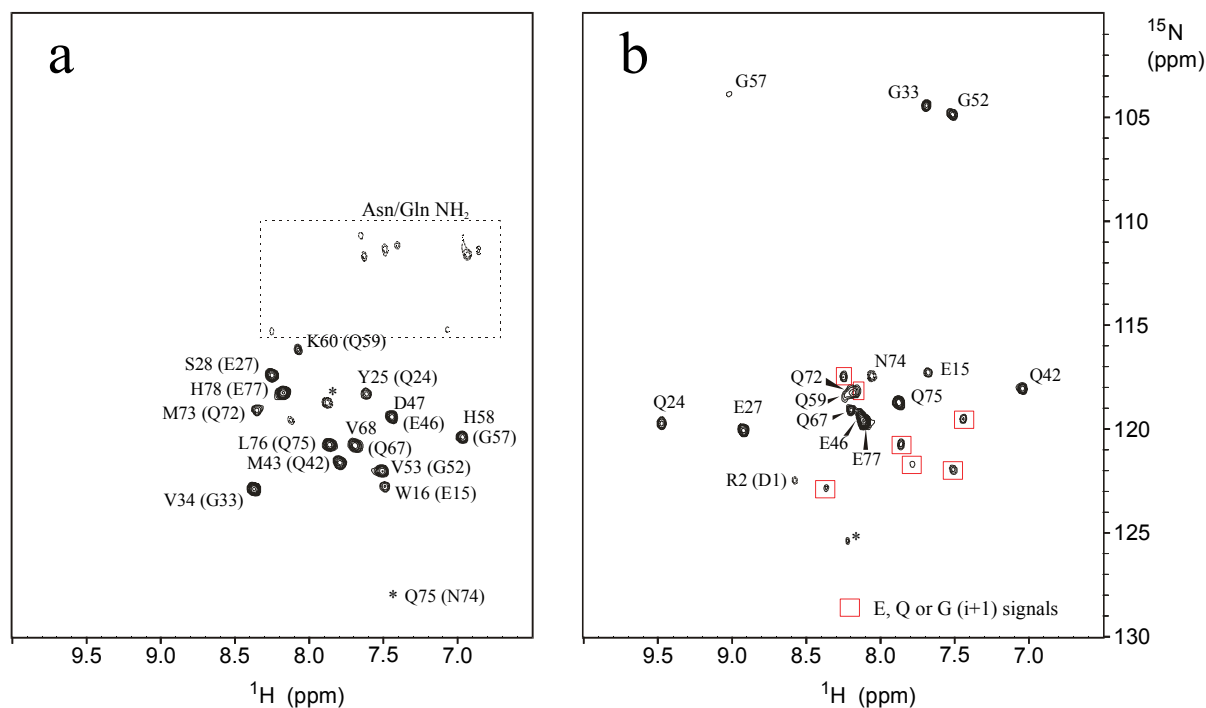
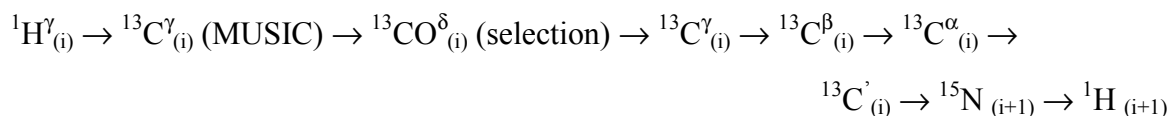


Fig. 4.26: Glu/Gln/Gly selective ^1H - ^{15}N -correlations of the SAM domain from EphB2. The SAM domain contains 4 Glu, 6 Gln and 3 Gly. (a) The EQG-(i+1)-HSQC was acquired in 3 h using 64 scans. It contains all four E-(i+1) and six Q-(i+1) signals. The suppression of D/N-(i+1) peaks is achieved by appropriate tuning of Δ_1 and Δ_3 . However, some negative breakthrough peaks of Asn/Asp (marked with *) and weak signals from side chain NDH-groups show up. (b) The EQG-(i,i+1)-HSQC was acquired in 3 h using 64 scans and contains all expected signals. Signals from Asp/Asn are suppressed except N74 which appears as a negative signal. Peaks marked with rectangles are signals of sequential neighbors and can be verified by the EQG-(i+1)-HSQC. A breakthrough peak from the His-tag is marked with an asterisk.

4.4.11 Selecting Glu: The E-(i+1)-, ED-(i+1)- and E-(i,i+1)-HSQC

Glutamate and aspartate are the only residues which contain a $\text{CH}_2\text{-CO}_2$ element. Just a C-terminal Gly would have the same motive. The only difference between Glu and Asp is the extension of the side chain by a CH_2 . The magnetization transfer in the E-(i+1)-HSQC can be described by:



The pulse sequences are similar to those of the D-(i+1)- and D-(i,i+1)-HSQC and are shown in Fig. 4.27. A relay step was added, the selective REBURP pulse within MUSIC was omitted and carefully tuned delays are used. The $\text{CH}_2\text{-CO}_2$ groups of Asp and Glu are selected as described in part 4.4.9. The ^{13}C chemical shifts of the C^{γ} and C^{β} Glu are not clearly separated and the use of a selective pulse effecting only one of the nuclei is not possible. Therefore transverse magnetization on C^{γ} can couple to the C^{β} during the MUSIC

element which makes the magnetization transfer to the $C'^{(\gamma)}$ less efficient. The transfer function till point a can be described by:

$$\sin^4(\pi J_{CH}2\tau_1) \sin(\pi J_{CC}2\tau_1) \cos(\pi J_{CC}2\tau_1) \quad [4.13]$$

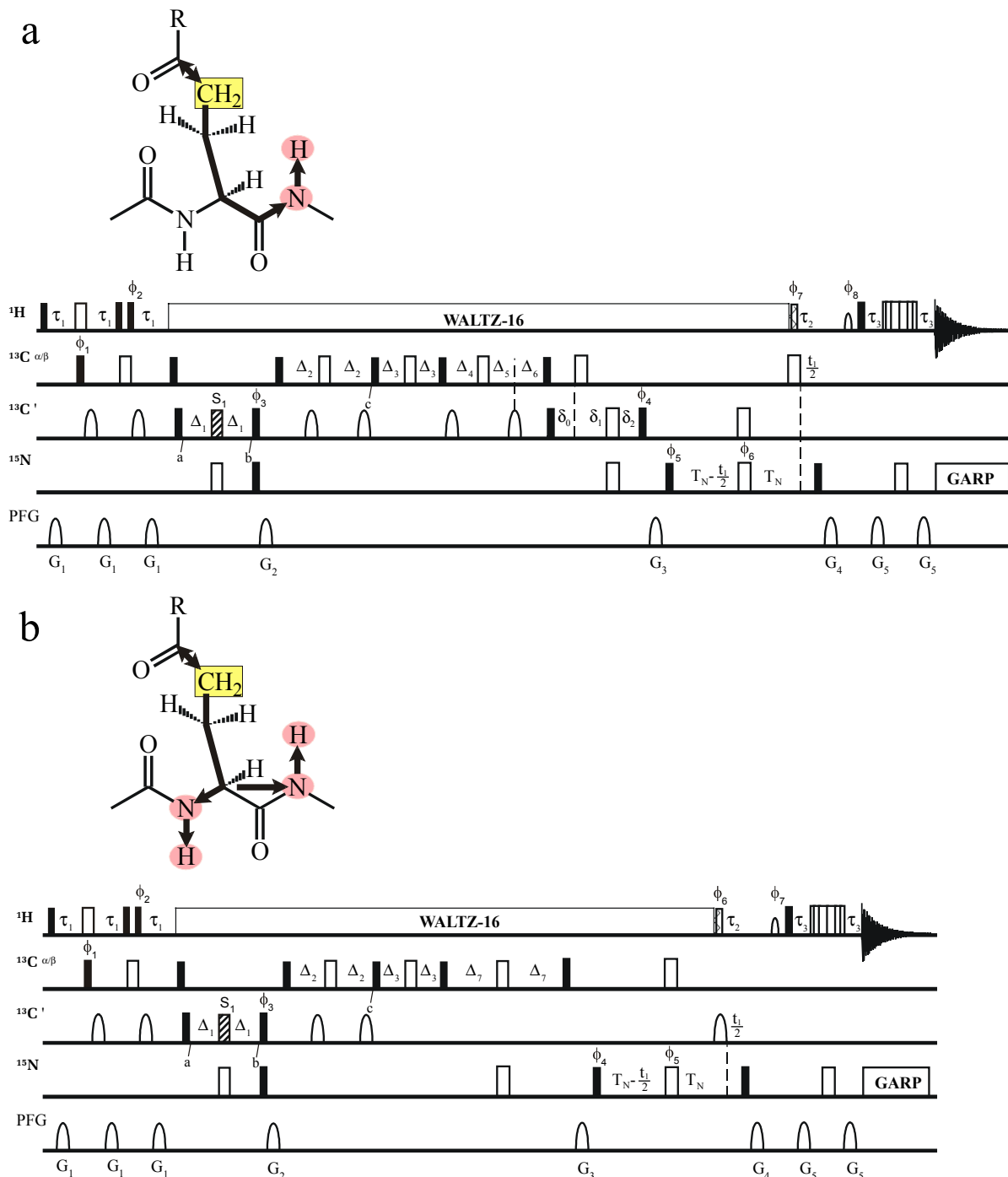


Fig. 4.27: Pulse sequences of the glutamate selective experiments: (a) E-(i+1)-, ED-(i+1)-HSQC and (b) E-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 16$ ms, $\Delta_2 = 4.2$ ms, $\Delta_3 = 4.5$ ms, $\Delta_4 = 9$ ms (4.5 ms for the ED-(i+1)-HSQC), $\Delta_5 = 4.5$ ms (0 ms for the ED-(i+1)-HSQC), $\Delta_6 = 4.5$ ms, $\Delta_7 = 9$ ms. The REBURP pulse S_1 with a duration of 409.6 μ s (at 600 MHz) is applied at 175 ppm. More experimental details are given in the appendix.

Setting of τ_1 to 3.5 ms leads to a sensitivity loss of 28% due to the cosine term. Antiphase magnetization $2C'_yC'_z$ of Glu and $2C'_yC^\beta_z$ of Asp is created at point a. During the CO_2 selection, all carbonyl carbons which can couple to a ^{15}N are suppressed. After point b magnetization $2C'_yC'_z$ of Glu and $2C'_yC^\beta_z$ of Asp is converted by the consecutive two 90° pulses to $2C'_zC'_y$ and $2C'_zC^\beta_y$, respectively. The following two COSY steps enable the magnetization transfer to the C^α . The transfer functions for Glu in the (i+1) and the (i,i+1)-HSQC are described by:

$$\sin(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_2) \sin^2(\pi J_{CC}2\Delta_3) \sin(\pi J_{CC}2\Delta_4) \sin(\pi J_{CC}2\Delta_6) \quad [4.14]$$

$$\sin(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_2) \sin^2(\pi J_{CC}2\Delta_3) \sin(\pi J_{CC}2\Delta_7) \sin(\pi J_{CN}2\Delta_7) \quad [4.15]$$

Signals of Asp can be suppressed using tuned delays as described in chapter 3.6. Two magnetization transfer pathways exist for Asp: either C^β_x magnetization stays in-phase during the first COSY step ($2\Delta_2$) and is evolving to $2C^\beta_yC^\alpha_z$ during the second COSY step ($2\Delta_3$) or $2C^\beta_yC^\alpha_z$ has already been formed during $2\Delta_2$, is converted to $2C^\beta_zC^\alpha_y$ and stays after refocusing in-phase on C^α . For the sake of clarity, only the two transfer functions for the (i+1)-HSQC are given:

$$\sin(\pi J_{CC}2\Delta_2) \cos(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_3) \sin(\pi J_{CC}2\Delta_4) \sin(\pi J_{CC}2\Delta_6) \quad [4.16]$$

$$\sin(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_3) \cos(\pi J_{CC}2\Delta_4) \sin(\pi J_{CC}2\Delta_6) \quad [4.17]$$

Asp signals are suppressed if equations [4.16] and [4.17] cancel each other out resulting in the E-(i+1)-HSQC. This is the case if $\Delta_2 + \Delta_4 = 1/2J_{CC}$ is fulfilled[‡]. The transfer functions for the E-(i,i+1)-HSQC are identical, except that Δ_4 and Δ_6 are replaced by Δ_7 and J_{CC} by J_{CN} . Nominal values for Δ_2 , Δ_3 , Δ_4 , Δ_5 and Δ_6 in the E-(i+1)-HSQC are 5.7, 7.1, 8.6, 4.1 and 4.5 ms, values for Δ_2 , Δ_3 , Δ_7 in the E-(i,i+1)-HSQC are 5.3, 7.1 and 9.0 ms.

An experimental optimization lead to values of $\Delta_2 = 4.5$ ms, $\Delta_3 = 4.5$ ms, $\Delta_4 = 9$ ms, $\Delta_5 = 4.5$ ms and $\Delta_6 = 4.5$ ms for the E-(i+1)-HSQC and $\Delta_2 = 4.5$ ms, $\Delta_3 = 4.5$ ms and $\Delta_7 = 9$ ms for the E-(i,i+1)-HSQC (the delays include the lengths of the G3 pulses). Signals resulting from Asp are largely but not completely suppressed however. In the case of the (i+1) experiment it is therefore recommended to sacrifice the Asp suppression and increase sensitivity by setting the delays Δ_2 , Δ_3 , Δ_4 , Δ_5 and Δ_6 to values of 4.5 ms, 4.5 ms, 4.5 ms, 0 ms and 4.5 ms, respectively, to obtain the ED-(i+1)-HSQC. Spectra of the E-(i+1)-HSQC, ED-(i+1)-HSQC and E-(i,i+1)-HSQC are shown in Fig. 4.28. Note that the ED-(i+1)-HSQC gives reasonable signal to noise ratios with $1/4$ of the measurement time as compared to the E-(i+1)-HSQC.

[‡] To help the reader: the correlation $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$ was used to derive this condition.

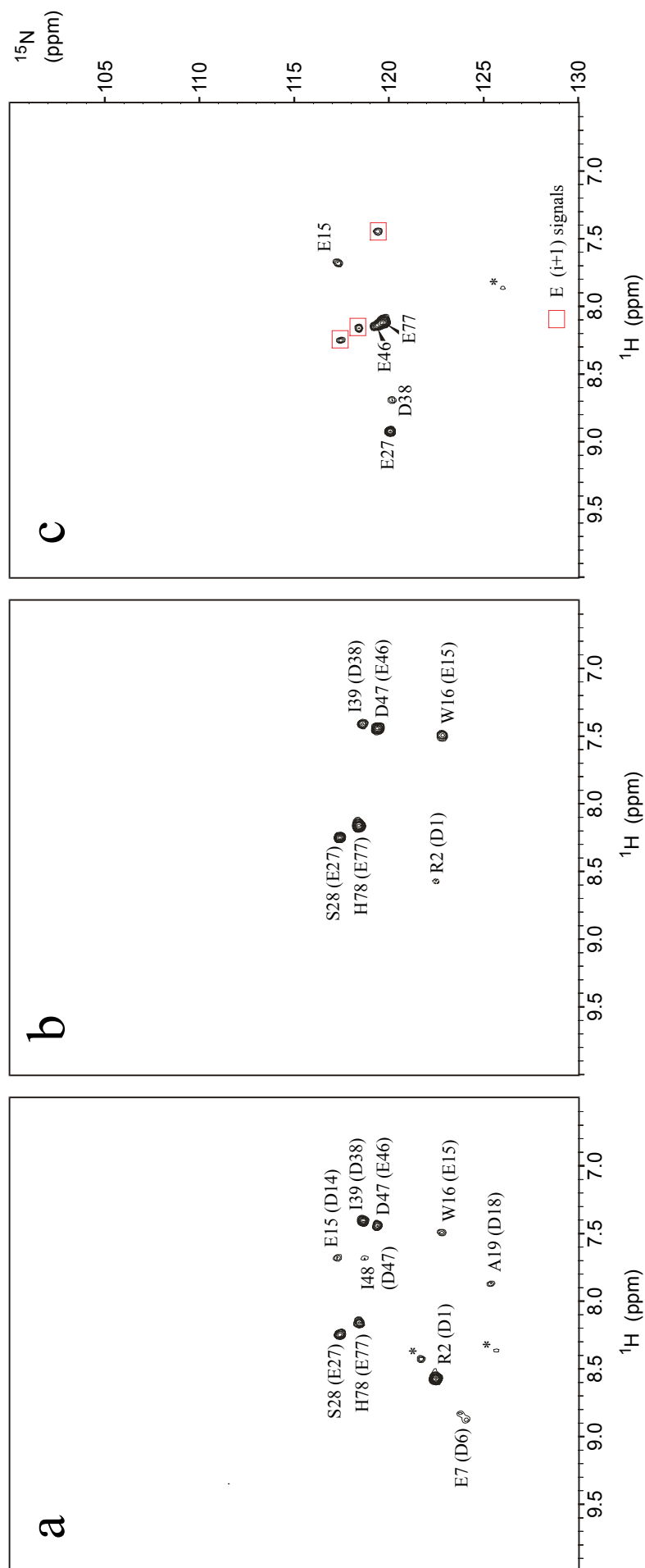
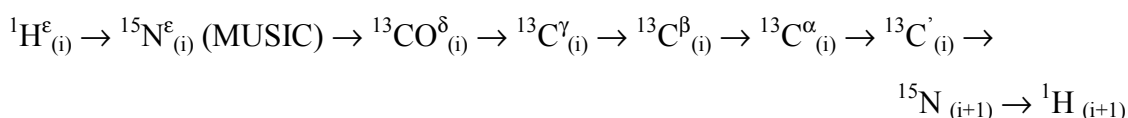


Fig. 4.28: Glu/Asp selective ^1H - ^{15}N -correlations of the SAM domain from EphB2. The SAM domain contains 6 Asp and 4 Glu. (a) The ED-(i+1)-HSQC was acquired 1 $\frac{1}{2}$ h using 32 scans. All expected signals are visible. Signals from a second conformation of the protein are marked with asterisks. (b) The E-(i,i+1)-HSQC was acquired in 6 h using 128 scans and contains signals of all four residues succeeding Glu. Except R2 and I39 all signals resulting from Asp are suppressed. The negative signal of R2 indicates that D1 has a smaller $^1J_{\text{CACB}}$ coupling constant than the remaining Asp residues. D38 has a higher $^1J_{\text{CACB}}$ coupling constant than the average. (c) The E-(i,i+1)-HSQC was acquired in 6 h using 128 scans and contains signals of all four Glu. All Asp residues except D38 are suppressed. Peaks marked with rectangles are signals of sequential neighbors and can be verified by the ED- or E-(i+1)-HSQC. An artefact from the His-Tag is marked with an asterisk.

4.4.12 Selecting Gln: The Q-(i+1)-, QN-(i+1)- and Q-(i,i+1)-HSQC

Glutamine and asparagine are the only amino acids containing a CONH_2 group. The only difference between Gln and Asn is the extension of the side chain by a CH_2 . The experiments are almost identical to the N-(i+1)- and N-(i,i+1)-HSQC but have been extended by an additional COSY step. The selected magnetization is transferred by a refocused INEPT from the N^ϵ to the carbonyl C^{δ} , by another refocused INEPT to the C^γ , followed by a COSY step which transfers it to the C^β . The last part of the sequences are identical to the CBCA(CO)NH and the CBCANH sequences which then transfer the magnetization from the C^β to the NH to yield the (i+1) and (i,i+1) version of a Gln and Asn selective experiment. The magnetization transfer of Gln in version (i+1) can be described by:



Asn can be suppressed to yield the Q-(i+1)- and Q-(i,i+1)-HSQC which are solely selective for Gln as was discussed in chapter 3.5 (case A and D). However, in the case of the Q-(i+1)-HSQC this suppression makes the experiment less sensitive so that a suppression of Asn does not always make sense. The transfer function for Gln between point a and b in the (i+1)-pulse sequence depicted in Fig. 4.29 a is:

$$\sin(\pi J_{\text{CC}}2\Delta_2) \sin(\pi J_{\text{CC}}2\Delta_2) \sin^2(\pi J_{\text{CC}}2\Delta_3) \sin(\pi J_{\text{CC}}2\Delta_4) \sin(\pi J_{\text{CC}}2\Delta_6) \quad [4.18]$$

For Asn, two magnetization pathways lead to detectable signals. After the magnetization was transferred to the C^β two COSY steps follow while it takes only one step to transfer the magnetization to the C^α . Either the magnetization stays in-phase on the C^β during $2\Delta_2$ and is not transferred to the C^α by the following 90° pulse, or magnetization is transferred to the C^α after $2\Delta_2$ and stays in-phase during $2\Delta_3$, which is described by the following two transfer functions for the case of the QN-(i+1)-HSQC:

$$\sin(\pi J_{\text{CC}}2\Delta_2) \cos(\pi J_{\text{CC}}2\Delta_2) \sin(\pi J_{\text{CC}}2\Delta_3) \sin(\pi J_{\text{CC}}2\Delta_4) \sin(\pi J_{\text{CC}}2\Delta_6) \quad [4.19]$$

$$\sin(\pi J_{\text{CC}}2\Delta_2) \sin(\pi J_{\text{CC}}2\Delta_2) \sin(\pi J_{\text{CC}}2\Delta_3) \cos(\pi J_{\text{CC}}2\Delta_4) \sin(\pi J_{\text{CC}}2\Delta_6) \quad [4.20]$$

Asn signals are suppressed if equations [4.19] and [4.20] cancel each other out resulting in the Q-(i+1)-HSQC. This is the case if $\Delta_2 + \Delta_4 = 1/2J_{\text{CC}}$. Note that the second part of the pulse sequences and the transfer functions [4.18-20] are identical to those of the E-(i+1)- and E-(i,i+1)-HSQC. With this condition nominal values of Δ_2 , Δ_3 , Δ_4 and Δ_6 are 5.7, 7.1, 8.6 and 4.5 ms, respectively. In the experiment best results were obtained with $\Delta_2 = 3.7$ ms, $\Delta_3 = 4.5$ ms, $\Delta_4 = 9.3$ ms, and $\Delta_6 = 4.5$ ms for the Q-(i+1)-HSQC. However, the suppression of Asn is not perfect and leads to a reduction of signal intensity. Because of the relatively low overall

sensitivity of the experiment it is recommended to sacrifice the Asn suppression. The QN-(i+1)-experiment works best with $\Delta_2 = 4.5$ ms, $\Delta_3 = 4.5$ ms, $\Delta_4 = \Delta_6 = 4.5$ ms and $\Delta_5 = 0$ ms.

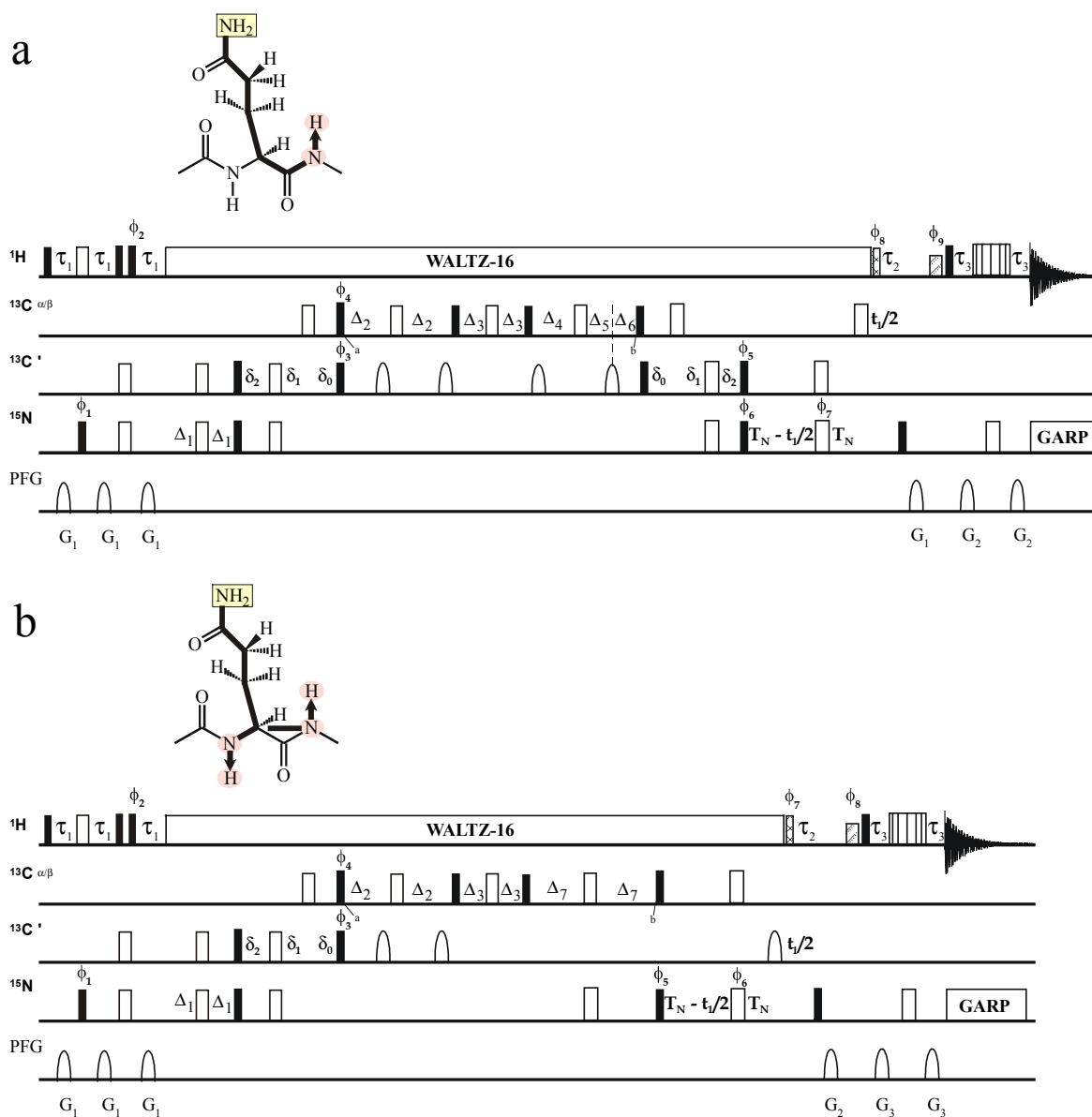


Fig. 4.29: Pulse sequences of the glutamine selective experiments Q-(i+1)-, QN-(i+1)- and Q-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 5.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ ms, $\Delta_2 = 3.7$ ms (4.5 ms for the QN-(i+1)-HSCQ), $\Delta_3 = 4.5$ ms, $\Delta_4 = 9.3$ ms (4.5 ms for the QN-(i+1)-HSCQ), $\Delta_5 = 4.8$ ms (0 ms for the QN-(i+1)-HSCQ), $\Delta_6 = 4.5$ ms, $\Delta_7 = 9.3$ ms. More experimental details are given in the appendix.

The transfer functions for the Q-(i,i+1) pulse are similar to [4.18], [4.19] and [4.20], only Δ_4 , Δ_6 and $J_{CC'}$ are replaced by Δ_7 (both) and J_{CN} , leading to the nominal values for Δ_2 , Δ_3 and Δ_7 of 5.3 ms, 7.1 ms and 9.0 ms, respectively. The values must be close to fulfilling the condition $\Delta_2 + \Delta_7 = 1/2J_{CC} = 14.2$ ms for the suppression of signals from Asn and Asp. The result of an experimental optimization was $\Delta_2 = 3.7$ ms, $\Delta_3 = 4.5$ ms and $\Delta_7 = 9.3$ ms.

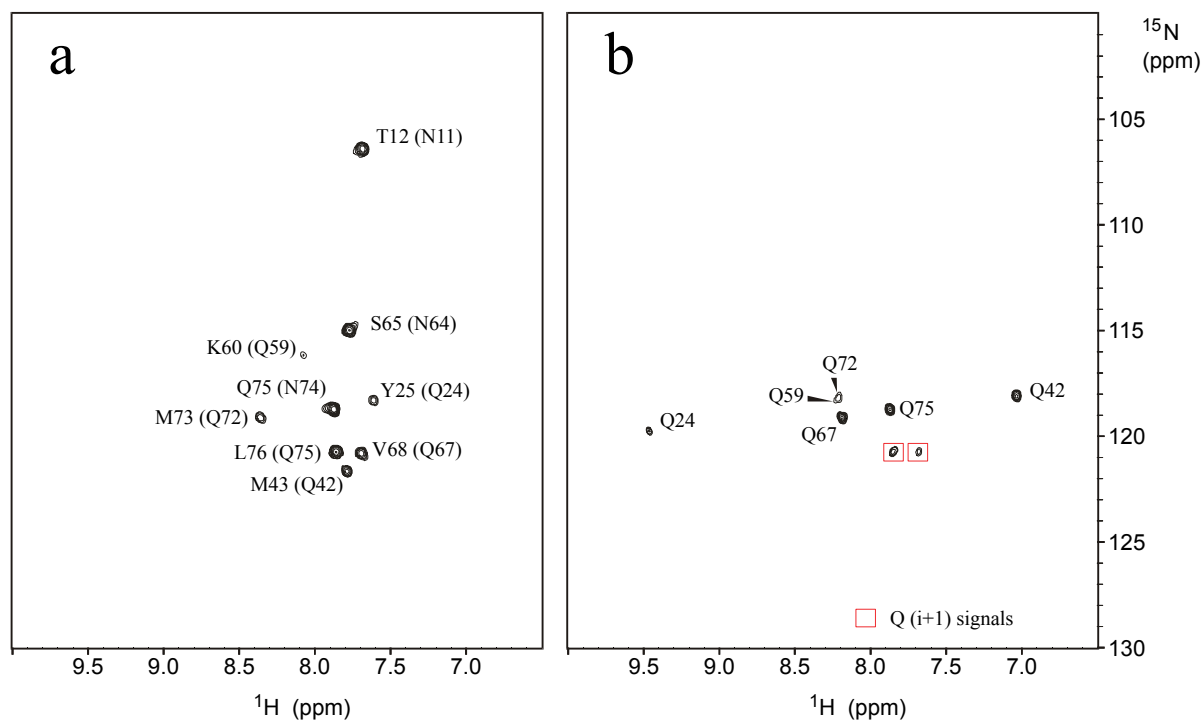


Fig. 4.30: Gln/Asn selective ^1H - ^{15}N -correlations of the SAM domain from EphB2. The SAM domain contains 3 Asn and 6 Gln. (a) The QN-(i+1)-HSQC was acquired in 6 h using 128 scans. All expected signals are visible. The signal of K60 (following Q59) is weak. (b) The Q-(i,i+1)-HSQC was acquired in 12 h using 256 scans and contains signals of all six Gln residues. The signal of Q59 is very weak, in addition signals of Q72 and Q59 overlap. Peaks marked with rectangles are signals of sequential neighbors and can be verified by the QN-(i+1)-HSQC. All Asn signals are suppressed.

Spectra of the QN-(i+1)- and Q-(i,i+1)-HSQC are shown in Fig. 4.30. Because of the higher sensitivity the QN-(i+1)-HSQC was recorded. The Q-(i+1)-HSQC is as insensitive as the Q-(i,i+1)-HSQC. For proteins larger than 10 kDa the EQG-HSQC experiments are recommended instead.

4.4.13 Selecting Arg: The R-(i+1)- and R-(i,i+1)-HSQC

The side chain of arginine offers a $\text{CH}_2\text{-N}$ element for selection, but because of the small coupling constant $^1J_{\text{CN}}$, the application of a $\text{CH}_2\text{-N}$ filter would be too inefficient. A more elegant method is to exploit the characteristic chemical shifts of C^γ and C^δ . The sufficiently distinct chemical shifts of these carbons allows the nuclei to be treated independently with selective pulses. The selectivity is based on the characteristic $\text{C}^\gamma/\text{C}^\delta$ carbon pair with chemical shifts of 27.3 ± 1.4 ppm and 43.1 ± 1 ppm, respectively (from the BMRB database: Seavey et al., 1991). Carbon pairs with similar chemical shifts are only found in Lys $\text{C}^\delta/\text{C}^\epsilon$ (28.8 ± 1.4 ppm and 41.7 ± 0.9 ppm) and Leu $\text{C}^\beta/\text{C}^\gamma$ (42.3 ± 1.8 ppm and 26.8 ± 1.3 ppm).

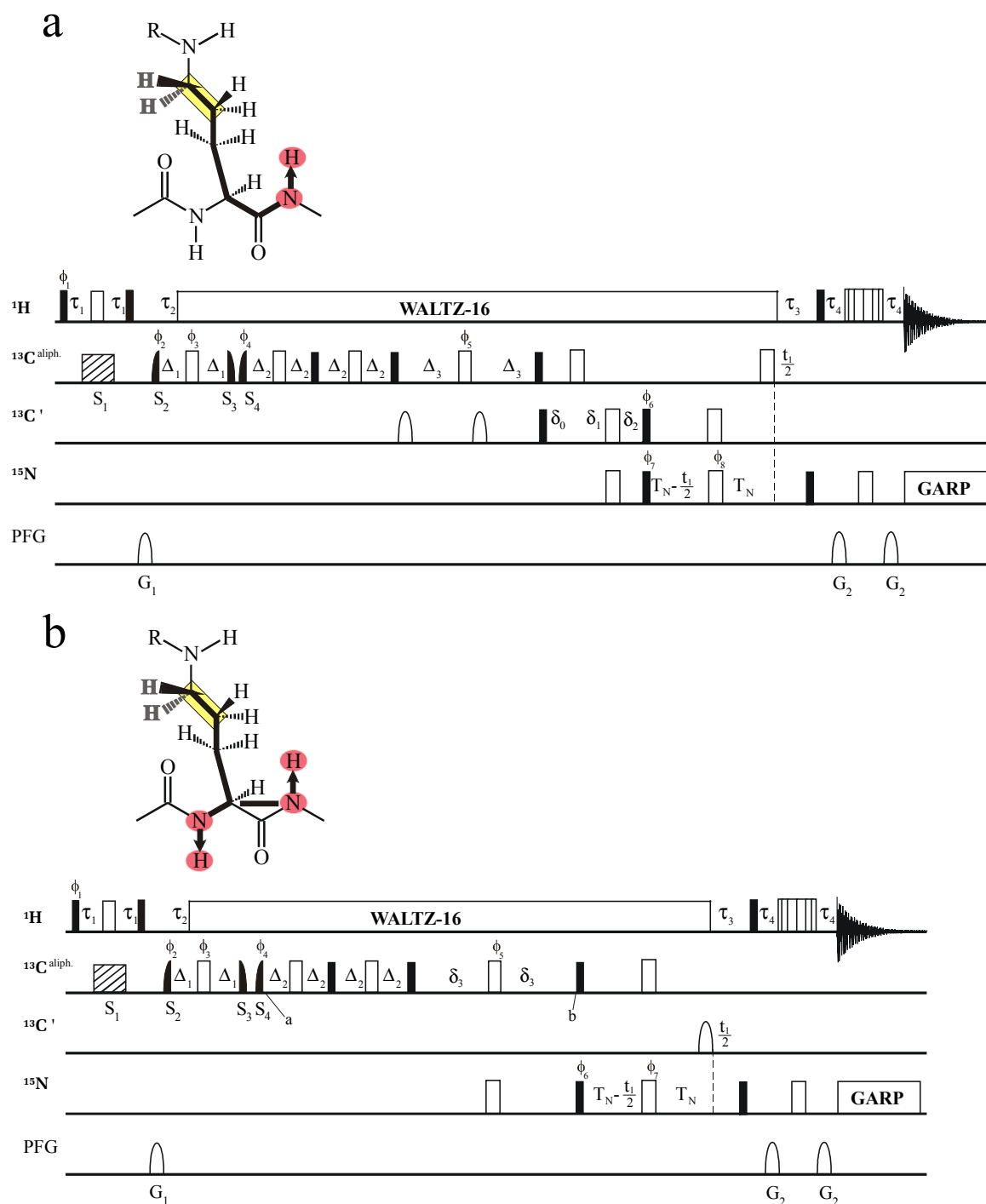
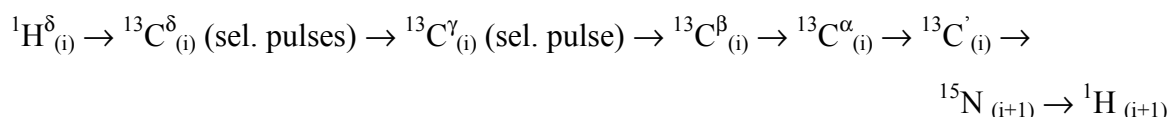


Fig. 4.31: Pulse sequences of the Arg selective ^1H - ^{15}N correlations (a) R-(i+1)-HSQC and (b) R-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 2.05$ ms (determined by pulse length S_1), $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $\delta_3 = 9$ ms, $T_N = 11$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms, $\Delta_3 = 4.5$ ms. The striped thick bar (S_1) stands for a band-selective 180° REBURP pulse ($4096 \mu\text{s}$) centered at 45 ppm.. Self-focusing selective LOS2-0 excitation pulses are represented by filled half-sine shapes ($S_2 - S_4$). The pulse lengths of $S_2 - S_4$ are $3072 \mu\text{s}$, the offsets 45 ppm, 45 ppm and 29 ppm, respectively. The second LOS2-0 pulse (S_3) has a time inverted shape. Pulse lengths are given for a 600 MHz spectrometer. More experimental details are given in the appendix.

The novel experiments are derived from the CBCA(CO)NH and the CBCANH. Two relay steps are implemented in the original sequences to obtain the R(i+1)- and R(i,i+1)-HSQC.

The desired selectivity for Arg is achieved via the implementation of soft pulses at the beginning of the sequence as described below. Despite an inevitable loss of magnetization due to relaxation during the additional delays, the intensity is reduced only slightly relative to the original sequences since the side chains of Arg are linear chains of carbon atoms with almost identical carbon-carbon one-bond couplings ($^1J_{\text{CC}}$). The pulse sequences are shown in Figure 4.31. The magnetization transfer in the R-(i+1)-HSQC can be described by:



In the initial INEPT-step both carbon pulses are replaced by pulses selective for the C^{δ} . The 180° pulse (S_1) is performed as a REBURP pulse with a bandwidth of 8 ppm (length of $4096 \mu\text{s}$ at 600 MHz) centered at 45 ppm. Note that this selective pulse extends throughout the complete delay $2\tau_1$. Only an effective J-coupling is therefore active between the proton and carbon spins. It was determined experimentally to be 90% of the true J-coupling (chapter 3.5) and the delay of $4096 \mu\text{sec}$ therefore has the proper length. Normally a IBURP pulse is used in such a case, because it has been designed for the inversion of z-magnetization. However during an IBURP pulse the effective coupling is smaller compared to a REBURP pulse due to its asymmetrical shape (chapter 3.5). It would thus have required a longer delay to evolve the coupling and a REBURP shape was used instead.

A self-refocusing LOS2-0 pulse (see chapter 3.5) with a bandwidth of 11 ppm (length of $3072 \mu\text{s}$ at 600 MHz) applied at 45 ppm is used as the 90° pulse (S_2). In the following delay $2\Delta_1$ the chemical shift evolution of the C^{δ} is refocused while the coupling between the C^{δ} and the C^{γ} evolves. This is accomplished using a non-selective rectangular 180° pulse in the center of the delay. The transfer of magnetization from the C^{δ} to the C^{γ} is then achieved by two subsequent selective 90° pulses (S_3 and S_4), one selective for the C^{δ} , the other for the C^{γ} . Both are performed as LOS2-0 pulses: S_3 has a time inverted shape and is centered at 45 ppm, S_4 is applied at 29 ppm. These two selective pulses which are phase cycled independently are of central importance for the new experiments. They transfer magnetization only between carbon spins that are connected via a one-bond coupling and have chemical shifts in the correct ranges. This is true for Arg (C^{δ} and C^{γ}) and Lys (C^{ϵ} and C^{δ}). Signals from the latter will not appear in the Arg spectra, since two relay steps are not sufficient to transfer magnetization from the C^{δ} to the C^{α} . In the case of leucine (C^{β} and C^{γ}) magnetization transfer is forced from the C^{β} to the C^{γ} , away from the backbone. Analyses of the transfer function result in a very low transfer amplitude for the transfer to C^{α} using the proposed delays. The resulting spectra in Figure 4.32 a and b demonstrate the clean selection achieved with the selective pulses, only signals from Arg are present.

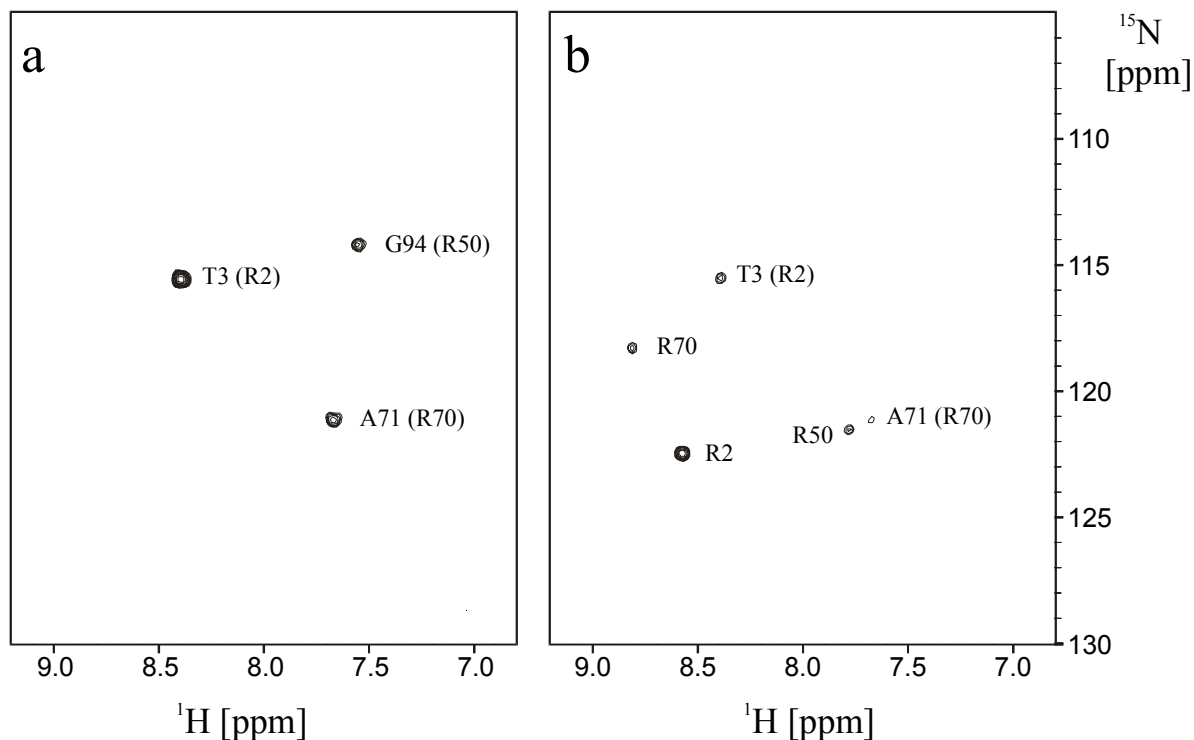


Fig. 4.32: Arg selective ^1H - ^{15}N -correlations of the chicken EphB2 receptor SAM domain, which contains 3 Arg. (a) The R-(i+1)-HSQC was acquired in 3 h using 64 scans. All expected signals are visible. (b) The (i,i+1)-HSQC was acquired in 4½ h using 96 scans. The spectrum contains signals of all Arg residues and smaller signals from the sequential neighbors.

4.4.14 Selecting Lys: The K-(i+1)-, KR-(i+1)- and K-(i,i+1)-HSQC

Lysine has the longest side chain of all natural amino acids. The linear carbon chain consists of 4 methylene groups and ends with a NH_3^+ group. Although the chemical shift of this ^{15}N nucleus is with 33.6 ± 0.6 ppm very characteristic, magnetization between C^ϵ and N^ζ would be very inefficient due to the small coupling constant $^1J_{\text{CN}}$. A much more sensitive method to select Lys (together with Arg) makes use of the characteristic chemical shifts of the C^δ and C^ϵ nuclei of 28.8 ± 1.4 ppm and 41.7 ± 0.9 ppm, respectively. The chemical shifts of these carbons are sufficiently distinct that the nuclei can be treated independently with selective pulses. Carbon pairs with similar chemical shifts are only found in Arg $\text{C}^\delta/\text{C}^\epsilon$ (27.3 ± 1.4 ppm and 43.1 ± 1 ppm) and Leu $\text{C}^\gamma/\text{C}^\beta$ (26.8 ± 1.3 ppm and 42.3 ± 1.8 ppm).

To record spectra selective for Lys the same pulse sequences as in the Arg experiments (chapter 4.14) are used but are extended by an additional relay step to allow magnetization transfer from the C^ϵ to the C^α . The selective pulses are of the same type and of the same length but are applied at slightly different frequency positions: S_1 is centered at 44 ppm, S_2 and S_3 at 45 ppm and S_4 at 29 ppm.

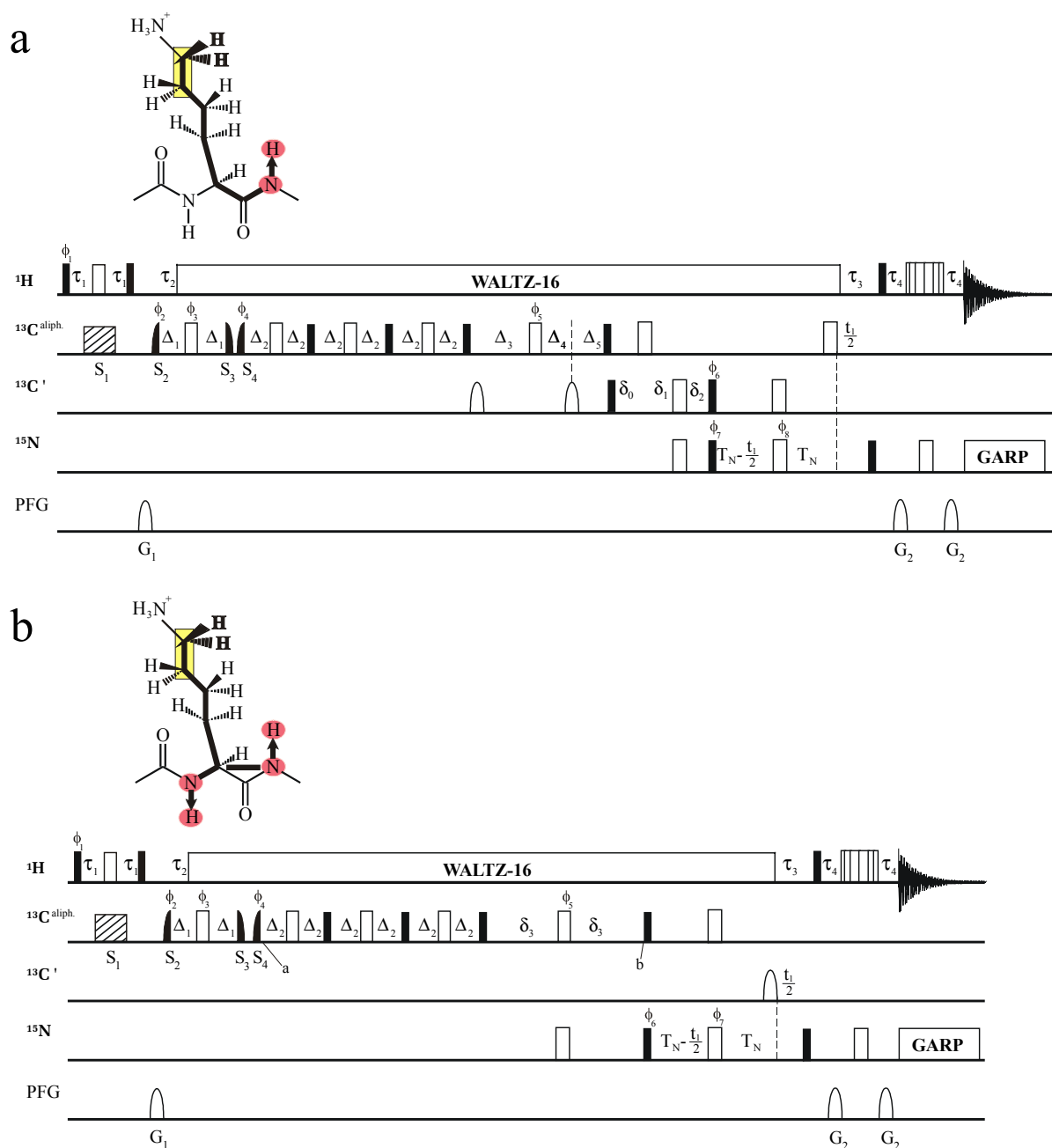
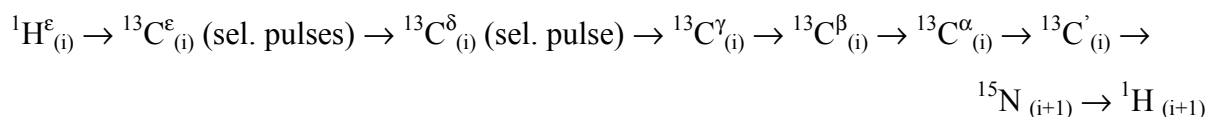


Fig. 4.33: Pulse sequences of the Lys selective ^1H - ^{15}N correlations (a) K-(i+1)-, KR-(i+1)-HSQC and (b) K-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 2.05$ ms (determined by pulse length S_1), $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $\delta_3 = 9$ ms, $T_N = 11$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms, $\Delta_3 = 8.8$ ms (4.5 ms for the KR-(i+1)-HSQC), $\Delta_4 = 4.3$ ms (0 ms for the KR-(i+1)-HSQC), $\Delta_5 = 4.5$ ms. The striped thick bar (S_1) stands for a band-selective 180° REBURP pulse (4096 μs) centered at 44 ppm. Self-refocusing selective LOS2-0 excitation pulses are represented by filled half-sine shapes ($S_2 - S_4$). The pulse lengths of $S_2 - S_4$ are 3072 μs , the offsets 45 ppm, 45 ppm and 29 ppm, respectively. The second LOS2-0 pulse (S_3) has a time inverted shape. Pulse lengths are given for a 600 MHz spectrometer. More experimental details are given in the appendix.

The magnetization transfer in the K-(i+1)-HSQC (Fig. 4.33 a) can be described by:



Signals from Arg can potentially appear in the Lys spectra. It is, however, desirable to obtain spectra that contain only signals from Lys. This can be achieved by tuning the delays in the relay steps to appropriate values as described in chapter 3.6 (three refocused COSY steps: case D and F). However, the transfer functions are slightly different since an additional passive spin (C^ϵ in Lys, C^δ in Arg) is present. The transfer function for Lys between point a and b in the (i,i+1)-pulse sequence depicted in Fig. 4.33 b is:

$$\sin^6(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\delta_3) \sin(\pi J_{CN}2\delta_3) \quad [4.21]$$

J_{CC} stands for all one-bond carbon-carbon couplings which are assumed to be uniform. During each of the three delays $2\Delta_2$ one carbon-carbon coupling is refocused and one is evolving and each of the three 90° pulses following the delay accomplishes a transfer of magnetization to the next carbon in the side chain.

For Arg three magnetization pathways are possible. There are three relay steps after the selective pulse S_4 while it takes only two steps to transfer the magnetization from the C^γ to the C^α . Consequently detectable signal only results if during one of the delays $2\Delta_2$ coupling to the previous carbon in the side chain is refocused while coupling to the next carbon in the chain does not evolve. No magnetization will then be transferred by the subsequent 90° pulse. Equations [4.22], [4.23] and [4.24] represent the situation if no transfer takes place after the first, second or third of the three delays $2\Delta_2$, respectively.

$$\sin(\pi J_{CC}2\Delta_2) \cos(\pi J_{CC}2\Delta_2) \times \cos(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_2) \times \sin^2(\pi J_{CC}2\Delta_2) \times \sin(\pi J_{CC}2\delta_3) \sin(\pi J_{CN}2\delta_3) \quad [4.22]$$

$$\sin^2(\pi J_{CC}2\Delta_2) \times \sin(\pi J_{CC}2\Delta_2) \cos(\pi J_{CC}2\Delta_2) \times \cos(\pi J_{CC}2\Delta_2) \sin(\pi J_{CC}2\Delta_2) \times \sin(\pi J_{CC}2\delta_3) \sin(\pi J_{CN}2\delta_3) \quad [4.23]$$

$$\sin^2(\pi J_{CC}2\Delta_2) \times \sin^2(\pi J_{CC}2\Delta_2) \times \sin(\pi J_{CC}2\Delta_2) \times \cos(\pi J_{CC}2\delta_3) \sin(\pi J_{CN}2\delta_3) \quad [4.24]$$

Note that the transfer functions [4.22] and [4.23] are identical. The delays Δ_2 and δ_3 can be tuned in a way that the signals from transfer functions [4.22], [4.23] and [4.24] cancel each other while the signal from transfer function [4.21] is optimal. After experimental optimization the best results were obtained with $\Delta_2 = 5.5$ ms and $\delta_3 = 9$ ms.

The same type of suppression can be used in the (i+1)-sequence. The delays Δ_3 , Δ_4 and Δ_5 are adjusted to 8.8, 4.3 and 4.5 ms, respectively, which is simultaneously optimal for the evolution of coupling between the C^α and the carbonyl and the cancellation of signals from Arg. Spectra of the K-(i+1)- and K-(i,i+1)-HSQC are shown in Fig. 4.34 a and b. All expected resonances are present while Arg signals are suppressed. If the delays Δ_3 , Δ_4 and Δ_5 are adjusted to 4.5, 0 and 4.5 ms, respectively, the sensitivity is higher but signals from Arg are not suppressed any more. The use of this version, the KR-(i+1)-HSQC should be preferred in the case of proteins >10 kDa.

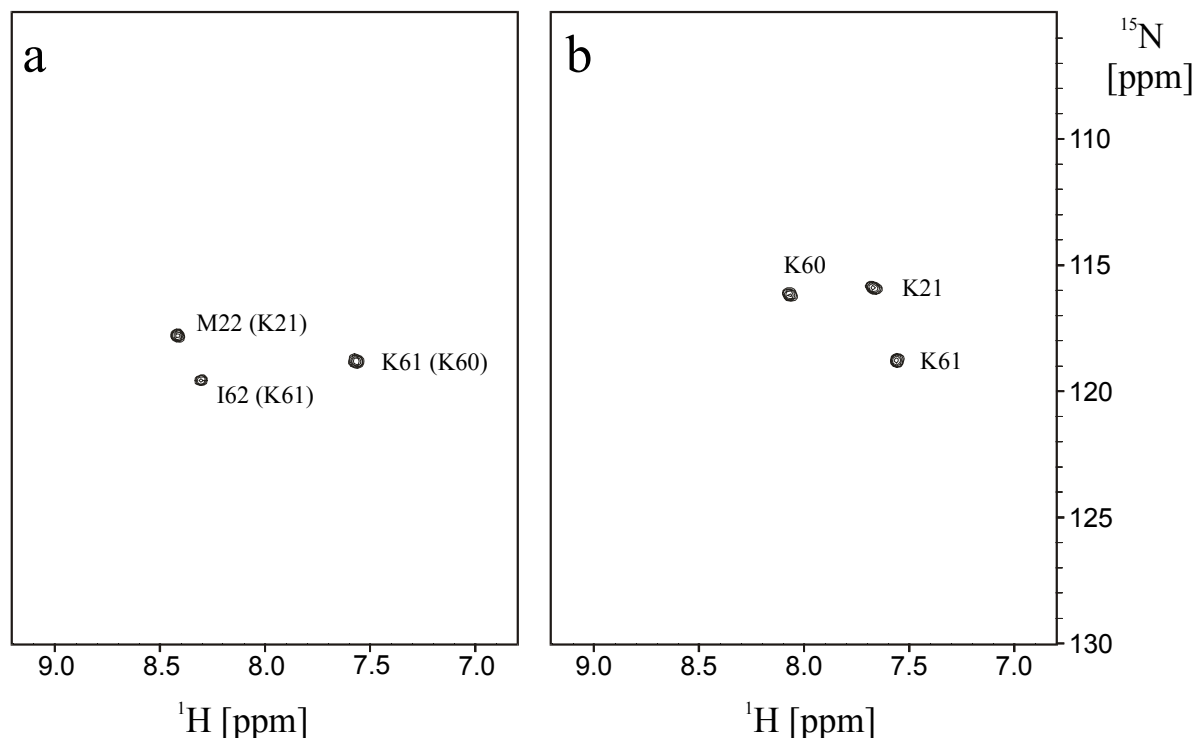


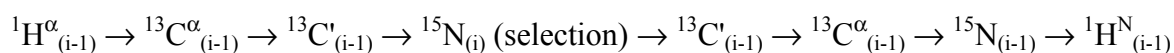
Fig. 4.34: Lys selective ^1H - ^{15}N -correlations of the chicken EphB2 receptor SAM domain (Smalla et al., 1999, PDB: 1SGG), which contains 4 Lys. (a) The K-(i+1)-HSQC was acquired in 6 h using 128 scans. It contains three out of four K-(i+1) signals. E27 the sequential neighbor of K26 is missing. (b) The K-(i,i+1)-HSQC was acquired in 6 h using 128 scans and contains three out of four Lys residues. K26 is missing.

4.4.15 Selecting neighbors of Pro: The P-(i-1)- and P-(i+1)-HSQC

Proline is different – it is the only amino acid which lacks a NH group and the nitrogen is connected to the C^δ to form a 5 membered ring. Proline residues hinder sequential assignment strategies based on a ^1H - ^{15}N correlation because the common NH-detected triple-resonance experiments do not show proline signals.

Therefore two proline selective experiments were designed which transfer the magnetization to the neighboring NH groups (i-1) and (i+1). The selection is based on the lack of a NH group analogous to the proline selective experiments described earlier (Olejniczak and Fesik, 1994; Bottomley, et al. 1999). Two ^{15}N nuclei are involved in the magnetization transfer, the one of Pro and one of the neighboring residue (i+1) or (i-1). Therefore either the ^{15}N of Pro or the ^{15}N of the NH group is recorded in a 2D experiment, leading to four experiments: the Pro-($\text{N}_{(i-1)}$, $\text{H}^{\text{N}}_{(i-1)}$)-HSQC, the Pro-(N_i , $\text{H}^{\text{N}}_{(i-1)}$)-HSQC, Pro-($\text{N}_{(i+1)}$, $\text{H}^{\text{N}}_{(i+1)}$)-HSQC and the Pro-(N_i , $\text{H}^{\text{N}}_{(i+1)}$)-HSQC.

The transfer of magnetization selected by the pulse sequences for the Pro-($\text{N}_{(i-1)}$, $\text{H}^{\text{N}}_{(i-1)}$)-HSQC and the Pro-(N_i , $\text{H}^{\text{N}}_{(i-1)}$)-HSQC can be described as:



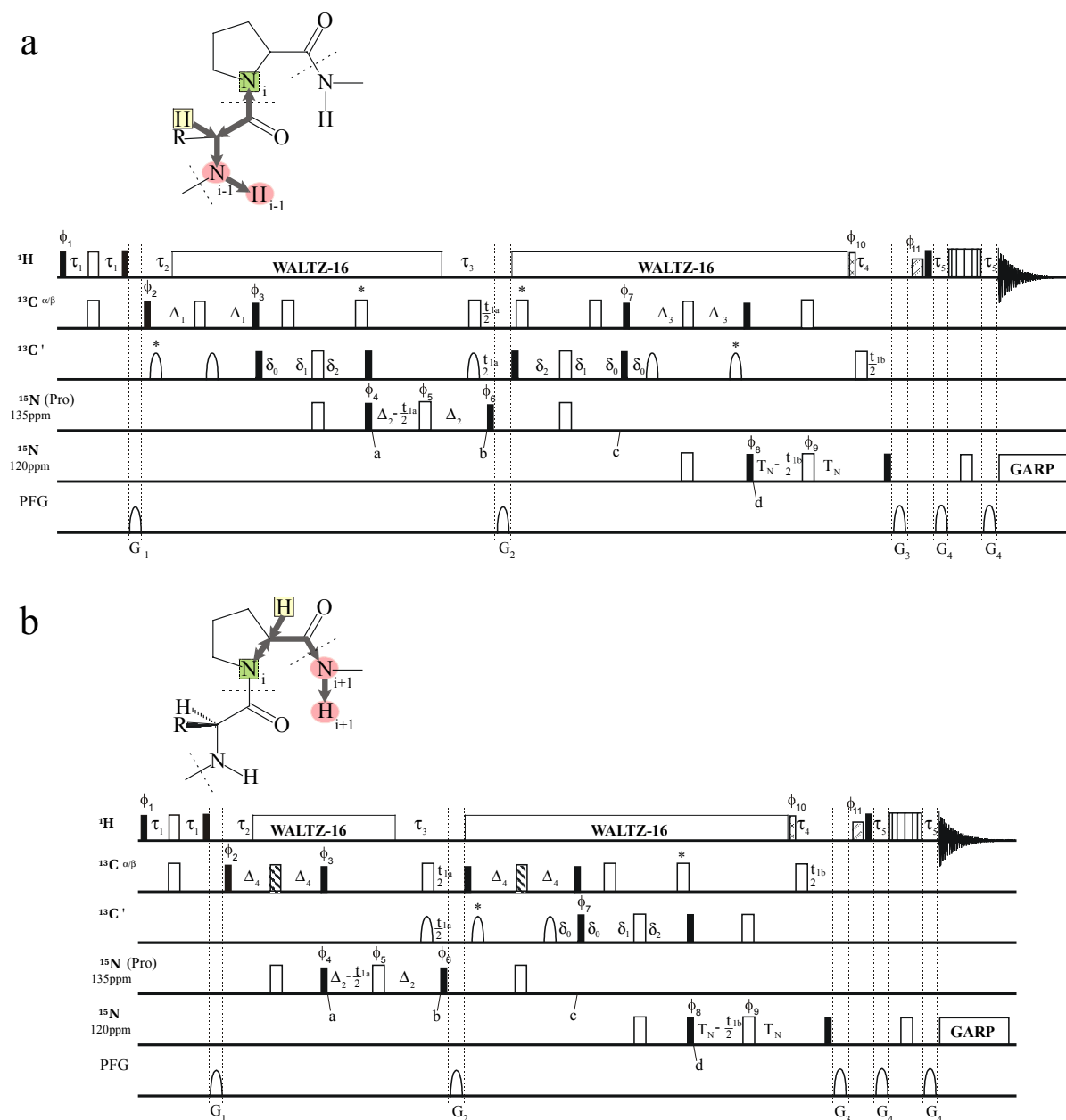
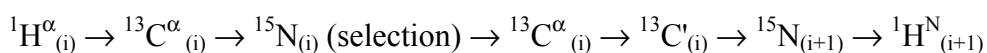


Fig. 4.35: Pulse sequences of the new proline neighbor selective ^1H - ^{15}N -correlations. The pulse sequences yield signal for residues preceding prolines (a) and residues following prolines (b). (a) $\text{Pro}(\text{N}_{(i-1)}, \text{H}_{(i-1)})$ -HSQC (t_{1b} is incremented) and $\text{Pro}(\text{N}_{(i)}, \text{H}_{(i-1)})$ -experiment (t_{1a} is incremented), (b) $\text{Pro}(\text{N}_{(i+1)}, \text{H}_{(i+1)})$ -HSQC (t_{1b} is incremented) and $\text{Pro}(\text{N}_{(i)}, \text{H}_{(i+1)})$ -experiment (t_{1a} is incremented). The following delays were used: $\tau_1 = 1.5$ ms, $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 5.5$ ms, $\tau_5 = 2.25$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms, $\Delta_3 = 14.2$ ms, $\Delta_4 = 7$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms. The black and white striped thick bars stand for band-selective 180° REBURP pulses at 56 ppm with a duration of 1024 μs (at 600 MHz). More experimental details are given in the appendix.

The pulse sequence is depicted in Figure 4.35 a. The magnetization originates from the $^1\text{H}^\alpha$ of the $(i-1)$ residue and is then transferred via the $^{13}\text{C}^\alpha$ and the $^{13}\text{C}'$ to the ^{15}N of proline (point a). During a delay τ_3 tuned to $1/2J_{\text{HN}}$ anti-phase magnetization is created for all residues where an amide proton is present (point b). This magnetization is destroyed by the subsequent composite pulse decoupling on the proton channel. Only proline magnetization,

where no anti-phase magnetization is created, is left. The magnetization is then transferred back to the $^{13}\text{C}'$ and the $^{13}\text{C}^\alpha$. Up to this point, the pulse sequence is identical with that of Olejniczak and Fesik. But instead of returning to the $^1\text{H}^\alpha$ in an out-and-back manner, the magnetization is transferred to the nitrogen (point d) and the amide proton in the new sequence. To accomplish the transfer from $^{13}\text{C}^\alpha_{(i-1)}$ to $^{15}\text{N}_{(i-1)}$ a delay of 28 ms ($1/J_{\text{CC}}$) is used (Δ_3 in Fig. 4.35 a). If the (i-1) residue turns out to be a glycine, the resulting peaks will have opposite sign compared to all other types of amino acid, since the coupling between $^{13}\text{C}^\alpha$ and $^{13}\text{C}^\beta$ is active during the delay for non-glycine amino acids. Two nitrogen nuclei are involved in the flow of magnetization. Depending on the nucleus that is allowed to evolve chemical shift, the resulting spectrum will show a correlation to the prolines own ^{15}N or to the ^{15}N of the (i-1) residue. To optimize the spectral window and the excitation of the nitrogen spins the position of the nitrogen carrier is changed from 135 ppm to 120 ppm during the experiment (point c) and switched back at the end of the sequence.

The transfer of magnetization that is selected by the pulse sequences for the $\text{Pro}(\text{N}_{(i+1)}, \text{H}^{\text{N}}_{(i+1)})$ -HSQC and the $\text{Pro}(\text{N}_{(i)}, \text{H}^{\text{N}}_{(i+1)})$ -HSQC can be described as:



The pulse sequence is depicted in Fig. 4.35 b. The magnetization originates from the $^1\text{H}^\alpha$ of proline and is then transferred via the proline $^{13}\text{C}^\alpha$ to the proline ^{15}N (point a). During the delay $2\Delta_4$ the carbon magnetization is in the transverse plane and a selective REBURP pulse is applied to the $^{13}\text{C}^\alpha$ region of the prolines to refocus the coupling between $^{13}\text{C}^\alpha$ and the $^{13}\text{C}^\beta$. The length of the delay Δ_4 can then be optimized independently of the $^1J_{\text{C}\alpha\text{C}\beta}$ coupling. The selection of the proline ^{15}N is accomplished in the same way as with the sequence described above (point b). Then the magnetization is transferred back to the $^{13}\text{C}^\alpha$ and from there via the $^{13}\text{C}'$ and the nitrogen of the following residue (point d) to the amino proton. Again two nitrogen nuclei are involved in the magnetization transfer pathway and consequently two types of spectra can be recorded depending on which nitrogen chemical shift is allowed to evolve. The same nitrogen carrier frequency jumps as in the other experiments are employed.

It should be noted that there is a potential drawback of the selection mechanism. The spectra are recorded in H_2O to allow for the suppression of amino acids other than proline. There is, however, an amount of 5 – 10 % of D_2O in the sample used for the internal lock. Consequently, 5 – 10 % of the amino acids have a deuteron bound to the nitrogen instead of a proton, irrespective of amino acid type. These residues can lead to low intensity breakthrough peaks (Fig. 4.36 b).

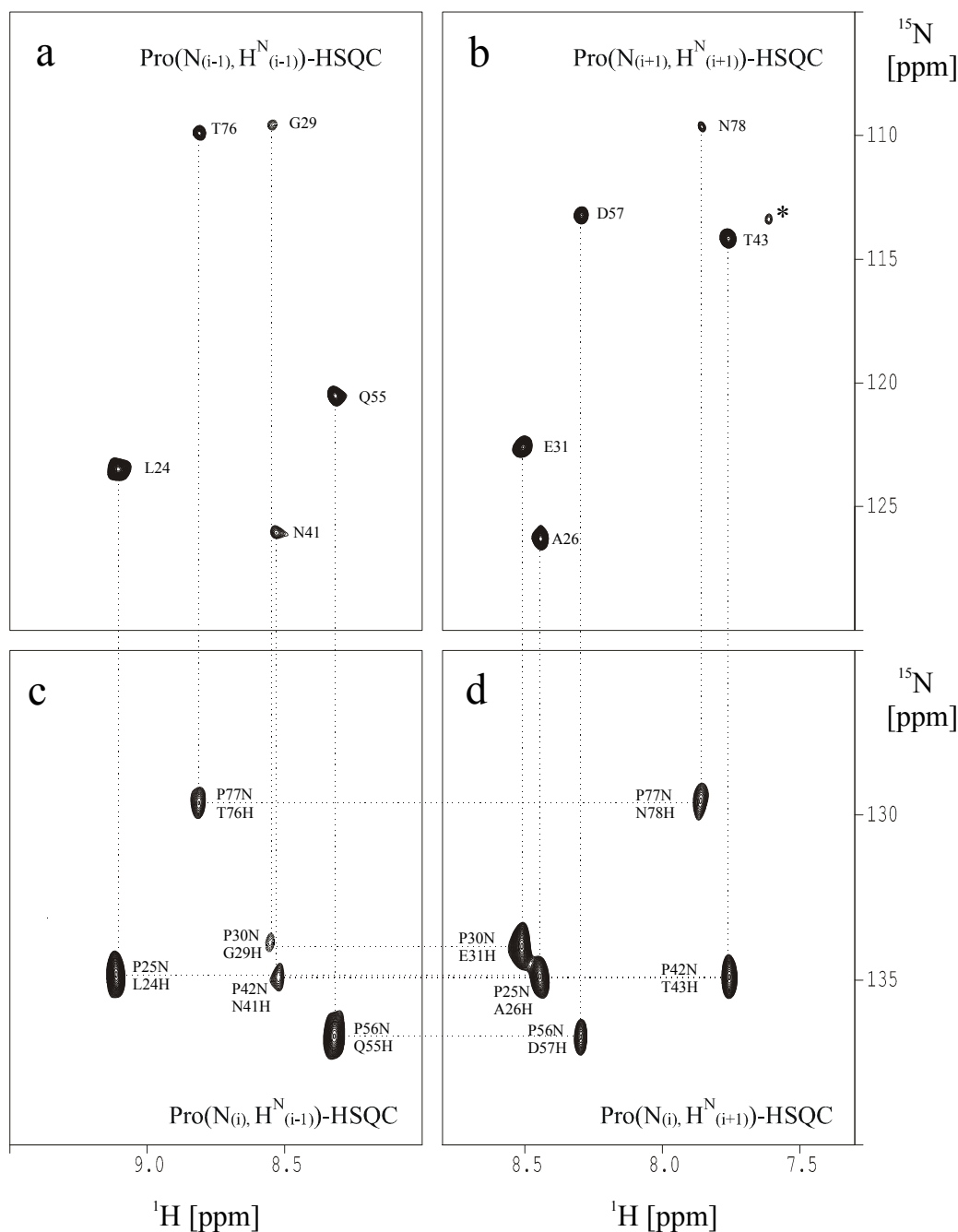


Fig. 4.36: Spectra from the proline-selective experiments of the EVH1 domain from VASP. The $\text{Pro}(\text{N}_{(i-1)}, \text{H}_{(i-1)}^{\text{N}})$ -HSQC spectrum is given in (a) the $\text{Pro}(\text{N}_{(i+1)}, \text{H}_{(i+1)}^{\text{N}})$ -HSQC spectrum in (b). Spectra of $\text{Pro}(\text{N}_{(i)}, \text{H}_{(i-1)}^{\text{N}})$ and $\text{Pro}(\text{N}_{(i)}, \text{H}_{(i+1)}^{\text{N}})$ are shown in (c) and (d). The EVH1 domain contains five prolines. (a) The $\text{Pro}(\text{N}_{(i-1)}, \text{H}_{(i-1)}^{\text{N}})$ -HSQC was acquired in 6 h using 128 scans. All expected signals are visible. The signal of the glycine preceding Pro 30 is negative (see text). (b) The $\text{Pro}(\text{N}_{(i+1)}, \text{H}_{(i+1)}^{\text{N}})$ -HSQC was acquired in 6 h using 128 scans. All five expected signals of residues following prolines are present. A peak from the C-terminal Gly 115, which shows a very intense peak in the ^1H - ^{15}N -HSQC, is visible as a breakthrough peak (marked with *). (c) The $\text{Pro}(\text{N}_{(i)}, \text{H}_{(i-1)}^{\text{N}})$ spectrum was acquired in 6 h using 256 scans. The spectrum shows correlations between the chemical shifts of $\text{H}_{(i-1)}^{\text{N}}$ and proline ^{15}N chemical shifts. All five $^{15}\text{N}_i/{}^1\text{H}_{(i-1)}$ correlations are present and the ^{15}N chemical shift of the prolines can be assigned to the corresponding peak in the P-(i-1)-HSQC. (d) The $\text{Pro}(\text{N}_{(i)}, \text{H}_{(i+1)}^{\text{N}})$ spectrum was acquired in 6 h using 256 scans. All five $^{15}\text{N}_i/{}^1\text{H}_{(i+1)}$ correlations are visible and the ^{15}N chemical shift of the prolines can be assigned to the corresponding peak in the $\text{Pro}(\text{N}_{(i+1)}, \text{H}_{(i+1)}^{\text{N}})$ -HSQC. Via the proline ^{15}N chemical shifts in spectra (c) and (d) a connection between residues directly preceding and following Pro can be established.

Spectra of the EVH1 domain are shown in Fig. 4.36. The protein contains 5 prolines (residues 25, 30, 42, 56 and 77). The sequential neighbors of proline are identified from the $\text{Pro}(\text{N}_{(i-1)}, \text{H}_{(i-1)}^{\text{N}})$ -HSQC (Fig. 4.36a) and $\text{Pro}(\text{N}_{(i+1)}, \text{H}_{(i+1)}^{\text{N}})$ -HSQC (Fig. 4.36b). The cross peaks of spectra (a) and (b) can be connected via the ^{15}N chemical shift of the proline common to both sets of experiments (a, b and c, d). This provides a method of linking the (i-1) and (i+1) residues either side of the proline. In the $\text{Pro}(\text{N}_{(i-1)}, \text{H}_{(i-1)}^{\text{N}})$ -HSQC only one peak with opposite sign appears. This identifies the residue immediately as Gly 29, since this is the only glycine in (i-1) position to a proline.

Note that the amino protons of the residues next to proline are required in these novel experiments. They are therefore not applicable to poly-proline sequences.

4.4.16 Selecting Cys and Ser: The CS-(i+1)- and CS-(i,i+1)-HSQC

Cysteine, alanine and serine are the only amino acids which contain a C^{β} but lack a C^{γ} carbon. Cys does not offer a unique topology or characteristic chemical shifts which could be used for a selection. It shares the CH-CH₂ motif with 14 other amino acids hence this is not sufficiently specific to have practical value. The chemical shift of the C^{β} carbon can vary between 22 and 57 ppm (D. Sharma and K. Rajarathnam, 2000) depending on the oxidation state (Fig. 2.10). Only the lack of a C^{γ} carbon might be used as selection criterion. To suppress signals of undesired amino acids, magnetization is generated on the C^{β} and then coupling to the C^{γ} is allowed to form undetectable double quantum coherence as described in the literature (Dötsch et al, 1996a and b). The design of such an experiment is rather difficult since the C^{γ} carbons can be aliphatic, aromatic or a carbonyl carbon and additionally the coupling constants $^1\text{J}_{\text{C}^{\beta}\text{C}^{\gamma}}$ vary from 30 to 55 Hz (Löhr et al, 2000). The delay during which the $^1\text{J}_{\text{C}^{\beta}\text{C}^{\gamma}}$ coupling can evolve has to be $1/2J$ in order to achieve a complete suppression of undesired signals. Since this is hardly possible with such a broad range of coupling constants the elimination of undesired signals will not be complete.

If the aliphatic, aromatic and carbonyl C^{γ} carbons are treated independently with selective pulses, the time of effective coupling between the C^{β} and C^{γ} can be adjusted independently for Asn/Asp, the aromatic amino acids His/Phe/Trp/Tyr and the remaining residues. However, variations of the $^1\text{J}_{\text{C}^{\beta}\text{C}^{\gamma}}$ coupling constants allow only a rough selection of Ser and Cys with a lot of breakthrough peaks. The pulse sequences for the CS-(i+1)- and CS-(i,i+1)-HSQC are shown in Fig. 4.37. Best results were obtained with effective $^1\text{J}_{\text{C}^{\beta}\text{C}^{\gamma}}$ coupling times of 9.49 ms for Asp/Asn, 9.96 ms for His/Phe/Trp/Tyr and 15.0 ms for the remaining amino acids, corresponding to theoretical coupling constants of 52.7, 50.2 and 33.3 Hz. The effective coupling between aliphatic carbons can evolve during the time $\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4 +$

S₁. Aromatic and aliphatic carbons effectively couple during $\Delta_1 - \Delta_2 + \Delta_3 + \Delta_4 - S_2$ and carbonyl carbons and aliphatic carbons during $\Delta_1 + \Delta_2 - \Delta_3 + \Delta_4 - S_3$.

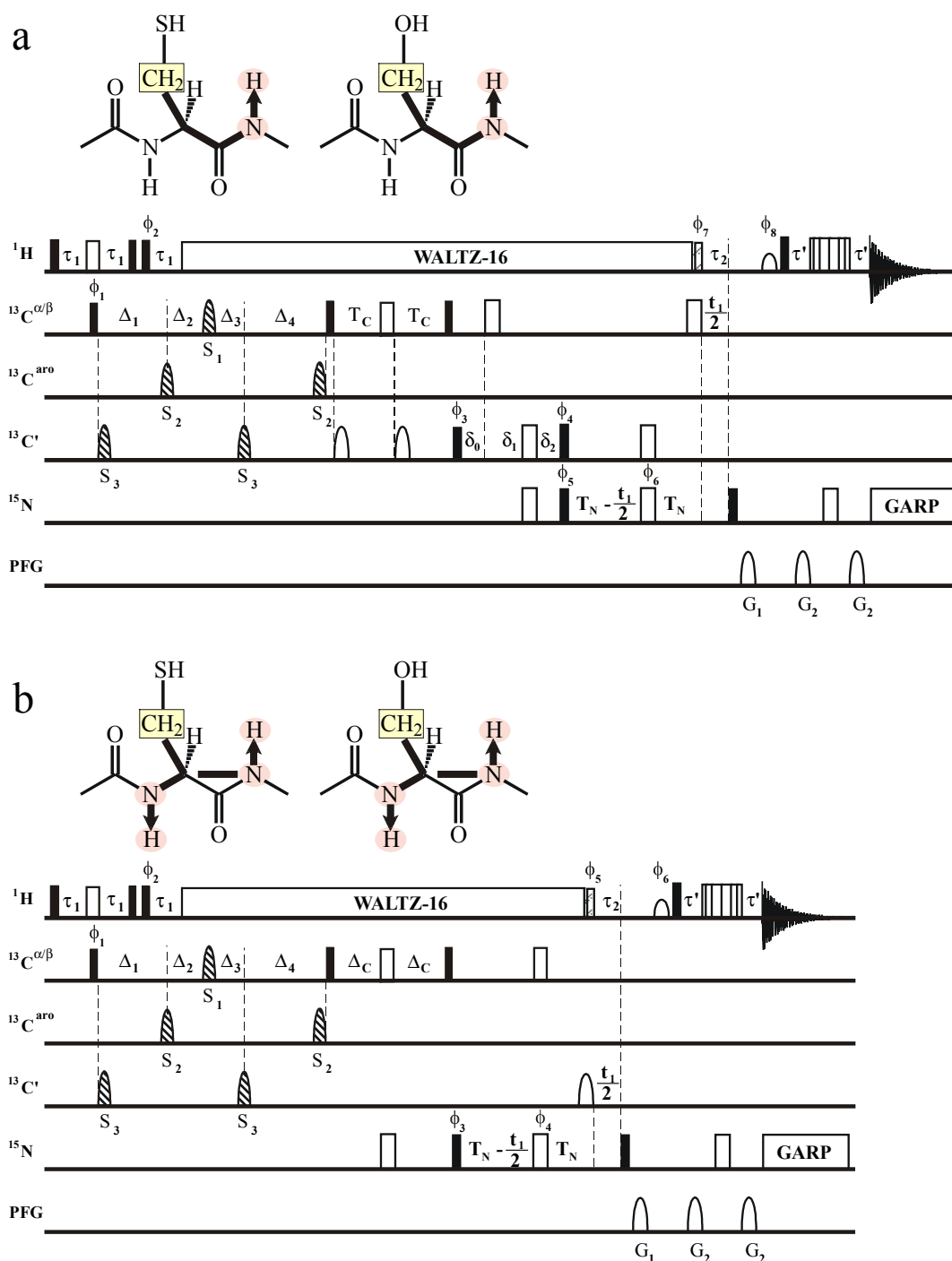


Fig. 4.37: Pulse sequence of the CS-(i+1)-HSQC. The selective pulses (black and white striped sine shapes) S₁, S₂ and S₃ were performed as REBURP pulses with lengths of 409.6 μs, 640 μs and 1024 μs, and offsets of 45 ppm, 120 ppm and 175 ppm respectively (for a 600 MHz spectrometer). The corresponding bandwidths are 65 ppm, 41 ppm and 26 ppm. The delays Δ₁, Δ₂, Δ₃ and Δ₄ were set to 5.62 ms, 1.68 ms, 1.53 ms and 5.77 ms, respectively. The remaining delays are set to the following values τ₁ = 3.5 ms, τ₂ = 5.5 ms, τ₃ = 2.25 ms, δ₀ = 4.5 ms, δ₁ = 6.9 ms, δ₂ = 11.4 ms, T_N = 11 ms, T_C = 3.7 ms, Δ_C = 9 ms. More experimental details are given in the appendix.

The resulting spectra are shown in Fig. 4.38. Among positive and negative breakthrough peaks the desired Cys and Ser signals are present with strong intensity. Because of variations of $^1J_{\text{C}\beta\text{C}\gamma}$ coupling constants, signals originating from undesired amino acid types like Met and Gln break through. Gly is, in principle, suppressed like Asn and Asp with its $^1J_{\text{C}\alpha\text{C}'}$ coupling but suppression is again not perfect. Ala residues are suppressed because the MUSIC- CH_2 selection suppresses the major part of CH_3 and additionally the selective pulse S_1 effects only partially the C^β .

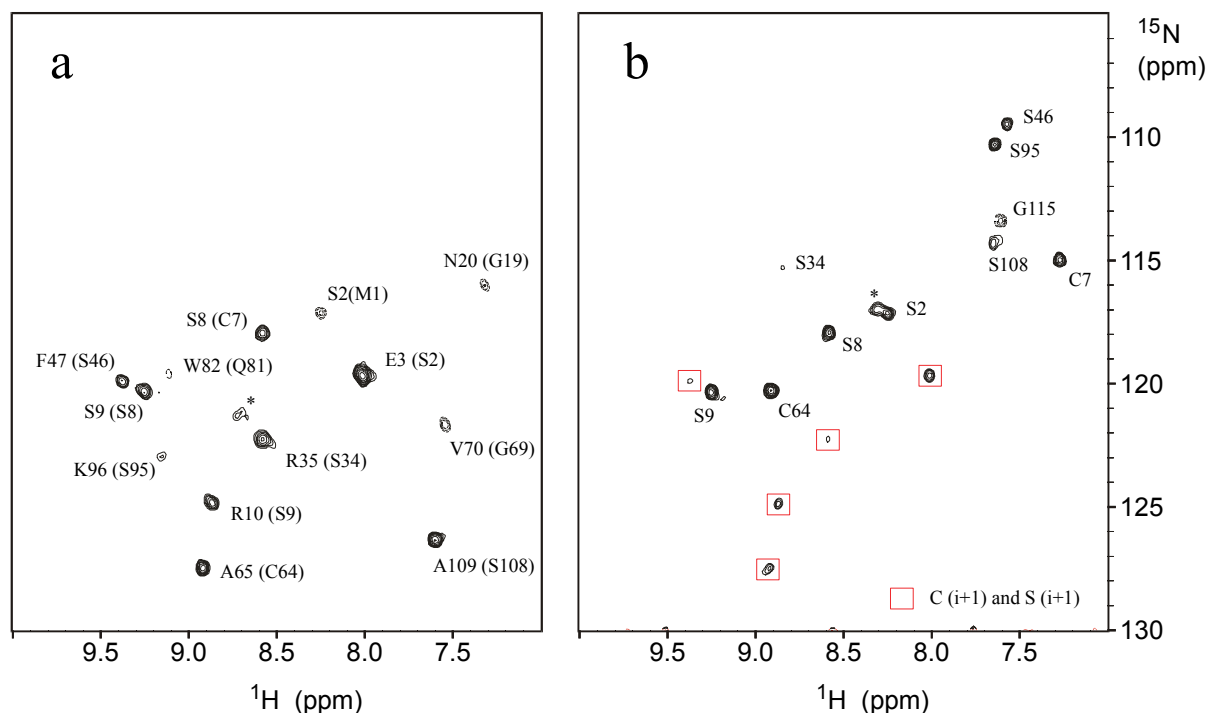


Fig. 4.38: Cys/Ser selective ^1H - ^{15}N correlations of the EVH1 domain from VASP. The EVH1 domain contains 2 Cys and 7 Ser. (a) The CS-(i+1)-HSQC was acquired in 1 ½ h using 32 scans. All expected signals of Cys and Ser neighbors are present. Four undesired signals (following Gly, Met and Gln) are not completely suppressed but they can be identified by their negative sign. A signal from a minor conformation is marked with an asterisk. (b) The CS-(i,i+1)-HSQC was acquired in 1 ½ h using 32 scans. All expected Cys and Ser signals are visible. A negative signal of the undesired G115 is present. A signal from a minor conformation is marked with an asterisk.

4.4.17 Attempts to select Met

Methionine contains two CH_2 groups, but this is not unique enough to select solely Met or Met among a small group of amino acid types. In principle, Met offers a methyl group which can be selected by MUSIC, but the $^{\epsilon}\text{CH}_3$ group is separated from the carbon chain by a sulfur atom. The coupling constants across the sulfur are very small (Bax et al., 1994) and a magnetization transfer from a selected C^{ϵ} to the C^{γ} and further to the backbone would be very inefficient. The chemical shifts of the ^{13}C nuclei of Met do not allow a filter based on

selective pulses. The chemical shifts of the C^β and C^γ are almost identical and strong coupling can occur which complicates magnetization transfer. An experiment with a MUSIC- CH_2 selection implemented at the beginning of a CBCA(CO)NH sequence extended by a relay step, lead to signals of almost all amino acids including Met. Attempts to make the experiment more specific for Met by suppressing undesired signals were not successful.

4.4.18 Selecting the indole NH of Trp: The $H^{\epsilon^1}N^{\epsilon^1}$ -HSQC

Signals of the $H^{\epsilon^1}-N^{\epsilon^1}$ group of Trp appear in the ^{15}N -HSQC in the same region as backbone NH groups which leads sometimes to overlap. An unambiguous assignment of the $H^{\epsilon^1}-N^{\epsilon^1}$ is important since NOE signals involving these nuclei provide often key distance restraints for structure calculations.

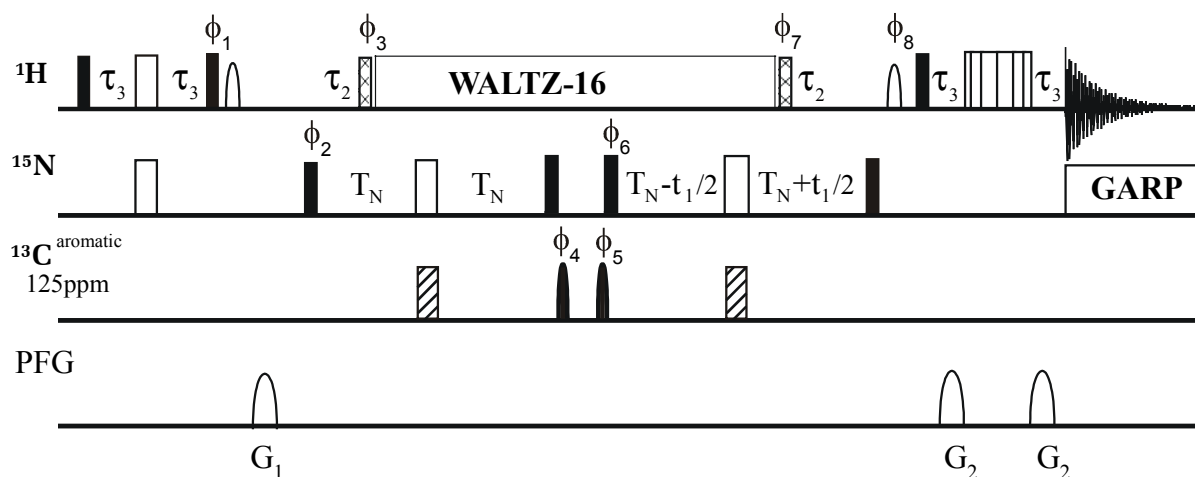


Fig. 4.39: Pulse sequence of the Trp $H^{\epsilon^1}N^{\epsilon^1}$ -HSQC correlation. The delays are set to $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$ ms. The carrier frequency for ^{13}C was centered at 125 ppm. 180° REBURP pulses (black and white striped thick bars) with a duration of 1024 μ s and 90° LOS2-0 pulses (filled sine shapes) with a duration of 512 μ s were used. More experimental details are given in the appendix.

The N^{ϵ^1} is scalar coupled to the adjacent aromatic carbons C^{δ^1} and C^{ϵ^2} . This coupling is used to transfer magnetization from the N^{ϵ^1} to the aromatic carbons. Because of their characteristic ^{13}C chemical shift they can be selected using shaped pulses. The pulse sequence is a 2D version of an HNCA-like experiment, however the ^{13}C offset is changed to 125 ppm (Fig. 4.39). The selectivity is achieved by using LOS2-0 pulses with a length of 512 μ s as 90° pulses and REBURP pulses with length of 1024 μ s as 180° pulses on the carbon channel. The experiment is quite sensitive and can be recorded in a very short time, a comparison between the conventional HSQC and the $H^{\epsilon^1}N^{\epsilon^1}$ -HSQC of the EVH1 domain is shown in Fig. 4.40. The signal intensity is 50% compared to a conventional ^{15}N -HSQC.

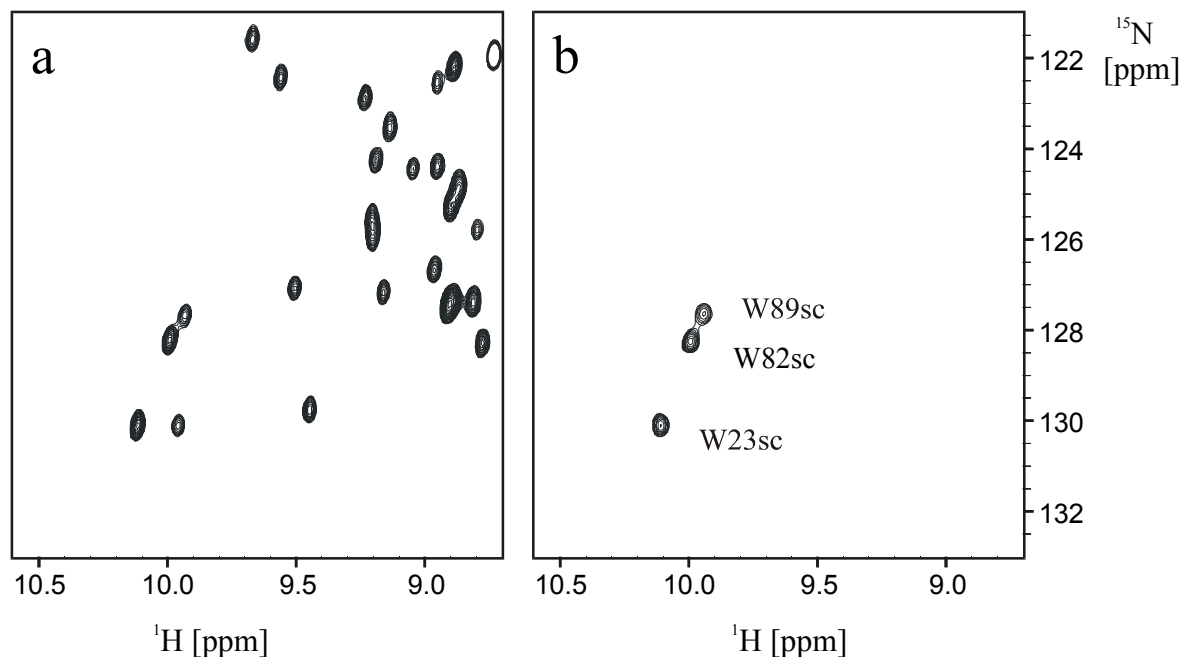


Fig. 4.40: Trp $\text{H}^{\text{e}1}\text{N}^{\text{e}1}$ -HSQC in comparison to a normal ^{15}N -HSQC of the EVH1 domain from VASP. (a) Region from the ^{15}N -HSQC. The spectrum was acquired in 12 min using 4 scans. (b) The Trp $\text{H}^{\text{e}1}\text{N}^{\text{e}1}$ -HSQC was acquired in 24 min using 8 scans. The three expected signals are present.

4.5 Conclusion and outlook

In total, 39 novel amino acid type-selective experiments were designed, which allow the identification of the amino acid types of signals in ^1H - ^{15}N correlation spectra. For 19 amino acid types (all except Met) selective experiments were developed, either for selecting a residue type alone or for selecting a group of amino acid types. An overview is given in table 4.1. Pairs of experiments for Gly, Ala, Thr/Ala, Val/Ile/Ala, Leu/Ala, Ser, Trp, Phe/Tyr/His, Asp/Asn/Gly, Asn, Asp, Glu/Gln/Gly, Gln, Glu, Arg, Lys, Cys/Ser which select either the $\text{NH}_{(i+1)}$ or $\text{NH}_{(i)}/\text{NH}_{(i+1)}$ amide signals were presented. Four Pro selective experiments were introduced allowing the linking of signals of the two neighbors ($i-1$) and ($i+1$) via the proline ^{15}N chemical shift. Finally, an experiment which selects Trp $\text{H}^{\text{e}1}\text{N}^{\text{e}1}$ groups was developed.

The significance of the presented experiments in terms of an automated assignment is discussed in the next chapter.

experiments		selection criteria	typical number of scans
G (i+1)	G (i,i+1)	MUSIC-CH ₂	8 / 16
A (i+1)	A (i,i+1)	MUSIC-CH ₃	48 / 48
TA (i+1)	TA (i,i+1)	MUSIC-CH ₃ , selective pulse	48 / 96
VIA (i+1)	VIA (i,i+1)	MUSIC-CH ₃ , selective pulse	48 / 96
LA (i+1)	LA (i,i+1)	MUSIC-CH ₃ , selective pulses, tuned delays	96 / 144
S (i+1)	S (i,i+1)	MUSIC-CH ₂ , selective pulses	32 / 32
DNG (i+1)	DNG (ii,+1)	MUSIC-CH ₂ , selective pulses	32 / 64
N (i+1)	N (i,i+1)	MUSIC-NH ₂ ,	64 / 128
D (i+1)	D (i,i+1)	MUSIC-CH ₂ , tuned delay to select CO ₂	32 / 64
EQG (i+1)	EQG (ii,+1)	MUSIC-CH ₂ , selective pulses, tuned delays	64 / 96
QN (i+1) or Q (i+1)	Q (i,i+1)	NH ₂ selection, tuned delays	128 / 256 256
ED (i+1) or E (i+1)	E (i,i+1)	MUSIC-CH ₂ , tuned delay to select CO ₂ , tuned delays	64 / 128 128
R (i+1)	R (i,i+1)	selecting C pair with selective pulses	64 / 96
KR (i+1) or K (i+1)	K (i,i+1)	selecting C pair with selective pulses, tuned delays	64 / 128 128
P (i+1)	P (i-1)	tuned delay to select N which lacks H,	128 / 128
FYH (i+1)	FYH (i,i+1)	MUSIC-CH ₂ selection, selective pulses	64 / 128
W (i+1)	W (i,i+1)	MUSIC-CH ₂ selection, selective pulses	64 / 128
CS (i+1)	CS (i,i+1)	MUSIC-CH ₂ selection, tuned delays	32 / 32
W-H ^{ε1} N ^{ε1}		selective pulses	8

Table 4.1: Overview of the 37 novel amino acid type-selective ¹H-¹⁵N correlations with a short description of the selection mechanism and typical number of scans (on the left for the (i+1), on the right for the (i,i+1) versions). Note that P-(i+1)- is identical to the Pro-(N_(i+1), H^N_(i+1))-HSQC, the P(i-1)- to the Pro-(N_(i-1), H^N_(i-1))-HSQC, respectively. The Pro-(N_i, H^N_(i-1))-HSQC and the Pro-(N_i, H^N_(i+1))-HSQC are not listed here.

All the experiments are open for further modifications, e. g. chemical shift evolution of any of the involved side chain nuclei or sensitivity enhancement can be implemented. Even the TROSY technique can be implemented, but this makes more sense in deuterated proteins in which specific protons essential for the experiments are not exchanged by deuterium, e. g. in ¹³C¹H₃/²H/¹³C/¹⁵N labeled samples.

References

- L. J. Ball, R. Kuhne, B. Hoffmann, A. Hafner, P. Schmieder, R. Volkmer-Engert, M. Hof, M. Wahl, J. Schneider-Mergener, U. Walter, H. Oschkinat and T. Jarchau, *EMBO J.* **19**, 4903-4914 (2000).
- A. Bax, R. H. Griffey and B. L. Hawkins, *J. Magn. Reson.* **55**, 301-315 (1983).
- A. Bax, F. Delaglio, S. Grzesiek and G. W. Vuister, *J. Biomol. NMR* **4**, 787-797 (1994).
- R. Bazzo, D. O. Cicero and G. Barbato, *J. Magn. Reson.* **136**, 15-21 (1999).
- F. J. Blanco, A. R. Ortiz and L. Serrano, *J. Biomol. NMR* **9**, 347 (1997).
- M. J. Bottomley, M. J. Macias, Z. Liu and M. Sattler, *J. Biomol. NMR* **13**, 381-385 (1999).
- W. Boucher, E. D. Laue, S. Campbell-Burk and P. J. Domaille, *J. Am. Chem. Soc.* **114**, 2262-2264 (1992a).
- W. Boucher, E. D. Laue, S. L. Campbell-Burk and P. J. Domaille, *J. Biomol. NMR* **2**, 631-637 (1992b).
- V. F. Bystrov, *Progr. NMR Spectrosc.* **10**, 41-81 (1976).
- T. Carlomagno, M. Maurer, M. Sattler, M. G. Schwendinger, S. J. Glaser and C. Griesinger, *J. Biomol. NMR* **8**, 161-170 (1996).
- G. Cornilescu, A. Bax and D. A. Case, *J. Am. Chem. Soc.* **122**, 2168-2171 (2000).
- C. Dalvit, *J. Magn. Reson. B.* **112**, 282-288 (1996).
- V. Dötsch, R. E. Oswald and G. Wagner, *J. Magn. Reson. B* **110**, 107-111 (1996a).
- V. Dötsch and G. Wagner, *J. Magn. Reson. B* **111**, 310-313 (1996b).
- V. Dötsch, R. E. Oswald and G. Wagner, *J. Magn. Reson. B* **110**, 304-308 (1996c).
- V. Dötsch, H. Matsuo and G. Wagner, *J. Magn. Reson. B* **112**, 95-100 (1996d).
- L. Emsley and G. Bodenhausen, *Chem. Phys. Lett.* **165**, 469-476 (1990).
- B.T. Farmer II and R. A. Venters, *J. Biomol. NMR* **7**, 59-71 (1996).
- H. Geen, R. Freeman, *J. Magn. Reson.* **93**, 93-141 (1991).
- K. Gehring and E. Guittet, *J. Magn. Reson. B* **109**, 206-208 (1995).
- S. Grzesiek and A. Bax, *J. Am. Chem. Soc.* **114**, 6291-6293 (1992a).
- S. Grzesiek and A. Bax, *J. Magn. Reson. B* **99**, 201-207 (1992b).
- S. Grzesiek and A. Bax, *J. Biomol. NMR* **3**, 185-204 (1993).
- S. Grzesiek and A. Bax, *J. Am. Chem. Soc.* **117**, 6527-6531 (1995).
- F. Löhr and H. Rüterjans, *J. Magn. Reson. B* **112**, 259-268 (1996).
- F. Löhr, C. Pérez, R. Köhler, H. Rüterjans and J. Schmidt, *J. Biomol. NMR* **18**, 13-22 (2000).
- D. Marion, M. Ikura, R. Tschudin and A. Bax, *J. Magn. Reson.* **85**, 393-399 (1989).
- D. R. Muhandiram, P. E. Johnson, D. Yang, O. Zhang, L. P. McIntosh and L. E. Kay, *J. Biomol. NMR* **10**, 283-288 (1997).
- L. Müller, *J. Am. Chem. Soc.* **101**, 4481-4484 (1979).
- A. Musacchio, M. Noble, R. Pauptit, R. Wierenga and M. Serraste, *Nature* **359**, 851-855 (1992).
- E. T. Olejniczak and S. W. Fesik, *J. Am. Chem. Soc.* **116**, 2215-2216 (1994).
- M. Pellecchia, G. Wider, H. Iwai and K. Wüthrich, *J. Biomol. NMR* **10**, 193-197 (1997a).
- M. Pellecchia, H. Iwai, T. Szyperski and K. Wüthrich, *J. Magn. Reson.* **124**, 274-278 (1997b).
- J. J. Prompers, A. Groenewegen, C. W. Hilbers and H. A. M. Pepermans, *J. Magn. Reson.* **130**, 68-75 (1998).
- N. S. Rao, P. Legault, D. R. Muhandiram, J. Greenblatt, J. L. Battiste, J. R. Williamson and L. E. Kay, *J. Magn. Reson. B* **113**, 272-276 (1996).
- P. Schmieder, M. Leidert, M. Kelly and H. Oschkinat, *J. Magn. Reson.* **131**, 199-202 (1998).
- B. R. Seavey, E. A. Farr, W. M. Westler and J. Markley, *J. Biomol. NMR* **1**, 217-236 (1991);
<http://www.bmrb.wisc.edu>
- A. J. Shaka, J. Keeler, T. Frenkiel and R. Freeman, *J. Magn. Reson.* **52**, 335-338 (1983).
- A. J. Shaka, P. B. Barker and R. Freeman, *J. Magn. Reson.* **64**, 547-552 (1985).

- V. Sklenar, M. Piotto, R. Leppik and V. Saudek, *J. Magn. Reson. A* **102**, 241-245 (1993).
- D. Sharma and K. Rajarathnam, *J. Biomol. NMR* **18**, 165-171 (2000).
- M. Smalla, P. Schmieder, M. Kelly, A. ter Laak, G. Krause, L. Ball, M. Wahl, P. Bork and H. Oschkinat, *Protein Sci.* **8**, 1954-1961 (1999).
- J. L. Sudmeier, E. L. Ash, U. L. Günther, X. Luo, P. A. Bullock and W. W. Bachovchin, *J. Magn. Reson. B* **113**, 236-247 (1996).
- D. Uhrín, S. Uhrínová, C. Leadbeater, J. Nairn, N. C. Price and P. N. Barlow, *J. Magn. Reson.* **142**, 288-293 (2000).
- H. Vis, R. Boelens, M. Mariani, R. Stroop, C. E. Vorgias, K. S. Wilson and R. Kaptein, *Biochemistry* **33**, 14858-14870 (1994).
- B. Whitehead, M. Tessari, P. Düx, R. Boelens, R. Kaptein and G. W. Vuister, *J. Biomol. NMR* **9**, 313-316 (1997).
- M. Wittekind, W. J. Metzler and L. Mueller, *J. Magn. Reson. B* **101**, 214-217 (1993).
- T. Yamazaki, J. D. Forman-Kay and L. E. Kay, *J. Am. Chem. Soc.* **115**, 11054-11055 (1993a).
- T. Yamazaki, M. Yoshida and K. Nagayama, *Biochemistry* **32**, 5656-5669 (1993b).
- T. Yamazaki, S. M. Pascal, A. U. Singer, J. D. Forman-Kay and L. E. Kay, *J. Am. Chem. Soc.* **117**, 3556-3564 (1995).
- C. Zwahlen, S. J. F. Vincent, K. H. Gardner and L. E. Kay, *J. Am. Chem. Soc.* **120**, 4825-4831 (1998).

Chapter 5

Amino acid type-selective experiments open up new strategies in automated assignment

Parts of this chapter are included in an application for a patent.

5.1 Introduction

The task of an automated assignment procedure is to extract a list of the chemical shift assignments of all ^{13}C , ^{15}N and ^1H nuclei from processed NMR spectra. The initial step is the sequential backbone assignment which delivers the chemical shift assignments of all backbone and $^{13}\text{C}^\beta$ nuclei. The state of the art in automated backbone assignment is reviewed in chapter 5.2.

In general, automated techniques should be as reliable as possible and manual interference should be minimized or not necessary at all. Novel technologies for automated resonance assignment require additional, more specific data in order to generate reliable results. Amino acid type-selective experiments provide such specific data and thus increase the reliability of the assignment output. A method to obtain ^1H - ^{15}N chemical shift assignments for unique sequential dipeptides in the protein sequence based solely on amino acid type-selective 2D experiments is introduced in chapter 5.3. This approach offers an additional independent route to ^1H - ^{15}N chemical shift assignment. Four potential concepts for the implementation of amino acid type-selective ^1H - ^{15}N correlations in automated assignment procedures are provided in chapter 5.4.

5.2 Automated backbone resonance assignment of $^{13}\text{C}/^{15}\text{N}$ labeled proteins - an overview

Sequential backbone resonance assignment is an essential requirement for protein structure determination by NMR which is outlined in Fig. 5.1. Efforts are ongoing to automate each of the steps (Moseley and Montelione, 1999). To automate the whole process all steps have to be linked and individual in- and output data have to be compatible. The sequential backbone assignment follows the processing of the spectra and delivers the input for the automated side chain and NOESY assignment. Automated procedures for backbone resonance assignment use, in general, the same analysis scheme which is shown in Fig. 5.2.

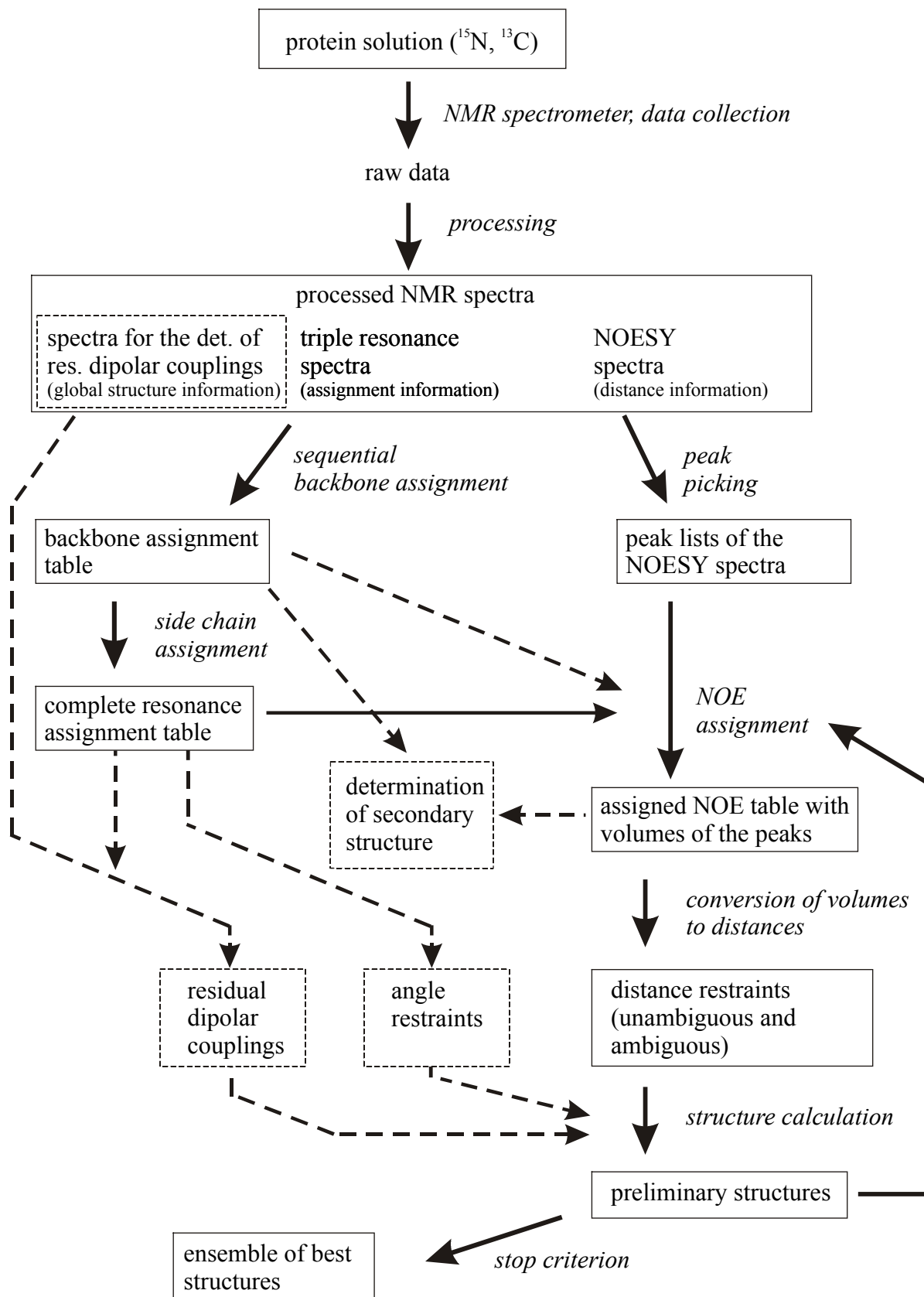


Fig. 5.1: Flow chart of the structure determination of $^{13}\text{C}/^{15}\text{N}$ labeled proteins by NMR. Dashed arrows represent the use of optional data.

The procedure starts with peak picking. In the next step peaks are filtered (filtering) and resonances from different spectra are related (referencing). Subsequently, resonances are grouped into generic spin systems (grouping, see chapter 2.3). Afterwards sequential generic spin systems are linked into segments (linking) and then the amino acid types of the spin systems within the segment are identified (typing). An alternative is to identify the amino acid types of each generic spin system before linking them into segments. Then the segments with the amino acid type information (usually probabilities) are mapped onto the protein sequence (mapping). To complete the backbone assignment the assigned segments have to be linked and extended until the entire sequence is assigned. Finally the assignment should be evaluated to ensure a reliable assignment table. However, the overall performance is dictated by the performance of the weakest step.

Various software packages already exist for automated/semiautomated backbone assignment which include several steps of the scheme in Fig. 5.2. They are reviewed in the literature (Moseley and Montelione, 1999). A program which contains all steps is not available so far.

Most of the programs use peak lists as input. Although several automated peak picking programs exist, the generation of reliable peak lists always requires some degree of manual interference (Moseley et al., 2001). Automated peak picking tools provided in Sparky (Goddard and Kneller, 1999), Felix (Molecular Simulation Inc.), NMRCompass (Molecular Simulation Inc.) or PIPP (Garett et al., 1991) can be used. Another approach is restricted peak picking (e.g. with Sparky) which searches for peaks only in certain areas in multidimensional spectra. These areas can, for example, be defined by the peak list of an ^{15}N -HSQC such that only peaks from certain strips are picked. The procedure is described in the literature (Moseley et al., 2001). Although the different automated assignment procedures can tolerate a certain percentage of artifacts, it is recommended to check the peak lists manually.

One of the few programs which use NMR spectra as direct input is Chatch23 (Croft et al., 1997). This program searches simultaneously in a variety of spectra for characteristic spin patterns (pattern search) and delivers a list of generic spin systems. The following steps (amino acid typing, linking and mapping) are not included.

Once peak lists of all necessary spectra are available, several programs like AutoAssign (Zimmerman et al., 1997), PASTA (Leutner et al., 1998), TATAPRO (Atreya et al., 2000) or the program from Lukin et al. (Lukin et al., 1997) deliver quite complete assignment lists with low error rates. However, the resulting peak lists have to be evaluated manually as long as the error rate (false assignment) is not zero. Note that missing assignments for certain residues cause only a minor problem. This is often not due to bad performance of the automation program, but is rather a direct consequence of missing signals.

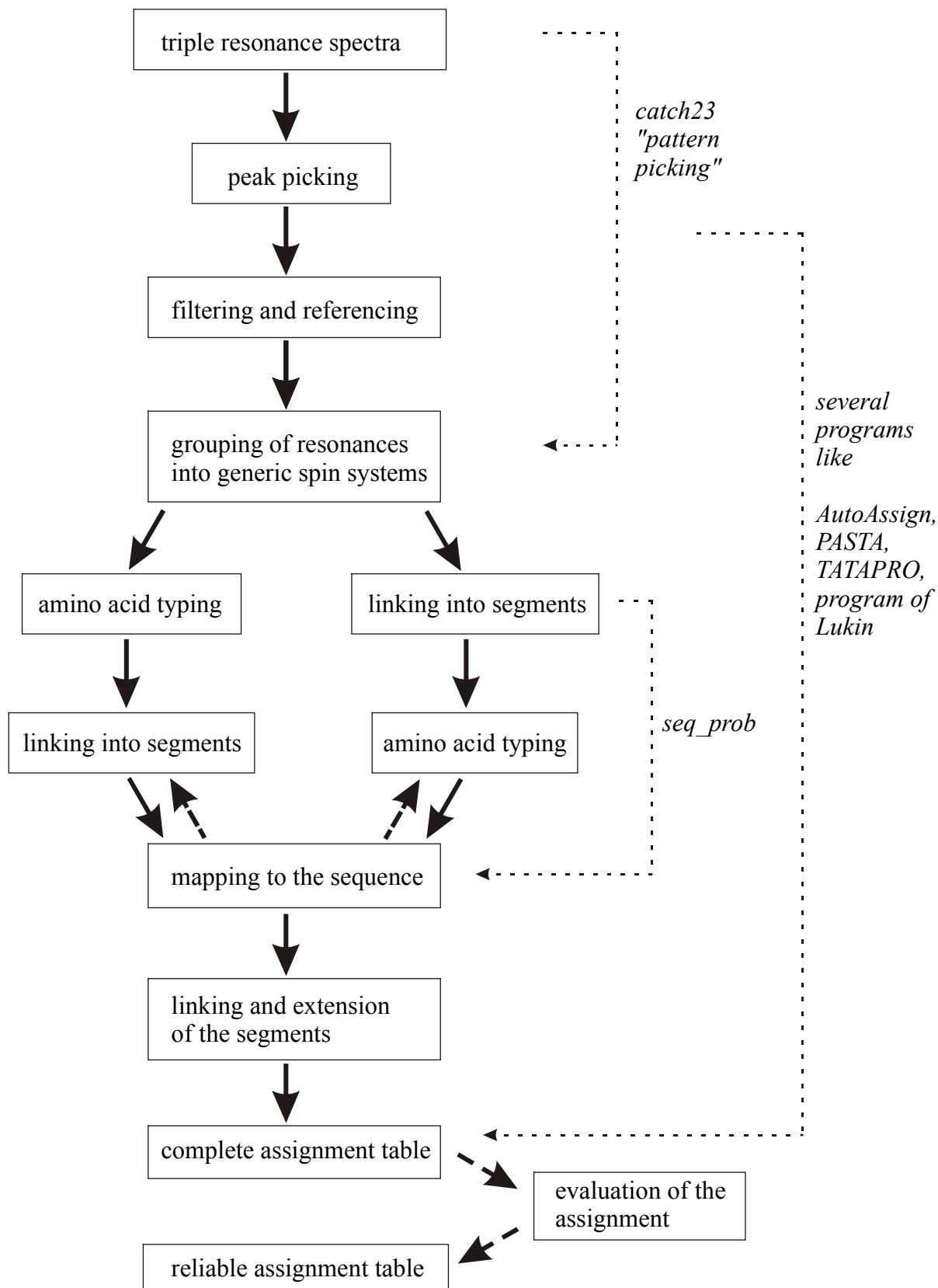


Fig. 5.2: General chart flow of automated backbone resonance assignment procedures. On the right available automated programs are listed and the steps they include. Except the program `seq_prob` (Grzesiek and Bax, 1993) all others are described in the text.

Although the available programs show a good performance starting from peak lists, the overall robustness can still be increased and false assignment should be excluded.

5.3 Problematic steps in automated assignment

Most problems occur during peak picking, grouping and amino acid typing. If peaks are not picked due to overlapped resonances or missing signals, problems appear during the grouping of signals into generic spin systems. Problems which may occur in spectra of the CBCA(CO)NH and CBCANH experiments (see chapter 2.3) are discussed as examples. The C^α and C^β frequencies are usually assigned using these experiments. If the chemical shifts of the C^α and C^β nuclei of a residue can not be unambiguously assigned, the amino acid type identification will be unreliable.

Typical peak patterns in the ^{13}C dimension of the CBCA(CO)NH and the CBCANH spectra are shown in Fig. 5.3. If no overlap of frequencies occur, the presence of Gly residues (G) and non-Gly residues (Z) lead to four possible patterns for sequential pairs of ZZ, ZG, GZ and GG (Fig. 5.2 a-d). In case (a), (c) and (d) an unambiguous assignment of the frequencies to C^α_i , C^α_{i-1} , C^β_i and C^β_{i-1} is straight forward. Overlapped signals in the ^{13}C dimension like overlapped frequencies of C^α_i , C^α_{i-1} , C^β_i and C^β_{i-1} in the CBCANH (Fig. 5.3 e-g, j, k) or cancellation of signals (Fig. 5.32 h and i) hinders the assignment of the C^α and C^β chemical shifts which are crucial for the amino acid typing. The chemical shifts of Ser C^β tend to fall in the range of C^α chemical shifts and cancellation may occur (Fig. 5.3 h). Typical C^α chemical shifts of Gly are similar to C^β chemical shifts of amino acids like Asp and Leu (Fig. 5.3 j). In case (b) and (e)-(l) the assignment of the ^{13}C chemical shifts of the generic spin systems will be ambiguous. Pattern (b) and (e) as well as (g) and (j) look the same and can not be easily distinguished without additional information although they result from different sequential pairs ZG and ZZ. Long repeats of the same kind of amino acids like LLL do often lead to case (g). Overlapped $^1\text{H}/^{15}\text{N}$ frequencies are even more severe because they lead to an additional set of C^α_i , C^α_{i-1} , C^β_i and C^β_{i-1} frequencies. It is difficult to assign the signals to the two spin systems unambiguously and the chances of overlap are more likely. Usually all possible combinations of shifts are used and amino acid typing gets very ambiguous.

Overlap of C^α and C^β frequencies is an intrinsic problem of the CBCA(CO)NH and CBCANH spectra, whereas degeneracy in ^1H and ^{15}N chemical shifts or in frequencies of sequential and non-sequential signals can potentially occur in most triple resonance experiments.

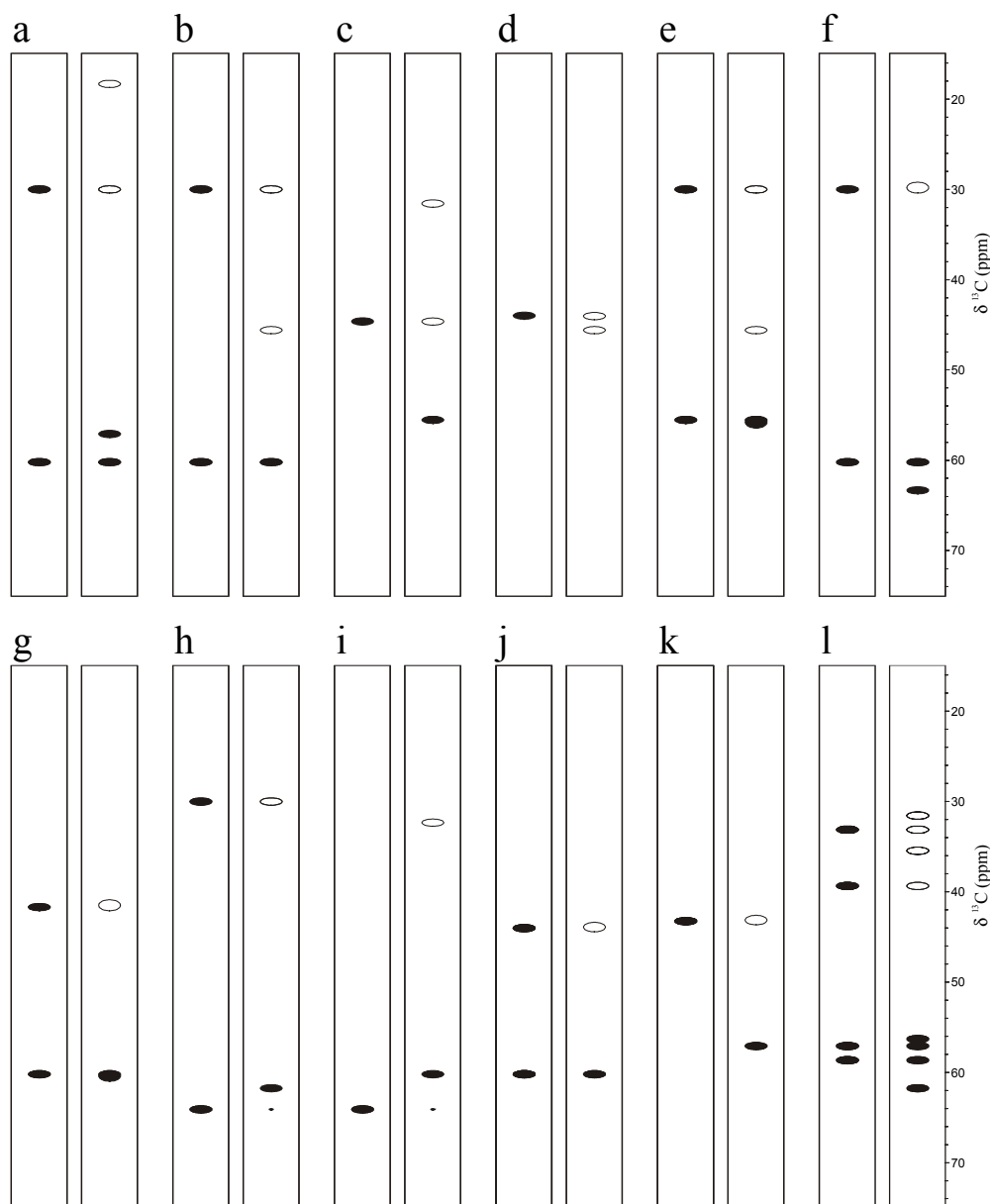


Fig. 5.3: Typical signal patterns occurring in strips of CBCA(CO)NH (left) and CBCANH (right) spectra. (a) most common pattern of two non-Gly residues with two signals in the CBCA(CO)NH originating from C^{α}_{i-1} and C^{β}_{i-1} and four signals in the CBCANH originating from C^{α}_{i-1} , C^{α}_i , C^{β}_{i-1} and C^{β}_i . The two C^{β} signals are negative. (b) common pattern if residue i is a Gly. The C^{α}_i signal of Gly is negative. (c) typical pattern if the preceding neighbor ($i-1$) is a Gly. (d) pattern for two neighbored Gly residues (i and $i-1$). (e) pattern of two non-Gly residues with overlapped frequencies of C^{α}_{i-1} and C^{α}_i . Note that this pattern is identical to pattern (b). (f) pattern of two non-Gly residues with overlapped frequencies of C^{β}_{i-1} and C^{β}_i . (g) pattern of two non-Gly residues with overlapped frequencies of C^{α}_{i-1} and C^{α}_i as well as of C^{β}_{i-1} and C^{β}_i . (h) pattern of a Ser residue i and a non-Gly residues ($i-1$) with a overlapped frequencies of C^{α}_{i-1} and C^{β}_i . The positive C^{α}_{i-1} signal and the negative C^{β}_i signal are canceling each other in the CBCANH. (i) pattern of a non-Gly residues i and a Ser residue ($i-1$) with a overlapped frequencies of C^{α}_{i-1} and C^{β}_{i-1} . In the CBCANH experiment the positive C^{α}_{i-1} signal and the negative C^{β}_{i-1} signal are canceling each other. (j) pattern of a Gly residues i and a non-Gly residues ($i-1$) with overlapped frequencies of C^{α}_i and C^{β}_{i-1} . This pattern is similar to pattern (g). (k) pattern of a non-Gly residues i and a Gly residues ($i-1$) with overlapped frequencies of C^{α}_{i-1} and C^{β}_i . (l) pattern with the doubled number of signals resulting from $^1\text{H}/^{15}\text{N}$ overlapped frequencies. The assignment of the signals to residue i and j as well as ($i-1$) and ($j-1$) is ambiguous. The probability of overlapped frequencies in the ^{13}C dimension is high and a variety of scenarios is possible.

In addition, signals in the ^{13}C dimension are sometimes missing and incomplete pattern occur which further complicate unambiguous assignment. Proton exchange or unfavorable dynamics may lead to the absence of several generic spin systems. Usually these signals are absent in the ^{15}N -HSQC too and can not be assigned.

Nevertheless, many of the automated assignment programs can handle several of the difficult cases mentioned above by matching the problematic spin systems to previously unambiguously assigned generic spin systems during the generation of segments or by using additional data. For example data from the HNCA and HN(CO)CA spectra can be used to assign overlapped frequencies of C^α and C^β . However, overlap and missing peaks are a problem for automated assignment procedures and with an increasing number of such critical spin systems they will fail. Especially an increasing number of overlapped $^1\text{H}/^{15}\text{N}$ chemical shift pairs and missing signals will result in failure at a certain point.

Another problematic step is the amino acid typing. As already outlined in chapter 2.3 the typical typing method involves statistical analysis of chemical shifts, specifically the C^α and C^β resonance shifts (Grzesiek and Bax, 1993). Bayesian statistical methods combined with BMRB database (Seavey et al., 1991) statistics are used in AutoAssign and the program of Lukin et al. to classify spin-system types. TATAPRO (Atreya et al., 2000) classifies the 20 amino acids into eight different categories, with each category having a characteristic C^α and C^β chemical shift range. TATAPRO is based on a deterministic rather than probabilistic approach.

Another method, the phase labeling of C-H and C-C spin system topologies in common 3D triple resonance experiments, also gives information about the amino acid type by different sign patterns (Tashiro et al., 1995; Feng et al., 1996; Rios et al., 1996; Gschwind et al., 1998). Modified 3D triple resonance experiments deliver patterns of positive and negative signals characteristic of groups of amino acid types. These data can be implemented in the program AutoAssign. However, only groups of amino acid types are identified, additional 3D experiments have to be recorded and the complex 3D spectra with all the difficulties described in Fig. 5.3 are not simplified.

During manual assignment, additional data are used for the identification of several amino acid types namely the chemical shifts of aliphatic γ , δ and ϵ carbons obtained from experiments like the (H)C(CO)NH-TOCSY (Grzesiek et al., 1993; Logan et al., 1993) and the identification of cross peaks between aliphatic and side-chain amide resonances of Asn and Gln in common 3D triple resonance experiments (Montelione et al., 1992; Lyons et al., 1993; Wittekind and Mueller, 1993). However, these data are usually not used for automated backbone assignment.

Amino acid type selective experiments like those described in chapter 4 can greatly facilitate the amino acid typing. Only $^1\text{H}/^{15}\text{N}$ frequency pairs of certain amino acid types (in

position i or $i-1$) occur in those 2D spectra. These data allow the amino acid typing of each $^1\text{H}/^{15}\text{N}$ frequency pair. The advantage over the other methods is the unambiguity of the amino acid typing.

Poly-Pro sequences can not be assigned with the state-of-the-art automated assignment procedures and an unambiguous assignment has to rely on manual assignment and on experiments especially designed for such cases. However, the protein sequence is known before the analysis and poly-Pro sequences are very rare in small globular protein domains.

5.4 Amino acid type identification of sequential pairs using the amino acid type-selective experiments

Amino acid type-selective ^{15}N -HSQCs do not only offer the amino acid type identification of a certain residue i but the amino acid identification of two sequential residues i and $i-1$ and thus a sequential pair of amino acid types. With the 20 common amino acids 400 different sequential pairs can occur in proteins. However, only a fraction of these pairs will be present in a particular protein.

The protein sequence has to be analyzed to see how frequently each sequential pair is present. The best way of doing this is a matrix with all 400 potential pairs with their number of occurrence as shown in Fig. 5.4. Such a matrix can be generated automatically by the program SELMA developed by R. Winter (Oschkinat group, unpublished results). Ile and Val as well as Phe, Tyr and His are grouped together (because they can not be distinguished using the amino acid type-selective experiments) which leads to 289 different types of pairs.

In order to determine the type of sequential pairs of each $^1\text{H}/^{15}\text{N}$ signal, several amino acid type-selective HSQCs have to be compared. For example the $^1\text{H}/^{15}\text{N}$ frequency pair of the Asn in a SN pair is present in a S-($i+1$)- as well as in the N-($i,i+1$)-HSQC (Fig. 5.5). The NH signal of Thr in a GT pair is present in the G-($i+1$)- and TA-($i,i+1$)-HSQC but not in the A-($i,i+1$)-HSQC. All experiments which have to be compared for each of the 289 potential pairs are summarized in table 5.1. Note that only the most selective experiments are included in the table, for example the N and D experiments instead of the DNG experiments.

Pro residues can not be detected on NH. However, an experiment for the identification of the amino acid type of the residues preceding Pro was introduced in chapter 4.4.15, i. e. residues i with Pro in the ($i+1$) position can be identified. The NH signal of Ser in a SP pair is present in the S-($i,i+1$)- and P-($i-1$)-HSQC (Fig. 5.6). Residues preceding Pro can follow any other amino acid type. In this particular case, not only sequential pairs, but triplets can be identified (X-Y-Pro).

SEIF~~T~~LLVEK VWNFDDLIMA INSKISNTHN NNISPITKIK YQDEDGDFVV
 LGSDE~~D~~WNVA KEMLAENNEK FLNIRLY

	A	D	E	G	K	L	N	Q	R	S	T	W	P	I/V	F/Y/H	C	M	total
A			1		1									1				3
D		1	2	1		1						1			1			7
E		2			2		1							1			1	7
G		1								1								2
K			1											3	2			6
L	1			1		1	1							2	1			7
N			1				3			1	1			3	1			10
Q		1																1
R						1												1
S		1	1		1		1						1					5
T					1	1									1			3
W							2											2
P														1				1
I/V	1		1		1	1	1		1	2	1	1		1	1		1	13
F/Y/H		1				1	1	1			1			1				6
C																		-
M	1					1												2
total	3	7	7	2	6	7	10	1	1	4	3	2	1	13	7		2	77

D54

SEIF~~T~~LLVEK VWNFDDLIMA INSKISNTHN NNISPITKIK YQDEDGDFVV
 LGSDE~~D~~WNVA KEMLAENNEK FLNIRLY

Fig. 5.4: Sequence analysis of the protein OPR (Ponting, 1996), on top the protein sequence is shown. The columns in the matrix represent the amino acid type of residue *i* (except Pro) and the rows the type of residue (*i*-1). Since Pro residues can not be detected on NH the "P" in the column represents Pro in (*i*+1) position. Within the matrix the number of occurrence of each pair is given. D54 the only Asp following a Ser is highlighted. The sequence contains 45 unique pairs out of 77 residues.

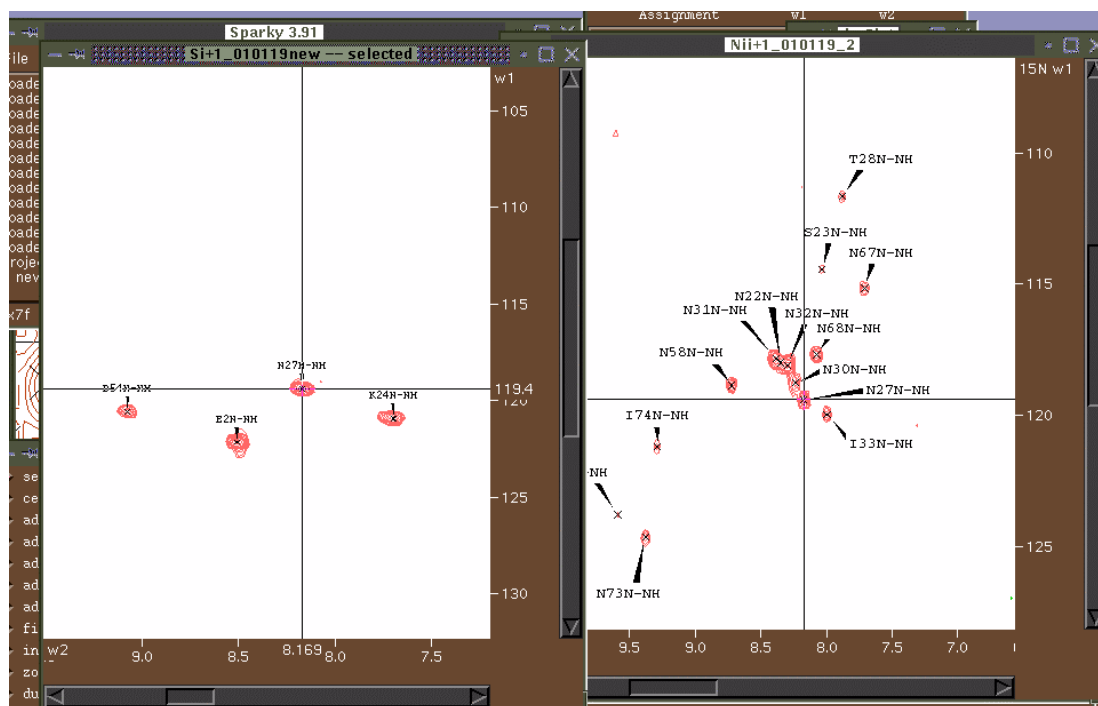


Fig. 5.5: An Asn ¹H/¹⁵N frequency pair with a preceding Ser is present in the S-(*i*+1)- as well as in the N-(*i*,*i*+1)-HSQC.

	A	D	E	G	K	L	N	Q	R	S
A		A(i+1) D(ii+1)	A(i+1) E(ii+1)	A(i+1) G(ii+1)	A(i+1) K(ii+1)	A(i+1), LA(ii+1) Intens.	A(i+1) N(ii+1)	A(i+1) Q(ii+1)	A(i+1) R(ii+1)	A(i+1) S(ii+1)
D	D(i+1) A(ii+1)		D(i+1) E(ii+1)	D(i+1) G(ii+1)	D(i+1) K(ii+1)	D(i+1), LA(ii+1) not A(ii+1), VIA(ii+1)	D(i+1) N(ii+1)	D(i+1) Q(ii+1)	D(i+1) R(ii+1)	D(i+1) S(ii+1)
E	ED(i+1) A(ii+1) not D(i+1)	ED(i+1) D(ii+1) not D(i+1)		ED(i+1) G(ii+1) not D(i+1)	ED(i+1) K(ii+1) not D(i+1)	ED(i+1), LA(ii+1) not D(i+1), A(ii+1), VIA(ii+1)	ED(i+1) N(ii+1)	ED(i+1) Q(ii+1) not D(i+1)	ED(i+1) R(ii+1) not D(i+1)	ED(i+1) S(ii+1) not D(i+1)
G	G(i+1) A(ii+1)	G(i+1) D(ii+1)	G(i+1) E(ii+1)		G(i+1) K(ii+1)	G(i+1), LA(ii+1) not A(ii+1), VIA(ii+1)	G(i+1) N(ii+1)	G(i+1) Q(ii+1)	G(i+1) R(ii+1)	G(i+1) S(ii+1)
K	KR(i+1) A(ii+1) not R(i+1)	KR(i+1) D(ii+1) not R(i+1)	KR(i+1) E(ii+1) not R(i+1)	KR(i+1) G(ii+1) not R(i+1)		KR(i+1), LA(ii+1) not R(i+1), A(ii+1), VIA(ii+1)	KR(i+1) N(ii+1) not R(i+1)	KR(i+1) Q(ii+1) not R(i+1)	KR(i+1) R(ii+1) not R(i+1)	KR(i+1) S(ii+1) not R(i+1)
L	LA(i+1) A(ii+1) not A(i+1), VIA(i+1)	LA(i+1) D(ii+1) not A(i+1), VIA(i+1)	LA(i+1) E(ii+1) not A(i+1), VIA(i+1)	LA(i+1) G(ii+1) not A(i+1), VIA(i+1)	LA(i+1) K(ii+1) not A(i+1), VIA(i+1)		LA(i+1) N(ii+1) not A(i+1), VIA(i+1)	LA(i+1) Q(ii+1) not A(i+1), VIA(i+1)	LA(i+1) R(ii+1) not A(i+1), VIA(i+1)	LA(i+1) S(ii+1) not A(i+1), VIA(i+1)
N	N(i+1) A(ii+1)	N(i+1) D(ii+1)	N(i+1) E(ii+1)	N(i+1) G(ii+1)	N(i+1) K(ii+1)	N(i+1), LA(ii+1) not A(ii+1), VIA(ii+1)		N(i+1) Q(ii+1)	N(i+1) R(ii+1)	N(i+1) S(ii+1)
Q	QN(i+1) A(ii+1) not N(i+1)	QN(i+1) D(ii+1) not N(i+1)	QN(i+1) E(ii+1) not N(i+1)	QN(i+1) G(ii+1) not N(i+1)	QN(i+1) K(ii+1) not N(i+1)	QN(i+1), LA(ii+1) not N(i+1), A(ii+1), VIA(ii+1)	QN(i+1) N(ii+1) not N(i+1)		QN(i+1) R(ii+1) not N(i+1)	QN(i+1) S(ii+1) not N(i+1)
R	R(i+1) A(ii+1)	R(i+1) D(ii+1)	R(i+1) E(ii+1)	R(i+1) G(ii+1)	R(i+1) K(ii+1)	R(i+1), LA(ii+1) not A(ii+1), VIA(ii+1)	R(i+1) N(ii+1)	R(i+1) Q(ii+1)		R(i+1) S(ii+1)
S	S(i+1) A(ii+1)	S(i+1) D(ii+1)	S(i+1) E(ii+1)	S(i+1) G(ii+1)	S(i+1) K(ii+1)	S(i+1), LA(ii+1) not A(ii+1), VIA(ii+1)	S(i+1) N(ii+1)	S(i+1) Q(ii+1)	S(i+1) R(ii+1)	
T	TA(i+1) A(ii+1), not A(i+1)	TA(i+1) D(ii+1), not A(i+1)	TA(i+1) E(ii+1), not A(i+1)	TA(i+1) G(ii+1), not A(i+1)	TA(i+1) K(ii+1), not A(i+1)	TA(i+1), LA(ii+1), not A(i+1), A(ii+1), VIA(ii+1)	TA(i+1) N(ii+1), not A(i+1)	TA(i+1) Q(ii+1), not A(i+1)	TA(i+1) R(ii+1), not A(i+1)	TA(i+1) S(ii+1), not A(i+1)
W	W(i+1) A(ii+1)	W(i+1) D(ii+1)	W(i+1) E(ii+1)	W(i+1) G(ii+1)	W(i+1) K(ii+1)	W(i+1), LA(ii+1) not A(ii+1), VIA(ii+1)	W(i+1) N(ii+1)	W(i+1) Q(ii+1)	W(i+1) R(ii+1)	W(i+1) S(ii+1)
P	P(i+1) A(ii+1)	P(i+1) D(ii+1)	P(i+1) E(ii+1)	P(i+1) G(ii+1)	P(i+1) K(ii+1)	P(i+1), LA(ii+1) not A(ii+1), VIA(ii+1)	P(i+1) N(ii+1)	P(i+1) Q(ii+1)	P(i+1) R(ii+1)	P(i+1) S(ii+1)
I/V	VIA(i+1) A(ii+1), not A(i+1)	VIA(i+1) D(ii+1), not A(i+1)	VIA(i+1) E(ii+1), not A(i+1)	VIA(i+1) G(ii+1), not A(i+1)	VIA(i+1) K(ii+1), not A(i+1)	VIA(i+1), LA(ii+1), not A(i+1), A(ii+1), VIA(ii+1)	VIA(i+1) N(ii+1), not A(i+1)	VIA(i+1) Q(ii+1), not A(i+1)	VIA(i+1) R(ii+1), not A(i+1)	VIA(i+1) S(ii+1), not A(i+1)
F/Y/H	FYH(i+1) A(ii+1)	FYH(i+1) D(ii+1)	FYH(i+1) E(ii+1)	FYH(i+1) G(ii+1)	FYH(i+1) K(ii+1)	FYH(i+1), LA(ii+1) not A(ii+1), VIA(ii+1)	FYH(i+1) N(ii+1)	FYH(i+1) Q(ii+1)	FYH(i+1) R(ii+1)	FYH(i+1) S(ii+1)
C	CS(i+1) A(ii+1) not S(i+1)	CS(i+1) D(ii+1) not S(i+1)	CS(i+1) E(ii+1) not S(i+1)	CS(i+1) G(ii+1) not S(i+1)	CS(i+1) K(ii+1) not S(i+1)	CS(i+1), LA(ii+1) not S(i+1), A(ii+1), VIA(ii+1)	CS(i+1) N(ii+1) not S(i+1)	CS(i+1) Q(ii+1) not S(i+1)	CS(i+1) R(ii+1) not S(i+1)	CS(i+1) S(ii+1) not S(i+1)
M										

Table 5.1: The matrix shows which experiments have to be compared in order to identify a certain type of sequential pairs in the protein sequence. The columns in the matrix represent the amino acid type of residue i (except Pro) and the rows the type of residue $(i-1)$. Since Pro residues can not be detected on NH the "P" in the column represents Pro in $(i+1)$ position. As an example, the identification of a ($^1\text{H}/^{15}\text{N}$) frequency pair of a Asp which succeeds Glu a signal has to be present in the ED- $(i+1)$ - as well as in the D- $(i,i+1)$ -HSQC, but not in the D- $(i+1)$ -HSQC. Note that the suppression of undesired signals in the CS and P experiments is not ideal.

T	W	P(i-1)	I/V	F/Y/H	C	M
A(i+1), TA(ii+1) Intens.	A(i+1) W(ii+1)	A(i,i+1) P(i-1)	A(i+1), VIA(ii+1) Intens.	A(i+1) FYH(ii+1)	A(i+1), CS(ii+1) not S(i,i+1)	
D(i+1), TA(ii+1) not A(ii+1)	D(i+1) W(ii+1)	D(i,i+1) P(i-1)	D(i+1), VIA(ii+1) not A(ii+1)	D(i+1) FYH(ii+1)	D(i+1), CS(ii+1) not S(i,i+1)	
ED(i+1), TA(ii+1) not D(i+1), A(ii+1)	ED(i+1) W(ii+1) not D(i+1)	E(i,i+1) P(i-1)	ED(i+1), VIA(ii+1) not D(i+1), A(ii+1)	ED(i+1) FYH(ii+1) not D(i+1)	ED(i+1), CS(ii+1) not D(i+1), S(i,i+1)	
G(i+1), TA(ii+1) not A(ii+1)	G(i+1) W(ii+1)	G(i,i+1) P(i-1)	G(i+1), VIA(ii+1) not A(ii+1)	G(i+1) FYH(ii+1)	G(i+1), CS(ii+1) not S(i,i+1)	
KR(i+1), TA(ii+1) not R(i+1), A(ii+1)	KR(i+1) W(ii+1) not R(i+1)	K(i,i+1) P(i-1)	KR(i+1), VIA(ii+1) not R(i+1), A(ii+1)	KR(i+1) FYH(ii+1) not R(i+1)	KR(i+1), CS(ii+1) not R(i+1), S(i,i+1)	
LA(i+1), TA(ii+1) not A(i+1), A(ii+1), VIA(i+1)	LA(i+1) W(ii+1) not A(i+1), VIA(i+1)	LA(i,i+1) P(i-1), not A(ii+1)	LA(i+1), VIA(ii+1) not A(i+1), A(ii+1), VIA(i+1)	LA(i+1) FYH(ii+1) not A(i+1), VIA(i+1)	LA(i+1), CS(ii+1) not A(i+1), VIA(i+1), S(i,i+1)	
N(i+1), TA(ii+1) not A(ii+1)	N(i+1) W(ii+1)	N(i,i+1) P(i-1)	N(i+1), VIA(ii+1) not A(ii+1)	N(i+1) FYH(ii+1)	N(i+1), CS(ii+1) not S(i,i+1)	
QN(i+1), TA(ii+1) not N(i+1), A(ii+1)	QN(i+1) W(ii+1) not N(i+1)	QN(i,i+1) P(i-1) not N(i+1)	QN(i+1), VIA(ii+1) not N(i+1), A(ii+1)	QN(i+1) FYH(ii+1) not N(i+1)	QN(i+1), CS(ii+1) not N(i+1), S(i,i+1)	
R(i+1), TA(ii+1) not A(ii+1)	R(i+1) W(ii+1)	R(i,i+1) P(i-1)	R(i+1), VIA(ii+1) not A(ii+1)	R(i+1) FYH(ii+1)	R(i+1), CS(ii+1) not S(i,i+1)	
S(i+1), TA(ii+1) not A(ii+1)	S(i+1) W(ii+1)	S(i,i+1) P(i-1)	S(i+1), VIA(ii+1) not A(ii+1)	S(i+1) FYH(ii+1)	S(i+1), CS(ii+1) not S(i,i+1)	
	TA(i+1) W(ii+1), not A(i+1)	TA(i,i+1) P(i-1), not A(ii+1)	TA(i+1), VIA(ii+1) not A(i+1), A(ii+1)	TA(i+1) FYH(ii+1) not A(i+1)	TA(i+1), CS(ii+1) not A(i+1), S(i,i+1)	
W(i+1), TA(ii+1) not A(ii+1)		W(i,i+1) P(i-1)	W(i+1), VIA(ii+1) not A(ii+1)	W(i+1) FYH(ii+1)	W(i+1), CS(ii+1) not S(i,i+1)	
P(i+1), TA(ii+1) not A(ii+1)	P(i+1) W(ii+1)		P(i+1), VIA(ii+1) not A(ii+1)	P(i+1) FYH(ii+1)	P(i+1), CS(ii+1) not S(i,i+1)	
VIA(i+1), TA(ii+1) not A(i+1), A(ii+1)	VIA(i+1) W(ii+1), not A(i+1)	VIA(i,i+1) P(i-1), not A(ii+1)		VIA(i+1) FYH(ii+1) not A(i+1)	VIA(i+1), CS(ii+1) not A(i+1), S(i,i+1)	
FYH(i+1), TA(ii+1) not A(ii+1)	FYH(i+1) W(ii+1)	FYH(i,i+1) P(i-1)	FYH(i+1), VIA(ii+1) not A(ii+1)		FYH(i+1), CS(ii+1) not S(i,i+1)	
CS(i+1), TA(ii+1) not S(i+1), A(ii+1)	CS(i+1) W(ii+1) not S(i+1)	CS(i,i+1) P(i-1) not S(i+1)	CS(i+1), VIA(ii+1) not S(i+1), A(ii+1)	CS(i+1) FYH(ii+1) not S(i+1)		

Some types of sequential pairs (dipeptides) can not be identified unambiguously: i.e. all dipeptides with the same amino acid type like AA, pairs of AL, AT, AI, AV, SC and all dipeptides including Met. For example the $^1\text{H}/^{15}\text{N}$ frequency pair of Thr in AT is present in the A-(i+1)-HSQC and TA-(i,i+1)-HSQC. In order to identify unambiguously Thr as residue i the presence of Ala in position i has to be excluded. This is usually done by comparing the TA-(i,i+1)-HSQC with the A-(i,i+1)-HSQC and if no signal is present in the latter the residue is a Thr. However, since it is a sequential signal of Ala (Ala in position i-1) a small signal usually appears in the A-(i,i+1)-HSQC. The absence of an Ala in position i can not be excluded since sequential signals (resulting from Ala in position i-1) and signals of Ala (i) can not be distinguished in a straightforward manner. Usually the intensities of both types of signals are different, however, a variety of factors may influence the signal intensities and hence these are not reliable for the use in automated assignment procedures.

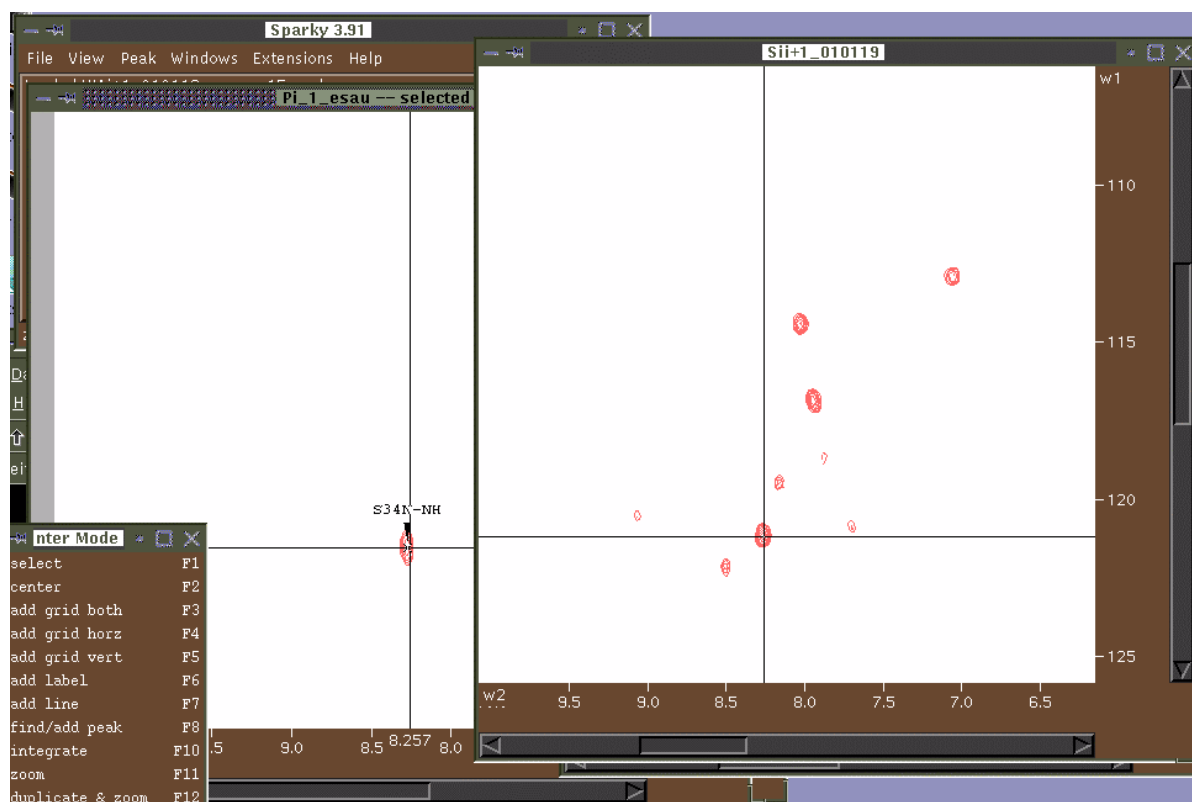


Fig. 5.6: A Ser $^1\text{H}/^{15}\text{N}$ frequency pair with a succeeding Pro is present in the P-(i-1)- as well as in the S-(i,i+1)-HSQC.

If a particular dipeptide (sequential pair) has been identified and occurs only once in the protein sequence (number of occurrence in Fig. 5.4 is "1"), the $^1\text{H}/^{15}\text{N}$ frequency pair can already be assigned unambiguously. The signal forms an unambiguous reference point and acts as an anchor point for further assignments. All unique pairs in the sequence which can be assigned unambiguously (white area in Fig. 5.4) are, in principle, potential reference points. Their number depends on the sequence. Usually a large part of the residues can be

unambiguously assigned using solely 2D amino acid type-selective ^{15}N -HSQC experiments. If the number of appearance is “2” the pair can be assigned to two different sites and will be an ambiguous reference point.

^1H - ^{15}N chemical shift degeneracies can hinder this strategy but this is not necessarily the case. The ^1H - ^{15}N frequencies of an Ala in a SA dipeptide, for example, can nevertheless be identified if a Trp in a DW dipeptide has the same frequencies because the two spectra of interest, the S-(i+1)- and the A-(i,i+1)-HSQC do not contain the Trp signal with the same ^1H - ^{15}N frequencies.

In other cases false results might be generated. This is illustrated with the following example. In a search for a Thr of a ST dipeptide a false positive result (additional to the correct one) is generated, if for example a Thr signal in a NT dipeptide has identical ^1H - ^{15}N chemical shifts as a Trp in a SW dipeptide. This is because the spectrum of the S-(i+1)-HSQC as well as the spectrum of the TA-(i,i+1)-HSQC contains a signal at these particular ^1H - ^{15}N frequencies. However, this is only a problem if this particular peptide (in this example ST) occurs in the sequence and if it is searched for it (here for a Thr signal in the S-(i+1)- and TA-(i,i+1)-HSQC spectra). Nevertheless, if two results are obtained for a unique dipeptide sequence, the result will be considered ambiguous and will not contribute to the unambiguous reference points.

Only a few unambiguous reference points, however, will already enhance the performance of automated assignment procedures drastically.

5.5 Potential strategies for implementing the amino acid type-selective experiments into automated assignment procedures

Several strategy to implement the amino acid type-selective experiments into automated assignment procedures are feasible.

The determination of unambiguous reference points gives a novel independent information and offers a parallel (independent) route additional to the flow chart in Fig. 5.2, leading directly to reliable $^1\text{H}/^{15}\text{N}$ frequency assignments of a large part of the sequence (Fig. 5.7). In the conventional procedure generic spin systems are generated as described in chapter 5.2. The residues which have been already assigned and the remaining generic spin systems are linked to form segments. The segments are then mapped onto the sequence with the help of the unambiguous reference points. The statistical analysis of the C^α and C^β frequencies can be used additionally and a mapping onto the sequence should independently confirm the results. For this strategy the complete set of amino acid type-selective ^1H - ^{15}N correlations is desirable but not necessary. An incomplete set just delivers less unambiguous reference points.

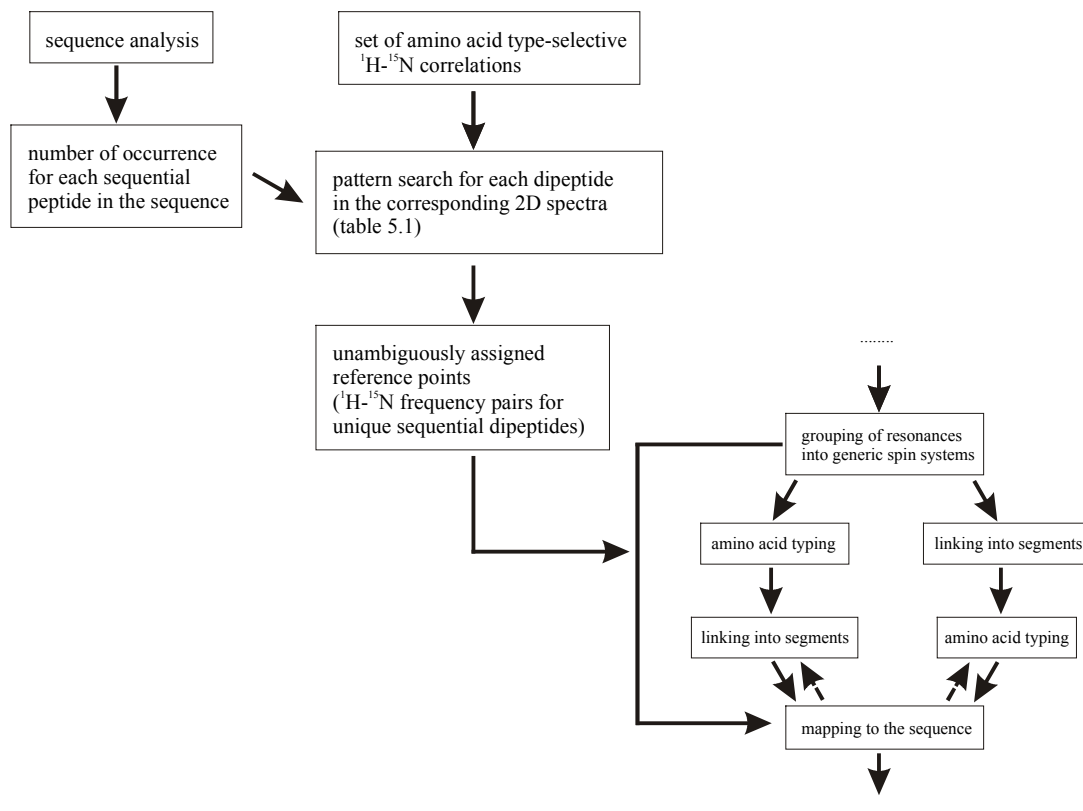


Fig. 5.7: First strategy of implementing the amino acid type-selective experiments into an automated assignment procedure. The generation of unambiguous reference points is an additional pathway for the assignment of ^1H - ^{15}N frequency pairs which is independent from the conventional pathway in Fig. 5.2.

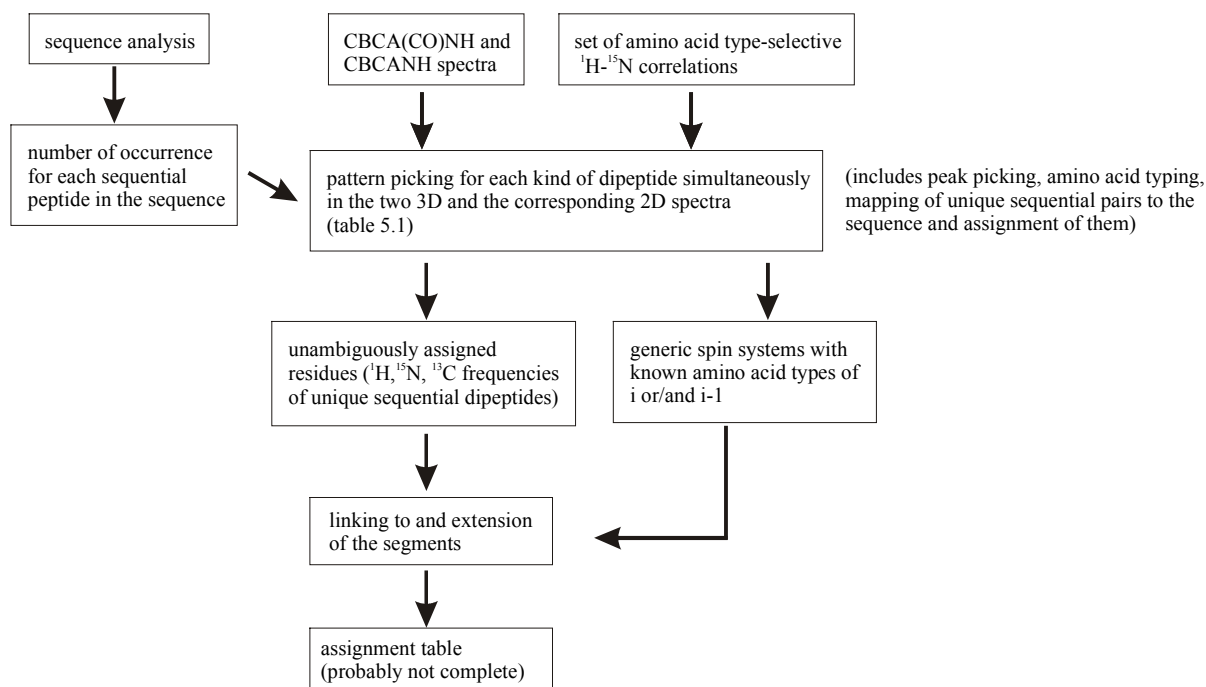


Fig. 5.8: Second strategy of implementing the amino acid type-selective experiments into an automated assignment procedure. A more or less complete assignment table is generated independently from the conventional route shown in Fig. 5.2. Missing signals in the selective 2D spectra lead to assignment gaps. Agreement between the results obtained from this strategy and results of another independent route gives the assignment a high reliability.

In an alternative strategy amino acid type-selective ^1H - ^{15}N correlations can be used together with 3D triple resonance experiments for restricted peak picking or for pattern search algorithms (like Chatch23) introducing the amino acid type information already before peaks are grouped (Fig.5.8).

The grouping of frequencies into generic spin systems is then facilitated because the assignment program can search in characteristic chemical shift areas for resonances of certain amino acid types. Ambiguities like those discussed in chapter 5.3 are thus reduced. The amino acid types of the generic spin systems (including residue i and $i-1$) are already unambiguously identified and the statistical analysis of the C^α and C^β frequencies can be omitted or used only to confirm the correct typing. Generic spin systems of dipeptides which are unique in the sequence can already be unambiguously assigned. The unambiguous reference points are linked with remaining unassigned generic spin systems. The assigned segments are extended until no unassigned generic spin systems remain. For this strategy, the complete set of amino acid type-selective ^1H - ^{15}N correlations is necessary. However, missing signals in the amino acid type-selective 2D experiments lead to missing generic spin systems. Therefore this strategy may lead to incomplete assignment tables. However, in conjunction with another independent route especially for generating generic spin systems this strategy is the most promising one.

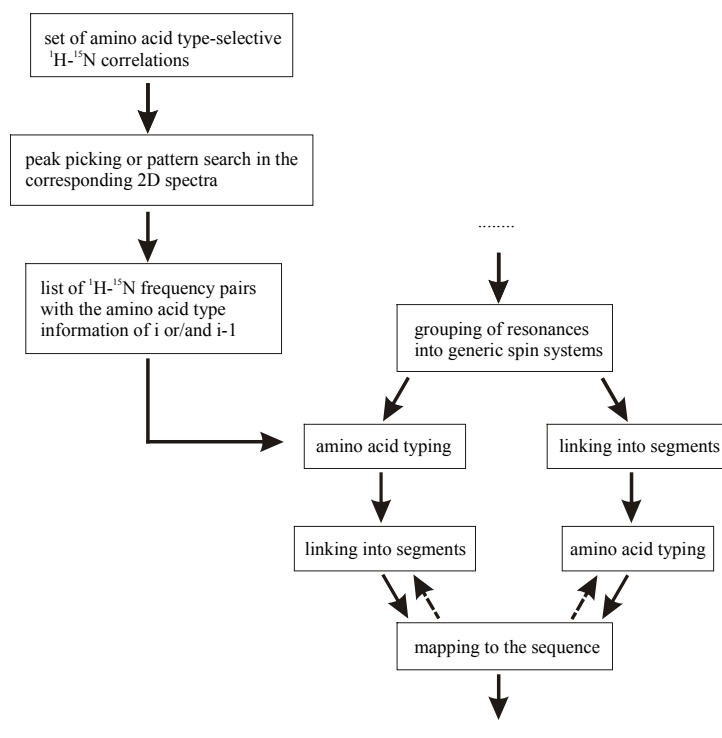


Fig. 5.9: Third strategy of implementing the amino acid type-selective experiments into an automated assignment procedure. The residue type information of the selective ^1H - ^{15}N correlations is used for the amino acid typing in the conventional scheme shown in Fig. 5.2.

A third potential strategy of implementing the amino acid type-selective ^1H - ^{15}N correlations into the automated assignment strategy is their use just for the amino acid typing (Fig. 5.9). Only a subset of experiments is necessary, e. g. only the more sensitive (i+1)-HSQC experiments which deliver the amino acid type of residue (i-1). The overall strategy described in chapter 5.2 remains. The statistical analysis of the C^α and C^β frequencies is either used for the non-sequential frequencies, for not yet identified residues or just used to confirm the correct typing.

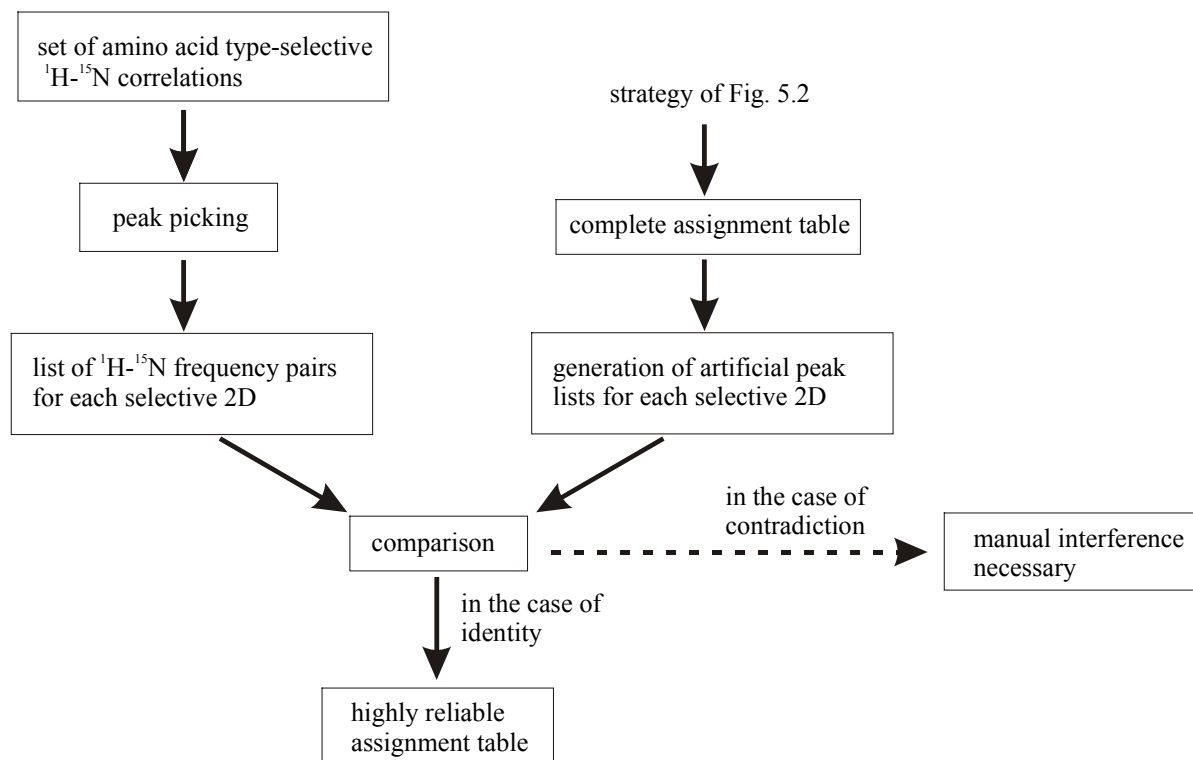


Fig. 5.10: Potential strategy for the evaluation of assignment results using the amino acid type-selective ^1H - ^{15}N correlations as additional information independent of the assignment pathway.

Since the information of the amino acid type-selective ^1H - ^{15}N correlations is achieved independently it can be used for the confirmation of assignments determined with conventional triple-resonance experiments. This can be accomplished for example by generating artificial peak lists for each selective ^1H - ^{15}N correlation from the final assignment table and by comparing those to the real peak lists (Fig. 5.10). Alternatively, pattern search algorithms deliver the information if an expected peak is there or not.

5.6 Conclusion and outlook

A set of 32 selective ^1H - ^{15}N correlations was proposed to give sufficient information about the amino acid type of almost every residue and of its preceding neighbor to be useful in automated assignment procedures. With spectra of appropriate quality, 14 amino acid

types can be identified unambiguously: Gly, Ala, Thr, Leu, Asn, Gln, Ser, Asp, Glu, Arg, Lys, Pro, Trp and Cys. Five amino acid types can be assigned into two groups: Ile/Val and Phe/Tyr/His. Only Met residues cannot be identified with the selective 2D spectra. However, where two successive Met residues (...MM...) occur in the sequence, no information about the amino acid type can be obtained. In any other case at least the amino acid type of either the residue itself (i) or of the preceding neighbor (i-1) can be obtained.

If the amino acid type of residue (i) and (i-1) can be determined an assignment can already be proposed. If the particular amino acid pair (i-1)/(i) is unique in the sequence it can be immediately assigned unambiguously. The protein sequence will determine how many such unambiguous reference points can be found.

Four different strategies to implement spectra of the amino acid type-selective ^1H - ^{15}N correlations into automated assignment procedures were proposed. They allow, together with the approaches discussed in chapter 5.2, at least two independent pathways to assign the resonances and thus increase the reliability of the output. The application of the proposed strategies has still to be tested extensively in conjunction with the development of a complete software package for automated resonance assignment.

References

- H. S. Atreya, S. C. Sahu, K. V. R. Chary and G. Govil, *J. Biomol. NMR* **17**, 125-136 (2000).
- D. Croft, J. Kemmink, K.-P. Neidig and H. Oschkinat, *J. Biomol. NMR* **10**, 207-219 (1997).
- W. Feng, C. B. Rios and G. T. Montelione, *J. Biomol. NMR* **8**, 98-104 (1996).
- D. S. Garrett, R. Powers, A. M. Gronenborn and G. M. Clore, *J. Magn. Reson.* **95**, 214-220 (1991).
- T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco (1999);
<http://www.cgl.ucsf.edu/home/sparky/>
- S. Grzesiek, J. Anglister and A. Bax, *J. Magn. Reson. B* **101**, 114-119 (1993).
- S. Grzesiek and A. Bax, *J. Biomol. NMR* **3**, 185-204 (1993).
- R. M. Gschwind, G. Gemmecker and H. Kessler, *J. Biomol. NMR* **11**, 191-199 (1998).
- M. Leutner, R. M. Gschwind, J. Liermann, C. Schwarz, G. Gemmecker and H. Kessler, *J. Biomol. NMR* **11**, 31-43 (1998).
- T. M. Logan, E. T. Olejniczak, R. X. Xu and S. W. Fesik, *J. Biomol. NMR* **3**, 225-231 (1993).
- J. A. Lukin, A. P. Gove, S. N. Talukdar and C. Ho, *J. Biomol. NMR* **9**, 151-166 (1997).
- B. A. Lyons, M. Tashiro, L. Cedergren, B. Nilsson and G. T. Montelione, *Biochemistry* **32**, 7839-7845, (1993).
- G. T. Montelione, B. A. Lyons, S. D. Emerson and M. Tashiro, *J. Am. Chem. Soc.* **114**, 10974-10975 (1992).
- H. N. Moseley and G. T. Montelione, *Curr. Opin. Struct. Biol.* **9**, 635-642 (1999) and references cited therein.
- H. N. B. Moseley, D. Monleon and G. T. Montelione, *Meth. Enzymol.* **339**, 91-108 (2001).
- C. P. Ponting, *Protein Sci.* **5**, 2353-2357 (1996).
- C. B. Rios, W. Feng, M. Tashiro, Z. Shang and G. T. Montelione, *J. Biomol. NMR* **8**, 345-350 (1996).
- B. R. Seavey, E. A. Farr, W. M. Westler and J. Markley, *J. Biomol. NMR* **1**, 217-236 (1991);
<http://www.bmrb.wisc.edu>
- M. Tashiro, C. B. Rios and G. T. Montelione, *J. Biomol. NMR* **6**, 211-216 (1995).

M. Wittekind and L. Mueller, *J. Magn. Reson. B* **101**, 201-205 (1993).

D. E. Zimmerman, C. A. Kulikowski, Y. Huang, W. Feng, M. Tashiro, S. Shimotakahara, C. Chien, R. Powers and G. T. Montelione, *J. Mol. Biol.* **269**, 592-610 (1997).

6.1 NMR of membrane proteins

Membrane proteins have a variety of functions in living cells, such as intracellular and transmembranal signal transduction, transport of ions and small molecules, electron transport and catalysis. Therefore it is not surprising that approximately 30% of all protein sequences decoded in eukaryotic genomes are membrane proteins (Wallin and van Heijne, 1998). In contrast to the vast amount of sequence and genetic information available for membrane proteins, little is known about their structures and mechanisms at the molecular level. Presently, fewer than 30 independent integral membrane protein structures have been solved which stands in contrast to more than 15,000 structures of soluble proteins determined by X-ray crystallography and NMR spectroscopy.

The few three-dimensional structures of membrane proteins known today were solved by X-ray crystallography, electron microscopy and very recently NMR spectroscopy. However, although X-ray crystallography has been relatively successful, many obstacles have to be overcome. It is very difficult to crystallize membrane proteins from detergent solutions and the search space for appropriate crystallization conditions is much larger than for soluble proteins. Electron microscopy of two-dimensional crystals of membrane proteins has progressed dramatically and some structures have already been solved to atomic resolution (Kuhlbrandt and Williams, 1999; Stahlberg et al., 2001). However, many two-dimensional crystals are not ordered well enough to give resolutions beyond 4 Å.

Recent progress in NMR spectroscopy allowed for the first time to solve complete structures of integral membrane proteins (reviewed by Arora and Tamm, 2001; Opella et al., 2001). In principle three NMR techniques have been used to investigate transmembrane proteins (Fig. 6.1). Some detergents which form micelles can solubilize membrane proteins in their native structure and liquid state NMR can be applied. Alternatively organic solvent mixtures are used for the solubilization of membrane proteins together with liquid state NMR. Membrane proteins which are integrated in lipid bilayers, an environment close to their natural one, can be studied by solid-state NMR.

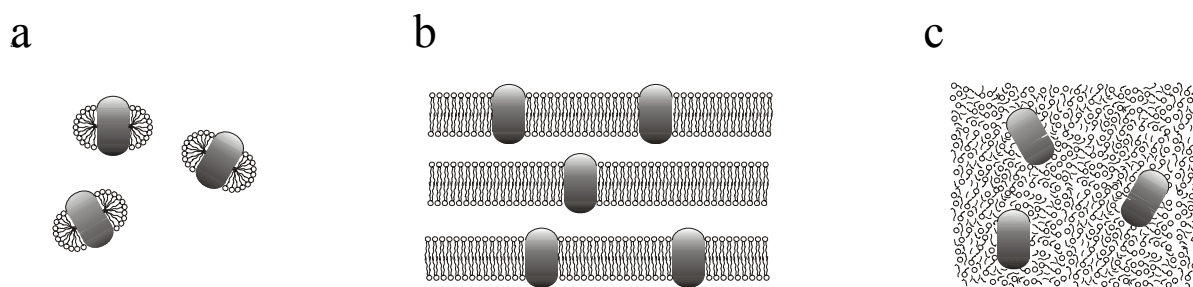


Fig. 6.1: Environments under which membrane proteins are investigated by NMR. (a) Solubilization with detergent micelles. (b) Membrane proteins in lipid bilayers between glass plates offer an environment close to natural membranes and a uniform global orientation. (c) Some membrane proteins are soluble in organic solvent mixtures without denaturation. Note that these conditions are far away from the natural membrane environment. In case (a) and (c) the proteins are solubilized and liquid state NMR methods can be applied. Because of the large particle size in case (a) resulting from of the protein and the detergent, methods like TROSY and sophisticated labeling techniques have to be used. Solid state NMR is used to investigate immobilized membrane proteins in lipid bilayers (b).

The first two approaches allow in principle the application of the conventional liquid state NMR strategies for structure determination as described in chapters 2.3, 2.4 and Fig. 5.1. However, protein containing micelles have a relatively slow tumbling time due to its large molecular mass (>50 kDa depending on the detergent and the protein). As outlined in chapter 2.4 novel methods such as TROSY and perdeuteration allow to extend conventional NMR techniques to such large particles. The study of membrane proteins solubilized in micelles seems to have a high potential, two three-dimensional structures of integral membrane proteins, namely of the outer membrane proteins OmpA (177 residues, 19 kDa) and OmpX (148 residues, 16 kDa), have been determined recently (Arora et al., 2001; Fernández et al., 2001). The detergents dodecylphosphocholine (DPC) and dihexanoylphosphatidylcholine (DHPC) which have been used are shown in Fig. 6.2. Some relevant properties of these detergents are given in table 6.1.

detergent	cmc *	aggregation number	aggregation molecular weight	Reference
DPC	1.5 mM	70-80	25-28 kDa	Arora and Tamm, 2001
DHPC	15.2 mM	ca. 35	16 kDa	Arora and Tamm, 2001; Tausk et al., 1974a and b
DM	0.1-0.6 mM	130‡	66-70 kDa	Timmins et al., 1988; Bhairi, 1997

Table 6.1: Properties of the detergents dodecylphosphocholine (DPC), dihexanoylphosphatidylcholine (DHPC) and dodecylmaltoside (DM). ‡ the aggregation number given by Bhairi is 98 (seems not to fit to the aggregation molecular weight).

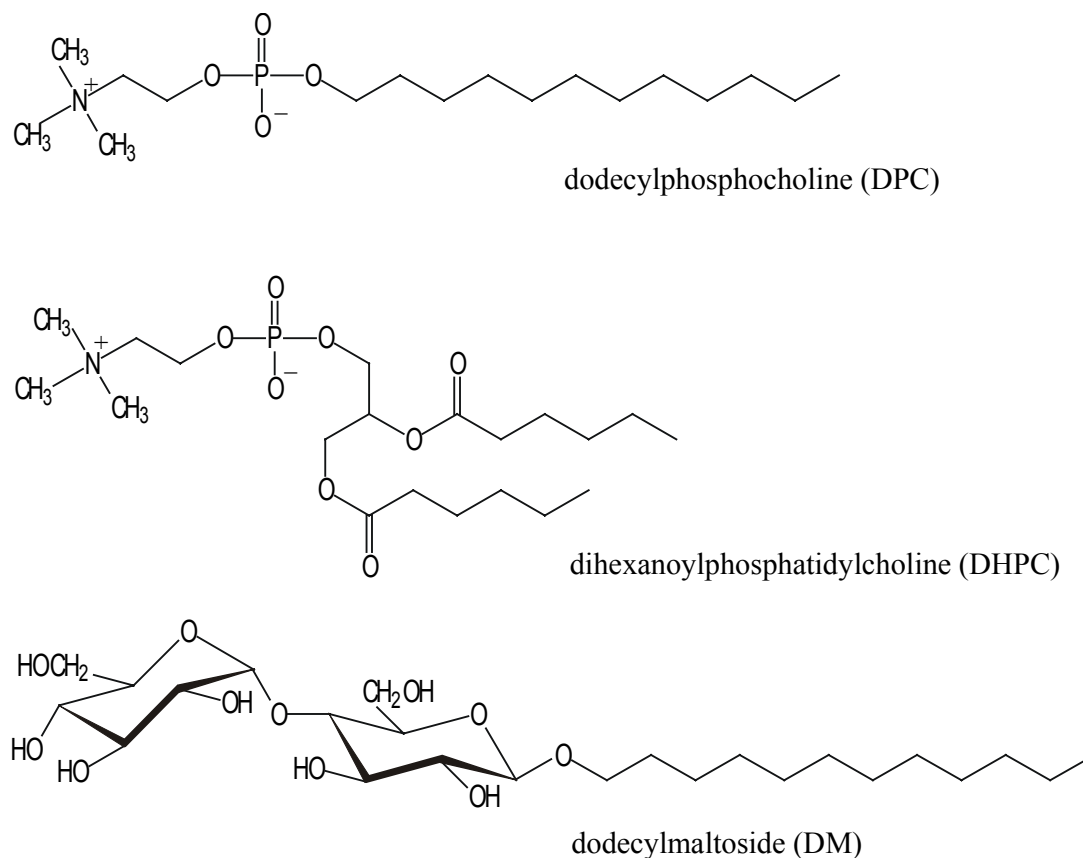


Fig. 6.2: Structures of the detergents dodecylphosphocholine, dihexanoylphosphatidylcholine and dodecylmaltoside.

The use of organic solvent mixtures for the solubilization of membrane proteins has the advantage that the molecular mass is not increased and the viscosity of the solvent is smaller which leads to a correlation time small compared to that of proteins solubilized in micelles. However, it has been questioned whether the observed protein structures are identical to their native structures and whether organic solvents are an appropriate environment for membrane proteins (reviewed by Arora and Tamm, 2001). The structures of subunits c and b of the *E. coli* F₀F₁ ATP synthase have been determined by NMR using organic solvent-water mixtures (Girvin et al., 1998; Dmitriev et al. 1999; Rastogi and Girvin, 1999). The NMR structure of subunit c fits quite well into the 3.9 Å electron density map calculated from X-ray diffraction data of the F₀F₁ ATP synthase (Stock et al., 1999). In the case of the multidrug transporter EmrE (110 residues, 12 kDa), however, the native tertiary structure seems to be disturbed by organic solvent molecules (Schwaiger et al., 1998) and the observation of long-range NOEs was difficult. Probably some solvent molecules penetrate the structure and thereby diminish tertiary contacts or induce slow conformational exchange.

Solid-state NMR spectroscopy is a rapidly developing field and recent progress is reviewed in the literature (Opella et al., 2001; de Groot, 2000; Fu and Cross et al., 1999). Despite few peptide structures determined so far, no complete protein structure determined

by solid state NMR has been published. However, novel technologies have recently allowed the first determination of a (globular) protein structure (Oschkinat group unpublished results). The spectra obtained with a 150 kDa membrane protein (LH2 light-harvesting nonameric complex) look promising (Egorova-Zatchernyuk et al., 2001).

6.2 Bacteriorhodopsin

Rhodopsins are a family of integral membrane proteins which consist of seven transmembrane helices, and a retinal covalently attached to a Schiff base located in a pocket formed by the helix bundle. They are found throughout the three domains of life: archaea (Oesterhelt, 1998), eubacteria (Béjà et al., 2000), and eukariota (Spudich et al., 2000). Their function - ion transport or photosensory signaling is based on a light-induced photoisomerization of the retinal. The visual rhodopsins are G-protein coupled receptors (GPCRs) which act as photoreceptors in the retina. The rhodopsins of haloarchaea have been extensively studied as model systems: Bacteriorhodopsin (BR) and halorhodopsin (HR) are light-driven ion pumps, transporting protons and chloride ions, respectively. Sensory rhodopsins I and II (SRI and SRII) are phototaxis receptors (Oesterhelt, 1998).

Bacteriorhodopsin is the best characterized system among the archaeal rhodopsins (for reviews see Lanyi, 2000a and b; Haupts et al., 1999). BR has been for a long time the only protein consisting of seven transmembrane helices whose structure was known at high resolution, and was therefore used as model system for GPCRs (Wess, 1997; Dohlman et al., 1991). The most recent crystal structures were determined at a resolution of 1.55 Å (Luecke et al., 1999) and 1.9 Å (Belrhali et al., 1999). The three-dimensional structures of halorhodopsin and sensory rhodopsin II have been determined recently by X-ray crystallography at resolutions of 1.8 Å (Kolbe et al., 2000) and 2.1 Å (Royant et al., 2001), respectively. Even the structure of bovine rhodopsin could be determined at a resolution of 2.8 Å (Palczewski et al., 2000).

Bacteriorhodopsin has been chosen as a model system for the present NMR investigation. The crystal structure of BR solved by Luecke et al. is shown in Fig. 6.3. It consists of seven well defined transmembrane helices and a short β -sheet. The loops are less well defined as indicated by higher B-factors. The structure of the EF-loop could not be determined, since the electron density was too diffuse. The reason for this might be local flexibility or multiple conformations.

The aim of the NMR study was to find out whether heteronuclear NMR techniques could deliver structural insights into the loop regions and thus provide complementary information to the X-ray crystallography data.

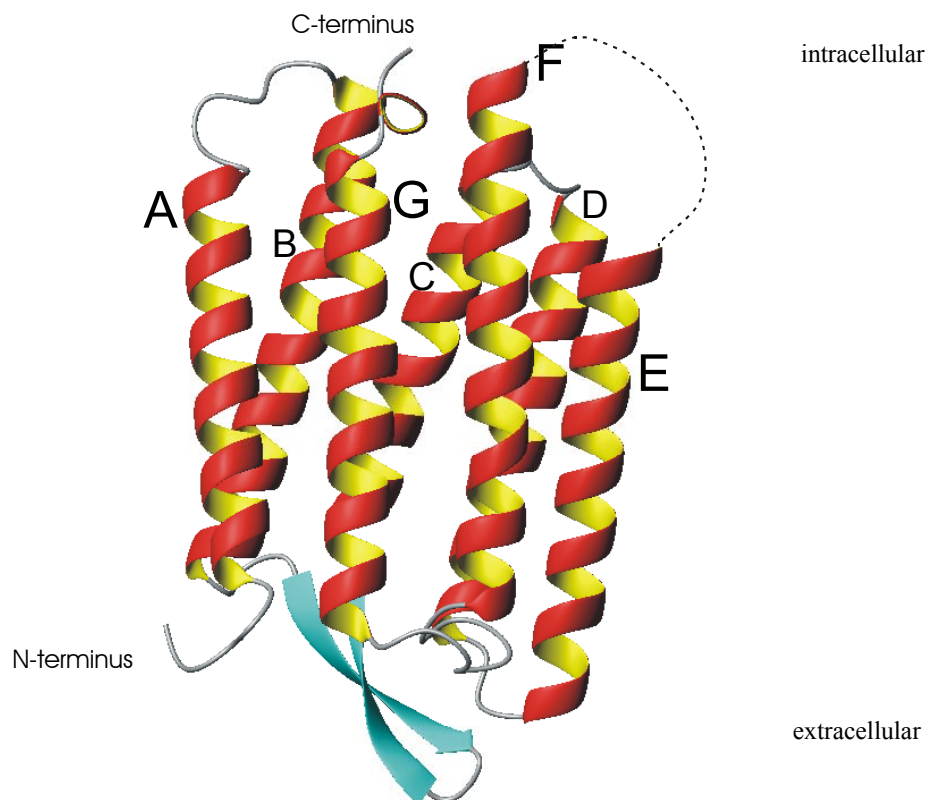


Fig. 6.3: The 1C3W crystal structure at resolution 1.55 Å (Luecke et al., 1999). Note that the residues 157-161 are missing because of low electron density. This figure was prepared with the program MOLMOL (Koradi et al., 1996).

6.3 Samples and NMR experiments

Two labeling patterns have been used for the NMR studies: a uniformly $^2\text{H}/^{15}\text{N}$ -labeled and a $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled sample solubilized in dodecylmaltoside micelles. Spin diffusion is reduced due to the high level of deuteration as explained in chapter 2.4. Dodecylmaltoside is an established detergent for the solubilization of bacteriorhodopsin, the stability of the samples and the quality of the spectra were demonstrated (Seigneuret et al., 1991; Seigneuret and Kainosho, 1993; Patzelt et al., 1997). Water is used as solvent and exchangeable deuterons are replaced by protons. Only amide groups and the retinal contain protons. However, the deuterium/proton exchange of the amide groups might not be complete especially in the rigid helical parts of the protein which might lead to missing NH signals in the spectra.

Almost all the experiments which were applied use the TROSY effect (chapter 2.4). Figure 6.4 shows a ^{15}N -HSQC and a ^{15}N -TROSY spectrum of $^2\text{H}/^{15}\text{N}$ labeled bacteriorhodopsin in DM micelles. The superior dispersion in the TROSY experiment due to narrower line widths which was discussed in chapter 2.4 is clear. However, chemical shift degeneracy in the center of the spectrum remains which hindered to some extent the sequential resonance assignment.

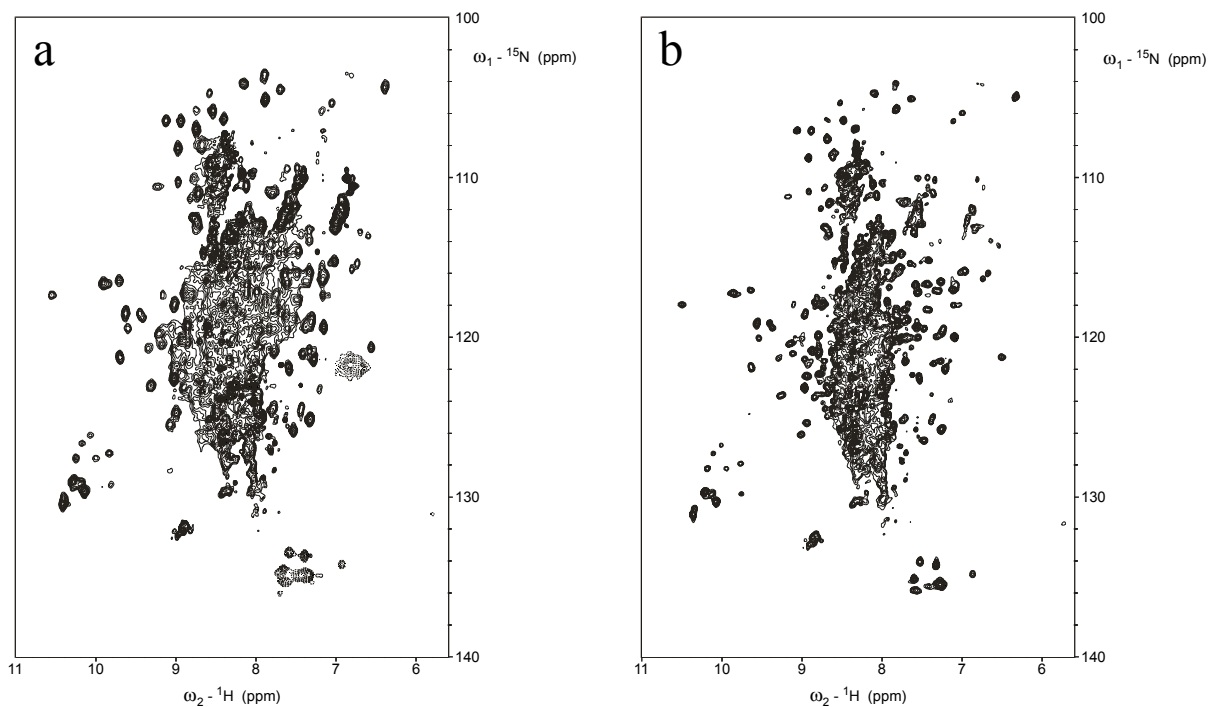


Fig. 6.4: Comparison between a ^{15}N -HSQC (a) and a ^{15}N -TROSY spectrum (b) of $^2\text{H}/^{15}\text{N}$ labeled bacteriorhodopsin. Both spectra are acquired in 15 h using 48 scans and are processed equally.

TROSY versions of the following 3D triple-resonance experiments were recorded using the $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled sample: HNCA, HN(CO)CA, HNCOC, HN(CA)CO, HNCACB and HN(CO)CACB (for references see chapter 2.4). Only the signal intensities in the HNCA and HNCOC experiments were satisfactory. However, very distinct signal intensities are present. The HNCACB was clearly less sensitive and some signals were missing (judged from a ^1H - ^{15}N projection compared to a projection of the HNCA). The HN(CA)CO spectrum contained only a fraction of the expected signals mainly from flexible parts like the termini. The HN(CO)CA and the HN(CO)CACB experiments were very insensitive and the spectra contained almost exclusively signals from the flexible termini.

It turned out that the magnetization transfer via the carbonyl carbon such as in the HN(CO)CA experiment is very inefficient within such large molecules (at high fields, e. g. 750 MHz magnet) due to fast relaxation caused by the large carbonyl chemical shift anisotropy (CSA). This has also been observed by others (Loria et al., 1999; Permi and Annala, 2001). Measuring at lower magnetic fields (500 or 600 MHz spectrometer) can reduce the CSA contribution to the relaxation, but the gain of sensitivity due to the TROSY effect as well as the higher resolution at 750 MHz would then be lost. An HN(CO)CA experiment recorded at 600 MHz using comparable parameters to that recorded at 750 MHz gave significantly worse spectra. The novel sequential HNCA experiment (Meissner and Sørensen, 2001) which provides spectra containing solely C^{α}_{i-1} should solve the problem, but this was not yet available at the time of these investigations. Unfortunately amino acid type-

selective experiments can not be applied because of the uniform deuteration (no CH₂ and CH₃ groups are available for selection).

The ²H/¹⁵N labeled sample was used to record (¹H, ¹⁵N, ¹H)-NOESY-TROSY (Zhu et al., 1999), (¹H, ¹⁵N, ¹H)-NOESY-HSQC (Clare and Gronenborn, 1993) and (¹⁵N, ¹⁵N, ¹H)-HMQC-NOESY-HSQC spectra.

6.4 Sequential assignment

The sequential backbone assignment was done semi-automatically with the program Sparky (Goddard and Kneller, 1999). Some of the peaks were picked automatically but others had to be picked manually due to the large variations in signal intensity and severe overlap of cross-peaks. Figure 6.5 illustrates the signal to noise ratios found in the HNCO spectrum which are representative for all other triple resonance spectra which have been recorded. The signal to noise ratios of the flexible termini are sometimes two orders of magnitude higher than those of well defined areas.

Generic spin systems (including resonances of ¹H^N, ¹⁵N, C^α_i, C^α_{i-1}, C'_{i-1} and if available, also of C^β_i, C'_i, C^β_{i-1}) were generated manually and temporary names were given. Unfortunately, the resonances of C^β_i, C^β_{i-1} and C'_i are often missing due to the low intrinsic intensity of the experiments TROSY-HNCACB and TROSY-HN(CA)CO. It was almost impossible to generate generic spin systems in the center of the ¹⁵N TROSY spectrum (Fig. 6.4), because of the severe degeneracy of the ¹H-¹⁵N frequency pairs.

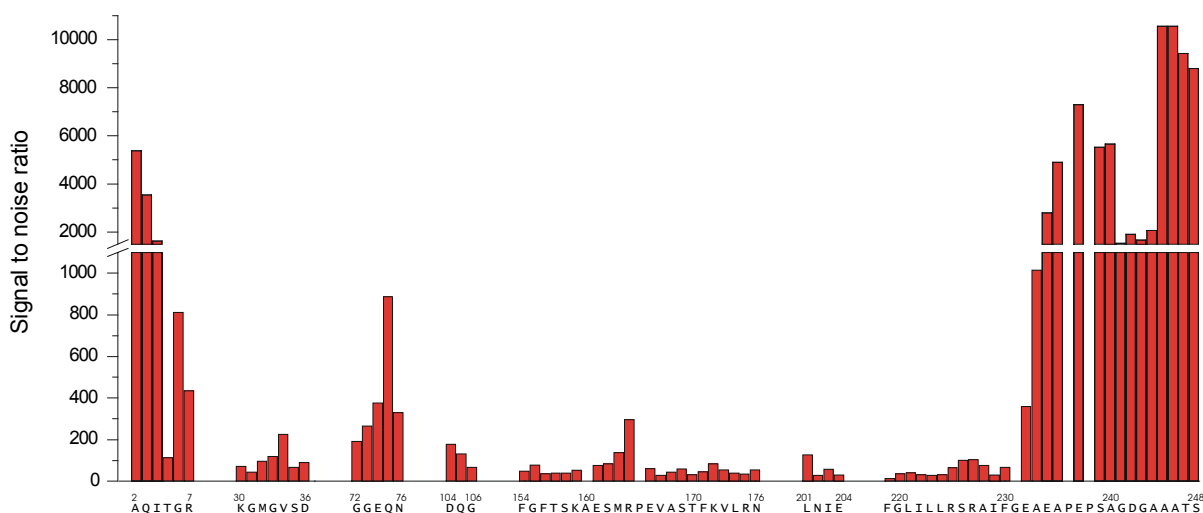


Fig. 6.5: Signal to noise ratios of the assigned signals in the TROSY-HNCO experiment.

Generic spin systems have been matched mainly using the TROSY-HNCA (Fig. 6.6). If the signals of interest were present in the TROSY-HN(CO)CA, this spectrum was used for the distinction between C^α_i and C^α_{i-1}, otherwise this was determined from the relative signal

intensities in the HNCA. Sequential peaks are normally less intense compared to intraresidue signals (see chapter 2.3).

Although the TROSY-HNCO is highly sensitive, it provides only one C'_{i-1} frequency per residue and therefore does not allow the establishment of connectivities between generic spin systems. The TROSY-HN(CA)CO contains sequential as well as intraresidue carbonyl chemical shifts (C'_{i-1} and C'_i) but is not very sensitive. However, in some cases it allowed sequential connectivities to be established as shown in Fig. 6.7.

Generic spin systems which could be unambiguously matched, using the automated matching tool in Sparky, were manually connected leading to chains of various lengths. All 3D NOESY spectra were of immense value for the evaluation of the correct linking of generic spin systems, the sequential $^1H^N$ - $^1H^N$ NOEs are seen in most cases. The obtained peptide segments were mapped to the protein sequence with the $^{13}C^\alpha$ and (if available) $^{13}C^\beta$ chemical shifts using the program "seq_prob" (Grzesiek and Bax, 1993). The program proposes five positions for the peptide segments along the protein sequence and calculates their probabilities. Only unambiguous assignments were taken into account and are presented in table 6.2.

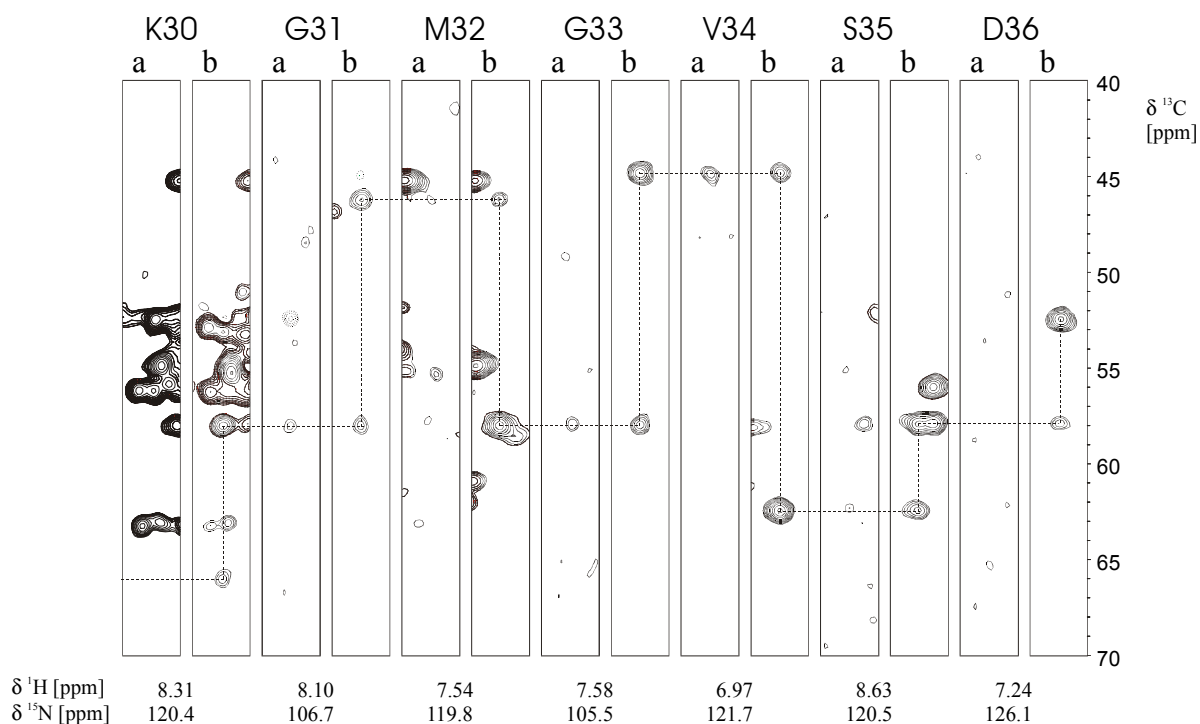


Fig. 6.6: Strip plot of the spectra (a) TROSY-HN(CO)CA and (b) TROSY-HNCA showing the region K30-D36. The sequential walk is indicated by dashed lines. Due to the intrinsic low sensitivity of the TROSY-HN(CO)CA experiment the C'_{i-1} signals are sometimes missing. Their presence allows the distinction between the intraresidue and the sequential signals in the TROSY-HNCA:

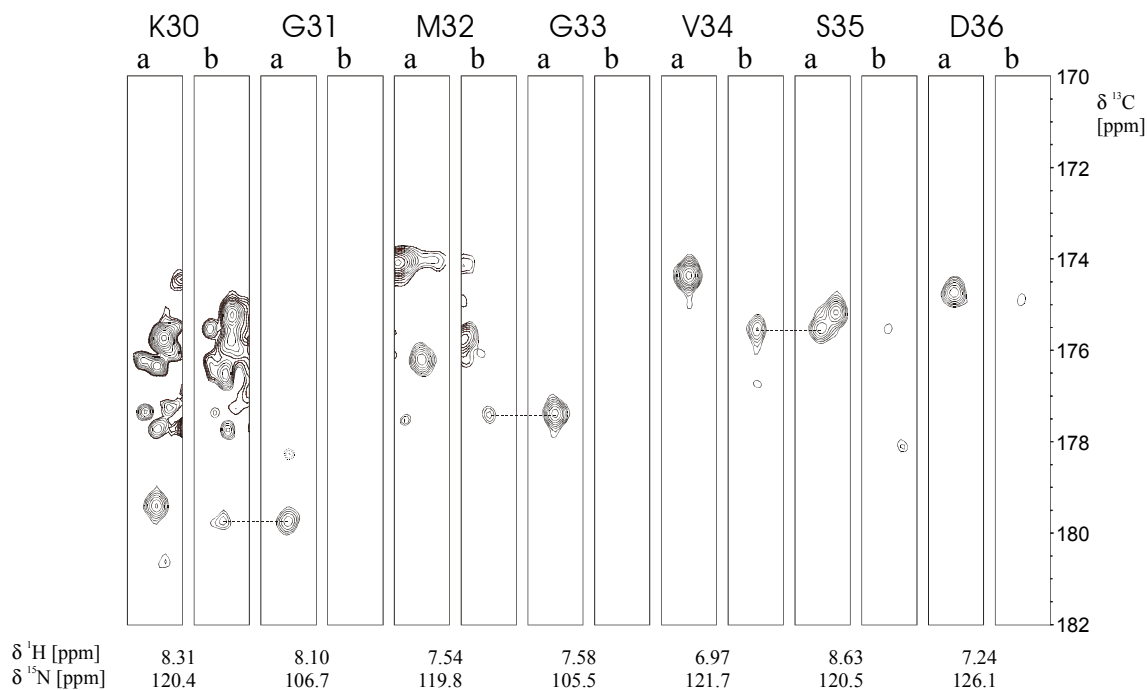


Fig. 6.7: Strip plot of the spectra (a) TROSY-HNCO and (b) TROSY-HN(CA)CO showing the region K30-D36. The low signal intensities in the HN(CA)CO only allowed to detect few sequential connectivities (marked with dashed lines).

The sequential backbone assignment was not possible for the entire sequence, because of missing signals and the severe degeneracy of the ^1H - ^{15}N frequencies. However, 33% of the sequence was assigned as shown schematically in Fig. 6.8. The complete EF-loop which is not defined in the 1C3W crystal structure was assigned. Two different sets of signals were found for F154-K159 which is illustrated in Fig. 6.9.

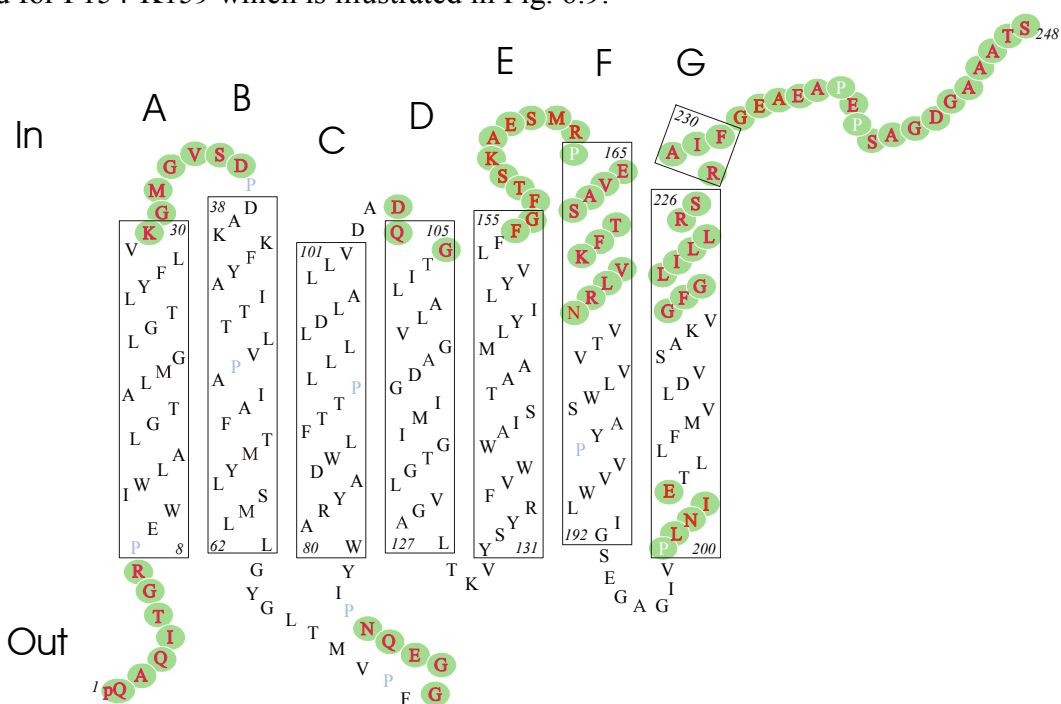


Fig. 6.8: Schematic presentation of the unambiguously assigned residues. 33% of the residues (81 out of 248) could be assigned, mostly loop regions.

Table 6.2: Chemical shift list of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled bacteriorhodopsin at 323.5 K referenced to H_2O according to Cavanagh et al., 1996 (p.175). The ^{15}N and $^1\text{H}^{\text{N}}$ chemical shifts are taken from the TROSY spectra.

residue	$\delta^{15}\text{N}$ [ppm]	$\delta^1\text{H}^{\text{N}}$ [ppm]	$\delta^{13}\text{CO}$ [ppm]	$\delta^{13}\text{C}^\alpha$ [ppm]	$\delta^{13}\text{C}^\beta$ [ppm]
E1	125.3	7.61	177.5	59.1	27.0
A2	124.1	8.12	177.5	52.1	18.4
Q3	119.7	8.29	176.4	55.6	28.7
I4	121.7	8.04	176.3	61.9	37.7
T5	113.5	7.65	174.8	61.7	68.4
G6	110.2	7.88	173.9	45.0	-
R7	120.7	7.72	175.0	53.3	30.0
V29	-	-	-	65.8	-
K30	120.4	8.31	179.6	57.8	-
G31	106.7	8.10	176.0	46.0	-
M32	119.8	7.54	177.2	57.7	-
G33	105.5	7.58	174.2	44.7	-
V34	121.7	6.97	175.3	62.3	31.9
S35	120.5	8.63	174.5	57.8	-
D36	126.1	7.24	-	52.0	43.2
F71	-	-	176.4	57.3	39.1
G72	112.9	8.48	175.0	45.8	-
G73	109.5	8.27	173.9	45.0	-
E74	119.8	7.66	175.5	54.7	30.6
Q75	120.1	8.26	175.0	55.1	29.0
N76	121.5	8.27	172.6	50.5	40.2
A103	-	-	176.2	51.7	18.8
D104	117.0	7.82	176.3	52.8	41.9
Q105	119.6	8.55	177.6	59.4	28.7
G106	107.3	8.45	-	47.3	-
F153a	-	-	177.6	61.3	-
F154a	117.5	8.59	177.1	59.7	37.7
G155a	110.9	7.83	176.8	47.1	-
F156a	121.1	8.73	178.0	57.4	34.3?
T157a	115.5	6.73	175.2	67.2	67.9
S158a	115.1	7.42	176.9	61.0	62.0
K159a	121.2	7.03	-	57.8	-
F153b	-	-	177.1	61.4	-
F154b	117.6	8.51	177.1	59.4	37.5
G155b	110.8	7.83	176.7	47.2	-
F156b	120.7	8.88	178.0	57.2	36.4?
T157b	115.7	6.72	175.2	67.2	68.0
S158b	115.3	7.53	177.0	61.0	-
K159b	121.3	7.11	-	57.9	-
A160	122.0	8.04	179.1	54.6	-
E161	113.2	7.72	177.3	57.9	-
S162	113.1	7.36	174.2	58.1	63.8
M163	122.3	7.29	174.8	54.8	33.4
R164	118.7	7.96	175.6	54.8?	28.5?
P165	-	-	178.8	65.7	-

Table 6.2: Chemical shift list of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled bacteriorhodopsin at 323.5 K referenced to H_2O according to Cavanagh et al., 1996 (p.175). The ^{15}N and $^1\text{H}^{\text{N}}$ chemical shifts are taken from the TROSY spectra.

residue	$\delta^{15}\text{N}$ [ppm]	$\delta^1\text{H}^{\text{N}}$ [ppm]	$\delta^{13}\text{CO}$ [ppm]	$\delta^{13}\text{C}^{\alpha}$ [ppm]	$\delta^{13}\text{C}^{\beta}$ [ppm]
E166	116.9	9.60	179.6	59.1	28.0
V167	121.2	7.47	176.9	65.8	-
A168	122.3	7.66	179.9	55.9	18.1
S169	111.3	8.43	177.1	61.0	62.1
T170	119.0	7.32	175.5	66.6	-
F171	120.1	8.91	175.6	62.5	-
K172	117.6	7.83	177.8	61.0	-
V173	116.8	7.07	177.6	66.2	30.7?
L174	118.9	7.57	180.0	57.1	-
R175	123.4	9.02	-	58.6	-
P200	-	-	177.7	61.5	-
L201	122.9	8.72	180.0	57.8	41.0
N202	118.8	9.33	175.6	57.8	-
I203	119.7	6.85	177.2	61.3	35.7
E204	122.8	8.45	-	59.3	-
G218	-	-	174.6	48.3	-
F219	119.6	7.05	176.2	60.7	-
G220	104.5	7.85	174.2	47.5	-
L221	122.0	8.54	178.6	57.8	-
I222	116.9	7.31	178.2	64.1	-
L223	118.0	7.58	178.6	57.8	-
L224	113.2	8.43	177.9	55.9	-
R225	116.2	7.02	176.0	55.4	-
S226	116.4	7.34	175.0	57.7	64.3
R227	120.6	8.55	178.0	55.8	28.6
A228	120.3	8.01	176.3	53.5	-
I229	104.6	6.09	173.8	61.3	-
F230	116.7	6.86	-	57.5	
G231					
E232	-	-	176.0	55.9	29.4
A233	125.1	8.17	177.5	52.0	18.5
E234	120.4	8.09	175.7	55.6	29.1
A235	126.8	8.08	175.5	49.9	-
P236	-	-	176.8	62.3	31.3
E237	122.1	8.13	174.9	53.6	28.8
P238	-	-	177.1	62.9	31.3
S239	115.9	8.12	174.5	57.8	63.3
A240	126.2	8.10	178.1	52.1	18.5
G241	108.5	8.06	174.1	44.9	-
D242	121.1	7.91	177.0	53.2	40.5
G243	109.7	8.18	174.2	45.1	-
A244	123.9	7.90	177.5	52.0	18.4
A245	124.1	-	177.5	51.9	18.3
A246	123.4	7.97	177.9	51.9	18.4
T247	113.1	7.87	174.0	61.1	69.5
S248	123.3	7.73	178.5	59.3	63.8

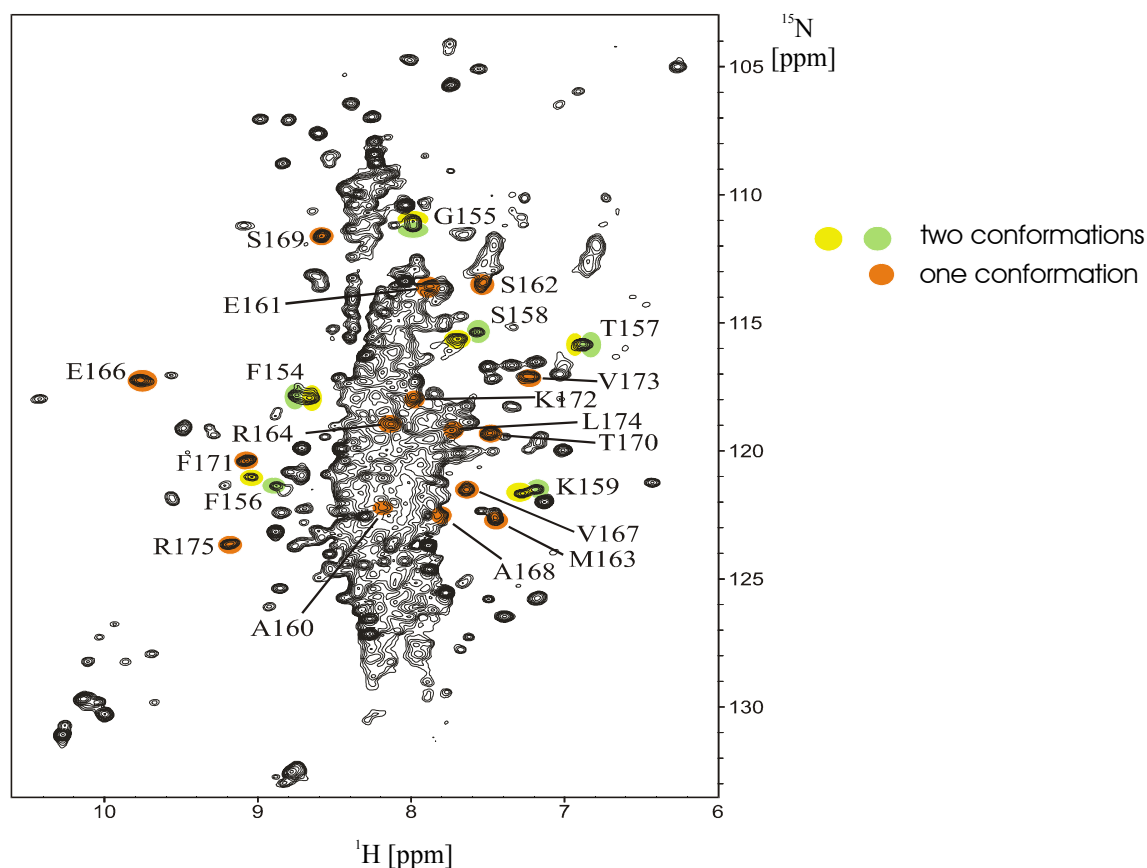


Fig. 6.9: ^{15}N -TROSY spectrum with the assignment of residues F154-R175. Two different sets of signals have been found for the residues 154-159. The signal set in green was named "a", the set in yellow was named "b" in table 6.2.

6.5 Secondary structure and obtained angle and distance restraints of the assigned regions

^{13}C chemical shift and NOE data provide important information on the secondary structure of a protein. The deviations of the observed $^{13}\text{C}'$, $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ chemical shifts from their residue dependent random coil values can be used for secondary structure prediction (Wishart et al., 1995) and thus ϕ and ψ angle predictions. The method was developed for $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled proteins and the observed ^{13}C values of $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled proteins have to be corrected for the deuterium effect (Farmer and Venters, 1999). The deviations in the $^{13}\text{C}'$, $^{13}\text{C}^{\alpha}$ and $^{13}\text{C}^{\beta}$ chemical shifts are shown in Fig. 6.10.

Due to the lack of many C^{β} chemical shifts, only C^{α} chemical shifts were used to derive ϕ and ψ angle constraints according to a method described in the literature (Luginbühl et al., 1995). This approach provides two categories of angle restraints. For $^{13}\text{C}^{\alpha}$ chemical shift deviations from the random coil values larger than + 2 ppm, the ϕ and ψ angles are restricted to values between -120° and -20° , and -100° and -0° , respectively. If the deviation is $1.5 \text{ ppm} < \Delta\delta(^{13}\text{C}^{\alpha}) \leq 2 \text{ ppm}$ ϕ angles in between -120° and 80° and ψ angles in between -100° and

60° are allowed. Corrected random coil values of residues preceding prolines were used (Wishart et al., 1995). The obtained angle restraints are shown schematically given in Fig. 6.11.

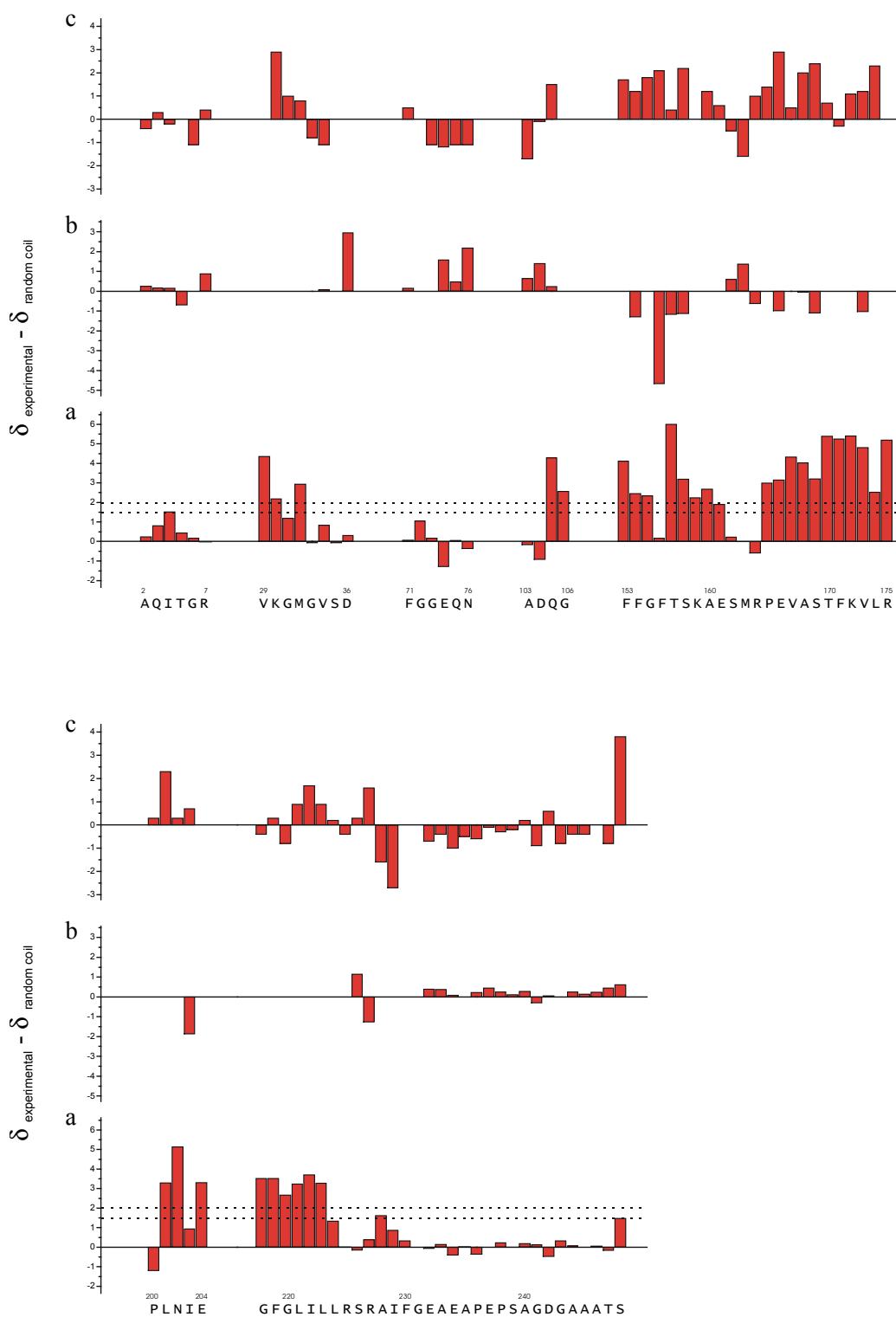


Fig. 6.10: Histogram showing the observed secondary shifts, defined here as the observed ^{13}C chemical shift (isotope effect of deuterium already corrected) minus the random coil ^{13}C chemical shift, for (a) C^α , (b) C^β and (c) C' chemical shifts. The ^{13}C chemical shifts of the two signal sets for residues 153-159 are very similar and signal set "a" was used. C^α chemical shifts (a) above the dotted lines allowed to derive angle restraints as explained in the text.

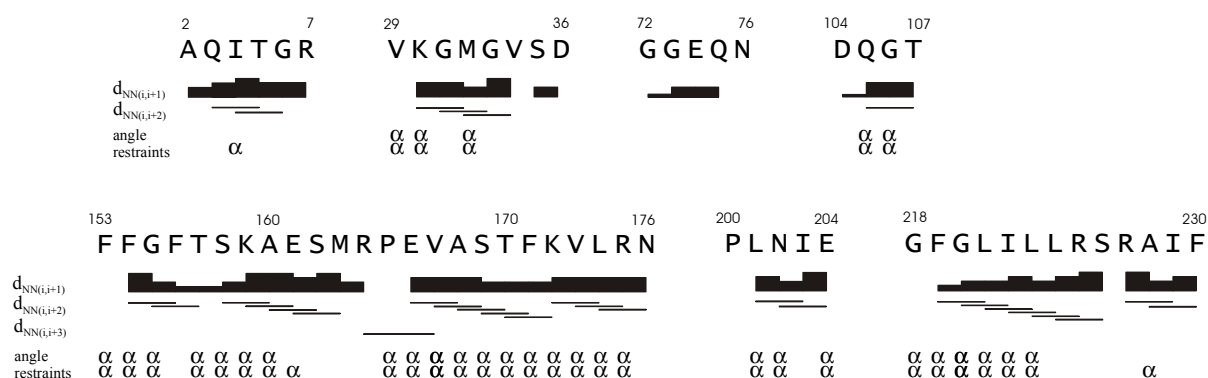


Fig. 6.11: Schematic presentation of the obtained angle and distance restraints (NOEs). Angle restraints, indicated by $2\times\alpha$ restrict the angles to $-120^\circ < \phi < -20^\circ$ and $-100^\circ < \psi < -0^\circ$, whereas those indicated by α restrict the angles to $-120^\circ < \phi < 60^\circ$ and $-100^\circ < \psi < 80^\circ$. In total 80 distance and 70 angle restraints were obtained. No NOEs or sufficient $^{13}\text{C}^\alpha$ chemical shift deviations were observed for residues 232-248.

The distance restraints which were observed are shown schematically in Fig. 6.11. 80 NOEs were unambiguously assigned using the 3D HNH-NOESY-TROSY, the 3D HNH-NOESY-HSQC and the 3D NNH-HMQC-NOESY-HSQC spectra. The appearance of $^1\text{H}^{\text{N}}$ - $^1\text{H}^{\text{N}}$ NOEs between two residues i and $i+2$ ($d_{\text{NN}}(i,i+2)$) are typical for α -helical conformations. Strips of the residues from 154 to 175 in the HNH-NOESY-TROSY are shown in Fig. 6.12.

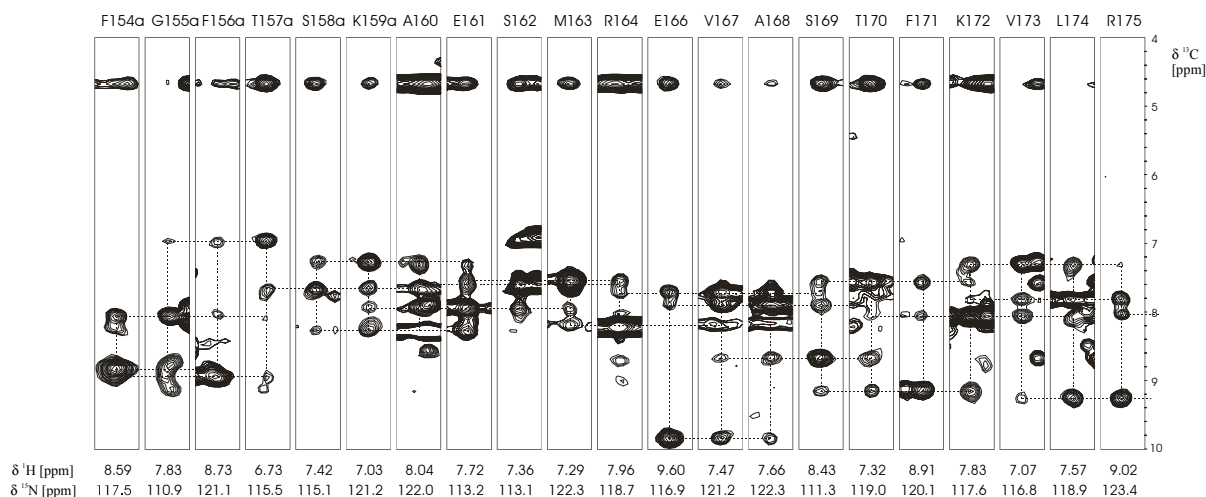


Fig. 6.12: Strips of residues 154-175 from the 3D NOESY-TROSY spectrum. Signal set "a" was used for residues 154-159. The NOE pattern for the other signal set was identical. The NOE connectivities are indicated by dashed lines. NOEs/exchange peaks from water are visible on top, they indicate water accessibility.

The secondary structure predicted by the C^α chemical shifts and $d_{\text{NN}}(i,i+2)$ NOEs corresponds widely with the X-ray structure 1C3W. However, the EF-loop is not defined in the 1C3W structure. The chemical shift and NOE data indicate that helix E is disrupted around F156, T157 by a kink / π -bulge, and another small helix is formed by the succeeding residues up to E161 or S162. The $d_{\text{NN}}(i,i+3)$ NOEs between R164 and V167 clearly indicates a turn. The data show that P165-R175 forms a helix which is consistent with X-ray data. The

NMR data indicate that residues V29-D36 represent the end of a helix in conjunction with a turn. The secondary structure prediction for G218-F230 is consistent with the X-ray structure, NOEs suggest that helix G is interrupted around S226 and a small helix fragment R227-F230 follows. The chemical shift data indicate an α -helix from 218 to L223. The structure of the succeeding residues, 224-230, may not be very stable and could be partially unfolded. Data for the short stretches A2-R7, G72-N76, D104-T107 and P200-E204 were too limited to allow conclusions about the secondary structure.

6.6 Structure calculations of the intracellular loops AB, EF and the C-terminal part of helix G

The angle and distance restraints allowed structure calculations of the loop AB and EF as well as the end of helix G. Note that these data define only the protein backbone and because of uniform deuteration no NOEs are observed to or from the side chains. Because the number of restraints per residue (1 to 4) were so limited, the AMBER force field (Case et al., 1997) was applied in addition.

The NOE intensities were classified into five categories corresponding to distance restraints of 0-2.3, 0-2.8, 0-3.2, 0-4.5 and 0-5.2 Å. Based on the NMR data, structures were calculated using the program AMBER 5.0 (Case et al., 1997), and the 1.55 Å crystal structure 1C3W (Luecke et al., 1999) as template. The missing loop, T157-E161, was constructed using the homology modeling facility of the SYBYL molecular modeling system (Tripos Inc.). All calculations were carried out in collaboration with Dr. Ronald Kühne.

A simulated annealing protocol was applied in which only the loop regions (see table 6.5) were allowed to move (elevating temperature to 1500 K at 0-10 ps, cooling down to 0 K at 10-20 ps with slowly increasing van der Waals repulsion terms from 0.001 to 1.0 between 0 and 10 ps in analogy to Cornell et al., 1995). Four hydrogen bonded water molecules observed in the crystal structure 1C3W were included in the calculations. All structures were minimized until an rmsd gradient lower than 0.05 was reached. Forces representing distance and torsion angle restraints were applied using a square-well function ($r_k = 50.0 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ and $r_k = 50.0 \text{ kcal mol}^{-1} \text{ deg}^{-2}$). Ensembles of 10 structures with the lowest energy were selected from 120 simulated annealing runs for further analysis.

Structure calculations of the AB loop, EF loop and the end of helix G were carried out. Structures of each region were calculated independently (calculations 1-3 in table 6.3) and in combination with each other (calculations 4-6 in table 6.3). The RMSD values of the ensembles (10 structures with the lowest energy) are given in table 6.3.

The result of calculation 6 (AB + EF + G) is shown in Fig. 6.13. The structure of the EF loop obtained with calculation 2 is presented in Fig. 6.14.

	loop regions which were allowed to move	residues which were allowed to move	number of angle restraints	number of distance restraints	RMSD of the calculated loop structures ‡
1	AB	K30-S35	6	8	0.229
2	EF	F153-R175	38	33	0.740
3	G	F219-G231	14	18	0.853
4	AB+EF	K30-S35, F153-R175	44	41	0.466
5	EF+G	F153-R175, F219-G231	52	51	1.129
6	AB+EF+G	K30-S35, F153-R175, F219-G231	58	59	1.341

Table 6.3: Overview of the different simulated annealing calculations with the regions which were allowed to move, the number of constraints and the RMSD of the resulting structures. ‡ the RMSD to the average structure of the C α atoms is given.

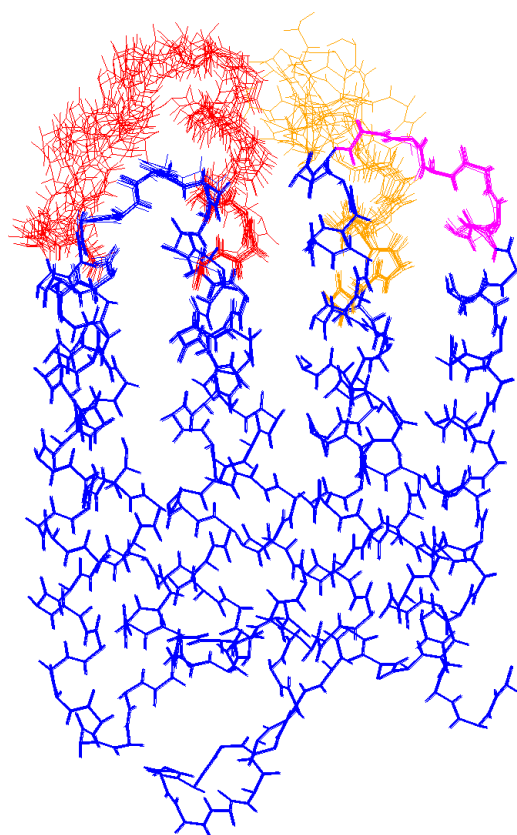


Fig. 6.13: Ensemble of the ten best structures with lowest energy of the structure calculation number 6 (table 6.3). This figure was prepared with the program MOLMOL (Koradi et al., 1996).

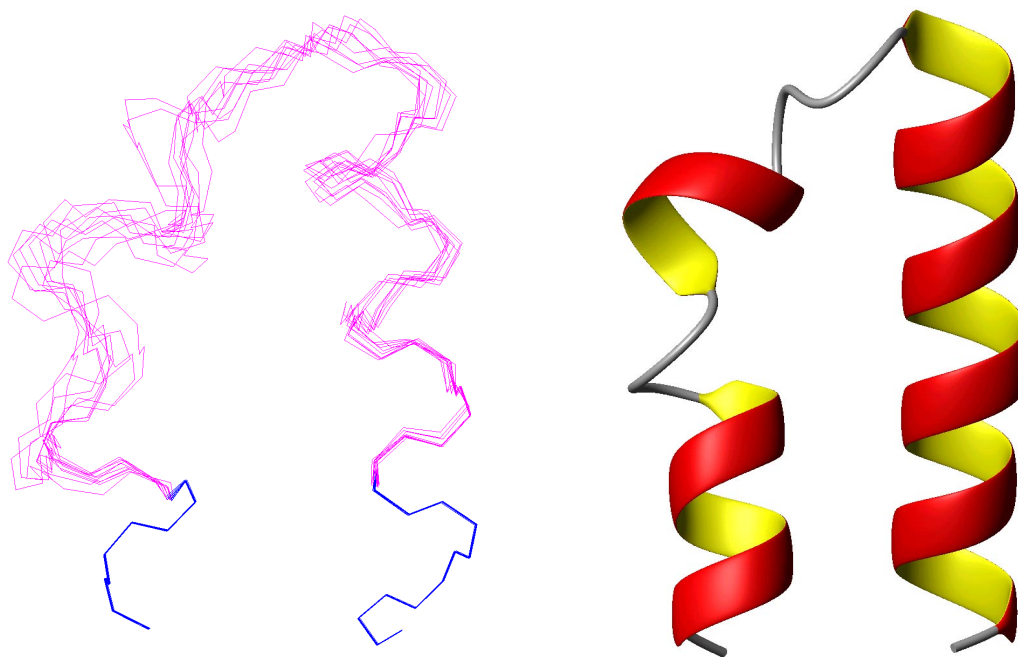


Fig. 6. 14: The structure of the EF-loop obtained from the calculation number 2 (table 6.3). An ensemble of the ten best structures with the lowest energy is shown on the left. Ribbon presentation of the structure (out of these ten) which is closest to the average. This figure was prepared with the program MOLMOL (Koradi et al., 1996).

6.7 Mobility

The mobility of the protein was studied using ^{15}N relaxation data. T_1 and T_2 relaxation measurements have been recorded using standard techniques (Farrow et al., 1994). The data were analyzed using the facility for relaxation analysis in the Sparky program (Goddard and Kneller, 1999). Due to the severe $^1\text{H}/^{15}\text{N}$ chemical shift degeneracies and the low signal intensities, the relaxation times T_1/T_2 of only few residues could be extracted. The T_1/T_2 ratios are shown in Fig. 6.13.

The residues shown in Fig. 6.14 can be grouped into three categories with different T_1/T_2 -ratios: the flexible termini with ratios of 0.5-4.7, residues in relatively rigid helices with ratios of 100-135 and residues of relatively flexible loops with T_1/T_2 -ratios in between 5-100. The residues G106, L201 and N202 show the highest T_1/T_2 -ratios which could be determined and are considered to represent the rigid parts of the protein. From the T_1/T_2 -ratios a value for the correlation time τ_c can be estimated (Fushman et al., 1994; Gryk et al., 1998)

The T_1/T_2 -ratio of the rigid part of the protein (ca. 120 ± 40) corresponds to an estimated correlation time $\tau_c = 35 \pm 5$ ns. However, contributions of local dynamics to the T_1/T_2 -ratio may lead in this estimation to a correlation time τ_c lower than the real one. Seigneuret et al. determined a correlation time τ_c of the BR containing micelle at 50°C of 33 ns based on viscosity measurements which corresponds to the result from NMR.

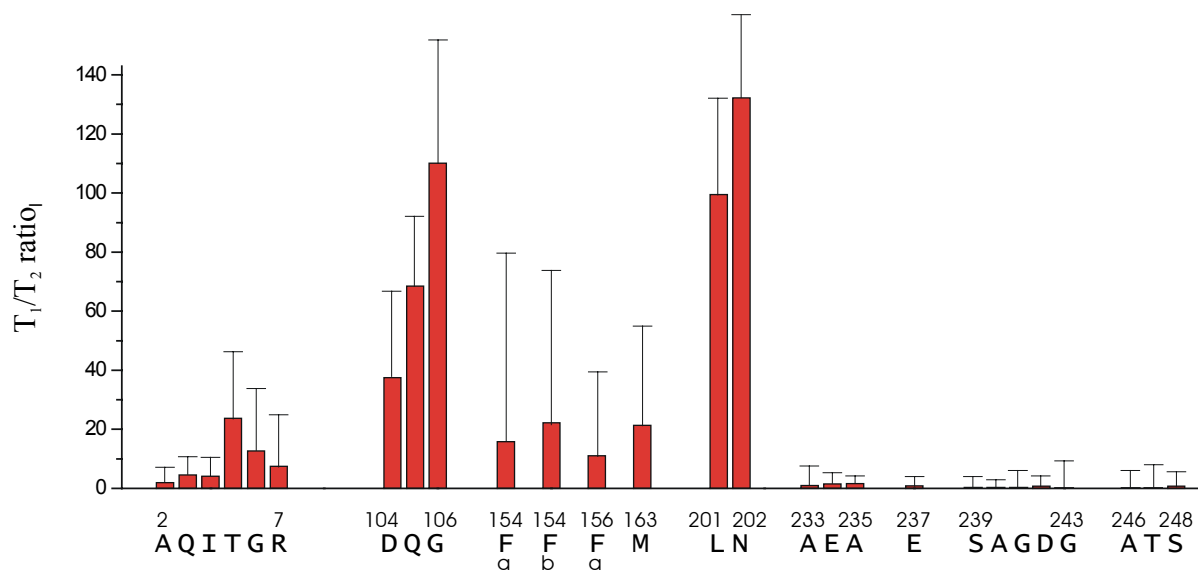


Fig. 6.15: T_1/T_2 ratios obtained from relaxation measurements at 600 MHz and 323.5 K.

The two different sets of signals for F154-K159 which have been already mentioned indicate two conformations at the end of helix E. The data do not allow any quantitative conclusions.

6.8 Discussion

The results of the structure calculations for the AB loop in calculations 1, 4 and 6 (Fig. 6.13) are fully consistent with the 1C3W crystal structure. However, only six residues were allowed to move and the flexibility was quite restricted by the two fixed ends of helix A and B during the simulated annealing protocol, so that the 14 NMR restraints contributed only partially to the structure which has been obtained.

The ensemble of structures obtained for the end of helix G (Fig. 6.13) agrees with the crystal structure. Residues 219-225 form a well defined helix. Around R225/S226 the helix is disturbed as in the 1C3W crystal structure. NMR data indicate that residues S226-F230 form a small helix fragment consistent with the X-ray structure. However, the global orientation of residues S226-F230 is not well defined.

Before discussing the structures of the EF-loop obtained from the simulated annealing calculations, the distance and angle restraints were compared to recent X-ray crystal structures. Violations between NMR restraints and the crystal structures are given in table 6.4.

The crystal structure 1C3W (Luecke et al., 1999), in which T157-E161 is not defined, shows only two significant violations to the NMR restraints around S162-R164, residues which are characterized by high B-factors in the crystal structure. No violations were

observed for residues 153-156 and 165-175, the NMR data for these residues are consistent with the X-ray structure at 1.55 Å resolution. This shows the reliability of the NMR data.

violated distance or angle restraints	restraints from NMR data	distances/angles in 1C3W (T157-E161 missing)	distances/angles in 1QHJ	distances/angles in 1BRR
F154NH-G155NH	0-2.3 Å			3.33 Å
S158NH-K159NH	0-3.2 Å			4.13 Å
K159NH-A160NH	0-2.3 Å		2.38 Å	2.50 Å
A160NH-E161NH	0-2.3 Å		2.68 Å	2.37 Å
E161NH-S162NH	0-2.8 Å			3.17 Å
S162NH-M163NH	0-2.3 Å	3.35 Å		
M163NH-R164NH	0-3.2 Å	4.81 Å	4.41 Å	3.43 Å
V167NH-A168NH	0-2.8 Å			3.00 Å
A168NH-S169NH	0-2.8 Å		2.84 Å	2.90 Å
K172NH-V173NH	0-2.8 Å	2.84 Å	2.83 Å	2.83 Å
G155phi	-120° to -20°		-130°	-170°

Table 6.4: Comparison of distance and angle restraints obtained from NMR data with the X-ray crystal structures 1C3W, 1QHJ and 1BRR. Only distances which violate the NMR restraints are shown.

The comparison of the crystal structure 1QHJ (Belrhali et al., 1999) and the NMR data reveal several small and two significant differences around residues 160/161 and 163/164. The crystal structure of a trimer 1BRR at 2.9 Å resolution (Essen et al., 1998) stays in contrast to several NMR restraints, the most severe differences are observed around F144/G155 and S158/K159. Larger differences are found by the comparison of the NMR data with the crystal structures 1AP9, 1QM8, 1AT9 and 2AT9.

A top view of the structure closest to the mean structure of an ensemble of the best 10 structures obtained from calculation number 2 is presented in Fig. 6.16. The side chain orientations are indicated by the C^α - C^β bonds. Note that no experimental data define the conformations of side chains during the calculations and atoms beyond C^γ are therefore not shown.

The results are consistent with recent EPR studies (Pfeiffer et al., 1999). However, one discrepancy between the model of the EF-loop proposed by Pfeiffer et al. exists. This model includes an extended helix E till S158 which does not agree with the NMR data.

The EPR data indicate that the side chains of residues F154, G155 and F156 face the protein interior and are less accessible to O_2 and chromium oxalate (CROX), while T157 and S158 are oriented towards the lipid phase (accessible to O_2 but not to CROX). This results

are consistent with the structure shown in Fig. 6.16. EPR data show two spectral components for residue F156 exactly in the same area where two different sets of signals are found by NMR. This is a striking finding. It may indicate two conformations of the Phe side chain as proposed by Pfeiffer et al. K159 is the first residue that faces the aqueous phase (accessible to CROX) which is consistent with the NMR structure.

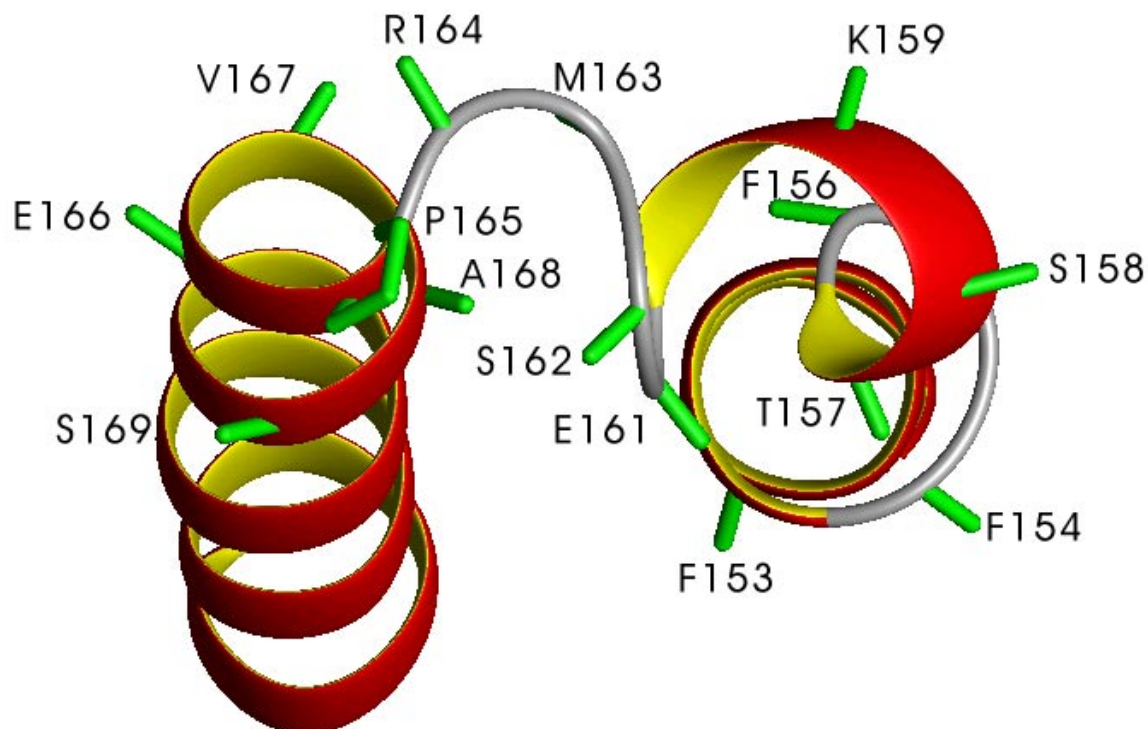


Fig. 6.16: Ribbon presentation of the EF-loop (the structure closest to the average of the ensemble of 10 best structures was selected) obtained from the simulated annealing calculation number 2 (table 6.3) with the orientations of the side chains indicated by the C^α - C^β bonds in green. This figure was prepared with the program MOLMOL (Koradi et al., 1996).

Positions 159, 161, 162, 164, 165 and 166 show a high degree of mobility in the EPR study; all side chains are found in the NMR structure either on the hydrophilic surface to the water or close to it. A160 and M163 are oriented towards the protein interior or towards the lipids and exhibit therefore less mobility and a reduced accessibility to CROX. A160 and E161 are accessible to O_2 , i. e. are facing the lipid phase which is consistent with the NMR structure. The side chains of P165 and E166 point into the water which is also indicated by the EPR study. Position 168 is still accessible for CROX whereas residues 167 and 169-171 are not accessible, as these point towards hydrophobic regions.

In addition to F156, residues S158, K159, A160, M163, P165 and A168 show two spectral components in the EPR spectra. NMR data present two different sets of signals for residues 154-159 with the largest deviations for F156, S158 and K159 which is strikingly

consistent with the EPR results. However, for the residues A160, M163 and A168, only one set of signals was found in the NMR spectra. This discrepancy might be caused by the different time scale of NMR observations (ms range versus ns/ μ s range) and by the different temperatures applied during the studies (323.5 K versus 293 K). Another difference between EPR and NMR studies which should be mentioned here is that the exchange of native amino acids by spin label side chains might influence the native structure.

6.9 Conclusion and Outlook

It was demonstrated that NMR techniques developed for larger globular proteins can be extended to membrane proteins solubilized in detergent micelles. 33% of the protein sequence of native bacteriorhodopsin was assigned. NOESY spectra and the ^{13}C chemical shifts allowed the derivation of 80 distance and 70 angle restraints. Two distinct sets of signals for F154-K159 indicate two conformations in the EF loop, however, the distance and angle restraints derived from both sets were very similar.

Structure calculations of the loops AB and EF as well as the end of helix G were carried out based on the NMR data together with the AMBER force field. A structure of the EF-loop which is poorly or not defined in X-ray crystal structures was determined. The results are in accordance with recent EPR studies. NMR data show that helix E is interrupted around F156/T157 by a kink or π -bulge which is a novel finding. The structures of the AB-loop and the end of helix G correspond to the X-ray structure 1C3W.

The use of additional techniques for further investigation such as selective ^{15}N labeling as described for the membrane protein OmpA (Arora et al., 2001) would greatly facilitate the backbone assignment and almost complete backbone assignments should be feasible. A partially deuterated sample will allow the assignment of ^1H frequencies of H^α and side chain protons. With the H^α chemical shifts (in conjunction with ^{13}C chemical shifts) the angle restraint prediction will get much more precise and reliable (Cornilescu et al., 1999). NOEs observed between H^α and H^{N} or other H^α protons will define the protein backbone more accurately.

References

- A. Arora and L. K. Tamm, *Curr. Opin. Struct Biol.* **11**, 540-547 (2001).
- A. Arora, F. Abildgaard, J. H. Bushweller and L. K. Tamm, *Nature Struct. Biol.* **8**, 334-338 (2001).
- O. Bèjà, L. Aravind, E. V. Koonin, M. T. Suzuki, A. Hadd, L. P. Nguyen, S. B. Jovanovich, C. M. Gates, R. A. Feldman, J. L. Spudich, E. N. Spudich and E.F. DeLong, *Science* **289**, 1902-1906 (2000).
- H. Belrhali, P. Nollert, A. Royant, C. Menzel, J. P. Rosenbusch, E. M. Landau and E. Pebay-Peyroula,

- Structure Fold. Des.* **7**, 909-917 (1999).
- S. M. Bhairi, "Detergents", Calbiochem-Novabiochem Corp. (1997).
- D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, D. M. Feguson, R. J. Radmer, G. L. Seibel, U. C. Singh, P. K. Weiner and P. A. Kollman, AMBER 5, University of California, San Francisco (1997).
- J. Cavanagh, W. J. Fairbrother, A. G. Palmer III and N. J. Skelton, "Protein NMR Spectroscopy", Academic Press, San Diego (1996).
- G. M. Clore and A. M. Gronenborn, "NMR of Proteins", CRC Press, Boca Raton (1993).
- W. D. Cornell, P. Cieplak, C. L. Bayly, I. R. Gould, K. M. Merz, D. M. Feguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollmann, *J. Am. Chem. Soc.* **117**, 5179-5197 (1995).
- G. Cornilescu, F. Delaglio and A. Bax, *J. Biomol. NMR* **13**, 289-302 (1999).
- O. Dmitriev, P. C. Jones, W. Jiang and R. H. Fillingame, *J. Biol. Chem.* **274**, 15598-15604 (1999).
- H. G. Dohlman, J. Thorner, M. G. Caron and R. J. Lefkowitz, *Annu. Rev. Biochem.* **60**, 653-688 (1991).
- T. A. Egorova-Zatchernyuk, J. Hollander, N. Fraser, P. Gast, A. J. Hoff, R. Cogdell, H. J. M. de Groot and M. Baldus, *J. Biomol. NMR* **19**, 243-253 (2001).
- L. O. Essen, R. Siegert, W. D. Lehmann, D. Oesterhelt, *Proc. Natl. Acad. Sci. USA* **95**, 11673-11678 (1998).
- B. T. Farmer II and R. A. Venters, *Biol. Magn. Reson.* **16**, 75-120 (1999).
- N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay and L. E. Kay, *Biochemistry* **33**, 5984-6003 (1994).
- C. Fernandez, K. Adeishvili and K. Wüthrich, *Proc. Natl. Acad. Sci. USA* **98**, 2358-2363 (2001).
- R. Fu and T. A. Cross, *Annu. Rev. Biophys. Biomol. Struct.* **28**, 235-268 (1999).
- D. Fushman, R. Weisemann, H. Thüring and H. Rüterjans, *J. Biomol. NMR* **4**, 61-78 (1994).
- M. E. Girvin, V. K. Rastogi, F. Abildgaard, J. L. Markley and R. H. Fillingame, *Biochemistry* **37**, 8817-8824 (1998).
- T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco (1999);
<http://www.cgl.ucsf.edu/home/sparky/>
- H. J. M. de Groot, *Curr. Opin. Struct. Biol.* **10**, 593-600 (2000).
- M. R. Gryk, R. Abseher, B. Simon, M. Nilges, H. Oschkinat, *J. Mol. Biol.* **280**, 879-96 (1998).
- S. Grzesiek and A. Bax, *J. Biomol. NMR* **3**, 185-204 (1993).
- U. Haupts, J. Tittor and D. Oesterhelt, *Annu. Rev. Biophys. Biomol. Struct.* **28**, 367-99 (1999).
- M. Kolbe, H. Besir, L. O. Essen and D. Oesterhelt, *Science* **288**, 1390-1396 (2000).
- M. Kolbe, Dissertation, LMU Munich (2001).
- R. Koradi, M. Billeter, and K. Wüthrich, *J. Mol. Graphics* **14**, 51-55 (1996);
<http://www.mol.biol.ethz.ch/wuthrich/software/molmol/>
- W. Kuhlbrandt and K. A. Williams, *Curr. Opin. Chem. Biol.* **3**, 537-547 (1999).
- J. K. Lanyi, *Biochim. Biophys. Acta.* **1460**, 1-3 (2000a) and all reviews in this issue.
- J. K. Lanyi, *J. Phys. Chem. B* **140**, 11441-11448 (2000b).
- J. P. Loria, M. Rance and A. G. Palmer III, *J. Magn. Reson.* **141**, 180-184 (1999).
- H. Luecke, B. Schobert, H.-T. Richter, J.-P. Cartailier and J. K. Lanyi, *J. Mol. Biol.* **291**, 899-911 (1999).
- P. Luginbühl, T. Szyperski and K. Wüthrich, *J. Magn. Reson. Ser. B* **109**, 229-233 (1995).
- A. Meissner and O. W. Sørensen, *J. Magn. Reson.* **150**, 100-104 (2001).
- D. Oesterhelt, *Curr. Opin. Struct. Biol.* **8**, 489-500 (1998).
- S. J. Opella, C. Ma and F. M. Marassi, *Methods Enzymol.* **339**, 285-313 (2001).
- K. Palczewski, T. Kumasaka, T. Hori, C. A. Behnke, H. Motoshima, B. A. Fox, I. Le Trong, D. C. Teller, T. Okada, R. E. Stenkamp, M. Yamamoto and M. Miyano, *Science* **289**, 739-745 (2000).

- H. Patzelt, A. S. Ulrich, H. Egbringhoff, P. Dux, J. Ashurst, B. Simon, H. Oschkinat, D. Oesterhelt, *J. Biomol. NMR* **10**, 95-106 (1997).
- P. Permi and A. Annala, *J. Biomol. NMR* **20**, 127-133 (2001).
- M. Pfeiffer, T. Rink, K. Gerwert, D. Oesterhelt and H.-J. Steinhoff, *J. Mol. Biol.* **287**, 163-171 (1999).
- V. K. Rastogi and M. E. Girvin, *Nature* **402**, 263-268 (1999).
- A. Royant, P. Nollert, K. Edman, R. Neutze, E. M. Landau, E. Pebay-Peyroula, J. Navarro, *Proc. Natl. Acad. Sci. USA.* **98**, 10131-10136 (2001).
- M. Schwaiger, M. Lebendiker, H. Yerushalmi, M. Coles, A. Gröger, C. Schwarz, S. Schuldinger and H. Kessler, *Eur. J. Biochem.* **254**, 610-619 (1998).
- M. Seigneuret, J. M. Neumann and J. L. Rigaud, *J. Biol. Chem.* **266**, 10066-10069 (1991).
- M. Seigneuret and M. Kainosho, *FEBS Lett.* **327**, 7-12 (1993).
- J. L. Spudich, C. S. Yang, K. H. Jung and E. N. Spudich, *Annu. Rev. Cell. Dev. Biol.* **16**, 365-92 (2000).
- H. Stahlberg, D. Fotiadis, S. Scheuring, H. Remigy, T. Braun, K. Mitsuoka, Y. Fujiyoshi and A. Engel, *FEBS Lett.* **504**, 166-172 (2001).
- D. Stock, A. G. Leslie and J. E. Walker, *Science* **286**, 1687-1688 (1999).
- R. J. M. Tausk, J. Karmiggelt, C. Oudshoorn and J. T. G. Overbeek, *Biophys. Chem.* **1**, 175-183 (1974a).
- R. J. M. Tausk, J. van Esch, J. Karmiggelt, G. Voordouw and J. T. G. Overbeek, *Biophys. Chem.* **1**, 184-203 (1974b).
- P. A. Timmins, M. Leonhard, H. U. Weltzien, T. Wacker and W. Welte, *FEBS Lett.* **238**, 361-368 (1988).
- E. Wallin and von Heijne, *Protein Sci.* **7**, 1029-1038 (1998).
- J. Wess, *FASEB J.* **11**, 346-354 (1997).
- D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges and B. D. Sykes, *J. Biomol. NMR* **5**, 67-81 (1995).
- G. Zhu, Y. Xia, K. H. Sze and X. Yan, *J. Biomol. NMR* **14**, 377-381 (1999).

Summary

This thesis describes the development and application of modern triple-resonance techniques in the context of protein structure determination by NMR. Novel heteronuclear NMR techniques for proteins up to 15 kDa which are selective for certain amino acid types have been developed. The application of state-of-the-art pulse sequences to the membrane protein bacteriorhodopsin solubilized in detergent micelles is demonstrated. Structures of the loop regions AB and EF as well as the end of helix G have been calculated on the basis of NMR data.

In **Chapter 2** a brief introduction into protein NMR spectroscopy is given. The sequential backbone resonance assignment, which is the basis for any structural investigation with NMR, is described in detail for $^{13}\text{C}/^{15}\text{N}$ and $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled proteins.

Basic NMR techniques used for the development of heteronuclear NMR pulse sequences are introduced in **Chapter 3**. The novel MUSIC (**M**ultiplicity **S**elective **I**n-phase **C**oherence transfer) technique which is able to select CH_2 , CH_3 or NH_2 groups is analyzed in detail. An introduction into shaped radio frequency pulses and the use of tuned delays for the selection of certain spin system topologies is provided.

The development of 39 novel amino acid type-selective ^1H - ^{15}N correlations based on MUSIC, selective pulses and tuned delays is described in **Chapter 4**. These two-dimensional experiments deliver spectra containing only signals resulting from selected amino acid types. Experiments for Gly, Ala, Thr/Ala, Val/Ile/Ala, Leu/Ala, Ser, Trp, Phe/Tyr/His, Asp/Asn/Gly, Asn, Asp, Glu/Gln/Gly, Gln, Glu, Arg, Lys, Cys/Ser which select either the $\text{NH}_{(i+1)}$ or $\text{NH}_{(i)}/\text{NH}_{(i+1)}$ amide signals are presented. Four Pro selective experiments are introduced allow the signals of the two neighbors ($i-1$) and ($i+1$) to be linked via the proline ^{15}N chemical shift. This set of experiments allows the identification of 16 amino acid types, namely Gly, Ala, Thr, Val/Ile, Leu, Ser, Trp, Phe/Tyr/His, Asp, Asn, Glu, Gln, Arg, Lys, Cys and Pro (all amino acids except Met). Val/Ile as well as Phe/Tyr/His are grouped together and can not be distinguished. Finally, an experiment which selects Trp $\text{H}^{\epsilon 1}\text{N}^{\epsilon 1}$ moieties is described.

Chapter 5 provides an introduction into automated assignment procedures. Strategies are discussed as to how the amino acid type-selective ^1H - ^{15}N correlations can be used to improve the performance of those procedures and the reliability of the results. An algorithm to derive ^1H - ^{15}N resonance assignments solely on the basis of selective 2D experiments has been

developed. The amino acid type-selective experiments either alone or together with three-dimensional triple resonance experiments allow an additional independent pathway for the resonance assignment.

It is demonstrated in **Chapter 6** that NMR techniques developed for larger globular proteins can be extended to membrane proteins solubilized in detergent micelles. Native bacteriorhodopsin was used as a model system. During the sequential backbone assignment, severe difficulties resulting from low sensitivity, missing signals and frequency degeneracies had to be overcome. Nevertheless, 33% of the sequence was assigned, including the EF-loop which is not resolved in the crystal structure at 1.55 Å resolution. NOESY spectra and the assigned ^{13}C chemical shifts allowed to derive distance and angle restraints. Two distinct sets of signals for F154-K159 indicate two conformations in the EF loop, however, the distance and angle restraints derived from both sets are very similar.

Structure calculations of the loops AB and EF as well as the end of helix G were carried out based on the NMR data. A structure of the EF-loop, which is poorly or not defined in X-ray crystal structures, could be determined. The NMR data show that helix E is interrupted around F156/T157 by a kink or π -bulge which is a novel finding. The structures of the AB-loop and the end of helix G correspond to the X-ray structure at 1.55 Å resolution.

Zusammenfassung

Die vorliegende Arbeit beschreibt die Entwicklung und Anwendung von modernen Triple-Resonanz Techniken im Rahmen der Strukturbestimmung von Proteinen mittels NMR-Spektroskopie. Neue heteronucleare NMR Experimente, welche selektiv für bestimmte Aminosäuretypen sind, wurden für Proteine bis 15 kDa entwickelt. Die Anwendung von kürzlich entwickelten Pulssequenzen auf das Membranprotein Bacteriorhodopsin, solubilisiert in Mizellen, wurde gezeigt. Strukturen der Loop-Regionen AB und EF sowie des Endes von Helix G wurden auf der Basis der erhaltenen NMR Daten bestimmt.

Kapitel 2 gibt eine kurze Einleitung in die NMR Spektroskopie von Proteinen. Die sequentielle Zuordnung, die Grundlage für alle strukturellen Untersuchungen mittels NMR Spektroskopie, wird ausführlich für $^{13}\text{C}/^{15}\text{N}$ und $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ markierte Proteine beschrieben.

NMR Techniken, die für die Entwicklung der heteronuclearen NMR Pulssequenzen verwendet wurden, werden in **Kapitel 3** vorgestellt. Die neu entwickelte MUSIC (**M**ultiplicity **S**elective **I**n-phase **C**oherence transfer) Technik welche CH_2 -, CH_3 - oder NH_2 -Gruppen selektieren kann, wird im Detail analysiert. Die Nutzung von selektiven Radiofrequenzpulsen und speziell eingestellten Wartezeiten für die Selektion von Spinsystem-Topologien wird ausführlich dargelegt.

Die Entwicklung von 39 neuen Aminosäure-selektiven ^1H - ^{15}N Korrelationen basierend auf MUSIC, selektiven Pulsen und speziell eingestellte Wartezeiten wird in **Kapitel 4** beschrieben. Die zwei-dimensionalen Experimente liefern Spektren, die lediglich Signale eines bestimmten Aminosäuretyps enthalten. Experimente für Gly, Ala, Thr/Ala, Val/Ile/Ala, Leu/Ala, Ser, Trp, Phe/Tyr/His, Asp/Asn/Gly, Asn, Asp, Glu/Gln/Gly, Gln, Glu, Arg, Lys, Cys/Ser welche entweder die $\text{NH}_{(i+1)}$ oder $\text{NH}_{(i)}/\text{NH}_{(i+1)}$ Amid Signale selektieren werden vorgestellt. Darüber hinaus wurden vier Prolin-selektive Experimente entwickelt, welche die Verknüpfung der zwei benachbarten Residien ($i-1$) und ($i+1$) unter Verwendung der ^{15}N chemischen Verschiebung des Prolins ermöglichen. Das gesamte Paket von Experimenten erlaubt die Bestimmung von 16 Aminosäuretypen: nämlich von Gly, Ala, Thr, Val/Ile, Leu, Ser, Trp, Phe/Tyr/His, Asp, Asn, Glu, Gln, Arg, Lys, Cys and Pro (alle Aminosäuren außer Met). Sowohl Val/Ile als auch Phe/Tyr/His werden zusammen in einer Gruppe identifiziert. Val kann nicht von Ile unterschieden werden. Dasselbe gilt für Phe, Tyr und His.








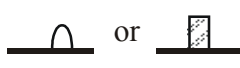




Kapitel 5 gibt einen Überblick über den derzeitigen Stand von automatischen Zuordnungsprozeduren. Strategien für die Implementierung der Aminosäure-selektiven ^1H - ^{15}N Korrelationen zur Verbesserung der Leistungsfähigkeit und Zuverlässigkeit solcher Prozeduren werden diskutiert. Ein Algorithmus für eine ^1H - ^{15}N Zuordnung, lediglich basierend auf selektiven 2D Experimenten, wurde entwickelt. Die Aminosäure-selektiven Experimente ermöglichen entweder allein oder zusammen mit 3D Triple-Resonanz Experimenten eine zusätzliche Route zu einer Signalzuordnung und erhöhen dadurch die Zuverlässigkeit von automatischen Routinen.

In **Kapitel 6** wird gezeigt, daß NMR Techniken, die für große globuläre Proteine entwickelt wurden, auch erfolgreich auf Membranproteine, solubilisiert in Mizellen, angewandt werden können. Dafür wurde natives Bacteriorhodopsin als Modellsystem verwendet. Während der sequentiellen Zuordnung der Signale mußten verschiedene Schwierigkeiten wie geringe Empfindlichkeit, fehlende Signale und Signalüberlagerungen überwunden werden. Trotz dieser Probleme konnten 33% der Proteinsequenz zugeordnet werden, einschließlich des EF-Loops, der in der Kristallstruktur mit einer Auflösung von 1.55 Å fehlt. Aus NOESY Spektren und den zugeordneten ^{13}C chemischen Verschiebungen wurden Abstands- und Winkel-Information gewonnen. Zwei unterschiedliche Signalsätze wurden für die Residuen F154-K159 gefunden, was auf zwei Konformationen des EF-Loops hindeutet. Sowohl die Winkel als auch Abstandsinformation unterscheiden sich kaum zwischen beiden Signalsätzen.

Strukturrechnungen der Loop-Bereiche AB und EF als auch des Endes der Helix G wurden basierend auf den NMR Daten durchgeführt. Eine Struktur des EF-Loops, welche in den verschiedenen Kristallstrukturen schlecht oder gar nicht definiert ist, konnte bestimmt werden. Die NMR Daten zeigen, daß Helix E bei F156/T157 durch einen Knick unterbrochen wird, was eine neue Erkenntnis ist. Die Strukturen des AB-Loops und des Endes von Helix G stimmen mit der Kristallstruktur mit einer Auflösung von 1.55 Å überein.

Appendix

Legend for the pulse sequences in chapter 3 and 4

	90° pulse with phase ϕ_2	¹ H 	¹ H decoupling
	180° pulse	¹⁵ N 	¹⁵ N decoupling
¹³ C 	shaped 180° pulse (G3) at 175 ppm	¹ H 	¹ H 90° flip back pulse
¹³ C 	shaped 180° pulse S ₁ for ¹³ C at a given offset	¹ H 	¹ H water selective 90° flip back pulse
¹³ C 	shaped 90° pulse S ₄ for ¹³ C at a given offset	¹ H 	¹ H 3-9-19 composite pulse within WATERGATE
PFG 	pulse field gradient G ₁	¹ H 	¹ H acquisition

Experimental part for chapter 4

General

All spectra were recorded on a 600 MHz spectrometer (Bruker-DRX600) in standard configuration using an inverse triple resonance probe equipped with three-axis self shielded gradient coils. Experiments were either recorded with a 1.35 mM sample of the EVH1 domain from VASP uniformly labeled with ¹⁵N and ¹³C or a 1.5 mM sample of the SAM domain from EphB2 uniformly labeled with ¹⁵N and ¹³C. The data were collected at 300 K. For both samples 5 mm ultra-precision sample tubes were used.

Proton hard pulses were applied with 25 kHz field strength; WALTZ-16 decoupling (Shaka et al., 1983) of ¹H spins was achieved using a field strength of 3.1 kHz. The same field strength was used for the subsequent 90° flip back ¹H pulses. The water-selective 90° square pulses had a duration of 1 ms. GARP-1 decoupling (Shaka et al., 1985) of ¹⁵N was achieved using a field strength of 830 Hz.

Water suppression was obtained using WATERGATE implemented with a 3-9-19 pulse (Sklenar et al., 1993). Pulses applied at the $^{13}\text{C}^\alpha$, $^{13}\text{C}^{\alpha/\beta}$ or ^{13}CO resonance frequencies were adjusted to provide a null at the corresponding ^{13}CO or $^{13}\text{C}^\alpha$ frequencies. Unless indicated otherwise, the rectangular $^{13}\text{C}^{\alpha/\beta}$ 90° and 180° pulses were set to $49\ \mu\text{s}$ and $44\ \mu\text{s}$, respectively. The rectangular $^{13}\text{C}^\alpha$ 90° and 180° pulses were set to $54\ \mu\text{s}$ and $48\ \mu\text{s}$, respectively. The rectangular ^{13}CO 90° and 180° pulses were set to $54\ \mu\text{s}$ and $108\ \mu\text{s}$, respectively. Shaped 180° ^{13}CO pulses were applied as G3 Gaussian cascade (Emsley and Bodenhausen, 1990) with a duration of $256\ \mu\text{s}$. The gradients were applied as a sinusoidal function from 0 to π . Unless indicated otherwise, the carrier frequencies were centered at $^1\text{H} = 4.8\ \text{ppm}$, $^{15}\text{N} = 120\ \text{ppm}$, $^{13}\text{C}^{\alpha/\beta} = 45\ \text{ppm}$, $^{13}\text{C}^\alpha = 55\ \text{ppm}$ and $^{13}\text{C}' = 175\ \text{ppm}$.

All spectra were recorded in an identical way with 64 (t_1) \times 512 (t_2) complex points and a spectral width of 3012 (^{15}N) \times $10000\ \text{Hz}$ (^1H). To achieve quadrature detection in the indirect dimension the States-TPPI-States protocol (Marion et al., 1989) was used in all experiments. XWINNMR (Bruker AG) was used for processing. The data were processed using a squared sinebell shifted by 90° as a window function in both dimensions. The ^{15}N t_1 interferograms were quadrupled in length by linear prediction using XWINNMR. The final spectra had a size of 512 (t_1) \times 1024 (t_2) real points.

All pulse programs, pulse shapes, gradient programs and datasets are available under:

<http://www.fmp-berlin.de/~schubert/pulseprograms.html>

a mirror is installed under

<http://userpage.chemie.fu-berlin.de/~mschub/pulseprograms.html>

Pulse Programs

Experiment	Name	Gradient Program	datasets	selective for	reference
G (i+1)	B g i+1	ma5crushwtg	B g i+1	Gly	JMR 1999
G (i,i+1)	B g ii+1	ma4crushwtg	B g i+1	Gly	JMR 1999
A (i+1)	B a i+1	ma4crushwtg	B a i+1	Ala	JMR 1999
A (i,i+1)	B a ii+1	ma4crushwtg	B a ii+1	Ala	JMR 1999
TA (i+1)	B ta i+1	ma4crushwtg	B ta i+1	Thr and Ala	JMR 1999
TA (i,i+1)	B ta ii+1	ma4crushwtg	B ta ii+1	Thr and Ala	JMR 1999
N (i+1)	B n i+1 <small>UPDATED</small>	ma4crushwtg	B n i+1	Asn	JMR 1999
N (i,i+1)	B n ii+1	ma4crushwtg	B n ii+1	Asn	JMR 1999
QN (i+1)	B qn i+1 <small>UPDATED</small>	ma4crushwtg	B qn i+1	Gln and Asn	JMR 1999
QN (i,i+1)	B qn ii+1	ma4crushwtg	B qn ii+1	Gln and Asn	JMR 1999
P (i+1)	B p i+1	ma3crushwtg	B p i+1	Pro	JBioNMR 2000
P (i-1)	B p i-1	ma3crushwtg	B p i-1	Pro	JBioNMR 2000
P ¹⁵ N (i)/ ¹ H(i+1)	b p i+1np	ma3crushwtg	B p i+1np	Pro	JBioNMR 2000
P ¹⁵ N (i)/ ¹ H(i-1)	b p i-1np	ma3crushwtg	B p i-1np	Pro	JBioNMR 2000
S (i+1)	B s i+1	ma4crushwtg	B s i+1	Ser	JMR 2001 a
S (i,i+1)	B s ii+1	ma4crushwtg	B s ii+1	Ser	JMR 2001 a
DNG (i+1)	B dng i+1	ma2crushwtg	B dng i+1	Asp, Asn and Gly	JMR 2001 a
DNG (i,i+1)	B dng ii+1	ma2crushwtg1	B dng ii+1	Asp, Asn and Gly	JMR 2001 a
ROG (i+1)	B rog i+1	ma2crushwtg	B rog i+1	Glu, Gln and Gly	JMR 2001 a

chapter 4.4.1

G-(i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = x, -x$; $\phi_4 = 50^\circ$; $\phi_5 = x$; $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = -y$; $\phi_8 = -x$; $\phi_{rec} = 2$ (x, 2 (-x), x), 4 (-x, 2 x, -x), 2 (x, 2 (-x), x). States-TPPI phase cycling was applied to ϕ_5 . Gradients had the following duration and strength: $G_1 = 800$ μ s (7 G/cm), $G_2 = 900$ μ s (28 G/cm), $G_3 = 1$ ms (21 G/cm).

G-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_1 = 7.0$ ms, $T_N = 11$ ms. The phase cycling was: $\phi_1 = x, -x$; $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = 8$ (x), 8 (-x); $\phi_4 = 4$ (x), 4 (y), 4 (-x), 4 (-y); $\phi_5 = 4$ (-y), 4 (y); $\phi_6 = (-x)$; $\phi_{rec} = x, 2$ (-x), x, 2 (-x, 2 (x), -x), x, 2 (-x), x. States-TPPI phase cycling was applied to phase ϕ_3 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 800$ μ s (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.2

A-(i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $T_C = 4.5$ ms. The phase cycling was: $\phi_1 = 24$ (x), 24 (-x); $\phi_2 = 2$ (30°), 2 (90°), 2 (150°), 2 (210°), 2 (270°), 2 (330°); $\phi_3 = x, -x$; $\phi_4 = 50^\circ$; $\phi_5 = x$; $\phi_6 = 12$ (x), 12 (y), 12 (-x), 12 (-y); $\phi_7 = -y$; $\phi_8 = -x$; $\phi_{\text{rec}} = 3$ (x, 2 (-x), x), 6 (-x, 2 (x), -x), 3 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

A-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_C = 9.3$ ms, $T_N = 11$ ms. The phase cycling was: $\phi_1 = 24$ (x), 24 (-x); $\phi_2 = 2$ (30°), 2 (90°), 2 (150°), 2 (210°), 2 (270°), 2 (330°); $\phi_3 = x, -x$; $\phi_4 = 12$ (x), 12 (y), 12 (-x), 12 (-y); $\phi_5 = -y$; $\phi_6 = -x$; $\phi_{\text{rec}} = 3$ (x, 2 (-x), x), 6 (-x, 2 (x), -x), 3 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_3 . The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.3

TAVI-(i+1)-HSQC (using normal rectangular $^{13}\text{C}^{\alpha\beta}$ 180° pulses for S_1 and S_2), **VIA-(i,i+1)-HSQC** (using a selective 180° REBURP pulse at 26.7 ppm for S_1) and **TA-(i+1)-HSQC** (using a selective 180° REBURP pulse at 68.5 ppm for S_2). The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_1 = 4$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $T_C = 4.5$ ms. The phase cycling was: $\phi_1 = 24$ (x), 24 (-x); $\phi_2 = 2$ (30°), 2 (90°), 2 (150°), 2 (210°), 2 (270°), 2 (330°); $\phi_3 = x, -x$; $\phi_4 = 50^\circ$; $\phi_5 = x$; $\phi_6 = 12$ (x), 12 (y), 12 (-x), 12 (-y); $\phi_7 = -y$; $\phi_8 = -x$; $\phi_{\text{rec}} = 3$ (x, 2 (-x), x), 6 (-x, 2 (x), -x), 3 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

TAVI-(i,i+1)-HSQC (using normal rectangular $^{13}\text{C}^{\alpha\beta}$ 180° pulses for S_1 and S_2), **VIA-(i,i+1)-HSQC** (using a selective 180° REBURP pulse at 26.7 ppm for S_1) and **TA-(i,i+1)-HSQC** (using a selective 180° REBURP pulse at 68.5 ppm for S_2). The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_1 = 4$ ms, $\Delta_C = 9.3$ ms, $T_N = 11$ ms. The phase cycling was: $\phi_1 = 24$ (x), 24 (-x); $\phi_2 = 2$ (30°), 2 (90°), 2 (150°), 2 (210°), 2 (270°), 2 (330°); $\phi_3 = x, -x$; $\phi_4 = 12$ (x), 12 (y), 12 (-x), 12 (-y); $\phi_5 = -y$; $\phi_6 = -x$; $\phi_{\text{rec}} = 3$ (x, 2 (-x), x), 6 (-x, 2 (x), -x), 3 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_3 . The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.4

LA-(i+1)-HSQC. The following delays were used: $\tau_1 = 4$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_1 = 3.6$ ms, $\Delta_2 = 7.1$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $T_C = 4.5$ ms. The phase cycling was: $\phi_1 = 24$ (x), 24 (-x); $\phi_2 = 2$ (30°), 2 (90°), 2 (150°), 2 (210°), 2 (270°), 2 (330°); $\phi_3 = x, -x$; $\phi_4 = 50^\circ$; $\phi_5 = x$; $\phi_6 = 12$ (x), 12 (y), 12 (-x), 12 (-y); $\phi_7 = -y$; $\phi_8 = -x$; $\phi_{\text{rec}} = 3$ (x, 2 (-x), x), 6 (-x, 2 (x), -x), 3 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

LA-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 4$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_1 = 3.6$

ms, $\Delta_2 = 7.1$ ms, $\Delta_C = 9$ ms, $T_N = 11$ ms. The phase cycling was: $\phi_1 = 24$ (x), 24 (-x); $\phi_2 = 2$ (30°), 2 (90°), 2 (150°), 2 (210°), 2 (270°), 2 (330°); $\phi_3 = x, -x$; $\phi_4 = 12$ (x), 12 (y), 12 (-x), 12 (-y); $\phi_5 = -y$; $\phi_6 = -x$; $\phi_{\text{rec}} = 3$ (x, 2 (-x), x), 6 (-x, 2 (x), -x), 3 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_3 . The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.5

S-(i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $T_C = 4.5$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = x, -x$; $\phi_4 = 50^\circ$; $\phi_5 = x$; $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = 4$ (-y), 4 (y); $\phi_8 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 x, -x), 2 (x, 2 (-x), x). States-TPPI phase cycling was applied to ϕ_5 . Gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

S-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau' = 2.25$ ms, $\Delta_C = 9$ ms, $T_N = 11$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = x, -x$; $\phi_4 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_5 = 4$ (-y), 4 (y); $\phi_6 = (-x)$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 x, -x), 2 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_3 . The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.6

FYH-(i+1)-HSQC (carrier frequency $^{13}\text{C}^\gamma = 120$ ppm) and **W-(i+1)-HSQC** (carrier frequency $^{13}\text{C}^\gamma = 120$ ppm). The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 6.7$ ms for FYH and 6.5 ms for W, $\Delta_2 = 4.5$ ms. The phase cycling was as follows: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = 32$ (x), 32 (-x); $\phi_4 = x, -x$; $\phi_5 = 50^\circ$; $\phi_6 = x$; $\phi_7 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_8 = 4$ (-y), 4 (y); $\phi_9 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_6 . The LOS2-0 pulses had a duration of 1024 μs and were applied at 135 ppm (FYH), 105 ppm (W), respectively. The rectangular 180° pulse marked with an asterisk was applied at 45 ppm in the FYH-(i+1)-HSQC and at 30 ppm in the W-(i+1)-HSQC. The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

FYH-(i,i+1)-HSQC (carrier frequency $^{13}\text{C}^\gamma = 120$ ppm) and **W-(i,i+1)-HSQC** (carrier frequency $^{13}\text{C}^\gamma = 120$ ppm). The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$ ms, $\Delta_1 = 6.7$ ms for FYH and 6.5 ms for W, $\Delta_3 = 9$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = 32$ (x), 32 (-x); $\phi_4 = x, -x$; $\phi_5 = 64$ (x), 64 (-x); $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = 4$ (-y), 4 (y); $\phi_8 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 4 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_5 . The same selective pulses as in the FYW-(i+1) and W-(i+1) experiment were used. The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 800 \mu\text{s}$ (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.7

DNG-(i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ ms (not including the pulse length of the LOS2-0 pulse), $\Delta_3 = 4.5$ ms. The phase cycling was as follows: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2(225°), 2(315°); $\phi_3 = x, -x$, $\phi_4 = 50^\circ$; $\phi_5 = x$, $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = 4$ (-y), 4 (y); $\phi_8 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 800$ μ s (28 G/cm), $G_3 = 1$ ms (21 G/cm).

DNG-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ ms (not including the pulse length of the LOS2-0 pulse), $\Delta_4 = 9$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = x, -x$; $\phi_4 = 32$ (x), 32 (-x); $\phi_5 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_6 = 4$ (-y), 4 (y); $\phi_7 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_4 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 800$ μ s (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.8

D-(i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 16$ ms, $\Delta_2 = 4.5$ ms, $\Delta_3 = 4.5$ ms. The phase cycling was as follows: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2(225°), 2(315°); $\phi_3 = x, -x$, $\phi_4 = 50^\circ$; $\phi_5 = x$, $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = 4$ (-y), 4 (y); $\phi_8 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 300$ μ s (17.5 G/cm), $G_3 = 500$ μ s (24.5 G/cm), $G_4 = 600$ μ s (31.5 G/cm), $G_5 = 1$ ms (21 G/cm).

D-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$ ms, $\Delta_1 = 16$ ms, $\Delta_2 = 4.5$ ms, $\Delta_4 = 9$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = x, -x$; $\phi_4 = 32$ (x), 32 (-x); $\phi_5 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_6 = 4$ (-y), 4 (y); $\phi_7 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_4 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 300$ μ s (17.5 G/cm), $G_3 = 500$ μ s (24.5 G/cm), $G_4 = 600$ μ s (31.5 G/cm), $G_5 = 1$ ms (21 G/cm).

chapter 4.4.9

N-(i+1)-HSQC. The following delays were used: $\tau_1 = 5.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ ms, $\Delta_2 = 4.5$ ms, $\Delta_4 = 4.5$ ms. The phase cycling was as follows: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2(225°), 2(315°); $\phi_3 = 310^\circ$, $\phi_4 = 32$ (x), 32 (-x); $\phi_5 = 50^\circ$; $\phi_6 = x, -x$; $\phi_7 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_8 = 4$ (-y), 4 (y); $\phi_9 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_6 . The gradients had the following duration and strength: $G_1 = 800$ μ s (7 G/cm), $G_2 = 800$ μ s (28 G/cm), $G_3 = 1$ ms (21 G/cm).

N-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 5.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$

ms, $\Delta_1 = 8$ ms, $\Delta_2 = 4.5$ ms, $\Delta_5 = 9.3$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = 310^\circ$, $\phi_4 = 32$ (x), 32 (-x); $\phi_5 = x$, -x; $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = 4$ (-y), 4 (y); $\phi_8 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 800$ μs (7 G/cm), $G_2 = 800$ μs (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.10

EQG-(i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 8.5$ ms, $\Delta_2 = 4.5$ ms, $\Delta_3 = 5.5$ ms. The phase cycling was as follows: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2(225°), 2(315°); $\phi_3 = x$, -x, $\phi_4 = 50^\circ$; $\phi_5 = x$, $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = 4$ (-y), 4 (y); $\phi_8 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 800$ μs (28 G/cm), $G_3 = 1$ ms (21 G/cm).

EQG-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ 7.2 ms, $\Delta_2 = 4.5$ ms, $\Delta_4 = 7.1$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = x$, -x; $\phi_4 = 32$ (x), 32 (-x); $\phi_5 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_6 = 4$ (-y), 4 (y); $\phi_7 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_4 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 800$ μs (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.11

E-(i+1)-HSQC and ED-(i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 16$ ms, $\Delta_2 = 4.2$ ms, $\Delta_3 = 4.5$ ms, $\Delta_4 = 9$ ms for the E-(i+1)-HSQC and 4.5 ms for the ED-(i+1)-HSQC, $\Delta_5 = 4.5$ ms for the E-(i+1)-HSQC and 0 ms for the ED-(i+1)-HSQC, $\Delta_6 = 4.5$ ms. The phase cycling was as follows: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2(225°), 2(315°); $\phi_3 = x$, -x, $\phi_4 = 50^\circ$; $\phi_5 = x$, $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = 4$ (-y), 4 (y); $\phi_8 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 300$ μs (17.5 G/cm), $G_3 = 500$ μs (24.5 G/cm), $G_4 = 600$ μs (31.5 G/cm), $G_5 = 1$ ms (21 G/cm).

E-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$ ms, $\Delta_1 = 16$ ms, $\Delta_2 = 4.2$ ms, $\Delta_7 = 9$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = x$, -x; $\phi_4 = 32$ (x), 32 (-x); $\phi_5 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_6 = 4$ (-y), 4 (y); $\phi_7 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_4 . The gradients had the following duration and strength: $G_1 = 1$ ms (7 G/cm), $G_2 = 300$ μs (17.5 G/cm), $G_3 = 500$ μs (24.5 G/cm), $G_4 = 600$ μs (31.5 G/cm), $G_5 = 1$ ms (21 G/cm).

chapter 4.4.12

QN-(i+1)- and Q-(i+1)-HSQC. The following delays were used: $\tau_1 = 5.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ ms, $\Delta_2 = 3.7$ ms for the Q-(i+1)-HSCQ and 4.5 ms for the QN-(i+1)-HSCQ, $\Delta_3 = 4.5$ ms, $\Delta_4 = 9.3$ ms for the Q-(i+1)-HSCQ and 4.5 ms for the QN-(i+1)-HSCQ, $\Delta_5 = 4.8$ ms for the Q-(i+1)-HSCQ and 0 ms for the QN-(i+1)-HSCQ, $\Delta_6 = 4.5$ ms, $\Delta_7 = 9.3$ ms. The phase cycling was as follows: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2(225°), 2(315°); $\phi_3 = 310^\circ$, $\phi_4 = 32$ (x), 32 (-x); $\phi_5 = 50^\circ$; $\phi_6 = x, -x$; $\phi_7 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_8 = 4$ (-y), 4 (y); $\phi_9 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_6 . The gradients had the following duration and strength: $G_1 = 800$ μs (7 G/cm), $G_2 = 800$ μs (28 G/cm), $G_3 = 1$ ms (21 G/cm).

Q-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 5.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$ ms, $\Delta_1 = 8$ ms, $\Delta_2 = 3.7$ ms, $\Delta_3 = 4.5$ ms, $\Delta_7 = 9.3$ ms. The phase cycling was: $\phi_1 = 16$ (x), 16 (-x); $\phi_2 = 2$ (45°), 2 (135°), 2 (225°), 2 (315°); $\phi_3 = 310^\circ$, $\phi_4 = 32$ (x), 32 (-x); $\phi_5 = x, -x$; $\phi_6 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_7 = 4$ (-y), 4 (y); $\phi_8 = -x$; $\phi_{\text{rec}} = 2$ (x, 2 (-x), x), 4 (-x, 2 (x), -x), 2 (x, 2 (-x), x), 2 (-x, 2 (x), -x), 4 (x, 2 (-x), x), 2 (-x, 2 (x), -x). States-TPPI phase cycling was applied to phase ϕ_5 . The gradients had the following duration and strength: $G_1 = 800$ μs (7 G/cm), $G_2 = 800$ μs (28 G/cm), $G_3 = 1$ ms (21 G/cm).

chapter 4.4.13

R-(i+1)-HSQC. The following delays were used: $\tau_1 = 2.05$ ms (determined by pulse length S_1), $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms, $\Delta_3 = 4.5$ ms. The phase cycling was as follows: $\phi_1 = y$; $\phi_2 = x, -x$; $\phi_3 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_4 = 2$ (x), 2 (-x); $\phi_5 = 4$ (x), 4 (-x); $\phi_6 = 50^\circ$; $\phi_7 = 4$ (x), 4 (-x); $\phi_8 = 8$ (x), 8 (-x); $\phi_{\text{rec}} = x, 2$ (-x), x, 2 (-x, 2 (x), -x), x, 2 (-x), x. States-TPPI phase cycling was applied to phase ϕ_7 . The gradients had the following duration and strength: $G_1 = 2$ ms (28 G/cm), $G_2 = 1$ ms (21 G/cm).

R-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 2.05$ ms (determined by pulse length S_1), $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 2.25$ ms, $\delta_3 = 9$ ms, $T_N = 11$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms. The phase cycling was: $\phi_1 = y$; $\phi_2 = x, -x$; $\phi_3 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_4 = 2$ (x), 2 (-x); $\phi_5 = 4$ (x), 4 (-x); $\phi_6 = 4$ (x), 4 (-x); $\phi_7 = 8$ (x), 8 (-x); $\phi_{\text{rec}} = x, 2$ (-x), x, 2 (-x, 2 (x), -x), x, 2 (-x), x. States-TPPI phase cycling was applied to phase ϕ_6 . The gradients had the following duration and strength: $G_1 = 2$ ms (28 G/cm), $G_2 = 1$ ms (21 G/cm).

chapter 4.4.14

K-(i+1)- and KR-(i+1)-HSQC. The following delays were used: $\tau_1 = 2.05$ ms (determined by pulse length S_1), $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms, $\Delta_3 = 8.8$ ms for the K-(i+1)-HSQC and 4.5 ms for the KR-(i+1)-HSQC, $\Delta_4 = 4.3$ ms for the K-(i+1)-HSQC and 0 ms for the KR-(i+1)-HSQC, $\Delta_5 = 4.5$ ms. The phase cycling was as follows: $\phi_1 = y$; $\phi_2 = x, -x$; $\phi_3 = 8$ (x), 8 (y), 8 (-x), 8 (-y); $\phi_4 = 2$ (x), 2 (-x); $\phi_5 = 4$ (x), 4 (-x); $\phi_6 = 50^\circ$; $\phi_7 = 4$ (x), 4 (-x); $\phi_8 = 8$ (x), 8 (-x); $\phi_{\text{rec}} = x, 2$ (-x), x, 2 (-x, 2 (x), -x), x, 2 (-x), x. States-TPPI phase cycling was applied to phase ϕ_7 . The gradients had the following duration and

strength: $G_1 = 2$ ms (28 G/cm), $G_2 = 1$ ms (21 G/cm).

K-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 2.05$ ms (determined by pulse length S_1), $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 2.25$ ms, $\delta_3 = 9$ ms, $T_N = 11$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms. The phase cycling was: $\phi_1 = y$; $\phi_2 = x, -x$; $\phi_3 = 8(x), 8(y), 8(-x), 8(-y)$; $\phi_4 = 2(x), 2(-x)$; $\phi_5 = 4(x), 4(-x)$; $\phi_6 = 4(x), 4(-x)$; $\phi_7 = 8(x), 8(-x)$; $\phi_{\text{rec}} = x, 2(-x), x, 2(-x, 2(x), -x), x, 2(-x), x$. States-TPPI phase cycling was applied to phase ϕ_6 . The gradients had the following duration and strength: $G_1 = 2$ ms (28 G/cm), $G_2 = 1$ ms (21 G/cm).

chapter 4.4.15

Pro($N_{(i-1)}, H^N_{(i-1)}$)-HSQC (t_{1b} is incremented) and **Pro($N_{(i)}, H^N_{(i-1)}$)-experiment** (t_{1a} is incremented). The following delays were used: $\tau_1 = 1.5$ ms, $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 5.5$ ms, $\tau_5 = 2.25$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms, $\Delta_3 = 14.2$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms. The phase cycling was: $\phi_1 = y$; $\phi_2 = 2(y), 2(-y)$; $\phi_3 = y$; $\phi_4 = x$; $\phi_5 = 32(x), 32(y), 32(-x), 32(-y)$; $\phi_6 = x, -x$; $\phi_7 = 16(x), 16(-x)$; $\phi_8 = 4(x), 4(-x)$; $\phi_9 = 8(x), 8(y), 8(-x), 8(-y)$; $\phi_{10} = -y$; $\phi_{11} = -x$; $\phi_{\text{rec}} = (x, 2(-x), x), 2(-x, 2x, -x), (x, 2(-x), x), (-x, 2x, -x), 2(x, 2(-x), x), 2(-x, 2x, -x), 2(x, 2(-x), x), (-x, 2x, -x), (x, 2(-x), x), 2(-x, 2x, -x), (x, 2(-x), x)$. States-TPPI phase cycling was applied to ϕ_4 and ϕ_8 in the Pro($N_{(i-1)}, H^N_{(i-1)}$)-HSQC and the Pro($N_{(i)}, H^N_{(i-1)}$)-experiment, respectively. Gradients had the following duration and strength: $G_1 = 1$ ms (y : 20 G/cm; z : 28 G/cm), $G_2 = 1$ ms (x : 25 G/cm; z : 35 G/cm), $G_3 = (z$: 31.5 G/cm), $G_4 = 900 \mu\text{s}$ (z : 21 G/cm).

Pro($N_{(i+1)}, H^N_{(i+1)}$)-HSQC (t_{1b} is incremented) and **Pro($N_{(i)}, H^N_{(i+1)}$)-experiment** (t_{1a} is incremented). The following delays were used: $\tau_1 = 1.5$ ms, $\tau_2 = 2.1$ ms, $\tau_3 = 5.5$ ms, $\tau_4 = 5.5$ ms, $\tau_5 = 2.25$ ms, $\Delta_1 = 3.5$ ms, $\Delta_2 = 5.5$ ms, $\Delta_4 = 7$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms. The phase cycling was: $\phi_1 = y$; $\phi_2 = 2(y), 2(-y)$; $\phi_3 = y$; $\phi_4 = x$; $\phi_5 = 32(x), 32(y), 32(-x), 32(-y)$; $\phi_6 = x, -x$; $\phi_7 = 16(x), 16(-x)$; $\phi_8 = 4(x), 4(-x)$; $\phi_9 = 8(x), 8(y), 8(-x), 8(-y)$; $\phi_{10} = -y$; $\phi_{11} = -x$; $\phi_{\text{rec}} = (x, 2(-x), x), 2(-x, 2x, -x), (x, 2(-x), x), (-x, 2x, -x), 2(x, 2(-x), x), 2(-x, 2x, -x), 2(x, 2(-x), x), (-x, 2x, -x), (x, 2(-x), x), 2(-x, 2x, -x), (x, 2(-x), x)$. States-TPPI phase cycling was applied to ϕ_4 and ϕ_8 in the Pro($N_{(i+1)}, H^N_{(i+1)}$)-HSQC and the Pro($N_{(i)}, H^N_{(i+1)}$)-experiment, respectively. Gradients had the following duration and strength: $G_1 = 1$ ms (y : 20 G/cm; z : 28 G/cm), $G_2 = 1$ ms (x : 25 G/cm; z : 35 G/cm), $G_3 = (z$: 31.5 G/cm), $G_4 = 900 \mu\text{s}$ (z : 21 G/cm).

chapter 4.4.16

CS-(i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $\delta_0 = 4.5$ ms, $\delta_1 = 6.9$ ms, $\delta_2 = 11.4$ ms, $T_N = 11$ ms, $\Delta_1 = 5.62$ ms, $\Delta_2 = 1.68$ ms, $\Delta_3 = 1.53$ ms, $\Delta_4 = 5.77$ ms, $T_C = 3.7$ ms. The phase cycling was: $\phi_1 = 16(x), 16(-x)$; $\phi_2 = 2(45^\circ), 2(135^\circ), 2(225^\circ), 2(315^\circ)$; $\phi_3 = x, -x$; $\phi_4 = 50^\circ$; $\phi_5 = x$; $\phi_6 = 8(x), 8(y), 8(-x), 8(-y)$; $\phi_7 = 4(-y), 4(y)$; $\phi_8 = -x$; $\phi_{\text{rec}} = 2(x, 2(-x), x), 4(-x, 2x, -x), 2(x, 2(-x), x)$. States-TPPI phase cycling was applied to ϕ_5 . Gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (28 G/cm), $G_2 = 1$ ms (21 G/cm).

CS-(i,i+1)-HSQC. The following delays were used: $\tau_1 = 3.5$ ms, $\tau_2 = 5.5$ ms, $\tau_3 = 2.25$ ms, $T_N = 11$ ms, $\Delta_1 = 5.62$ ms, $\Delta_2 = 1.68$ ms, $\Delta_3 = 1.53$ ms, $\Delta_4 = 5.77$ ms, $\Delta_C = 9$ ms. The phase cycling was: $\phi_1 = 16(x), 16(-x)$; $\phi_2 = 2(45^\circ), 2(135^\circ), 2(225^\circ), 2(315^\circ)$; $\phi_3 = x, -x$; $\phi_4 = 8(x), 8(y), 8(-x), 8(-y)$;

$\phi_5 = 4 (-y), 4 (y); \phi_6 = (-x); \phi_{\text{rec}} = 2 (x, 2 (-x), x), 4 (-x, 2 x, -x), 2 (x, 2 (-x), x)$. States-TPPI phase cycling was applied to phase ϕ_3 . The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (28 G/cm), $G_2 = 1 \text{ ms}$ (21 G/cm).

chapter 4.4.18

Trp $\text{H}^{\text{e1}}\text{N}^{\text{e1}}$ -HSQC correlation. The delays are set to $\tau_2 = 5.5 \text{ ms}$, $\tau_3 = 2.25 \text{ ms}$, $T_N = 11 \text{ ms}$. The phase cycling was: $\phi_1 = y; \phi_2 = 8 (x), 8 (-x); \phi_3 = y; \phi_4 = x, -x; \phi_5 = 4 (x), 4 (-x); \phi_6 = 2 (x), 2 (-x); \phi_7 = -y; \phi_8 = -x; \phi_{\text{rec}} = x, 2 (-x), x, 2 (-x, 2 (x), -x), x, 2 (-x), x$. States-TPPI phase cycling was applied to phase ϕ_6 . The carrier frequency for ^{13}C was centered at 125 ppm. The gradients had the following duration and strength: $G_1 = 800 \mu\text{s}$ (28 G/cm), $G_2 = 1 \text{ ms}$ (21 G/cm).

Experimental part for chapter 6

Sample preparation

The expression and purification of all samples were carried out in the group of Prof. Oesterhelt at the Max Planck Institut of Biochemistry in Martinsried. Two labeling patterns have been used - uniform $^2\text{H}/^{15}\text{N}$ - and uniform $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeling. The $^2\text{H}/^{15}\text{N}$ labeled sample was prepared as described by Patzelt et al., 1997. For the preparation of the $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled sample the method was modified slightly and is described by Kolbe, 2001. 100% D_2O , $^{15}\text{NH}_4\text{Cl}$ and $^{13}\text{CO}_2$ were used as solely sources of hydrogen, carbon and nitrogen to produce the $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled peptone with *Scenedesmus obliquus*. The peptone was the basis of the medium on which halobacterium salinarum was grown. The retinal negative strain JW5 was used. Unlabeled retinal (^1H and ^{12}C) was biosynthetically incorporated. The protein was solubilized with dodecylmaltoside (with a deuterated dodecyl moiety). The final samples contained 12-15 mg/ml bacteriorhodopsin, 1% partially deuterated DM, 95 % H_2O and 5 % D_2O . The pH was adjusted to 5.6 using a phosphate buffer (25 mM Na/KP_i). Exchangeable deuterons are replaced by protons. Only amide groups and the retinal contain protons. However, the deuterium/proton exchange of the amide groups might not be complete especially in the rigid helical parts of the protein which leads to missing signals.

NMR data acquisition

The spectra were recorded on a 750 MHz spectrometer (Bruker-DMX750) in standard configuration using an inverse triple resonance probe equipped with three-axis self shielded gradient coils. The data were collected at 323.5 K. For both samples 5 mm ultra-precision sample tubes were used.

Proton hard pulses were applied with 25 kHz field strength. Pulses applied at the $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\alpha\beta}$ or ^{13}CO resonance frequencies were adjusted to provide a null at the corresponding ^{13}CO or $^{13}\text{C}^{\alpha}$ frequencies. The rectangular $^{13}\text{C}^{\alpha}$ 90° and 180° pulses were set to 43 μs and 86 μs , respectively. The rectangular $^{13}\text{C}^{\alpha\beta}$ 90° and 180° pulses were set to 40 μs and 35 μs , respectively. The rectangular ^{13}CO 90° and 180° pulses were set to 43 μs and 86 μs , respectively. The gradients were applied as a

sinusoidal function from 0 to π . Unless indicated otherwise, the carrier frequencies were centered at $^1\text{H} = 4.5$ ppm, $^{15}\text{N} = 120$ ppm, $^{13}\text{C}^{\alpha/\beta} = 45$ ppm, $^{13}\text{C}^{\alpha} = 55$ ppm and $^{13}\text{C}' = 175$ ppm.

The uniformly $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -labeled sample was used to record triple-resonance experiments with the parameters given in table A1. A measurement time of approximately 7 days was used for all spectra, except HNCO.

NMR triple-resonance experiments	data size * [pts] $t_1 \times t_2 \times t_3$	spectral width [Hz] $\omega_1 \times \omega_2 \times \omega_3$	ns ‡
HNCA (TROSY)	50×62×1024	6250×3788×12500	24
HNCO (TROSY)	64×64×1024	3012×3788×12500	16
HNCACB (TROSY)	50×60×1024	12500×3788×12500	24
HN(CA)CO (TROSY)	50×60×1024	3012×3788×12500	24
HN(CO)CA (TROSY)	50×62×1024	6250×3788×12500	24
HN(CO)CACB (TROSY)	56×62×1024	12500×3788×12500	24

Table A1: Acquisition parameters of the triple-resonance experiments applied to the $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ labeled sample. All experiments were recorded at 750 MHz and 323.5 K. * The number of complex points of each dimension is given. ‡ The number of scans for one increment is given.

The uniformly $^2\text{H}/^{15}\text{N}$ -labeled sample was used to record three 3D-NOESY experiments with the parameters given in table A2.

NOESY experiments	data size * [pts] $t_1 \times t_2 \times t_3$	spectral width [Hz] $\omega_1 \times \omega_2 \times \omega_3$	ns ‡	mixing time [ms]
HNH-NOESY-TROSY	128×80×1024	10000×3788×12500	8	80
HNH-NOESY-HSQC	80×128×1024	10000×3788×12500	8	80
NNH-HMQC-NOESY-HSQC	64×64×1024	3788×3788×12500	24	50

Table A2: Acquisition parameters of the triple-resonance experiments applied to the $^2\text{H}/^{15}\text{N}$ labeled sample. All experiments were recorded at 750 MHz and 323.5 K. * The number of complex points of each dimension is given. ‡ The number of scans for one increment is given.

A series of experiments have been recorded for T_1 and T_2 measurement (without TROSY). The spectra were recorded at 600 MHz (Bruker-DRX600). A series of experiments using relaxation delays of 5, 25, 50, 75, 100, 150, 200, 300, 400, 1000, 2500 ms were

recorded for the T_1 determination. Relaxation delays of 5, 10, 20, 30, 40, 50, 100, 150, 200, 250, 300 ms were used for the T_2 determination.

References

- L. Emsley and G. Bodenhausen, *Chem. Phys. Lett.* **165**, 469-476 (1990).
- M. Kolbe, Dissertation, LMU Munich (2001).
- H. Patzelt, A. S. Ulrich, H. Egbringhoff, P. Dux, J. Ashurst, B. Simon, H. Oschkinat, D. Oesterhelt, *J. Biomol. NMR* **10**, 95-106 (1997).
- A. J. Shaka, J. Keeler, T. Frenkiel and R. Freeman, *J. Magn. Reson.* **52**, 335-338 (1983).
- A. J. Shaka, P. B. Barker and R. Freeman, *J. Magn. Reson.* **64**, 547-552 (1985).
- V. Sklenar, M. Piotto, R. Leppik and V. Saudek, *J. Magn. Reson. A* **102**, 241-245 (1993).

Publications

- (8) **M. Schubert**, H. Oschkinat, and P. Schmieder, MUSIC and Aromatic Residues: Amino Acid Type-Selective ^1H - ^{15}N Correlations, III, *J. Magn. Reson.* **153**, 186-192 (2001).
- (7) **M. Schubert**, H. Oschkinat, and P. Schmieder, Amino acid type-selective ^1H - ^{15}N -correlation's for Arg and Lys, *J. Biomol. NMR* **20**, 379-384 (2001).
- (6) **M. Schubert**, H. Oschkinat, and P. Schmieder, MUSIC, Selective Pulses and Tuned Delays: Amino Acid Type-Selective ^1H - ^{15}N Correlations, II, *J. Magn. Reson.* **148**, 61-72 (2001).
- (5) A. Schulz, K. Bruns, P. Henklein, G. Krause, **M. Schubert**, T. Gudermann, V. Wray, G. Schultz, and T. Schöneberg, Requirement of Specific Intrahelical Interactions for Stabilizing the Inactive Conformation of Glycoprotein Hormone Receptors, *J. Biol. Chem.* **275**, 37860-37869 (2000).
- (4) **M. Schubert**, L. J. Ball, H. Oschkinat, and P. Schmieder, Bridging the gap: A set of selective ^1H - ^{15}N -correlations to link sequential neighbors of prolines, *J. Biomol. NMR* **17**, 331-335 (2000).
- (3) **M. Schubert**, M. Smalla, P. Schmieder, and H. Oschkinat, MUSIC in Triple-Resonance Experiments: Amino Acid Type-Selective ^1H - ^{15}N Correlations, *J. Magn. Reson.* **141**, 34-43 (1999).
- (2) M. Edgar, **M. Schubert**, H. H. Limbach, and C. G. Göltner, NMR Investigations of Nanostructured Ceramics, *Ber. Bunsenges. Phys. Chem.* **101**, 1769-1772 (1997).
- (1) S. Kleinhenz, **M. Schubert**, and K. Seppelt, Preparation and Structures of Hexaphenyl- and Hexatolyltantalates(V), *Chem. Ber./Recueil* **130**, 899-902 (1997).

Curriculum Vitae

Name	Mario Schubert
Date of Birth	25.06.1972
Place of Birth	Chemnitz, Germany
Nationality	German

Education

Mar 1998 - Present	Forschungsinstitut für Molekulare Pharmakologie (FMP) Berlin, PhD thesis in the group of Prof. Dr. H. Oschkinat
Apr 1999 - Jun 1999	Novartis Pharma AG Basel Internship, Dr. W. Jahnke
Feb 1998	Freie Universität Berlin "Diplom" in Chemistry
Jun 1997 - Jan 1998	Freie Universität Berlin and Consejo Superior de Investigaciones Científicas, Madrid Diploma thesis "Synthesis of ^{15}N 3,5-Dimethylpyridine and Solid State ^{15}N -NMR Study of the Hydrogen Bonded Complex formed with 3,5-Dinitrobenzoic Acid", Prof. Dr. H. H. Limbach and Dr. J. Elguero
Okt 1994 - Feb 1998	Freie Universität Berlin Diplom-program Chemistry
Okt 1992 - Jul 1994	Universität Heidelberg Undergraduate studies in Chemistry

