Aspects and Applications of Pulse Sequence Design for Solution-State Nuclear Magnetic Resonance Spectroscopy

Peter Würtz

Institute of Biotechnology and
Laboratory of Physical Chemistry
Department of Chemistry
Faculty of Science
University of Helsinki

ACADEMIC DISSERTATION

To be presented with the permission of
the Faculty of Science of the University of Helsinki
for public criticism in Auditorium 1041 of Biocenter 2 (Viikinkaari 5)
on November 7th 2008 at 12 noon

Helsinki 2008

Supervisor:

Dr. Perttu Permi
Institute of Biotechnology
University of Helsinki
Finland

Pre-examiners:

Prof. Jukka Jokisaari
Department of Physics
University of Oulu
Finland

Prof. Reino Laatikainen
Department of Chemistry
University of Kuopio
Finland

Opponent:

Prof. Michael Sattler
Department Chemie
Technische Universität München
Germany

Custos:

Prof. Markku Räsänen
Department of Chemistry
University of Helsinki
Finland

ISSN 1795-7079
ISBN 978-952-92-4571-0
ISBN 978-952-10-5023-7 (PDF)
http://ethesis.helsinki.fi
Yliopistopaino
Helsinki 2008

Abstract

NMR spectroscopy enables the study of biomolecules from peptides and carbohydrates to proteins at atomic resolution. The technique uniquely allows for structure determination of molecules in solution-state. It also gives insights into dynamics and intermolecular interactions important for determining biological function.

Detailed molecular information is entangled in the nuclear spin states. The information can be extracted by pulse sequences designed to measure the desired molecular parameters. Advancement of pulse sequence methodology therefore plays a key role in the development of biomolecular NMR spectroscopy.

A range of novel pulse sequences for solution-state NMR spectroscopy are presented in this thesis. The pulse sequences are described in relation to the molecular information they provide. The pulse sequence experiments represent several advances in NMR spectroscopy with particular emphasis on applications for proteins. Some of the novel methods are focusing on methyl-containing amino acids which are pivotal for structure determination. Methyl-specific assignment schemes are introduced for increasing the size range of ^{13}C,^{15}N labeled proteins amenable to structure determination without resolving to more elaborate labeling schemes. Furthermore, cost-effective means are presented for monitoring amide and methyl correlations simultaneously.

Residual dipolar couplings can be applied for structure refinement as well as for studying dynamics. Accurate methods for measuring residual dipolar couplings in small proteins are devised along with special techniques applicable when proteins require high pH or high temperature solvent conditions.

Finally, a new technique is demonstrated to diminish strong-coupling induced artifacts in HMBC, a routine experiment for establishing long-range correlations in unlabeled molecules. The presented experiments facilitate structural studies of biomolecules by NMR spectroscopy.

Keywords: NMR, methyl assignment, time shared evolution, epitope mapping, residual dipolar coupling, strong coupling, HMBC

List of Original Publications

I P. Würtz, M. Hellman, H. Tossavainen, and P. Permi. Towards unambiguous assignment of methyl-containing residues by double and triple sensitivity-enhanced HCCmHm-TOCSY experiments.
J. Biomol. NMR **36**, 13-26 (2006).

II P. Würtz*, O. Aitio*, M. Hellman, and P. Permi. Simultaneous detection of amide and methyl correlations using a time shared NMR experiment: application to binding epitope mapping.
J. Biomol. NMR **39**, 97-105 (2007).

* Equal author contribution.

III P. Würtz and P. Permi. SESAME-HSQC for simultaneous measurement of NH and CH scalar and residual dipolar couplings.
Magn. Reson. Chem. **45**, 289-295 (2007).

IV P. Würtz, K. Fredriksson, and P. Permi. A set of HA-detected experiments for measuring scalar and residual dipolar couplings.
J. Biomol. NMR **31**, 321-330 (2005).

V P. Würtz, P. Permi, N.C. Nielsen, and O.W. Sørensen. *Clean* HMBC: Suppression of strong-coupling induced artifacts in HMBC spectra.
J. Magn. Reson. **194**, 89-98 (2008).

List of Abbreviations

BIRD	Bilinear Rotation Decoupling
CSA	Chemical Shift Anisotropy
COSY	COrrelation SpectroscopY
CT	Constant-Time
DD	Dipole-Dipole
E.COSY	Exclusive COrrelation SpectroscopY
FID	Free Induction Decay
H2BC	Heteronuclear 2-Bond Correlation
HMBC	Heteronuclear Multiple Bond Correlation
HSQC	Heteronuclear Single Quantum Correlation
INEPT	Insensitive Nuclei Enhanced by Polarization Transfer
IPAP	In-Phase Anti-Phase
LPJF	Low Pass J Filter
MBP	Maltose Binding Protein
MD	Molecular Dynamics
NMR	Nuclear Magnetic Resonance
NOE(SY)	Nuclear Overhauser Effect (SpectroscopY)
PDB	Protein Data Bank
RDC	Residual Dipolar Coupling
rf	Radiofrequency
TOCSY	TOtal Correlation SpectroscopY
TPPI	Time Proportional Phase Incrementation
TROSY	Transverse Relaxation Optimized SpectroscopY
SAXS	Small-Angle X-ray Scattering
SESAME-HSQC	Spin statE Selective All Multiplicity Edited-HSQC
SW	Signal Width

Contents

Chapter 1

Introduction

Biological molecules are the building blocks of life. The molecules can take different kinds of structural forms depending on their biological function. In order to understand the molecular mechanisms of biomolecules it is important to know their structure at atomic resolution. As the molecules are too small to be seen using microscopy indirect methods are required. Nuclear magnetic resonance (NMR) spectroscopy provides means to study a wide range of molecules by linking the fine-structured energies of the nuclei to the structure of the molecule as a whole. Careful combination of the spectral information allows the determination of three-dimensional structures of proteins and other macromolecules. In addition, dynamical characterization of molecules as well as information on site-specific molecular interactions are obtainable by NMR spectroscopy. In contrast to X-ray crystallography, the main alternative for structural determination at atomic resolution, the molecular parameters can be measured in solution state which ensures a native-like environment for soluble molecules. For carbohydrates and lipids NMR spectroscopy is an indispensable analytical tool to understand their structure, function and biosynthesis.

The intricate information available in NMR spectra must be disentangled to provide the required structural parameters. This is achieved by manipulating the nuclear spins to separate the information into manageable parts, which are subsequently combined as jigsaw puzzle pieces towards the structure determination. The spin states can be tailored to reveal the desired parameters by designing pulse sequences, the topic of this study. This thesis presents several advances in pulse sequence design for solution state NMR spectroscopy which facilitate structural analysis and studies of protein-ligand interactions. The em-

phasis is on protein NMR methodology, but an important improvement to a routine experiment for small molecules is also presented.

Structure determination of proteins by NMR spectroscopy is currently fairly routine for small proteins. Difficulties inherent to the procedure, however, grow rapidly with increasing molecular size due to spectral crowding and broader lineshapes. These aspects complicate the assignment process *i.e.*, linking a given resonance frequency to the corresponding nucleus in the molecule. The increase in complication is particularly severe for side chain assignment, which is commonly a requirement for structure determination. Experiments introduced in Article I provide a powerful strategy to overcome this problem without the need of protein perdeuteration.

An attractive feature of NMR spectroscopy is the ease with which it can be used for studying molecular interactions such as ligand binding sites of proteins. This is likely to play an increasing role in proteomics as focus shifts from static protein structures towards protein function and interactions. Article II describes a cost-effective methodology where methyl and amide correlations are simultaneously detected for this purpose.

Advances in NMR methodology continue to pave the way for extending the molecular size limit, and facilitate and automate spectral analysis. Pulse sequence design plays a vital role in this respect. Pulse sequence development often goes hand in hand with hardware advances or the introduction of novel labeling protocols. For instance, the introduction of weak alignment for inducing residual dipolar couplings (RDCs) has triggered the development of a myriad of experiments for coupling measurements in proteins (Tjandra and Bax, 1997). This allows for local and long-range structure refinement and, in addition, provides valuable dynamical information. Articles III and IV present accurate and versatile methods for measuring several scalar coupling constants and RDCs.

Structure elucidation of molecules at natural isotope abundance requires somewhat different spectroscopic techniques than those used for protein NMR spectroscopy. The molecules studied are often smaller than proteins. Small molecules tumble faster in solution, thus spin relaxation is of less concern in this case. However, spectral assignment is more complicated since it is necessary to use multiple-bond scalar couplings for coherence transfer, instead of solely one- and two-bond scalar couplings as in the case of protein NMR. Long-range correlations link the chemical shift information from direct correlations by providing hinges for the pieces in the spectral analysis puzzle. Article V describes an improved technique for obtaining artifact free long-range correlations spectra without compromising sensitivity.

The thesis is organized as follows. The following chapter introduces the

basic NMR phenomenon and outlines the spectroscopic methods for protein as well as small molecule structure determination. The goals set for the study are summarized in the subsequent chapter. I proceed to a discussion of each of the experiment types introduced and present additional relevant material which did not find its way into the publications. I conclude with remarks on the applicability of the presented methods. Detailed descriptions of the experiments including coherence flows and experimental demonstrations are presented in the appended articles.

Chapter 2

Nuclear Magnetic Resonance

The energy of a spin in a nucleus splits into separate levels when a molecule experiences a magnetic field. NMR spectroscopy is about finding the energy it takes for a spin to flip between these levels. The energy differences between the levels can in turn be related to the molecular structure in the neighborhood of the nucleus. NMR is thus a nuclear phenomenon, but it provides information on the molecule as a whole.

A charged spinning particle constitutes a magnetic dipole. The associated magnetic dipole moment $\boldsymbol{\mu}$ is proportional to the angular momentum of the particle. By analogy, a particle with a spin has a magnetic dipole moment which is proportional to its spin angular momentum $\boldsymbol{I}$,

$$\boldsymbol{\mu} = \gamma \boldsymbol{I} \tag{2.1}$$

with γ being the gyromagnetic ratio. When placed in a magnetic field $\boldsymbol{B}$, the magnetic dipole experiences a torque which tends to align it in the direction of the magnetic field. This interaction gives rise to an energy with an associated Zeeman Hamiltonian

$$H_{\text{Zeeman}} = -\boldsymbol{\mu} \cdot \boldsymbol{B} = -\mu_z B_0 = -\gamma B_0 I_z = \omega_0 I_z. \tag{2.2}$$

The magnetic field is by convention taken to be along the z axis and $\omega_0 = -\gamma B_0$ is defined as the Larmor frequency. The magnetic resonance phenomenon arises

because of this interaction between the nuclear magnetic dipole moment and the external magnetic field.

For a spin $\frac{1}{2}$ particle the Zeeman energy splits into two states when the particle is in a magnetic field. The spin angular momentum vector is, except from a factor of $\hbar/2$, composed of the Pauli spin matrices. The corresponding eigenstates of H_{Zeeman} are denoted $|\alpha>$ (spin up) and $|\beta>$ (spin down) with eigenfrequencies $\omega_0/2$ and $-\omega_0/2$, respectively. The general solution Ψ to the time-dependent Schrödinger equation is a superposition of the eigenstates

$$|\Psi(t)>= ae^{-i\omega_0 t/2}|\alpha> + be^{i\omega_0 t/2}|\beta> . \tag{2.3}$$

The complex constants a and b can be determined from the initial conditions and fulfills $|a|^2 + |b|^2 = 1$. Assuming a and b are real we can write $a = \cos(\theta/2)$ and $b = \sin(\theta/2)$. By calculating the expectation values of the magnetic moment we get

$$<\mu_x>= \tfrac{\hbar}{2}\gamma \sin\theta \cos(-\omega_0 t) \quad <\mu_y>= -\tfrac{\hbar}{2}\gamma \sin\theta \sin(-\omega_0 t)$$

$$<\mu_z>= \tfrac{\hbar}{2}\gamma \cos\theta = \tfrac{\hbar}{2}\gamma(a^2 - b^2). \tag{2.4}$$

These results imply that the magnetic moment vector is tilted at an angle θ from the z-axis, and precess around the static magnetic field at the negative of the Larmor frequency $-\omega_0 = \gamma B_0$. The magnitude of this precessional frequency is equal to the excitation frequency where the transition rate between the α and β states is maximum.

The time-dependence of the x and y components of the magnetic moment cause the ensemble average of these expectation values to vanish. In contrast, the z-component is time-independent and the ensemble average can be rewritten as $<\overline{\mu_z}> = \hbar/2\gamma(P_\alpha - P_\beta)$, where P_α and P_β are the populations associated with the $|\alpha>$ and the $|\beta>$ states, respectively. The difference in populations of the equilibrium states is denoted as the polarization. In the case of no external field interaction the bulk magnetization is proportional to the polarization. The magnetization can thus be depicted as precessing around the static field.

According to Faraday's law of induction a time-varying magnetic field induces an electric field in a coil. At equilibrium the bulk magnetization is parallel to the static field so no signal is generated in the coil. To create a signal the magnetization must be brought to the transversal plane by generating some non-vanishing phase relation between the superposition of eigenstates in the ensemble. This is known as coherence and accomplished by applying a magnetic field orthogonal to the static magnetic field in form of a radiofrequency (rf)

pulse oscillating at the magnitude of the Larmor frequency. The effect of an on-resonance 90° pulse is to equalize the probabilities of the $|\alpha>$ and $|\beta>$ states and hereby rotate the axis along which the magnetization is aligned. By applying the pulse to all spins simultaneously the bulk magnetization is transferred to the transversal plane and the subsequently induced current can be measured. After Fourier transforming the resulting signal it is possible to extract the desired energy differences between the spin states as spectral frequencies.

Interacting spins The power of NMR spectroscopy lies in the detailed and versatile relation between the spectral energies and molecular properties. The actual precession frequency of magnetization ω is sensitive to slight perturbations from the static Larmor frequency ω_0. The variation in resonance frequencies is called chemical shift. The chemical shifts are caused by the electronic charges around the nuclei which lead to local shielding of the external magnetic field in the neighborhood of the nuclei. The delicate sensitivity of chemical shifts offer the possibility to distinguish spins in different chemical environments.

The α and β spin states are coupled to other spins on separated nuclei. This is mediated through the scalar coupling J which introduces fine structure in the energy levels and allows for coherence transfer between spins. For two interacting spins A and B the spin Hamiltonian takes the form

$$H_{\mathrm{AB}} = \omega_{\mathrm{A}} A_z + \omega_{\mathrm{B}} B_z + 2\pi J_{\mathrm{AB}} \mathbf{A} \cdot \mathbf{B}. \tag{2.5}$$

For most cases encountered in highfield solution state NMR spectroscopy the off-diagonal terms $2\pi J_{\mathrm{AB}}(A_x B_x + A_y B_y)$ can be neglected. This simplification is known as the weak coupling limit and holds as a good approximation when $|\omega_A - \omega_B| >> 2\pi J_{\mathrm{AB}}$. In the weak coupling limit all terms within the Hamiltonian commute, and the product basis of Cartesian rotations is identical to the eigenbasis (Sørensen *et al.*, 1983). Invoking the weak coupling approximation drastically simplifies analysis of pulse sequences and reproduces most experimental results quantitatively well. However, as shown in Article V salient features of coherence transfer by *rf* pulses can lead to undesired effects in the spectra which may be overlooked when neglecting the strong coupling interaction.

The Hamiltonian for multiple spin systems is additive and of similar form as Eq. 2.5 when relaxation effects are not taken into account. Electrons effectively shield the nuclei from molecular collisions so the coherent spin states can have rather long lifetimes which make them easy to manipulate and observe. This is accomplished by modifying the Hamiltonian in a desired manner by a series of *rf* pulses. The present thesis describes some new methods and improvements

of pulse sequences for obtaining molecular information of proteins and small molecules using this methodology.

2.1 Protein NMR Spectroscopy

Proteins are essential parts of the organism and take part in all cellular processes. Proteins perform a vast variety of functions such as catalyzing biochemical reactions, constituting the cytoskeleton, being molecular transporters, and receptors for cell signaling. Biological functions and activities depend on protein fold, the specific semi-rigid spatial configuration they adopt. Thus to understand protein function and molecular interactions, it is essential to know the three-dimensional structures of proteins.

Structure determination of proteins on atomic scale is feasible by solution state NMR spectroscopy. The technique provides a complementary alternative to protein structure determination by X-ray crystallography and NMR determined structures currently constitute $\sim$15% of the depositions in the Protein Data Bank (PDB). There is, however, a strong over representation of small structures in the PDB due to intrinsic complications associated with structure determination of large proteins. The main reasons are increased signal overlap and faster relaxation rates caused by the longer rotational correlation times τ_c of larger proteins. The combination of broad lines and numerous resonances in the same spectral range complicate the assignment process and the determination of long-distance restraints obtained from NOESY spectra.

The first solution state protein structures were determined by homonuclear spectroscopy (Williamson *et al.*, 1985; Kaptein *et al.*, 1985). In these pioneering efforts the resonance assignment was conducted by identifying spins belonging to a specific amino acid type and subsequently mapping the spins onto the primary sequence by identifying adjacent residues from analysis of NOESY spectra. Long-range distance restraints were obtained in an iterative process and incorporated into structure calculation by distance geometry methods or molecular dynamics (MD). The complexity of the approach puts severe limitations on the maximum protein size amenable for structure determination.

Since these first landmarks continuous methodological progress has established NMR spectroscopy as a central tool in structural biology. Advances in magnet and probe design, pulse sequences, data analysis, and computational methods in addition to development in molecular biology methods facilitate studies of ever more complex systems.

Especially the introduction of isotopic labeling schemes, which enable hete-

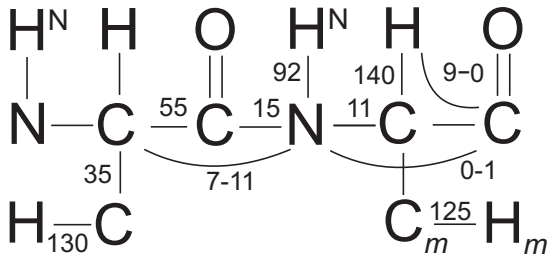


Figure 2.1: Peptide bond structure with scalar coupling constants given in Hz. m indicates methyl group.

ronuclear triple resonance experiments has altered the possibilities of answering interesting biological questions. Overlap problems in homonuclear methods are hereby circumvented by separating the signals into independent dimensions which yield higher signal dispersion. Triple resonance experiments rely solely on one- and two-bond scalar couplings for coherence transfer. These couplings are largely independent of protein conformation, in contrast to the three-bond scalar couplings required for homonuclear spectroscopy. Sequential assignment is feasible in unambiguous manners using the redundant information provided by the independent coherence transfer pathways.

The scalar coupling constants used for magnetization transfer in the peptide backbone are indicated in Fig. 2.1 along with some of the scalar coupling constants measured in Articles III and IV. The relatively large magnitudes of ${}^1J$ and ${}^2J$ coupling constants ensure effective coherence transfer which is particularly favorable for larger proteins.

The sensitivity of triple resonance experiments becomes problematic for high molecular-weight systems, ~ 30 kDa and above, due to increased transverse relaxation rates $R_2 = 1/T_2$. Relaxation times T_2 can be prolonged substantially using deuterated samples and hereby enable studies of larger biological systems. The main source of relaxation, the dipole-dipole relaxation between protons directly bound to carbon or nitrogen spins, is scaled by a factor of 0.02, due to the much smaller value of the deuterium gyromagnetic ratio. In addition, proton spin diffusion is drastically limited for fully or even partly deuterated samples (Sattler and Fesik, 1996). The procedure has some drawbacks though. Unfolding and refolding of deuterated proteins may be necessary for H/D exchange to occur for amide protons buried in the hydrophobic core of the protein. In

addition, deuteration is mostly beneficial for backbone assignment experiments since the proton density in side chains is lowered by fractional deuteration. Most importantly, random fractional deuteration cause ^{2}H isotope shifts leading to significant linebroadening which counterbalances the improved sensitivity and resolution.

The problem of increasing linewidth due to slower molecular tumbling with increasing molecular size is partly alleviated by exploiting the TROSY effect (Pervushin *et al.*, 1997). This technique selects the most slowly relaxing spin multiplet where the chemical shift anisotropy (CSA) and dipole-dipole (DD) interactions counterbalance each other. TROSY is close to optimal for NH moieties where the principal symmetry axis of the CSA tensor is almost parallel to the N-H dipole vector. Although the size of the ^{1}H CSA is ca. 10 times smaller than the ^{15}N CSA in NH, this is balanced by the higher gyromagnetic ratio. Thus TROSY linenarrowing is effective in both dimensions for similar magnetic field strengths, which happen to be optimal around 1 GHz Larmor frequency. The TROSY effect is most pronounced when combined with deuterated proteins since remote protons account for 95% of the residual transverse relaxation of amide ^{1}H and 75% of residual transverse relaxation on ^{15}N (Permi and Annila, 2004). TROSY pulse elements have been incorporated into all important triple-resonance experiments for sequential assignment (Salzmann *et al.*, 1998). Additional spectral simplification for studying high-molecular weight systems can be achieved by specific labeling, in particular of methyl groups, as discussed in Chapter 4.1. The specific labeling methods can be combined with methyl TROSY techniques to study side chains of large supramolecular systems (Sprangers and Kay, 2007b).

Although the procedure of protein structure determination is well established spectral analysis often requires a good deal of manual intervention. The main steps towards the three-dimensional structure are outlined in Fig. 2.2 and explained in further detail below. Importantly, not the whole structure determination process is necessarily required for characterization of protein dynamics or studies of intermolecular interactions. The latter topic is briefly described in the subsequent section.

2.1.1 Structure Determination

NMR is not like a microscope that would directly produce an image of a protein with atomic resolution. Instead, the technique yields indirect information in terms of structural restraints from which a three-dimensional structure can be modeled. The first step in structural analysis is the resonance assignment pro-

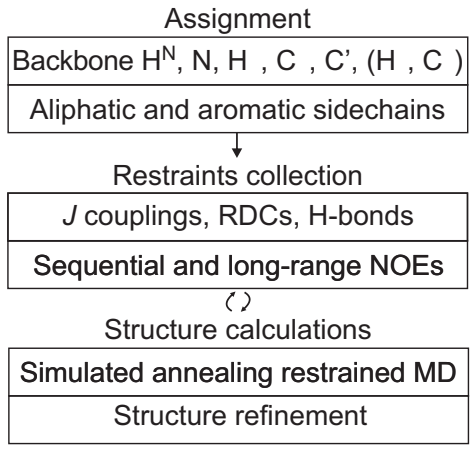


Figure 2.2: Schematic indication of the main steps in the structure determination protocol for ^{13}C,^{15}N,(^{2}H) labeled proteins. In automated structure calculation protocols collection of NOE restraints and the structure determination step are combined in iterative loops.

cess. This determines which frequency, or energy difference, corresponds to a given nucleus in the molecule. The procedure is divided into two separate steps: backbone and side chain assignment. Backbone assignment is rather straightforward for smaller folded proteins using uniformly double labeled samples and the technique readily lends itself to automation. A typical set of sequential assignment experiments is shown in Fig. 2.3a. All backbone assignment experiments link the information of carbon resonances to the amide NH correlations, which provide high resolution and resonate far from the water signal.

The most common strategy for sequential backbone assignment is to link resonances from the intraresidue carbon and the preceding residue carbon to the NH correlation in each amino acid. The two resonance types can be distinguished by measuring both resonances in one experiment and solely the preceding residue in a complementary experiment (Ikura *et al.*, 1990). This is accomplished by transferring magnetization through one- and two-bond J couplings which are of similar magnitude in the case of $J_{\mathrm{NC}^{\alpha}}$, but very different magnitude in the case of $J_{\mathrm{NC}'}$, *cf.* Fig. 2.1. For instance, the HNCA experiment correlates both intra- and interresidual C^{α} chemical shifts whereas the

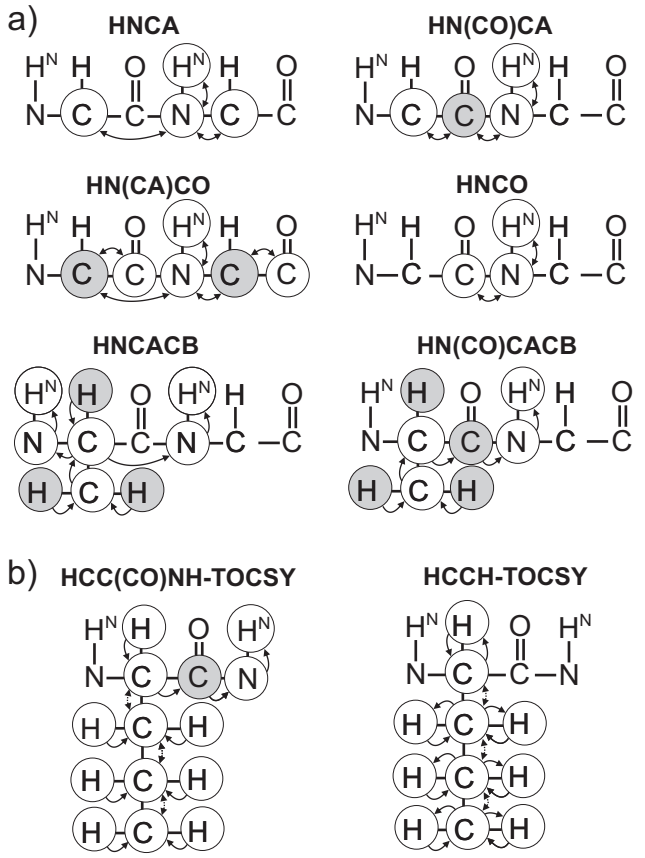


Figure 2.3: Heteronuclear experiments for protein assignment. Open circles indicate the correlations recorded in a given experiment and shaded circles indicate coherence steps which are used but where the nuclei are not correlated. Solid arrows specify the scalar couplings used for coherence transfer and dashed arrows TOCSY-transfer. *a*) 3D triple resonance experiments for backbone chemical shift assignment. *b*) 3D or 4D experiments for side chain assignment.

HN(CO)CA experiment only correlates the latter by doing a stop-over at carbonyl carbon where $^2J_{\mathrm{NC'}}$ is vanishingly small. Intra- and interresidue correlations can subsequently be connected by sequential walk through the amino acid sequence. Using the set of complementary experiments HNCA and HN(CO)CA, HN(CA)CO and HNCO, and additionally HNCACB and HN(CO)CACB, ambiguities due to chemical shift degeneracies are overcome. In challenging cases of partly folded structures, additional information from other coherence transfer pathways such as intraresidual HNCA may be required (Permi, 2002a).

A series of connectivities are obtained for the H^N, H^α, N, C', C^α, and C^β chemical shifts from the set of backbone assignment experiments. The ^{13}C chemical shifts depend strongly on amino acid type and by exploiting this, it is easy to align the spectral information to the primary sequence of the protein. Only a few connectivities are required for each segment so problems with missing signals and gaps in segments due to proline residues are easily resolved.

With assignment of C^α and C^β as a starting point, the subsequent side chain assignment proceeds one amino acid spin system at the time. This step is less robust than backbone assignment but is a prerequisite for obtaining long-range distance restraints. The two main approaches commonly used for recording side chain chemical shifts are illustrated in Fig. 2.3b. HCC(CO)NH is basically an isotropic mixing extension of CBCA(CO)NH whereas HCCH-TOCSY detects the aliphatic protons after shuffling the magnetization along the side chain. Both variants employ ^{13}C-^{13}C TOCSY transfer for effective coherence transfer linking connectivities throughout the side chain. For small proteins COSY transfer instead of TOCSY transfer sometimes yields complementary information in cases of missing signals.

High spectral resolution, good sensitivity, and unequivocal spectral interpretation along with effective water suppression are the main requirements for side chain assignment experiments. Water suppression is especially important in HCCH-TOCSY. It is effectively achieved by pulsed field gradients combined with coherence order selective coherence transfer (COS-CT) to retain the highest possible sensitivity (Sattler *et al.*, 1995a). Problematic signal overlap can be circumvented by extending the spectral editing to four dimensions. Applying COS-CT throughout a 4D HCCH-TOCSY experiment unfortunately results in undesired multiple-bond correlations and special attention is required to avoid their generation without sacrificing too much sensitivity.

The HCC(CO)NH experiment is favorable to HCCH-TOCSY for the smallest proteins since assignment is performed more conveniently by linking side chain resonances to NH correlations directly. Obtaining sufficient sensitivity is nevertheless problematic due to extensive coherence transfer steps which limit

applications to proteins larger than 15-20 kDa. The HCCH-TOCSY experiment is more sensitive and thus applicable to higher molecular weight proteins, but the corresponding spectrum is rather tedious to assign due to the low acquired resolution. Furthermore, the side chain assignment process becomes less susceptible to automated analysis with increasing protein weight. Thus thorough manual analysis of the spectra is often required and this currently constitutes a major bottleneck for rapid protein structure determination. Although the problems can be alleviated to some extent by deuteration or specific labeling, this requires more laborious procedures in the protein production. By using methyl targeted HCCmHm-TOCSY experiments it is feasible to extend the size-range of double labeled proteins amenable for structure determination. This is demonstrated in Article I where a set of 3D and 4D experiments enabling unambiguous assignment of methyl-containing residues is presented. The two types of side chain experiments are compared in Chapter 4.1.1 along with a description of the noted extension of HCCH-TOCSY methodology.

Structural restraints With the assignment in hand, the subsequent step in the structure determination process is collection of restraints for the structure calculation. A good deal of structural information is already available from the chemical shifts, for instance, the secondary structure of proteins can be predicted since C^{α}, H^{α}, C^{β} and C' chemical shifts depend on backbone conformation (Wishart and Sykes, 1994). Also dihedral angles can be estimated directly from chemical shifts by comparison with the ones from a database of protein structures. Recently, exciting developments have demonstrated the possibility of determining protein structures using backbone chemical shifts as sole experimental restraints (Cavalli *et al.*, 2007; Shen *et al.*, 2008). The structural resolution is, however, lower than that which is conceivable by conventional methods and the procedure is limited to small proteins.

The by-far most important restraint for determining overall protein fold is the distance information contained in the nuclear Overhauser effect (NOE). The NOE arises due to dipole-dipole cross relaxation and therefore scales by the inverse sixth power of the distance between two spins. Application of NOEs as distance restraints is complicated by spin diffusion and the nonlinear dependence of cross relaxation rates in the crosspeak intensity for long mixing times. Distance separations can only be obtained semi-quantitatively and NOEs are instead grouped into categories based on intensity with upper bound separation. NOEs are obtained from NOESY spectra which usually are incorporated into HSQC type pulse sequences such as 3D NOESY-HSQC or 4D ^{15}N-^{13}C

edited HMQC-NOESY-HMQC (Griesinger *et al.*, 1989; Kay *et al.*, 1990a). The vast majority of long-range ^{1}H-^{1}H NOEs are found in methyl-containing and aromatic residues. A high degree of assignment of these residues is therefore essential for obtaining a sufficient number of distance restraints. The precision of structure determination is highly related to the number of unambiguous NOE restraints obtained with 20-25 restraints required per residue for obtaining the highest possible resolution. The NOESY peaks are assigned automatically in an iterative process during the initial phase of the structure calculation process.

Scalar coupling constants, RDCs, and amide proton-solvent exchange rates are the most common restraints to complement the local structural information contained in chemical shifts and long-range information in NOE restraints. Three-bond scalar coupling constants ^{3}J depend on the dihedral angle formed by the covalent bonds according to the Karplus relation

$$^{3}J = A\cos^{2}\theta + B\cos\theta + C, \tag{2.6}$$

where A, B, and C are semiempirically determined constants dependent on the nuclei involved in the bond. Also one- and two-bond scalar coupling constants correlate to some extent with protein conformation. Since these coupling constants are needed to determine the corresponding RDCs their use in structure determination is expected to become more widespread than hitherto. Nevertheless, scalar couplings seldom find application for standard structure determination since dihedral angle geometry is easily obtainable by comparing chemical shifts and sequence homology to a database of assignments and high-resolution structures (Cornilescu *et al.*, 1999).

Hydrogen bond restraints are important for the conformation and precision of the final structure. In favorable cases scalar coupling constants over hydrogen bonds can be observed to identify hydrogen bond restraints (Dingley and Grzesiek, 1998). Similar information is obtainable from $^{1}H^{N}$ solvent exchange rates but less ambiguously and without information of the acceptor group for the hydrogen bond. H/D exchange rate restraints are either obtained from saturation transfer experiments or by following the reduction in $^{1}H^{N}$ intensities after transferring the protein from H_2O to D_2O (Forsén and Hoffman, 1963; Koskela *et al.*, 2007). Slow exchange rates indicate that the amide proton may be involved in a hydrogen bond.

Residual dipolar couplings provide structural information relating bond vector orientations as well as report on protein dynamics (Prestegard *et al.*, 2000). In isotropic solutions the dipolar coupling interaction is effectively averaged to zero due to Brownian motion. However, by introducing a weak anisotropy in the

sample a small fraction of the dipolar coupling effect can be obtained without obscuring the spectral quality (Tjandra and Bax, 1997). Weak alignment can be induced by steric or electrostatic effects from *e.g.*, bicelles, bacteriophages, or polyacrylamide gels, and the degree of alignment is tunable by varying concentration or gel stretching.

The advantage of RDCs is that, in addition to constraining local geometry, RDCs also contain global ordering character as they restrain all bond orientations relative to a common principal axis system. This allows for local and long-range structure refinement and validation. For refinement purposes, RDCs are incorporated into molecular force fields to tighten dihedral geometry and improve Ramachandran plots by orienting the bond vectors in accordance with the RDCs. In addition, RDCs may be used as indicators of protein dynamics. They are particularly attractive for filling the gap between the ps-ns and ms timescales unobservable by conventional relaxation measurements (Tolman and Ruan, 2006). Incorporation of the average angular dependence of RDCs in determination of structure and dynamics is, however, far from straightforward due to a high degree of degeneration. Data interpretation is further complicated as the information is a convolution of structural and dynamical contributions.

A minimum of five independent RDCs per domain are required to disentangle the angular dependency of RDC restraint towards the alignment frame and additional RDCs are needed to extract dynamical information. In this respect, the tetrahedral coordination of the C^{α} nucleus is superior to the peptide plane since the planarity of the peptide bond makes the RDCs from this site intrinsically susceptible to redundancies. The C^{α} site can also be beneficial for proteins that require relatively high pH solvent conditions, since amide protons are amenable to fast exchange rates with water under those conditions. When several RDCs are required from both the amide and aliphatic parts of the amino acids, it is advantageous to measure them simultaneously. This also ensures an equal effect of alignment fluctuations on the various couplings. Two new methods for measuring scalar and residual dipolar couplings featuring these qualities are presented in Articles III and IV.

Structure calculation Data obtained from NMR spectroscopy does not uniquely define the three-dimensional structure of a protein. The experimental restraints are ranges of allowed parameter values and as such not applicable to a single protein conformation. The objective of the structure calculation is to find the coordinates for the atoms that satisfy the input restraints. A range of conformations are compatible with the input restraints. The structure calcula-

tion is therefore repeated many times to find an ensemble of structures which are consistent with the NMR restraints, either for individual conformations or for the whole ensemble. The experimental restraints are supplemented by chemically imposed restraints, such as covalently allowed bond lengths and angles, and chirality, when fed into structure calculation programs.

Structure calculations are commonly conducted by molecular dynamic simulations with the restraints incorporated as weighted penalty terms in the force field energy function. Global fold conformation is searched for by performing simulated annealing with simplified force fields at initially high temperatures and slowly lowering the temperature of the protein during the calculations. The resulting conformations are refined using complete force fields and inclusion of solvent models. The procedure is repeated to obtain an ensemble of low-energy protein conformations. Structure calculations can be combined with automatic analysis of NOESY spectra in an iterative manner, which dramatically speeds up the time-consuming structure determination process (Güntert, 1998).

Computational analysis of NMR spectra is conceptually simple, but noisy and sparse data cause difficulties in the automation process. A low signal-to-noise ratio makes peak identification cumbersome and spin diffusion complicates the determination of NOEs. Errors in assignment easily propagate through the data analysis and make subsequent analysis harder, a problem which also holds for manual analysis. Current structure determination protocols are robust and can be fast for small proteins, but for larger proteins spectral analysis is still complicated. Methodological advances for overcoming this problem are continuously pushing the limits of protein structure determination by NMR spectroscopy.

In addition to structure determination, NMR spectroscopy makes a unique contribution to structural biology by providing detailed information on macromolecular dynamics on a broad span of time scales (Palmer, 2004). The dynamical effects can be incorporation into the structure determination process for accurate description of conformational fluctuations. This holds promise for improving our understanding of biological functions such as enzymatic recognition, conformational selection and structural flexibility (Lindorff-Larsen *et al.*, 2005; Boehr *et al.*, 2006; Lange *et al.*, 2008).

2.1.2 Intermolecular Interactions

A great part of protein functionality arises from specific interactions with different types of ligands, including RNAs and other proteins. The binding interaction can either inhibit or promote a biological process by changing the conformational

energy landscape or surface electrostatic properties. The exquisite sensitivity of chemical shifts makes NMR spectroscopy a versatile method for determining molecular interactions (Takeuchi and Wagner, 2006). In comparison to other assays for studying protein activity or binding, the NMR techniques are fast and sensitive to weak changes induced in a protein when a ligand binds.

If the ligand is tightly bound to the protein target it is possible to determine the structure of the complex at atomic resolution. This is achieved by similar means as described above with additional filtering experiments incorporated to separate intermolecular NOEs from intramolecular NOEs. A much faster approach is suitable for screening ligands for binding activity and to identify interaction sites in the weak binding regime. This is commonly studied by following perturbations in chemical shifts upon ligand titration.

Signals showing chemical shift changes are commonly assumed to be located near the binding interface. The results are, however, often difficult to interpret conclusively due to overlap or missing peaks. Most importantly, chemical shift perturbations alone do not reveal whether the change stems from conformational changes caused by the proximity of the ligand or second order effects, such as reorientation of nearby aromatic ring. In addition, backbone amide sites are hydrophilic and, therefore, frequently not located on the intermolecular interfaces.

A possible remedy for these issues is to map interfaces in protein-ligand complexes by saturation transfer difference NOE experiments (Meyer and Peters, 2003). Alternatively, the binding interface can be resolved by cross-saturation transfer (Takahashi *et al.*, 2000). This technique is less equivocal, but requires cumbersome sample preparation due to the requirement of deuteration. Many of the ambiguities for chemical shift perturbation studies are overcome less elaborately by tracking changes in amide backbone and methyl side chain correlations simultaneously. Article II describes new experiments which are optimized to this end.

Protein-ligand interaction studies are mostly performed with isotopically enriched proteins whereas the ligand is nonlabeled. To determine the unknown structure of ligands at natural abundance isotope levels, small molecule NMR techniques are used. Spectroscopic methods for this purpose are described in the following section.

2.2 Small Molecule NMR

The diversity of small biomolecules, such as carbohydrates and lipids, involved in cellular processes is astronomical. The molecules play essential roles in recognition processes, energy storage, and disease development, to name just a few functions. The detailed molecular mechanisms of these events are still poorly understood. Understanding how the functions are imparted through interactions with proteins requires analysis of molecular structure and dynamics. NMR spectroscopy is used intensively for this purpose. Along and in combination with mass spectroscopy NMR spectroscopy provides unique methods for structure determination of small biomolecules such as carbohydrates and natural products (Duus *et al.*, 2000; Fukushi, 2006). In addition, the information richness of the technique also makes NMR spectroscopy the method of choice for structure verification of products in the organic synthesis laboratory.

For small molecule studies isotope labeling may either not be necessary or even feasible. The latter is typical in the case of lipids and carbohydrates where no over-expression mechanism is available. The natural abundance level of ^{13}C and even ^{15}N allows for correlating protons to a single heteronucleus when sufficient sample amounts are available, but correlating several heteronuclei is rarely feasible for biomolecules. Thus alternative spectroscopic techniques than those used for protein NMR spectroscopy are required. The more diverse building blocks and assembly possibilities of carbohydrates and lipids than for proteins significantly complicate spectral analysis. The final target of structure determination is therefore different. While the primary structure is always known when studying proteins, often the task for small molecule NMR studies is to determine the covalent connectivities in order to solve the unknown structural composition.

Standard one-dimensional ^{1}H and ^{13}C techniques are still the most commonly used experiments due to the sample concentration requirements and the spectral simplicity of the molecules in question. In principle, all the necessary structural information is often present from the one-dimensional spectra, but due to resonance overlap and complex coupling patterns the interpretation may be too complicated. Assistance to alleviate this problem is achievable from a quantum chemical spin system fitting to the spectrum. This is also useful for verifying spectral interpretations by comparing the experimental and the theoretically predicted spectra (Laatikainen *et al.*, 1996). A long standing goal for small molecule NMR spectroscopy is automated structure determination without the interference of the scientist. This is feasible in the case of simple structures and for pure samples by comparing experimental data to spectral libraries

and theoretical spectral predictions, but in most cases some interpretation is required by the spectroscopist.

For manual spectral interpretation more unambiguous structure determination is obtainable from two-dimensional homo- and heteronuclear experiments. The standard package for structure determination of unlabeled biomolecules includes classical experiments such as COSY, TOCSY, and NOESY (Aue *et al.*, 1976; Braunschweiler and Ernst, 1983; Kumar *et al.*, 1980). For obtaining one-bond heteronuclear correlations HMQC or HSQC experiments are indispensable (Bodenhausen and Ruben, 1980). Combinations of these experiments with homonuclear experiments, such as HSQC-TOCSY and HSQC-NOESY, are useful for assigning and connecting spin systems (Kövér *et al.*, 1998). Pulsed field gradient enhanced versions of the experiments ensure efficient artifact and solvent suppression. This enables the experiments to be acquired using low sample concentrations. In addition to information based on chemical shifts and correlation connectivities, also scalar coupling constants are vital restraints for determining stereochemistry and conformation specificity.

An essential part in structure elucidation of unlabeled biomolecules is obtaining heteronuclear correlations over several bonds. For instance, the antennary structure of oligosaccharides makes linking sugar residues over the glycosylic bond crucial. A key experiment for this purpose is HMBC, which yields multiple-bond correlations using the ${}^{n}J_{\mathrm{CH}}$ couplings ($n \geq 2$) for coherence transfer (Bax and Summers, 1986). For discriminating between correlations over two-bond from three-bond correlations or higher the H2BC experiment is most helpful (Nyberg *et al.*, 2005).

While the intensities of direct ^{13}C-^{1}H correlations are rather equal, signal intensities in long-range correlation experiments are lower and less uniform. To avoid overlap and allow for spectral simplification in long-range correlation experiments it is worthwhile to suppress one-bond correlations. A convenient implementation to this end is the Low Pass J Filter (LPJF) which suppresses magnetization arising from protons directly bound to ^{13}C (Kogler *et al.*, 1983). The pulse element is based on matching the relatively large ${}^{1}J$ coupling to a magnetization transfer delay and subsequent purging of antiphase magnetization. Nevertheless, even with exact LPJF matching, one-bond correlations often occur in HMBC spectra. This causes resonance overlap and renders spectral interpretation less equivocal. The origins of these artifacts are analyzed in Article V and a method for their suppression without compromising sensitivity is devised. Details on this experiment are given in Chapter 4.

Although spectral analysis of small molecules is often routine procedure, provided the sample is pure and sufficiently concentrated for recording 2D experi-

ments, it is still challenging to study mixtures of compounds, such as metabolic products. The widespread use of NMR techniques and continuous challenges in spectral analysis of new products makes the effort of improving existing techniques by pulse sequence design worth pursuing.

Chapter 3

Aims of Study

The aim of this study is to improve modern NMR spectroscopic methods in order to facilitate structural and interaction studies of biomolecules, in particular proteins. This is achieved by designing pulse sequences targeted to overcome problematic issues related to resonance overlap, low spectral resolution, and incomplete spectral filtration.

Article I describes pulse sequences aimed at increasing the size range of ^{13}C,^{15}N labeled proteins amenable to side chain assignment. This is achieved by extending the spectral editing of methyl targeted HCCH-TOCSY experiments to four dimensions while retaining high sensitivity. In addition, methods are devised to circumvent artifact generation that arises whilst increasing the number of spectral dimensions. The introduced methodology is a cost-effective approach for proteins in the range of 15-35 kDa and readily lends itself for automated assignment procedures.

Optimal sensitivity for detection of methyl-containing residues is also the focus in Article II, where enhanced methods for simultaneous detection of amide and methyl correlations are presented. The aim is to maximize sensitivity within a given measurement time when both correlation types are monitored. The method provides less ambiguous binding epitope mapping data and represents a powerful approach for ligand interaction screening by using cost-effective random fractional ^{13}C labeling.

Simultaneous detection of aliphatic and amide correlations is applied for the study of Article III as well. Here a convenient and accurate method is introduced for simultaneous measurement of several scalar coupling constants and RDCs in small proteins. If the protein is dissolved in D_2O or requires high pH solvent

conditions, conventional detection of amide protons is no longer feasible. A set of HCACO experiments for measuring scalar and residual dipolar couplings is presented in Article IV to circumvent this problem. The methodology is also favorable in applications with proline-rich proteins.

Finally, the purposes of the small molecule NMR study are to trace the origin of one-bond correlation artifacts in HMBC spectra and devise means to overcome the problem. Strong coupling effects are numerically and analytically shown to be the root cause of the artifacts. An effective approach for the artifact suppression is proposed and experimentally verified in Article V.

Chapter 4

Applications

4.1 Methyl Targeted Experiments

Methyl-containing amino acids are hydrophobic and, along with aromatic amino acids, form the cores of folded proteins. Ligand-binding sites are often localized in hydrophobic cavities formed by methyl groups, which also provide sites for many enzymatic reactions. Methyl groups are particularly valuable from a protein structure determination perspective as they provide the majority of long-range NOE connectivities. In addition, methyl groups are important probes of dynamical information and for studying protein-ligand interactions even in large complexes.

Experiments designed for detecting methyl resonances are attractive for several reasons. First of all, the magnetic equivalence of methyl protons due to rapid rotation results in high sensitivity. Secondly, the methyl correlations are more dispersed than the rest of the aliphatic carbons. In addition, methyl carbon and proton transverse relaxation rates are dominated by ^{1}H-^{1}H and ^{13}C-^{1}H dipole-dipole interactions whereas the effect of ^{13}C chemical shift anisotropy on relaxation is small. Also proton spin flips contribute to relaxation, but as the methyl group rotates rapidly about its symmetry axis, transitions that contribute to methyl crosspeaks are generally long lived. Finally, since methyl groups are located in the side chain peripheries they are most often dynamic which in turn leads to slower relaxation rates. All these favorable properties enable extensive spin manipulations and results in high resolution ^{13}Cm-^{1}Hm correlation spectra.

A great number of pulse sequences has been designed for achieving optimal resolution for methyl correlation in especially HSQC-NOESY type experiments (Diercks *et al.*, 1998; Zwahlen *et al.*, 1998). The relatively slow relaxation rates and high crosspeak dispersion make methyl groups particularly favorable for studying high molecular weight proteins. However, transverse relaxation rates rapidly increase when protein size increases. This impairs the necessary assignment of side chains including methyl-containing residues. For very large proteins the HMQC experiment is preferable for obtaining methyl correlation spectra since in this sequence an effect analogous to TROSY is optimal for methyl groups. In fact, the intramethyl dipolar relaxation interactions cancel completely for a single resonance during multiple quantum evolution due to destructive interference between the ^{13}C-^{1}H and ^{1}H-^{1}H autocorrelated and cross-correlated dipolar contributions (Tugarinov *et al.*, 2003).

Random fractional deuteration can diminish the problems of fast relaxation rates. This approach, however, also reduces proton density and causes line-broadening due to ^{2}H isotope shifts. To overcome the size limitations a number of specific methyl labeling protocols, which enable structural and dynamical studies of very large proteins have been developed. The idea is to use uniformly ^{13}C labeled ketoacid precursors in the protein growth media for selective methyl protonation in an otherwise deuterated background. An example of such labeling scheme for Leu, Val and δ1-Ile is shown in Fig. 4.1a. This labeling scheme allows structural and dynamical studies of very large proteins and have enabled the global fold determination of a 82 kDa monomeric protein (Tugarinov and Kay, 2005). The same labeling scheme is also useful for minimizing intermethyl relaxation interactions, which benefit the methyl TROSY effect in a similar manner as conventional TROSY benefits from deuteration.

While specific deuteration combined with uniform ^{13}C labeling certainly benefit the determination of protein structure and dynamics, it may not be necessary for ligand screening or interactions studies. Methyl targeted experiments are also useful for these purposes since methyl correlations provide considerably higher sensitivity than the corresponding amide correlations. In addition, methyl-containing residues are more likely to be in close proximity of the cavities where ligand binding often occurs. An approach exploiting this for high-throughput screening is selective ^{13}C labeling of methyl groups as shown in Fig. 4.1b (Hajduk *et al.*, 2000). This labeling protocol is also based on ketoacid precursors, but the low degree of ^{13}C isotopes and avoidance of specific ^{2}H labeling make the approach cost-effective for screening. When combined with uniform ^{15}N labeling this labeling scheme is applicable to the NC*m*-HSQC experiment introduced in Article II, even if only a small fraction of methyl carbons is ^{13}C

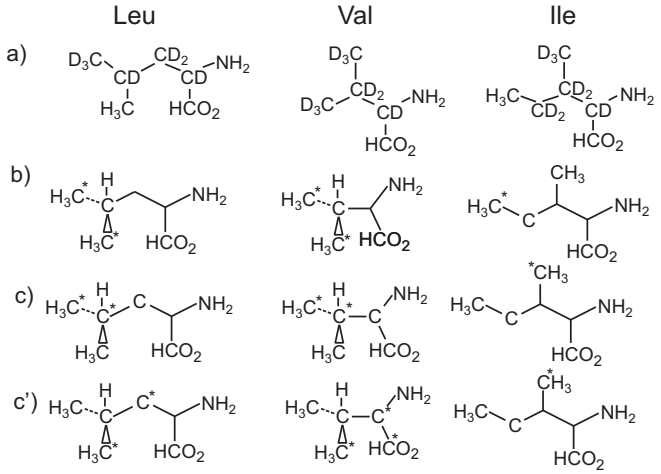


Figure 4.1: Selective labeling schemes for methyl groups. All carbons are ^{13}C labeled in *a* while asterisks indicate ^{13}C in *b*, *c* and *c*′. *a*) Methyl proton labeling scheme with uniform ^{13}C isotopes and deuterated background (Tugarinov and Kay, 2003). *b*) Methyl selective ^{13}C labeling suited for screening of protein-ligand interactions (Hajduk *et al.*, 2000). *c*) and *c*′) 10% Random fractional labeling for stereospecific assignment (Neri *et al.*, 1989). *pro*-R methyl correlations from Leu and Val (*c*) appear as doublets while the *pro*-S counterparts (*c*') appear as singlets since there is no ^{13}C-^{13}C coupling.

labeled.

An attractive labeling alternative for the NC*m*-HSQC experiment described in Article II is to use random fractional labeling with 10-20% level of ^{13}C isotopes (Neri *et al.*, 1989). Random fractional labeling is based on a mixture of uniformly ^{13}C labeled and unlabeled glucose for biosynthetic protein production. The metabolic pathways result in isotope labeling for Leu, Val, and Ile residues as depicted in Fig. 4.1c and c'. In addition to being valuable for studying protein interactions, this labeling protocol is also useful for stereospecific assignment of methyl groups, a step which significantly improves qualities of structures. If a ^{15}N labeled sample is produced for assessing spectral quality or for ligand titration studies, it can be worthwhile to include 10% fractional ^{13}C labeling. This enables recording of the most sensitive backbone experiments for checking structure determination possibilities and the additional costs are relatively

low. While the methyl selective labeling scheme is more attractive for high-throughput screening purposes, random fractional labeling is more widespread in research laboratories since it is required for stereospecific assignment purposes.

The labeling schemes for specific methyl protonation in an otherwise deuterated background and the pulse sequences developed in this context are pushing the size limits of systems feasible for NMR studies. Notwithstanding the assets of the method, it is desirable to be able to use uniformly double labeled sample whenever it is possible. This circumvents the time-consuming and often costly production of more elaborate labeling schemes. In addition, it does not require protein refolding which is occasionally necessary in deuterated samples for complete amide proton back substitution.

4.1.1 Side Chain Assignment by HCCmHm-TOCSY

Side chain assignment is essential for protein structure determination, but is often a limiting step. The procedure can be time-consuming and is less automated than backbone assignment. There are two main assignment strategies for side chain assignment, HCC(CO)NH and HCCH experiments, *cf.* Fig. 2.3. The HCC(CO)NH-TOCSY experiment provides excellent resolution by linking the side chain resonances to the ^{15}N-^{1}H correlation of the preceding residue (Montelione *et al.*, 1992; Grzesiek *et al.*, 1993). Solvent suppression is unproblematic in this method since the amide protons resonate far from the water signal. However, the inherent sensitivity of the experiment is low and there are no direct ^{1}H-^{13}C connectivities in the 3D versions of the experiment. Significantly better sensitivity is provided by the HCCH-TOCSY experiment where assignment is facilitated from the direct connectivities (Fesik *et al.*, 1990; Bax *et al.*, 1990). Nonetheless, this experiment can be tedious to analyze due to low resolution. Combinations of the two experiments are most likely to yield complete assignment for small proteins.

The NH-detected HCC(CO)NH-TOCSY is more sensitive when applied to larger, deuterated proteins, where the experiment benefits from the TROSY effect. Nonetheless, deuteration leads to lower proton density and longer T_1 relaxation times. In addition, for large proteins the carbonyl CSA relaxation is prohibitively fast at the highest fields necessary for TROSY linenarrowing. To circumvent the need for deuterium labeling or expensive selective methyl protonation it is desirable to use double labeled samples whenever feasible. A powerful strategy for proteins of mid-range size (∼15-35 kDa) is represented by the 4D HCCmHm-TOCSY experiments for methyl-containing side chain

assignment introduced in Article I.

The HCCmHm-TOCSY pulse sequences are derived from the conventional HCCH-TOCSY versions, but are targeted specifically to methyl-containing residues for the purpose of achieving high resolution and sensitivity within one experiment. This allows to overcome some of the drawbacks of the conventional side chain assignment strategies. By detecting only the slowly relaxing methyl groups it is possible to apply a long constant-time (CT) period for methyl frequency labeling which is key to achieve high resolution (Uhrín *et al.*, 2000a; Permi *et al.*, 2004). During the CT period proton spin flip induced relaxation is quenched by proton decoupling. In contrast to the conventional HCCH-TOCSY approach, the relatively small signal width of methyl carbons results in a low number of points required for long acquisition times. Moreover, the need for substantial spectral aliasing in 4D experiments is overcome. In extent to the degeneracy of methyl protons high sensitivity is obtained by optimizing polarization transfer to I_3S multiplicities and applying COS-CT throughout the sequence for optimal combination with pulsed field gradients. The latter is especially efficient in combination with TOCSY isotropic mixing, since this transfer is inherently coherence order selective, and thus leads to signal loss if amplitude modulated data are recorded.

Schematically the coherence transfer in the 4D version of the HCCmHm-TOCSY pulse sequences is

$$\begin{aligned} {}^{1}\mathrm{H}(\mathrm{SCT}{-}t_1) &\xrightarrow{{}^{1}J_{\mathrm{CH}}} {}^{13}\mathrm{C}(t_2) \xrightarrow{{}^{1}J_{\mathrm{CC}} \quad \mathrm{TOCSY}} \\ {}^{13}\mathrm{C}_m(\mathrm{CT}{-}t_3) &\xrightarrow{{}^{1}J_{\mathrm{CH_3}}} {}^{1}\mathrm{H}_m(t_4) \end{aligned} \tag{4.1}$$

where CT and SCT denote constant-time and semi-constant-time, respectively. In the 3D H(C)CmHm-TOCSY experiment the frequency labeling period prior to the TOCSY mixing period is omitted.

The high resolution warranted by the HCCmHm-TOCSY experiment facilitate assignment of the resulting spectra. Only a small signal range is browsed to link the methyl resonances to the C^{α}, C^{β}, H^{α}, and H^{β} resonances that are available from the backbone assignment. The complete view of all resonances in each side chain spin system provided by the 4D experiment variants makes the assignment unambiguous. This is in contrast to 3D HCCH-TOCSY where many methyl resonances often cannot be assigned due to poor resolution in both indirect dimensions and poor dispersion of ${}^{1}H^{\alpha}$,${}^{13}C^{\alpha}$ and ${}^{1}H^{\beta}$,${}^{13}C^{\beta}$ crosspeaks. For instance, by resolving into four spectral dimensions the likely resonance overlap of Thr residues is circumvented.

The sensitivity of the HCCmHm-TOCSY experiments is excellent and allows for data acquisition using only a single scan per free induction decay (FID), especially when recorded on a spectrometer equipped with cryogenically cooled probehead. The methyl signals resonate far from the water, thus even a two step phase cycle for cancellation of residual solvent signal is not required. This serves well in combination with the 4D versions of the experiment to keep experimental time at levels comparable to many 3D experiments. If the experimental time is limiting, it would, in principle, be possible to apply nonlinear sampling methods. This approach would be beneficial as sensitivity is not the limiting issue of the experiment.

The superior sensitivity of the HCCmHm-TOCSY experiments in comparison to HCC(CO)NH-TOCSY is demonstrated in Fig. 4.2. It is illustrated for coactosin (16 kDa) with the (H)CCmHm-TOCSY spectrum recorded at 15°C in D_2O where the rotational correlation time of coactosin is $\tau_c = 14.5$ ns. This corresponds to a 30 kDa protein in H_2O at room temperature. The (H)CC(CO)NH-TOCSY spectrum was recorded at 800 MHz frequency and in H_2O at 25°C where coactosin's correlation time is $\tau_c = 10$ ns. A similar sensitivity difference can be anticipated between the H(CCCO)NH-TOCSY and H(C)CmHm-TOCSY experiments and in the 4D versions of the experiments.

Although the assignment strategy is demonstrated here for a relatively small protein of 16 kDa molecular weight, it is also applicable for considerably larger proteins. Indeed, a closely related approach using multiple quantum coherence evolution for the first carbon dimension in a 3D (H)CCmHm-TOCSY version has been demonstrated to provide almost complete assignment of methyl-containing residues in double labeled 42 kDa Maltose Binding Protein (MBP) (Yang *et al.*, 2004). In combination with 4D ^{13}C,^{15}N-edited NOESY this has proved a powerful strategy for structure determination of large proteins without deuteration (Xu *et al.*, 2006).

In addition to facilitate manual assignment, the excellent resolution of the 4D HCC$_m$H$_m$-TOCSY experiments allows for improved automated side chain assignment. In the potentially challenging cases of isoleucine and leucine assignment two methyl signals are detected within the same residue and the redundant information enables cross-validating ambiguities. A robust computer aided implementation of the side chain assignment would help to overcome one of the major bottlenecks encountered in protein structure determination.

The four dimensional extensions of HCCmHm-TOCSY experiments to improve resolution can also be incorporated into general HCCH-TOCSY side chain assignment experiments by including COS-CT throughout the pulse sequence. Implementing an additional dimension with gradient coherence selection into the

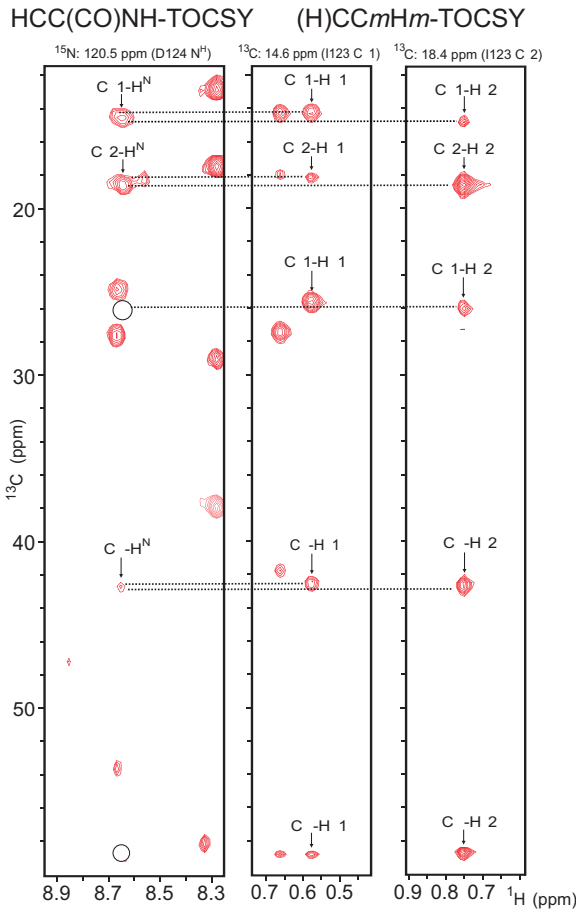


Figure 4.2: Comparison of (H)CC(CO)NH-TOCSY and (H)CCmHm-TOCSY spectra. Circles indicate the missing $^{13}C^{\alpha}$ and $^{13}C^{\gamma}$ signals in (H)CC(CO)NH-TOCSY. The strips are from Ile123 in 0.8 mM coactosin dissolved in D_2O at 15°C recorded at 600 MHz for (H)CCmHm-TOCSY and in H_2O at 25°C at 800 MHz for (H)CC(CO)NH-TOCSY. The total experimental time was 16 h for (H)CCmHm-TOCSY and 23 h for (H)CC(CO)NH-TOCSY.

double sensitivity enhanced versions of HCCH-TOCSY (Sattler *et al.*, 1995a) would significantly improve sensitivity and solvent suppression in comparison to previously suggested 4D experiment variants (Olejniczak *et al.*, 1992).

We tested this type of experiment but found the performance of the methyl targeted counterpart to be superior. A concern with the extension to 4D in conventional HCCH-TOCSY experiments was the occurrence of COSY-type correlations. These are crosspeaks between ^{13}C and ^{1}H two-bonds away which arise due to evolution of J_{CC} couplings during the mixing steps. This problematic issue is demonstrated in Fig. 4.3 which shows excerpts from 4D HCCH-TOCSY. The illustrated spectra were acquired with a pulse sequence featuring COS-CT in the final two indirect dimensions. Also a triple sensitivity enhanced pulse sequence variant was tested with similar results with respect to COSY peaks. The intensity of the COSY peaks (indicated by asterisks in Fig. 4.3) is substantial and occasionally higher than that of the direct correlations, as in the case of L100δ2. Although the COSY peaks also arise in the methyl targeted experiments in connection with triple sensitivity enhancement, they can be purged by alternative coherence transfer in the mixing period between t_1 and t_2 as described in Article I. In contrast, due to real time chemical shift evolution in the second ^{13}C dimension, COSY correlations are not purged using INEPT between the first two indirect dimension as evident from Fig. 4.3. In addition to COSY peaks causing spectral complications, the large ^{13}C spectral width in HCCH-TOCSY requires substantial folding in both ^{13}C dimensions in order to maximize resolution and minimize experimental time. Even though this usually does not cause spectral overlap it renders spectral analysis cumbersome. Furthermore, efficient water suppression is more critical than is the case where only methyl protons are detected. A minimum of two scans per FID is therefore required.

The unambiguous assignment provided by the HCCmHm-TOCSY experiments can be supplemented by conventional side chain assignment experiments for proteins in the midrange size interval (up to 30 kDa). For larger monomeric proteins complete assignment of all side chains is unlikely to be feasible with double labeled samples. However, as the methyl groups provide the majority of long-range NOE connectivities it is possible to determine global fold of larger proteins using only methyl and amide distance restraints, as in the case of selective methyl protonation in deuterated background. The resolution obtainable from this approach is substantially higher than the $\sim$6 Å resolution obtained from completely deuterated samples.

For ligand screening or other interaction studies complete structure determination may not be intended. Nevertheless, methyl groups constitute vital parts

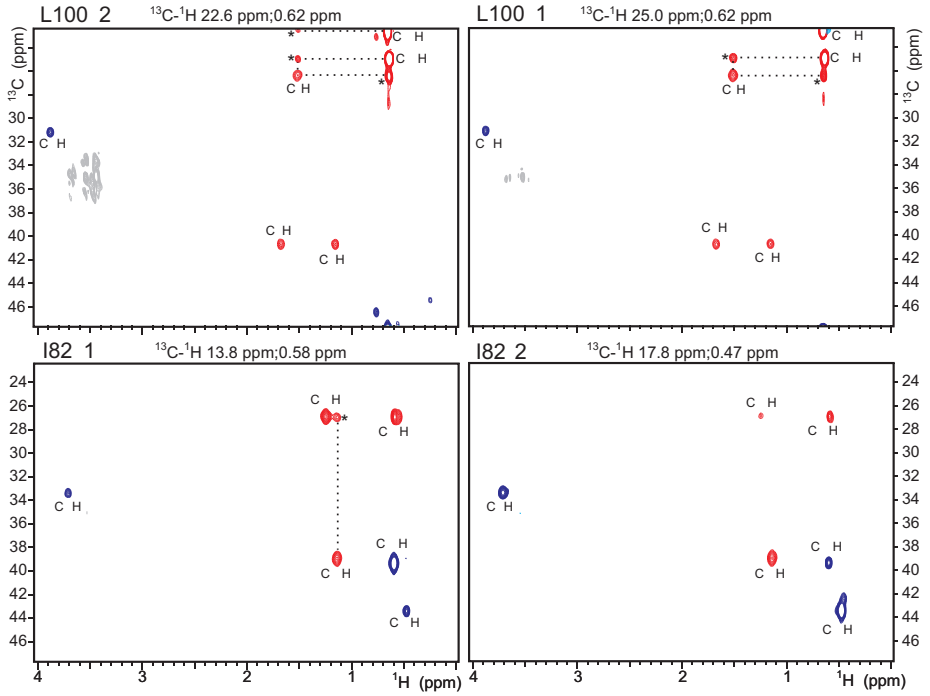


Figure 4.3: Spectral excerpts from a preliminary version of 4D double sensitivity enhanced HCCH-TOCSY. The F_1-F_2 (^{1}H-^{13}C) planes of two representative residues are shown for direct comparison with the methyl optimized counterpart in Fig. 4a-d of Article I. COSY type peaks are marked with asterisks. The C^{α} region and a part of the methyl region are folded onto the ^{13}C axis and appear with negative contours (blue). The gray-scaled signals are due to incompletely suppressed buffer. The spectra were acquired on a 0.8 mM sample of coactosin (16 kDa) dissolved in D_2O. The experiment is conceptually similar to the pulse sequence in Fig. 32b of Sattler *et al.* (1999) but extended with an additional dimension. Two scans were recorded per increment with 32, 20, 20, 512 points in dimensions 1-4, respectively. This corresponds to 7.8, 5.2, 5.2, and 51.1 ms of acquisition times, resulting in a total measurement time of 62 h.

of ligand binding sites which can be exploited to study how proteins change in response to ligand binding. Methyl correlations are also useful for characterizing the dynamical properties of proteins, which are likely to be important for protein function. Assignment of methyl carbon and proton chemical shifts is required for these purposes. The experiments introduced in Article I provide swift and unambiguous means to obtain this. Methods for exploiting methyl correlations for ligand-protein interaction studies are described in the following section.

4.1.2 Epitope Mapping by Time-Shared NCm-HSQC

Protein-ligand interactions play vital roles for the biological function of biomolecules. Identification of binding interfaces in protein complexes provide valuable information for rational design of functional molecules with numerous applications in drug development. The specificity for molecular interactions is one of the forces of NMR spectroscopy beyond structure determination. Especially, binding epitope mapping by following the perturbations in chemical shifts upon ligand titration constitute an attractive way of performing functional assays. This is highly useful from a drug discovery point of view for screening of low-affinity ligands (Shuker *et al.*, 1996). Also protein-protein interactions can be studied in a similar manner with only one of the proteins being isotopically labeled.

The changes in protein chemical shifts induced by the ligand are commonly monitored using ^{15}N-^{1}H HSQC or TROSY experiments. This approach allows good signal dispersion, high sensitivity and can be performed as soon as backbone assignment is in hand. Yet, the binding site information obtained from amide correlation maps can be rather ambiguous since the chemical shift parameter does not differ between perturbations caused by conformational changes near the binding site and second order changes in the shielding further away. Moreover, it is known that molecular binding studies based on amide chemical shift perturbations are not always in line with related information obtained from H/D exchange studies or X-ray crystallography data of complexes. Although most of the ambiguities can be circumvented by cross-saturation transfer techniques (Takahashi *et al.*, 2000), this procedure is so laborious sample production-wise that it is only justifiable to large protein-protein complexes were no other method suffices. The alternative, saturation transfer difference experiments, is only applicable to small molecules, but can yield binding information from the ligand's point of view.

A likely reason for the discrepancy in the binding interface between the

different experimental methods is that backbone amide correlations are often not in close proximity of the epitope binding site. Ligand binding is more likely to involve side chains located in the hydrophobic core of the protein (Hajduk *et al.*, 2000; Takahashi *et al.*, 2006). Therefore it is favorable to perform epitope mapping by monitoring the aliphatic side chain correlations. This approach is, however, impaired by the same ambiguities with respect to second order changes in chemical shifts. To gain more accurate binding interaction data we propose to record both amide and methyl correlations when conducting binding epitope mapping. A set of pulse sequences optimized to this end is described in Article II.

Mapping aliphatic chemical shift perturbations is rather straightforward but it requires that the protein is ^{13}C labeled and side chain assignment has been performed. In addition, the chemical shift dispersion of ^{13}C is intrinsically lower than that of ^{15}N and considerable overlap is likely, especially for methylene and methine moieties. The signal dispersion for methyl correlations is more favorable than for the other side chain moieties. The magnetic equivalence of the methyl protons results in high sensitivity, thus lower sample concentration is required than that for recording all aliphatic correlations. Furthermore, most methyl groups are buried within the hydrophobic core of the protein and thus more likely to be close to binding epitope as noted above. Importantly, complete side chain assignment is not a requirement. Instead, the necessary assignment of methyl-containing residues is easily and unambiguously obtained using the HCCmHm-TOCSY experiments described in the previous section.

Sample production costs are up to ten-fold higher for uniformly ^{13}C labeled proteins in comparison to the cost of uniform ^{15}N labeling. Although lower sample amounts are required when monitoring methyl correlations than amide correlations, resolved J_{CC} couplings decrease sensitivity. Selective methyl labeling in Ile, Leu, and Val residues (Fig. 4.1b) was introduced for ligand screening purposes with this in mind, and the labeling scheme requires less sample amounts than conventional amide correlation binding epitope mapping methods.

Random fractional ^{13}C labeling (Fig. 4.1c) constitutes an alternative low-cost approach, especially when combined with uniform ^{15}N labeling. The signal to noise ratio of ^{15}N-^{1}H and ^{13}Cm-^{1}Hm correlations in HSQC are similar at around 15% of ^{13}C labeling. At this labeling degree most spectral crowding due to ^{13}C-^{13}C couplings is avoided and it favorably coincides with the labeling degree appropriate for stereospecific assignment. The simultaneous ^{15}N,^{13}Cm HSQC experiment (NCm-HSQC) is demonstrated using this labeling scheme in Article II.

When both backbone amide and methyl correlations are measured, it is ad-

vantageous to conduct the measurement simultaneously (Sørensen, 1990; Sattler *et al.*, 1995b). In principle, a gain of $\sqrt{2}$ in sensitivity per unit time is obtainable in comparison to recording the experiments separately. The gain is directly convertible into time saving or lower sample concentration requirements.

The NCm-HSQC experiment is based on the concept of time-shared evolution. Pulses and delays are tailored to optimize both coherence transfer pathways without interference between the two. Additional refocusing of $^1J_{\mathrm{CH}}$ couplings is required to account for the different magnitudes between $^1J_{\mathrm{NH}}$ and $^1J_{\mathrm{CH}}$. Since ^{1}H pulses affect both the ^{15}N and the ^{13}C coherence transfer pathways, care must be taken not to obscure the spin state selection for the TROSY line in the NCm-TROSY version of the experiment. Furthermore, the wider spectral width for ^{13}C requires smaller t_1 increment size than that of ^{15}N. These aspects lead to signal loss due to relaxation. Along with the additional number of pulses required to obtain both coherence transfers simultaneously, this renders the sensitivity gain somewhat lower than the theoretical maximum of $\sqrt{2}$.

The parallel flow of coherences during the NCm-HSQC experiments is as follows

$$\begin{aligned} {}^1\mathrm{H}_m + {}^1\mathrm{H}^{\mathrm{N}} &\xrightarrow{{}^1J_{\mathrm{CH}}+{}^1J_{\mathrm{NH}}} {}^{13}\mathrm{C}_m(t_1) + {}^{15}\mathrm{N}(\kappa t_1) \\ &\xrightarrow{{}^1J_{\mathrm{CH}}+{}^1J_{\mathrm{NH}}} {}^1\mathrm{H}_m(t_2) + {}^1\mathrm{H}^{\mathrm{N}}(t_2) \end{aligned} \tag{4.2}$$

As alternative to the 'sensitivity enhanced' coherence order selective backtransfer in HSQC, a TROSY backtransfer step for ^{15}N,^{1}H correlations is also feasible, as described in Article II.

For methyl correlations the conventional INEPT transfer yields higher sensitivity than COS-CT by a factor 1.14 (Schleucher *et al.*, 1994). A Cartesian backtransfer step is therefore implemented, although it requires the use of selective pulses for the methyl protons in order to avoid dephasing from the pulsed field gradient. This is no concern with respect to unintended inversion of amide protons since they resonate far away, but one must pay attention not to perturb water magnetization for maximal preservation of the detected signal.

The time-shared implementation in NCm-HSQC is conceptually similar to the SESAME-HSQC experiment described in Article III (*vide infra*). Yet, it is suitable for a larger range of protein molecular weights, since there is no constant-time implementation and only methyl correlations are detected.

The time saving afforded by the time-shared approach in NCm-HSQC is demonstrated for both amide and methyl correlations in Article II. The benefit of the approach is decreasing with increasing molecular size due to prolonga-

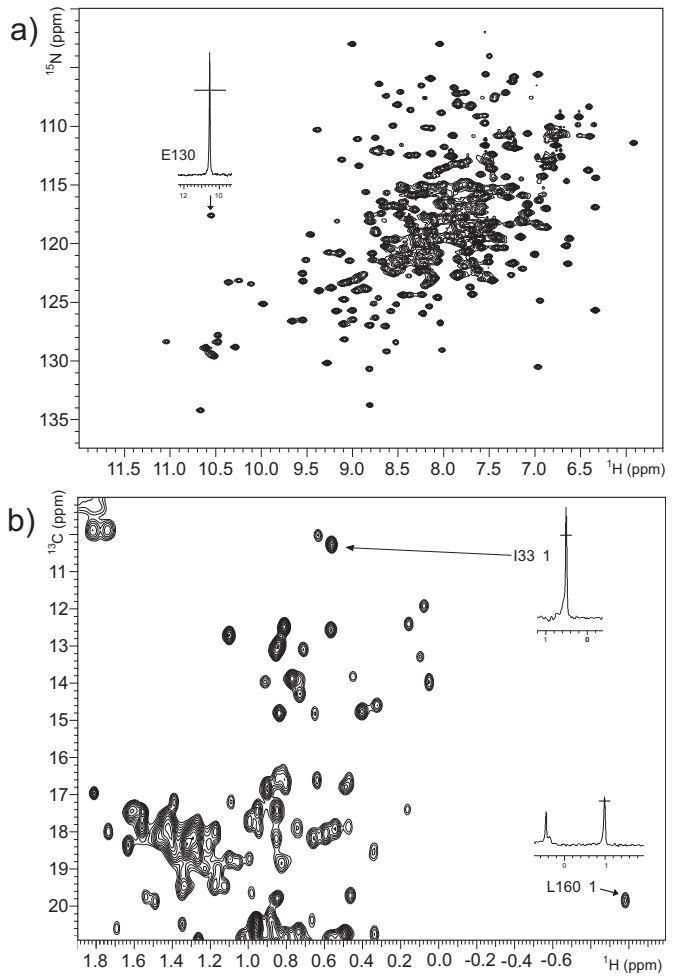


Figure 4.4: Demonstration of time-shared NC*m*-HSQC spectrum of MBP in complex with cyclodexterin recorded at 600 MHz. *a*) ^{15}N-^{1}H HSQC part. *b*) ^{13}C-^{1}H HSQC part. Trace insets of selected resonances show the sensitivity gain obtained in comparison to measuring both experiments separately using the same total amount of measurement time. The figure is modified from Supplementary material of Article II where experimental details are given.

tion of the pulse sequence with respect to conventional ^{15}N and ^{13}C HSQC experiments. Nonetheless, a noticeable sensitivity gain can still be obtained for many amide as well as methyl correlations in the case of 42 kDa MBP as evident from Fig. 4.4. For large proteins, as the example demonstrated in the figure, the spectral crowding due to J_{CC} in Ile and pro-R sites of Leu and Val of fractionally ^{13}C labeled proteins can be problematic. In that case the selective methyl ^{13}C labeling in Ile, Leu, and Val may be more favorable due to higher sensitivity and less spectral crowding (Hajduk *et al.*, 2000). Although chemical shift perturbations are monitored for fewer residues with this approach, the likelihood that a labeled methyl group is within close proximity of a ligand is still higher than that for amide protons. For even larger proteins higher resolution can be anticipated when using selective methyl carbon and proton labeling in deuterated background in combination with ^{13}C HMQC and ^{15}N TROSY.

The time-shared pulse elements presented in Article II are also applicable for NOESY experiments targeted to NH and CH_3 groups (Uhrín *et al.*, 2000b). When incorporated into multidimensional heteronuclear NOESY experiments they improve sensitivity per unit time in comparison to currently available methods. The time saving afforded this way is highly desired since overall acquisition times of these experiments tend to be prohibitively long. An application related to NCm-HSQC for time-shared recording of 4D NOESY spectra for protonated methyl groups in deuterated background has recently been proposed, albeit with suboptimal sensitivity for CH_3 groups (Frueh *et al.*, 2006).

The recording of both amide and methyl correlation maps decreases ambiguities in binding epitope mapping data. The benefits of time-shared recording are also applicable to experiments for measuring scalar and residual dipolar couplings as described in the following section.

4.2 Methods for Measuring Scalar and Residual Dipolar Couplings

Orientational restraints are important for improving the resolution of protein structures beyond the global fold obtained from NOE distance restraints. Scalar coupling constants contain local geometry information related to dihedral angles. Three-bond J coupling constants are especially important in this respect. The angular dependence of other coupling constants can also be parameterized and included into molecular force fields to decrease Ramachandran plot violations. Inclusion of scalar coupling constants in structure determination is most

applicable for backbone dihedral angles since dynamic averaging limits usage of J couplings for side chain conformations. In addition to structure calculation restraints, J couplings can be used for stereospecific assignment of diastereotopic atoms.

Alternatively, restraints based on the angular dependence of cross-correlated relaxation rates can be exploited to give direct information on bond vector angles without the need for parameterization (Reif *et al.*, 1997). The relaxation effects are manifested as differential linewidths or line intensities. Accurate extraction of the structural parameters is, however, difficult and applications of this technique have remained limited.

Residual dipolar couplings also contain angular and distance dependencies, both local as well as long-range information. The latter is of particular importance in complementing NOE restraints for refined structure determination. The most important RDCs are those arising from the backbone which find many applications in structure determination. In contrast, side chain RDCs are mainly used in protocols that generate ensembles based on constraints that include the effects of dynamics, such as methyl group order parameters.

There are numerous applications of RDCs in extension to force field methods for improving dihedral geometry. For instance, protein folds have been recognized on the basis of RDC patterns and complete protein models can be constructed from fragments with orientations based on RDCs (Annila *et al.*, 1999; Delaglio *et al.*, 2000). Furthermore, long-range domain orientations obtained from subunits can be used to construct models of macromolecular complexes and can be combined with small-angle X-ray scattering (SAXS) data to determine structures of multidomain proteins (Mattinen *et al.*, 2002; Grishaev *et al.*, 2005).

The requirement of five independent RDCs to define domain orientations sets a need for measuring several dipolar couplings between different bond vectors. Higher data accuracy is obtained by measuring additional RDCs which also allow dynamical information to be extracted. Since the alignment tensor depends on the given alignment medium, less degenerate RDC data is obtained by measuring the RDCs in multiple alignment media (Bax *et al.*, 2001).

Experiments for measuring scalar coupling constants have seen a revival after the introduction of RDCs. A variety of experiments have been developed for measuring basically every possible short-range residual dipolar coupling (Prestegard *et al.*, 2000). The RDC interaction between two spins takes the same form as the weak scalar coupling Hamiltonian, thus the RDC contribution simply adds to the scalar coupling splitting. RDCs are thus measured in the same way as scalar couplings and the contributions are disentangled by measuring in both

isotropic and weakly anisotropic media. As in the case of scalar coupling experiments, pulse sequence design for obtaining RDCs is focused on maximizing accuracy and precision. There are, however, certain considerations to be noticed when measuring in aligned media. Linewidths are increased and lineshapes appear less symmetrical due to additional dipolar couplings arising in anisotropic media. ^{1}H-^{1}H dipolar couplings are especially enhanced and become comparable in size to geminal scalar coupling constants. Measuring RDCs from the indirect dimension can therefore be attractive despite the lower digital resolution than that obtained from the direct dimension.

Experimental techniques for coupling constant measurements Scalar couplings constants can be measured from the resonance splitting resulting when a coupling evolves during frequency labeling. Common methods include omitting decoupling or applying the E.COSY principle to extract a small unresolved coupling by resolving a large coupling in another dimension (Griesinger *et al.*, 1986). Alternatively, J couplings can be obtained from intensity ratios by quantitative J correlation (Vuister and Bax, 1993). In all cases the accuracy of the measured couplings is dependent on the degree of interference between different relaxation mechanisms. Means to equalize relaxation rate dependencies of signals split by a coupling are therefore of importance even for small proteins.

Measurement of scalar and residual dipolar couplings results in spectral crowding due to the requirement of the couplings to evolve in frequency-resolved measurements. This undesirable doubling of the number of crosspeaks may bias precision or render the couplings immeasurable. The resonance overlap can be circumvented by resolving to 3D experiments, albeit with the cost of substantially lengthened measurement time. An attractive method to diminish the spectral crowding is spin state selection where crosspeaks are edited into separate subspectra corresponding to the attached nucleus being either α or β state (Yang and Nagayama, 1996; Meissner *et al.*, 1997). A simple way of spin state selection known as IPAP (InPhase AntiPhase) is achieved by coadding inphase and antiphase subspectra where a coupling has been allowed to evolve (Andersson *et al.*, 1998; Ottiger *et al.*, 1998).

In aligned media the uniformity of the spectral splitting $J + D$, where D is the RDC contribution, is lowered. For complete separation of spin states in the subspectra it worthwhile to apply spin state selective filters with high J mismatch tolerance or use unequal weighting of inphase and antiphase datasets. In rapidly relaxing systems it is desirable to equalize the relaxation rates in the two subspectra by inversion of the spin state of one the nuclei. Also TROSY-

type spin state selection can be useful for measuring the couplings albeit the full TROSY effect is not obtained as the couplings have to be measured between rapidly and slowly relaxing components. By combining deuteration or selective labeling techniques with TROSY and spin state selection, RDCs can be obtained in high molecular weight systems (Permi and Annila, 2000; Sprangers and Kay, 2007a).

Both new methods for measuring scalar and residual dipolar couplings presented in Articles III and IV feature extensive use of spin state selection for spectral simplification. The novel experiments are introduced below.

4.2.1 Spin statE Selective All Multiplicity Edited-HSQC

The principle of time-shared recording described for the NCm-HSQC experiment in section 4.1.2 can also be applied for scalar and residual dipolar coupling measurements. High spectral resolution combined with good sensitivity indeed make HSQC experiments suitable for measuring coupling constants in smaller proteins.

The SESAME-HSQC (Spin statE Selective All Multiplicity Edited-HSQC) experiment presented in Article III takes advantage of the sensitivity gain afforded by time-shared evolution in a manner optimized for simultaneously extracting couplings in a number of spin moieties with high accuracy. The pulse sequence is tailored to achieve the high resolution required for coupling measurements by a constant-time period for aliphatic carbons. Furthermore, it features the recently developed IPAP-filter for spin state selective editing in all IS_n multiplicities (Nolis *et al.*, 2006).

Although the principle of the SESAME-HSQC experiment is similar to NCm-HSQC there are important differences. First of all, the experiment is targeted to detect all aliphatic carbons correlations in uniformly double labeled proteins. For optimal sensitivity in $^{13}C^{\alpha}$-$^{1}H^{\alpha}$ correlations, which are the least intense, both amide and aliphatic magnetization transfer pathways are coherence order selective. This in turn means that all proton magnetization must pass the final refocusing gradient before detection. A stronger gradient is therefore employed to transverse nitrogen magnetization than to carbon magnetization. Simultaneous COS-CT thus circumvents the need for selective pulses on the ^{1}H channel prior to detection, as was required in NCm-HSQC. Secondly, SESAME-HSQC employs constant-time evolution in the ^{13}C dimension for removing all passive coupling interactions during the frequency labeling. This is a prerequisite to obtain the high crosspeak resolution that allows to extract couplings from a sufficient number amino acids residues. Especially J_{CC} couplings are hamper-

ing spectral quality and the performance of selective pulses for decoupling this interaction is usually inadequate for coupling measurement experiments.

Schematically the magnetization transfer during the SESAME-HSQC experiment is as follows

$$^{1}\mathrm{H}_i + {}^{1}\mathrm{H}_j^{\mathrm{N}} \xrightarrow{{}^{1}J_{\mathrm{CH}}+{}^{1}J_{\mathrm{NH}}} {}^{13}\mathrm{C} + {}^{15}\mathrm{N}(\kappa t_1) \rightarrow$$

$$^{13}\mathrm{C}(\mathrm{CT}-t_1) + {}^{15}\mathrm{N} \xrightarrow{{}^{1}J_{\mathrm{CH}}+{}^{1}J_{\mathrm{NH}}} {}^{1}\mathrm{H}_i(\alpha/\beta)(t_2) + {}^{1}\mathrm{H}_j^{\mathrm{N}}(\alpha/\beta)(t_2) \qquad (4.3)$$

with indices $i = \{1, 2, 3\}$ and $j = \{1, 2\}$ denoting the spin multiplicities.

The spin state filter incorporated into the coherence backtransfer achieves separation of α and β spin states in the direct dimension in a coherence order selective manner. This combines the benefits of COS-CT HSQC and the IPAP principle for spin state selection in all spin moieties of interest. One-bond J_{NH}, J_{CH}, $J_{\mathrm{NH_2}}$, $J_{\mathrm{CH_2}}$, and $J_{\mathrm{CH_3}}$ scalar coupling constants and corresponding RDCs can hereby be measured, as demonstrated in Fig. 3 of Article III. With a slight alteration of the backtransfer step, also geminal $^{2}J_{\mathrm{HH}}$ couplings can also be measured from zero/double quantum splittings (Permi, 2002b).

The sensitivity improvement afforded by concomitant measurement of amide and aliphatic correlations in SESAME-HSQC is slightly lower than that gained in NCm-HSQC. This is partly due to the CT-period required for obtaining sufficient resolution and partly because the full aliphatic signal width (SW) is larger than the nitrogen signal width, *i.e.*, the incrementation scaling factor κ =SW(^{13}C)/SW(^{15}N) is larger than it is in NCm-HSQC. This in turn extends the duration κt_1 where the magnetization arising from aliphatic protons is relaxing as longitudinal two-spin order $2H_zC_z$. These effects, along with the limited resolution afforded in two dimensional correlation experiments, limit the applicability of the SESAME-HSQC experiment to small proteins.

Although the sensitivity of the experiment would improve with real-time evolution in the indirect ^{13}C dimension, the resolution is insufficient for accurate measurement of residual dipolar couplings. This was confirmed by testing a COS-CT ^{15}N,^{13}C-HSQC experiment with the spin state selective filter described above. The resolution obtained in this version proved insufficient for aliphatic correlations. This is evident in a non-CT version of the SESAME-HSQC experiment (Nolis and Parella, 2007).

The most problematic artifacts in the SESAME-HSQC experiment are the phase distortions caused by ^{1}H-^{1}H dipolar couplings, in particular for CH_2 groups. Although this issue is partly counterbalanced by the high digital resolution in the direct dimension, it may diminish the measurement accuracy in

case of high alignment degree. The problem can be circumvented by conducting the coupling measurements in the indirect dimension, albeit at the cost of lowered sensitivity and spin state selection becoming only possible in a single spin moiety. Furthermore, when protons are in fast exchange with the solvent proton spin states are mixed. This render the apparent $^1J_{\mathrm{NH}}$ coupling constant magnitude smaller than the actual magnitude when couplings are extracted in the indirect dimension. An additional concern occurs when measuring at high field strengths where the large spectral width of aliphatic carbons may lead to significant sensitivity losses at off-resonance frequencies. Therefore, it is recommended to apply broadband off-resonance compensated ^{13}C inversion pulses (Skinner *et al.*, 2006) throughout the SESAME-HSQC experiment.

Overall, the SESAME-HSQC is an accurate and convenient experiment suitable for measuring numerous scalar and residual dipolar couplings in small proteins. Experiments suitable for proteins and flexible polypeptides in conditions where H/D exchange is particular problematic are described in the following section.

4.2.2 Residual Dipolar Couplings Around the C^α Site

For small proteins the crosspeak dispersion in ^{15}N HSQC or ^{13}C CT-HSQC can be sufficient for scalar and residual dipolar coupling measurements. Accurate coupling measurements in larger or partly unfolded proteins spectral overlap necessitates the higher resolution afforded by triple resonance experiments. The most attractive experiment type for this purpose is HNCO since this experiment is the most sensitive of the triple resonance experiments. The high crosspeak dispersion warranted by H^N detection along with the possibility to exploit the TROSY effect are advantageous for larger proteins. There are, however, circumstances where the HNCO approach is not favorable for coupling constant measurements.

First of all, if the amide proton exchange rate with water is very high it limits the sensitivity of H^N detected experiments. In addition, this also underestimates the coupling size measured in the indirect dimension. The situation arises when relatively high pH solvent conditions are required to dissolve the protein or if experiments are conducted at high temperatures. Secondly, proline-rich polypeptides require alternatives to HNCO, since in this experiment it is not possible to measure couplings in residues preceding proline. Finally, the tetrahedral coordination of the C^α site is superior for determining the alignment tensor, since it is less susceptible than the planar peptide plane to redundant RDC data. This issue is especially important for unstructured proteins since their dynamic

nature impedes the accuracy of local alignment determination and their intrinsically low signal dispersion suffers from the broadened lineshapes caused by H/D exchange. In all these cases it becomes favorable to measure RDCs around the C^{α} site using HCACO type experiments as described in Article IV.

The HCACO experiment was originally developed to correlate intraresidue resonances for backbone assignment (Kay *et al.*, 1990b). The experiment has been improved several times to include constant-time evolution, coherence order selection by pulsed field gradients, and enhanced water suppression. For measuring coupling constants it is also worthwhile to include a spin state selective filter in the HCACO sequence in order to minimize spectral crowding (Andersson *et al.*, 1998).

The coherence flow in all the HCACO type experiments introduced in Article IV can be described schematically as follows

$$\begin{aligned} {}^{1}\mathrm{H}^{\alpha} \xrightarrow{{}^{1}J_{\mathrm{CH}}} {}^{13}\mathrm{C}^{\alpha} \xrightarrow{{}^{1}J_{\mathrm{C'C}^{\alpha}}} {}^{13}\mathrm{C'}(t_1) \xrightarrow{{}^{1}J_{\mathrm{C'C}^{\alpha}}} \\ {}^{13}\mathrm{C}^{\alpha}(\mathrm{CT}{-}t_2) \xrightarrow{{}^{1}J_{\mathrm{CH}}} {}^{1}\mathrm{H}^{\alpha}(t_3) \end{aligned} \tag{4.4}$$

The pulse sequences are designed for measuring respectively ${}^{1}(J+D)_{\mathrm{H}^{\alpha}\mathrm{C}^{\alpha}}$, ${}^{1}(J+D)_{\mathrm{C}^{\alpha}\mathrm{C'}}$, and ${}^{2}(J+D)_{\mathrm{C'H}^{\alpha}}$ scalar and residual dipolar couplings. All the experiments feature coherence order selection by gradients in the final indirect dimension and spin state selection.

The main differences between the proposed experiments are which coupling interactions are active during frequency labeling and the spin states selective methods employed in connection here to. For instance, the HCACO(α/β-J-HACA) experiment depicted in Fig. 1a of Article IV features spin state selection in the detected dimension in a coherence order selective manner without the need for inphase/antiphase combinations. In contrast, the HCACO(α/β-J-C'HA) pulse sequence employs the E.COSY principle for accurately resolving the small ${}^{2}J_{\mathrm{C'H}^{\alpha}}$ coupling in addition to conventional spin state selection. Coherence transfer pathways of the individual experiments are described in detail in Article IV.

RDCs from additional bond vectors improve data accuracy for both structural and dynamical applications. With this in mind, we tested a pulse sequence for measuring ${}^{1}(J+D)_{\mathrm{C}^{\alpha}\mathrm{N}}$ couplings by an H^{α} detected method to supplement the set determined by the pulse sequences in Article IV. The HCAN type pulse sequence is shown in Fig. 4.5. Coherence transfer in this experiment is somewhat different from the HCACO approach with its schematic magnetization flow

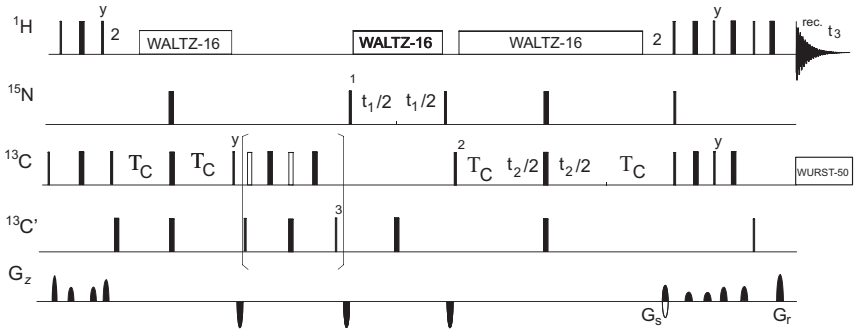


Figure 4.5: Spin state selective HCAN based pulse sequence for measurement of ${}^1(J+D)_{\mathrm{C}^\alpha\mathrm{N}}$ scalar and residual dipolar couplings in ^{13}C,^{15}N labeled proteins. Pulse sequence parameters are as given in Fig. 1 of Article IV. Delay settings $\tau = 1/(4J_{\mathrm{CH}}) =$ 1.7 ms, $\Delta = 1/(4J_{\mathrm{C}^\alpha\mathrm{C}'}) = 4.5$ ms, and $T_{\mathrm{C}} = 1/(4J_{\mathrm{C}^\alpha\mathrm{C}^\beta}) = 13.5$ ms. Phase settings: ϕ_{rec}=$\{x,-x,-x,x\}$, ϕ_1=$\{x,-x\}$+TPPI, ϕ_2=x+TPPI where TPPI denotes time proportional phase incrementation. Spin state selection is obtained by recording inphase and antiphase datasets during the filter marked by brackets. The filled 180° pulses on ${}^{13}\mathrm{C}^\alpha$ apply for the inphase dataset along with ϕ_3=x and the unfilled pulses for antiphase dataset with ϕ_3=y. Quadrature detection in F_1 is obtained by altering the phase of ϕ_1 according to States-TPPI and in F_2 by gradient selected coherence order selective transfer with the sign of Ψ=x inverted along with the sign of gradient strength G_S. Recommended experimental settings are as given in Article II, except for acquisition parameters which are 2 scans per FID and 160, 120, and 820 points in F_1-F_3, respectively. This corresponds to acquisition times of 80 ms, 24 ms, 51.2 ms in t_1, t_2, and t_3, respectively, with a total experiment time of 100 h.

being

$$ {}^1\mathrm{H}^\alpha \xrightarrow{{}^1J_{\mathrm{CH}}} {}^{13}\mathrm{C}^\alpha \xrightarrow{{}^1J_{\mathrm{C}'\mathrm{C}^\alpha}+{}^1J_{\mathrm{NC}^\alpha}+{}^2J_{\mathrm{NC}^\alpha}} {}^{13}\mathrm{C}' \xrightarrow{{}^1J_{\mathrm{C}'\mathrm{C}^\alpha}} {}^{13}\mathrm{C}'(\alpha/\beta) \rightarrow $$

$$ {}^{15}\mathrm{N}(t_1)(i,i+1) \xrightarrow{{}^1J_{\mathrm{C}'\mathrm{C}^\alpha}+{}^1J_{\mathrm{NC}^\alpha}+{}^2J_{\mathrm{NC}^\alpha}} {}^{13}\mathrm{C}^\alpha(\mathrm{CT}-t_2) \xrightarrow{{}^1J_{\mathrm{CH}}} {}^1\mathrm{H}^\alpha(t_3) \qquad (4.5) $$

Although magnetization arising from the sequential residue also gives rise a crosspeak in the ^{15}N dimension, we are only interested in the more intense intraresidual crosspeak.

The HCAN pulse sequence was tested experimentally, but the internuclear vector directions determined from the dipolar couplings were poorly correlating with those from the RDC refined solution structure of ubiquitin. It was later

realized that resolution enhancement using linear prediction diminished the accuracy of the couplings. Nevertheless, it can be anticipated that the depicted HCAN pulse sequence is useful for obtaining the one-bond scalar and residual dipolar couplings between C^{α} and N in small proteins.

The set of HCACO experiments presented in Article IV is tailored to produce coupling data with high accuracy. This is particularly the case in the pulse sequences where the couplings are measured in the ^{13}C' dimension which provides small linewidth and where ^{1}H-^{1}H dipolar couplings are not a concern. In extent to applications in structure determination and refinement of small proteins, this will prove useful for studying flexible systems. In addition, all three scalar coupling constants measured in Article III are known to correlate with backbone dihedral angles ϕ (between C'-N-C^{α}-C') and especially ψ (between N-C^{α}-C'-N) which can be used as energy penalty function for restrained molecular dynamics (Mierke *et al.*, 1992; Vuister and Bax, 1992; Cornilescu *et al.*, 2000).

4.3 Heteronuclear Long-range Correlation Experiments

Heteronuclear long-range connectivities provide a wealth of structural and assignment information in unlabeled molecules. The correlations are indispensable for linking the direct correlations obtained from HSQC. For instance in oligosaccharides, anomeric correlations yield a starting point for the assignment and when combined with long-range correlations open a pathway to the low-dispersed bulk region of ^{13}C-^{1}H correlation spectra.

The standard experiment for acquiring long-range connectivities is HMBC, which correlates protons to carbons separated by two or more bonds (Bax and Summers, 1986). Coherence transfer is performed through the small and often non-uniform $^{n}J_{\mathrm{CH}}$ couplings ($n \geq 2$) which render the experiment rather insensitive in comparison to direct-correlation experiments. The problem of non-uniform excitation over the long-range coupling magnitude span can be alleviated by using an ACCORDION principle or by coadding magnitude-mode spectra recorded with a range of excitation delays (Hadden *et al.*, 2000; Meissner and Sørensen, 2000).

While the large number of correlations obtained in HMBC spectra is desirable, it also causes ambiguities since there is no way to tell how many bonds the correlations are over. Furthermore, not all long-range correlations will be visible in the HMBC spectrum since some heteronuclear couplings may be vanishingly

small. Methods to overcome these intrinsic problems of HMBC have been developed based on variable scaling of peak tilt angles or alternative coherence transfer pathways (Krishnamurthy *et al.*, 2000; Sprang and Bigler, 2004). In the H2BC experiment coherence transfer is achieved through the $^{n}J_{\mathrm{HH}}$ ($n \geq 3$) couplings (Nyberg *et al.*, 2005). The H2BC spectrum corresponds to a HMQC-TOCSY pattern cut off after the first ^{1}H-^{1}H transfer step, thus correlations separated by two-bonds are almost solely observed. Overlaying the direct correlations obtained from HSQC with the two-bond correlations of H2BC enables assignment-walk. The direct ^{13}C-^{1}H correlations are here unequivocally connected through their two-bond correlations. This allows for tracing out the entire carbon skeleton of the molecule, *e.g.*, the intraring spin system of a carbohydrate residue in an oligosaccharides (Petersen *et al.*, 2006).

Although the H2BC experiment circumvents many of the problems associated with HMBC, it does by no means replace HMBC. The different coherence transfer pathways make the two experiments highly complementary. For instance, for establishing quaternary carbon correlations or linking over the glycosylic bond in carbohydrates, the HMBC experiment is indispensable. It is therefore still highly justified to improve this widely used experiment. A novel method for this purpose is presented below.

Key peaks in long-range correlation spectra can be of very low intensity and artifact suppression is therefore a central issue. It is desirable to suppress one-bond correlations since they are likely to cause overlap with multiple-bond correlations and complicate spectral analysis. Direct correlations are instead preferably obtained from HSQC spectra and subsequently overlaid with the long-range correlation spectra. Gradient selection of coherence transfer pathways ensures effective suppression of protons not coupled to ^{13}C, but one-bond correlations cannot be directly suppressed in this manner. In both HMBC and H2BC experiments the suppression is achieved by employing low pass J filters (Kogler *et al.*, 1983). Inphase ^{1}H magnetization is allowed to evolve into antiphase and the multiple quantum coherence generated by the 90°(^{13}C) pulse is subsequently purged. As the magnitude of $^{n}J_{\mathrm{CH}}$ is small relative to $^{1}J_{\mathrm{CH}}$, long-range magnetization remains almost solely inphase during the time delay $\tau = 1/(2\,{}^{1}J_{\mathrm{CH}})$ required for the transfer. By repeating the filter element with different delay settings, one-bond correlation suppression can in principle achieved over a span of ^{1}J magnitudes to any desired degree. While H2BC spectra do usually not suffer from incomplete suppression, it has been experimentally observed that one-bond correlations often occur in HMBC spectra despite accurate J filter tuning. The origin of these undesired artifacts is related to strong coupling as described below and in further detail in Article V.

4.3.1 Strong Coupling Artifacts Suppression by *Clean* HMBC

Strong coupling modulations to 1D pulse-acquisition spectra are well known. Especially peak intensities are distorted in comparison to the weak coupling case. For 2D experiments the effect is typically less problematic since assignment connectivities are established by correlations with more simple peak patterns and the peak intensities are less central to the spectral analysis. However, 2D J-spectra are known to be susceptible to strong coupling artifacts and also, for instance, interpretation of NOE distance relations is affected by the strong coupling modulations (Thrippleton *et al.*, 2005; Kay *et al.*, 1986). Although high magnetic field strengths and intensive use of heteronuclear experiments often make strong coupling artifacts a minor concern, they can be problematic when filtration of certain correlations are required. This is exactly the case of heteronuclear long-range correlation experiments.

Generation of strong coupling artifacts during the HMBC pulse sequence can mainly be ascribed to two culprits. The first is incomplete suppression by LPJF in the presence of strongly coupled spins, here called LPJF leakage. The second is unintended coherence mixing caused by the 180° pulse in the middle of the evolution period. The effects can be depicted in simple form for two strongly coupled spins A and B, respectively short-range and long-range coupled to a heteronucleus X, as following

$$\mathrm{A} + \mathrm{B} \xrightarrow{\mathrm{LPJF}} \eta_1 \mathrm{A} + \mathrm{B} \xrightarrow{180^\circ(^1\mathrm{H})} (\eta_1 + \eta_2)\mathrm{A} + \eta_1\eta_2\mathrm{B}. \tag{4.6}$$

η_1 indicates the degree of LPJF leakage and η_2 coherence transferred between spin B and spin A during the mixing period. In contrast to the HMBC experiment, LPJF is employed immediately prior to acquisition in the H2BC experiment, and unintended mixing of coherence between A and B spins is therefore suppressed. This fact is central to the suppression of strong coupling induced artifacts in *clean* HMBC.

To get a clear description of the origin of strong coupling artifacts magnetization transfer during the HMBC pulse sequence is analyzed for a three spin system in the following. Both the weak coupling (AMX spin system) and the strong coupling (ABX spin system) cases are analyzed using single-element basis operators in order to emphasize the parallel between the two cases. The preparation part of the pulse sequence is only briefly discussed using product operators. Incomplete one-bond correlation suppression due to LPJF leakage is described in the following section.

Coherence transfer in an AMX spin system during an m^{th} order LPJF, when neglecting homonuclear J_{AB} coupling evolution, is

$$A_z + M_z \xrightarrow{\frac{\pi}{2}(A_x+M_x),(H^{\text{AMX}}\,\tau_i,\frac{\pi}{2}X)^m} -A_y \prod_{i=1}^{n} \cos(\pi J_{\text{AX}}\tau_i) - M_y \prod_{i=1}^{n} \cos(\pi J_{\text{MX}}\tau_i).$$

Magnetization originating from the directly-coupled spin A is effectively suppressed when τ_i matches the coupling J_{AX} while long-range coupled magnetization M_y is virtually unaffected.

The HMBC experiment features multiple quantum evolution during frequency labeling. Long-range coupled magnetization is transferred to antiphase coherence during the long polarization transfer delay Δ. This delay is set as a compromise of maximizing the intensity of the broad range of coupling magnitudes while minimizing relaxation losses. In operator terms, the preparation period when focusing only on detected terms is

$$-M_y \xrightarrow{2\pi J_{\text{MX}} M_z X_z \Delta} 2M_x X_z \xrightarrow{\frac{\pi}{2}X_x} -2M_x X_y \xrightarrow{G_z} -2M^+X^+.$$

G_z indicates coherence order selection by pulsed field gradients. Although pure absorption phases are not obtainable, since ^{1}H chemical shift is not refocused, phase-modulated echo/antiecho processing results in superior resolution when compared to recording only one of the echo parts.

The double quantum product operator can be expressed in terms of single-element operators $M^+X^+ = A^\alpha M^+X^+ + A^\beta M^+X^+$. The two terms correspond, respectively, to the density matrix elements σ_{15} and σ_{48} in the weak coupling limit. Evolution of single-element operators is easily calculated during free precession according to the density matrix evolution under the weak coupling limit Hamiltonian H^{AMX}

$$\sigma_{ij} \xrightarrow{H^{\text{AMX}}\,t} \sigma_{ij} e^{i(\omega_j-\omega_i)t} \equiv \sigma_{ij} e_{ij}^{t}. \tag{4.7}$$

For the analysis during the evolution period we focus on the term $A^\beta M^+X^+$. This term evolves during t_1 as

$$A^\beta M^+X^+ \xrightarrow{H^{\text{AMX}}\,t_1/2,\pi(A_x+M_x),H^{\text{AMX}}\,t_1/2}$$

$$A^\alpha M^-X^+ e_{48}^{t_1/2} e_{32}^{t_1/2} = A^\alpha M^-X^+ e^{i(-\omega_X+\pi J_{\text{AM}})t_1}.$$

The final expression is obtained by inserting the AMX eigenfrequencies given in Article V. During the gradient echo only homonuclear evolution occurs so the

backtransfer amounts to the 90°(X) pulse. The frequency modulation during acquisition is given by

$$A^\alpha M^- X^+ \xrightarrow{\frac{\pi}{2} X_x} \tfrac{1}{2} A^\alpha M^- X^\alpha - \tfrac{1}{2} A^\alpha M^- X^\beta \xrightarrow{H^{\mathrm{AMX}} t_2}$$

$$\tfrac{1}{2} A^\alpha M^- X^\alpha e^{i(\omega_M + \pi J_{\mathrm{AM}} + \pi J_{\mathrm{MX}}) t_2} - \tfrac{1}{2} A^\alpha M^- X^\beta e^{i(\omega_M + \pi J_{\mathrm{AM}} - \pi J_{\mathrm{MX}}) t_2}.$$

The two terms arising at $\omega_M + \pi J_{\mathrm{AM}}$ split by πJ_{MX} correspond to the lower two peaks simulated in Fig. 4a [V]. In the case of a strongly coupled spin system additional peaks appear in HMBC spectra. Calculation of these peaks is slightly more complicated as exemplified below.

Strong coupling case The eigenbase is no longer identical to the product operator base in the strong coupling case. Rather the eigenfunctions of the strong coupling Hamiltonian are linear combinations of the weak coupling eigenfunctions with same magnetic quantum number and cannot be ascribed to a single nucleus. The degree of mixing between two spins A and B is characterized by the strong coupling parameter θ. In an ABX spin system this is given by

$$\tan(2\theta_{\alpha/\beta}) = \frac{2\pi J_{\mathrm{AB}}}{\omega_{\mathrm{A}} - \omega_{\mathrm{B}} \pm \pi(J_{\mathrm{AX}} - J_{\mathrm{BX}})}, \tag{4.8}$$

where the α/β subscripts refer to the spin state of the X spin. Small $\theta_{\alpha/\beta}$ values mean that the spins are rather weakly coupled. The strong coupling eigenbase goes into the weak coupling product operator eigenbase in the limit $\theta_{\alpha/\beta} \to 0$.

The effect of pulses is complicated by the strong coupling perturbation to the eigenfunctions. A 180° pulse no longer only inverts the sign of coherences and switches α and β spin states as in the weak coupling case. Rather the pulse is also causing coherence transfer between transitions, which in the weak coupling limit are associated with different spins (Ernst *et al.*, 1987).

Analogously to the weak coupling case, the indirect evolution period starts with the coherence term $\{B^+X^-\}$. The curly bracket notation makes explicit connection of the operators to their weak coupling counterparts (Thrippleton *et al.*, 2005). In the single-element notation the term equals $\{A^\alpha B^+ X^-\} + \{A^\beta B^+ X^-\}$, representing respectively σ_{15} and σ_{48} in the ABX density matrix. In parallel to the weak coupling case we obtain for the evolution of the latter

$$\{A^\beta B^+ X^+\} \xrightarrow{H^{\mathrm{ABX}} t_1/2} \{A^\beta B^+ X^+\} e_{48}^{t_1/2}.$$

The effect of the 180° pulse is

$$\{A^{\beta}B^{+}X^{+}\}e_{48}^{t_1/2} \xrightarrow{\pi(A_x+B_x)} \cos 2\theta_{\alpha}\{A^{\alpha}B^{-}X^{+}\}e_{48}^{t_1/2} - \sin 2\theta_{\alpha}\{A^{-}B^{\alpha}X^{+}\}e_{48}^{t_1/2}.$$

The two resulting terms evolve under free precession with different evolution frequencies

$$\cos 2\theta_{\alpha}\{A^{\alpha}B^{-}X^{+}\}e_{48}^{t_1/2} - \sin 2\theta_{\alpha}\{A^{-}B_{\alpha}X^{+}\}e_{48}^{t_1/2} \xrightarrow{H^{\mathrm{ABX}}\, t_1/2}$$

$$\cos 2\theta_{\alpha}\{A^{\alpha}B^{-}X_{+}\}e_{48}^{t_1/2}e_{32}^{t_1/2} - \sin 2\theta_{\alpha}\{A^{-}B_{\alpha}X^{+}\}e_{48}^{t_1/2}e_{42}^{t_1/2}.$$

Finally, by inserting the ABX eigenfrequencies from Article V we get

$$\cos 2\theta_{\alpha}\{A^{\alpha}B^{-}X^{+}\}e^{i(-\omega_X+\pi J_{\mathrm{AB}})t_1} - \sin 2\theta_{\alpha}\{A^{-}B^{\alpha}X^{+}\}e^{i(-\omega_X+\pi J_{\mathrm{AB}}+\Lambda_{\alpha})t_1}.$$

The first term is analogous to the weak coupling limit result, whereas the second term arises due to strong coupling. The 'strong coupling' term evolves at a Λ_{α}-shifted frequency in t_1. Most importantly, the term $\{A^{-}B^{\alpha}X^{+}\}$ corresponds to coherence primarily associated with spin A.

The 'weak coupling' term is the most intense for all ABX spin systems. This term evolves during the backtransfer and detection period as

$$\cos 2\theta_{\alpha}\{A^{\alpha}B^{-}X^{+}\} \xrightarrow{\frac{\pi}{2}X_x, H^{\mathrm{ABX}}\, t_2} \cos 2\theta_{\alpha}\left[\{A^{\alpha}B^{-}X^{\alpha}\}e_{31}^{t_2} + \{A^{\alpha}B^{-}X^{\beta}\}e_{52}^{t_2}\right]$$

$$\approx \cos 2\theta_{\alpha}\left[\{A^{\alpha}B^{-}X^{\alpha}\}e^{i(\omega_B+\pi(J_{\mathrm{AB}}+J_{\mathrm{BX}}))t_2} + \{A^{\alpha}B^{-}X^{\beta}\}e^{i(\omega_B+\pi(J_{\mathrm{AB}}-J_{\mathrm{BX}}))t_2}\right],$$

where the approximation of the eigenfrequencies holds true in the weak coupling limit. In contrast, the 'strong coupling' term evolves as

$$\sin 2\theta_{\alpha}\{A^{-}B^{\alpha}X^{\alpha}\} \xrightarrow{\frac{\pi}{2}X_x, H^{\mathrm{ABX}}\, t_2} \sin 2\theta_{\alpha}\left[\{A^{-}B^{\alpha}X^{\alpha}\}e_{41}^{t_2} + \{A^{-}B^{\alpha}X^{\beta}\}e_{62}^{t_2}\right]$$

$$\approx \sin 2\theta_{\alpha}\left[\{A^{-}B^{\alpha}X^{\alpha}\}e^{i(\omega_A+\pi(J_{\mathrm{AB}}+J_{\mathrm{AX}}))t_2} + \{A^{-}B^{\alpha}X^{\beta}\}e^{i(\omega_A+\pi(J_{\mathrm{AB}}-J_{\mathrm{AX}}))t_2}\right].$$

Hence the signals appear at the position where spin A correlations would occur in the weak coupling limit and with the characteristic large J_{AX} splitting. The final two coherence terms are the ones numbered artifact *1 in Fig. 4b [V]. Table 2 in Article V was derived in this way.

The overall result, that strong coupling induced artifacts are evolving under the effect of ${}^1J_{\mathrm{AX}}$ after acquisition, is used for achieving artifact suppression in *clean* HMBC. By applying LPJF immediately prior to acquisition the artifacts are suppressed to a high degree and what remains is only asymmetric LPJF leakage which cannot be suppressed by the devised method.

Low pass J filter leakage Coherence mixing due to the 180°(^{1}H) pulse is not the only cause of strong coupling induced one-bond coupled signals in HMBC. Low pass J filters have inherent limitations for strongly coupled spin systems. Incomplete LPJF suppression is due to the eigenfunctions being mixtures of the product operator states which in turn evolve under the effect of different couplings. Exact results for resonance intensities after a 1$^{\text{st}}$ order LPJF are given in Article V. These results are obtained from density matrix calculations by using the product base ϕ^{AMX} for transformations under *rf* pulses and the eigenbase ϕ^{ABX} for free precession. The combination of basis sets can be represented for a 1$^{\text{st}}$ order LPJF as follows

$$\phi_0^{\text{AMX}} \xrightarrow{U} \phi_0^{\text{ABX}} \xrightarrow{H^{\text{ABX}}\tau} \phi_\tau^{\text{ABX}} \xrightarrow{U^{-1}} \phi_\tau^{\text{AMX}} \xrightarrow{\frac{\pi}{2}X_z} \phi_{\pi/2}^{\text{AMX}} \xrightarrow{U} \phi_{\pi/2}^{\text{ABX}}. \tag{4.9}$$

U denotes the unitary transformation from the AMX product operator base to the ABX eigenbase, and ϕ_0^{AMX} and ϕ_0^{ABX} are the density matrices after the initial excitation pulse in their respective basis sets. Density matrix calculations are feasible in a similar way throughout the sequence, but as evident from this example, the analysis easily becomes too complicated to track the relevant terms.

An alternative, more intuitive picture of LPJF leakage can be visualized by decomposing the single-element operators in the eigenbase into their product operator states. For instance, the resonance line (41) in Fig. 2a [V], corresponding to $\{A^-B^\alpha X^\alpha\}$, arises from the transition

$$\begin{aligned} |4><1| &= \cos\theta_\alpha|\beta\alpha\alpha><\alpha\alpha\alpha| + \sin\theta_\alpha|\alpha\beta\alpha><\alpha\alpha\alpha| \\ &= \cos\theta_\alpha A^-B^\alpha X^\alpha + \sin\theta_\alpha A^\alpha B^- X^\alpha. \end{aligned} \tag{4.10}$$

The first term corresponds to the product operator evolving under the effect of $^1J_{\text{AX}}$ and is thus suppressed by the LPJF. Similarly, the second term is the product operator evolving under $^nJ_{\text{BX}}$ and therefore virtually unaffected by the filter. It is this remaining part with intensity on the order $\sim\sin\theta_\alpha$ we see as a result of imperfect filtration. Although the strong coupling effect slightly perturbs the effective coupling magnitude, this qualitative picture also holds true for the X^β states where the resulting intensity is $\sim\sin\theta_\beta$. Exact results from this picture can be obtained using the ABX product operator formalism (Kay and McClung, 1988; Nakai and McDowell, 1994).

Applying successive J filters is effective for achieving good suppression when ^1J does not match the τ delay setting, but is ineffective for suppression of LPJF leakage signals. This is because the leaking magnetization is still evolving

under the effect of the long-range coupling in successive filters. The transfer between evolution under J_{BX} and J_{AX} occurs with magnitude $\Lambda_{\alpha/\beta}$ (Nakai and McDowell, 1994). Hence for the spin system in Fig. 2a [V] higher order suppression is mainly affecting the X^{β} lines which are already well suppressed, while the additional effect on the X^{α} lines is negligible.

Low pass J filter implementation There are several ways to construct filters for suppression of protons directly coupled to a heteronucleus, here assumed to be carbon. The idea is to exploit the difference in magnitude between J couplings over one and over several bonds. The issue is related to isotope filtering and the differentiation of correlations can be achieved by half-filters or BIRD-pulse implementations (Otting *et al.*, 1986; Garbow *et al.*, 1982). Indeed, suppression of strong coupling induced correlations is possible through half-filters as has been demonstrated for HSQC spectra (Kövér and Batta, 1999).

When performing filtration of correlations steaming from a single molecule, heteronuclear long-range coupling evolution must be taken into account. The magnetization transferred from inphase to antiphase or vice versa is discarded which results in signal loss. The sensitivity loss per filter order is $\sin(\pi\, {}^{n}J\tau)$. Thus, sensitivity losses mainly occur for the most intense correlations where long-range coherence transfer is efficient, whereas the loss is negligible for the least intense correlations.

Most suppression filter types purge antiphase magnetization by phase cycling or gradients. However, if the desired magnetization is of antiphase character a non-destructive suppression filter is required. An example of such an implementation is depicted in Fig. 4.6a. The composite 180° pulse inverts long-range bound magnetization $B^{-}X_{z}$ in every other transient, while evolution of directly-coupled magnetization into A^{-} is rendering this term unaffected by the pulse. Combination of the separately stored transients enables separation of long-range and short-range correlations post-acquisitionally.

The 1^{st} order suppression outlined in Figure 4.6a works well in some cases. Unfortunately, for correlations not affected by strong coupling it turned out to be necessary to retain an initial 3^{rd} order LPJF to achieve the same suppression level as in conventional HMBC. Moreover, for molecules with a large spread in ${}^{1}J$ magnitudes it is necessary to perform higher order terminal filtering for efficient suppression of strong coupling artifacts. A possible remedy is the four scan 2^{nd} order filter implementation illustrated in Fig. 4.6b. The effective pulse rotations and delay combinations are shown explicitly in Fig. 4.6c. Since it is preferable to have a constant number of pulses applied in each scan, the

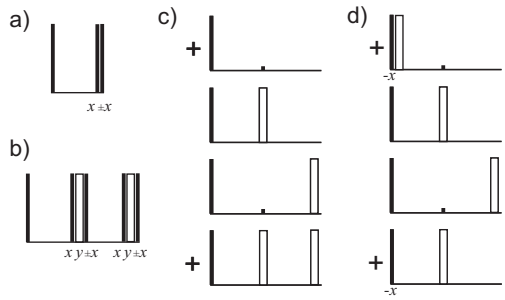


Figure 4.6: Examples of non-destructive LPJF versions applicable as terminal filter for *clean* HMBC. The first 90°(^{13}C) pulse is the backtransfer pulse and not part of the filter as such. *a*) 1$^{\text{st}}$ order filter with effective 0° and 180° pulses applied alternately with post-acquisitional subtraction of the two transients. *b*) 2$^{\text{nd}}$ order filter using four combinations with two pulses in every scan. *c*) The explicit combinations of *b* in terms of their effective rotations. Signs indicate the linear combination required for extracting long-range correlations post-acquisitionally. *d*) MBOB-type LPJF: the equivalent of *c* using a single inversion pulse in each scan.

composite inversion pulse version is preferred. This procedure for extending the LPJF order is, however, increasingly costly in terms of hampering off-resonance effects on the ^{13}C channel, which can have a very large spectral width. Although composite pulses can diminish the problem some loss is inevitable. To resolve these issues we turned to the MBOB procedure for non-destructive J filtering (Meissner and Sørensen, 2000). This is illustrated in Fig. 4.6d. The MBOB implementation yields the equivalent effect on the spin system as the procedure in Fig. 4.6c, but using only a single 180° pulse in each scan. The phase of the backtransfer 90° pulse can be combined with the sign of the linear combinations. Thus an equivalent effect is achieved by dropping phase shifting and coadding all transients. To facilitate processing, spectral transients can be combined in real time on the spectrometer albeit with the cost of only selecting a single data combination; the suppressed correlations can therefore not be extracted using this approach.

The effect of the 2$^{\text{nd}}$ order non-destructive LPJF shown in Fig. 4.6d is equivalent to a 2$^{\text{nd}}$ order destructive LPJF. In the first transient magnetization evolves under J_{AX} during the time $\tau_1 + \tau_2$, whereas in the second transient it is evolving during $\tau_1 - \tau_2$. The scans are coadded and the second order intensity scaling

follows

$$\cos \pi J(\tau_1 + \tau_2) + \cos \pi J(\tau_1 - \tau_2) = 2 \cos \pi J \tau_1 \cos \pi J \tau_2. \qquad (4.11)$$

An identical effect results from coaddition of the third and fourth transients. However, any magnetization having evolved into inphase during the filter delays occurs with the opposite sign of the latter combinations. Coaddition of all four transients thus eliminates this 1^{st} order term artifact.

In principle, a similar suppression over a range of J couplings is achievable by applying an adiabatic inversion pulse and exploiting the roughly linear correlation between scalar coupling constants $^1J_{CH}$ and ^{13}C chemical shift (Kupče, 2001). This approach would limit the total numbers of scans required for effective suppression, but optimizing the pulse for a given J span at every usage would be more cumbersome. If the increased minimal number of scans required for the *clean* HMBC is problematic it is possible to create a broadband one-scan non-destructive filter by sophisticated composite rotation pulses (Sørensen, 1989; Stuart *et al.*, 1999). The *clean* HMBC artifact suppression scheme is directly applicable to the multitude of HMBC modifications, such as multiplicity edited and broadband HMBC (Nyberg and Sørensen, 2006).

Experimental demonstration on molecule with broad ^1J span *Clean* HMBC is shown to provide efficient strong coupling artifact suppression in a monosaccharide in Article V. The small $^1J_{CH}$ span in the bulk region of carbohydrates, however, raises the question whether the suppression is effective over a broader range of scalar coupling constants. Therefore we tested *clean* HMBC on the more challenging case of androsterone (Fig. 4.7) where the spin pattern is a myriad of IS, I_2S and I_3S systems and the J coupling span is $\sim$50 Hz.

The wider applicability of *clean* HMBC is demonstrated by comparison with conventional HMBC in Fig. 4.8. Incompletely suppressed one-bond correlations were verified by overlaying t_2-coupled HSQC with the HMBC spectra. Evidently, the suppression obtained by *clean* HMBC is superior to conventional HMBC.

Several one-bond correlations complicate spectral analysis in the conventional experiment, but are effectively suppressed by the improved LPJF implementation in *clean* HMBC. While some artifacts are easily identified as one-bond correlations in Fig. 4.8a, *e.g.* $^{13}C_{4\beta}$-$^1H_{4\beta}$ and the intense methyl signal $^{13}C_{18}$-$^1H_{18}$, most of them are overlapping with long-range correlations and may be difficult to assign as artifacts. For more complex molecules, such as oligosaccharides, substantial one-bond correlation artifacts are inevitable and the *clean* HMBC can prove most helpful in these cases. In addition, automated spectral

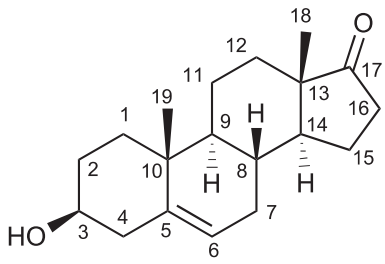
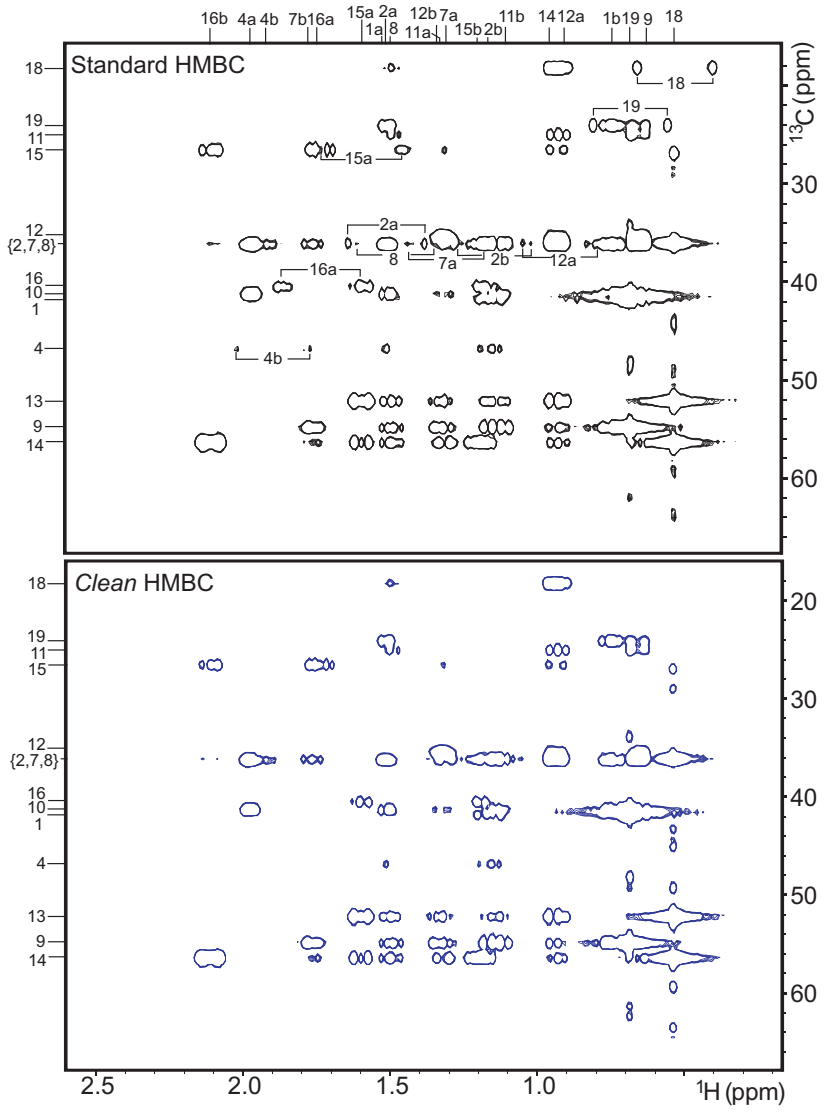


Figure 4.7: Structure of *trans*-dehydroisoandrosterone (5-androsten-3β-ol-17-one) with numbering referring to the correlations indicated in Fig. 4.8.

analysis protocols benefit from less artifact-prone spectra since data analysis becomes less complicated.

Figure 4.8 (following page): Excerpts of HMBC spectra from 150 mM dehydroisoandrosterone in $CDCl_3$. *a*) Conventional HMBC with 3rd order initial LPJF. Incompletely suppressed one-bond correlations are indicated. *b*) *Clean* HMBC with 1st order initial and 2nd order terminal LPJF. Although only the aliphatic region is shown, the full spectrum was recorded without noticeable signal difference between the two variants. Both spectra were recorded with 16 transients per FID and 256 t_1 increments. Long-range transfer delay was Δ=65 ms, and the low pass J filter delay span set for 120 Hz $< J_{CH} <$ 175 Hz. The data matrix covering 25 kHz and 5 kHz in the indirect, respectively the direct dimension, were apodized prior to Fourier transformation using squared cosinebell in t_1 and a $\pi/4$-shifted square sinebell in t_2. The ringing artifacts of the intense methyl correlations (18 and 19) along F_1 could be suppressed by applying stronger apodization functions.


Standard HMBC
Clean HMBC
13C (ppm)
1H (ppm)
2.5
2.0
1.5
1.0
16b 4a 4b 7b 16a 15a 1a 2a 8 12b 11a 7a 15b 2b 11b 14 12a 1b 19 9 18
18
19
11
15
12
{2,7,8}
16
10
1
4
13
9
14
15a
2a
8
7a
2b
12a
16a
4b
20
30
40
50
60

Chapter 5

Conclusion

A variety of pulse sequence techniques has been proposed in this thesis to facilitate structure determination of biomolecules by solution state NMR spectroscopy. The introduced methods may also find applications in studies of protein dynamics and protein-ligand interactions. Emphasis has been put on developing cost-effective methods which take advantage of standard protein production protocols as well as devising methods with improved applicability for automated structure determination. These issues remain bottlenecks for solution state structure determination.

Overcoming the size limitations of proteins amenable to solution state structure determination is a continuing challenge in protein NMR spectroscopy. As described in the previous chapters, versatile techniques are available for studying large proteins. They are, however, more elaborate and costly than the routine methods applicable to double labeled proteins. The experiments for assignment of methyl-containing residues described in this thesis provide a powerful alternative for side chain assignment. This circumvents the need for specific labeling or deuteration of mid-range molecular weight proteins.

The field of NMR spectroscopy continues to lag behind X-ray crystallography for protein structure determination. Nonetheless, the unique capabilities of NMR spectroscopy beyond solving single conformation structures are becoming increasingly important as the research emphasis in proteomics shifts from protein structures towards biological function and systems biology. The target focus and emphasis on structure/activity relationship in drug discovery also increase the use of NMR methods in the pharmaceutical industry (Pellecchia *et al.*, 2002). The experimental building blocks for simultaneous recording of methyl

and amide correlations help to minimize the sample amount requirements for screening of protein-ligand binding and for collecting multidimensional NOESY spectra.

Residual dipolar couplings continue to find new applications in structural biology, in particular for studying multiconformational dynamics and partially folded proteins (Fredriksson *et al.*, 2004; Jensen *et al.*, 2008). The proposed experiments for measuring RDCs will be useful for studying weakly structured proteins and provide insight into protein folding mechanisms. These areas are likely to receive much attention in the future.

Proteins and nucleic acids have received the limelight in structural biology. Nevertheless, numerous important biological functions are related to macromolecular carbohydrates and lipids. Structural analysis of these molecules is found challenging as the biosynthetical methods for their production and isotope labeling are limited. The spectral analysis relies to a high extent on long-range connectivities. The *clean* HMBC experiment is the method of choice for measuring heteronuclear multiple-bond correlations. The improved artifact suppression in the experiment presents a great asset when facing structure elucidation of complex unlabeled biomolecules.

Continuous progress in biochemical sample preparation techniques, along with improvements in hardware and data analysis methods hold promise to push the versatile applications of NMR spectroscopy. Increasingly robust and fast methods continue to appear in pulse sequence methodology. The new techniques are likely to alter the way NMR spectroscopy is performed in the future and widely extend its application range. For instance, dynamic nuclear polarization has been demonstrated to dramatically increase signal-to-noise ratios (Maly *et al.*, 2008). A deeper understanding of the relation between chemical shifts and global geometry may circumvent the need to acquire distance restraints (Shen *et al.*, 2008). Meanwhile, it is important to improve the reliable and robust NMR methods that exist today. I hope that the present work will benefit this effort and find applications for a broad range of studies towards increased understanding of biomolecules.

Acknowledgments

First of all, I wish to thank my supervisor Dr. Perttu Permi, the head of the Finnish National Biological NMR Center, for support and guidance throughout the study. Your expertise has been indispensable for realizing this project, both while designing the pulse sequences and when 'fighting' with the spectrometers to actually make the experiments work. My gratitude also goes to professor Arto Annila for introducing me to the fascinating field of biomolecular NMR spectroscopy and for inspiring me with diverse matters.

I have been fortunate to be surrounded by skillful colleagues who have made my working days pleasant. Sincere thanks to Kai Fredriksson, Maarit Hellman, Helena Tossavainen, Tero Pihlajamaa, and Olli Aitio for helping me all the way with data analysis and in particular with sample preparation. For me as a spectroscopist, it has been a true privilege to have a 'drive-in' protein production facility next door. Warm thanks to all my other colleagues during the time in Helsinki: Martti, Renee, Ville, Jani, Tapio, Eetu, Ndegwa, Kimmo, Paula, Petri, Tuomas, Harri, Anne, Juho, Elina, Raila, Sampo, Vytas, Hannu, and Hideo. Especially I want to thank my office mates for all the fun we have had through the years.

I am grateful to professors Niels Chr. Nielsen and Ole W. Sørensen for your contributions to initiating and finalizing the small molecule project. Your creativity and persistence improved the quality of the data analysis and the final experiment. In the early parts of my studies you taught me much of what I know about NMR spectroscopy.

The present study was conducted at the Institute of Biotechnology of the University of Helsinki. I wish to thank research director Mart Saarma for providing the excellent working facilities at the Institute of Biotechnology. This has allowed me to truly focus on the research work. Sincere thanks to professor Markku Räsänen from the Department of Chemistry for welcoming me into the Laboratory of Physical Chemistry. I would also like to acknowledge professors Jukka Jokisaari and Reino Laatikainen for kindly reviewing my thesis work.

I wish to thank the Finnish National Graduate School in Informational and Structural Biology and its director Mark Johnson for financial support during my graduate studies. The Academy of Finland is gratefully acknowledged for funding in the initial phase of the project.

Finally, my deepest thanks to my wife Marjukka. Your love and support has borne me through this project and your smile is a blessing every day.

Helsinki, October 2008 Peter Würtz

Bibliography

Andersson, P, Weigelt, J, and Otting, G. 1998. Spin-state selection filters for the measurement of heteronuclear one-bond coupling constants. *J. Biomol. NMR*, **12**, 435–441.

Annila, A, Aitio, H, Thulin, E, and Drakenberg, T. 1999. Recognition of protein folds via dipolar couplings. *J. Biomol. NMR*, **14**, 223–230.

Aue, WP, Bartholdi, E, and Ernst, RR. 1976. Two-dimensional spectroscopy. Application to nuclear magnetic resonance. *J. Chem. Phys.*, **64**, 2229–2246.

Bax, A, and Summers, MF. 1986. ^{1}H and ^{13}C assignments from sensitivity-enhanced detection of heteronuclear multiple quantum NMR. *J. Am. Chem. Soc.*, **108**, 2093–2094.

Bax, A, Clore, GM, and Gronenborn, AM. 1990. ^{1}H-^{1}H correlation via isotropic mixing of ^{13}C magnetization, a new three-dimensional approach for ^{1}H and ^{13}C spectra of ^{13}C-enriched proteins. *J. Magn. Reson.*, **88**, 425–431.

Bax, A, Kontaxis, G, and Tjandra, N. 2001. Dipolar couplings in macromolecular structure determination. *Methods Enzymol.*, **339**, 127–174.

Bodenhausen, G, and Ruben, DJ. 1980. Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy. *Chem. Phys. Letters*, **69**, 185–189.

Boehr, DD, Dyson, HJ, and Wright, PE. 2006. An NMR perspective on enzyme dynamics. *Chem. Rev.*, **106**, 3055–3079.

Braunschweiler, L, and Ernst, RR. 1983. Correlation transfer by isotropic mixing: application to proton correlation spectroscopy. *J. Magn. Reson.*, **53**, 521–528.

Cavalli, A, Salvatella, X, Dobson, CM, and Vendruscolo, M. 2007. Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. USA*, **104**, 9615–9620.

Cornilescu, G, Delaglio, F, and Bax, A. 1999. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.

Cornilescu, G, Bax, A, and Case, DA. 2000. Large variations in one-bond $^{13}C^{\alpha}$-$^{13}C^{\beta}$ J couplings in polypeptides correlate with backbone conformation. *J. Am. Chem. Soc.*, **122**, 2168–2171.

Delaglio, F, Kontaxis, G, and Bax, A. 2000. Protein structure determination using molecular fragment replacement and NMR dipolar couplings. *J. Am. Chem. Soc.*, **122**, 2142–2143.

Diercks, T, Schwaiger, M, and Kessler, H. 1998. HSQC-based methyl group selection via gradients in multidimensional NMR spectroscopy of proteins. *J. Magn. Reson.*, **130**, 335–340.

Dingley, AJ, and Grzesiek, S. 1998. Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide $^{2}J_{\mathrm{NN}}$ couplings. *J. Am. Chem. Soc.*, **120**, 8293–8297.

Duus, J, Gotfredsen, CH, and Bock, K. 2000. Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. *Chem. Rev.*, **100**, 4589.

Ernst, RR, Bodenhausen, G, and Wokaun, A. 1987. *Principles of nuclear magnetic resonance in one and two dimensions*. Clarendon press.

Fesik, SW, Eaton, HL, Olejniczak, ET, Zuiderweg, ERP, McIntosh, LP, and Dahlquist, FW. 1990. 2D and 3D NMR spectroscopy employing ^{13}C-^{13}C magnetization transfer by isotropic mixing. Spin system identification in large proteins. *J. Am. Chem. Soc.*, **112**, 886–888.

Forsén, S, and Hoffman, RA. 1963. Study of moderately rapid chemical exchange reactions by means of nuclear magnetic double resonance. *J. Chem. Phys.*, **39**, 2892–2901.

Fredriksson, K., Louhivuori, M., Permi, P., and Annila, A. 2004. On the Interpretation of Residual Dipolar Couplings as Reporters of Molecular Dynamics. *J. Am. Chem. Soc.*, **126**, 12646–12650.

Frueh, DP, Vosburg, DA, Walsh, CT, and Wagner, G. 2006. Determination of all nOes in ^{1}H-^{13}C-Me-ILV-U-^{2}H-^{15}N proteins with two time-shared experiments. *J. Biomol. NMR*, **34**, 31–40.

Fukushi, E. 2006. Advanced NMR approaches for a detailed structure analysis of natural products. *Biosci. Biotechnol. Biochem*, **70**, 1803–1812.

Garbow, JR, Weitekamp, DP, and Pines, A. 1982. Bilinear rotation decoupling of homonuclear scalar interactions. *Chem. Phys. Lett.*, **93**, 504–509.

Griesinger, C, Sørensen, OW, and Ernst, RR. 1986. Correlation of connected transitions by two-dimensional NMR spectroscopy. *J. Chem. Phys.*, **85**, 6837.

Griesinger, C, Sørensen, OW, and Ernst, RR. 1989. Three-dimensional Fourier spectroscopy. Application to high-resolution NMR. *J. Magn. Reson.*, **84**, 14–63.

Grishaev, A, Wu, J, Trewhella, J, and Bax, A. 2005. Refinement of multidomain protein structures by combination of solution small-angle X-ray scattering and NMR data. *J. Am. Chem. Soc.*, **127**, 16621–16628.

Grzesiek, S, Anglister, J, and Bax, A. 1993. Correlation of backbone amide and aliphatic side-chain resonances in ^{13}C/^{15}N-enriched proteins by isotropic mixing of ^{13}C magnetization. *J. Magn. Reson. B*, **101**, 114–119.

Güntert, P. 1998. Structure calculation of biological macromolecules from NMR data. *Q. Rev. Biophys.*, **31**, 145–237.

Hadden, CE, Martin, GE, and Krishnamurthy, VV. 2000. Constant time inverse-detection gradient accordion rescaled heteronuclear multiple bond correlation spectroscopy: CIGAR-HMBC. *Magn. Reson. Chem.*, 143–147.

Hajduk, PJ, Augeri, DJ, Mack, J, Mendoza, R, Yang, J, Betz, SF, and Fesik, SW. 2000. NMR-based screening of proteins containing ^{13}C-labeled methyl groups. *J. Am. Chem. Soc.*, **122**, 7898–7904.

Ikura, M, Kay, LE, and Bax, A. 1990. A novel approach for sequential assignment of ^{1}H, ^{13}C, and ^{15}N spectra of proteins: heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry*, **29**, 4659–4667.

Jensen, MR, Houben, K, Lescop, E, Blanchard, L, Ruigrok, RW, and Blackledge, M. 2008. Quantitative conformational analysis of partially folded proteins from residual dipolar couplings: application to the molecular recognition element of Sendai virus nucleoprotein. *J. Am. Chem. Soc.*, **130**, 8055–8061.

Kaptein, R, Zuiderweg, ER, Scheek, RM, Boelens, R, and van Gunsteren, WF. 1985. A protein structure from nuclear magnetic resonance data. lac repressor headpiece. *J. Mol. Biol.*, **182**, 179–182.

Kay, LE, and McClung, RED. 1988. A product operator description of AB and ABX spin systems. *J. Magn. Reson.*, **77**, 258–273.

Kay, LE, Holak, TA, Johnson, BA, Armitage, IM, and Prestegaard, JH. 1986. Second-order effects in two-dimensional cross-relaxation spectra of proteins: investigation of glycine spin systems. *J. Am. Chem. Soc.*, **108**, 4242–4244.

Kay, LE, Clore, GM, Bax, A, and Gronenborn, AM. 1990a. Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1 beta in solution. *Science*, **249**, 411–414.

Kay, LE, Ikura, M, Tschudin, R, and Bax, A. 1990b. Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.*, **89**, 496–514.

Kogler, H, Sørensen, OW, Bodenhausen, G, and Ernst, RR. 1983. Low-pass J filters. Suppression of neighbor peaks in heteronuclear relayed correlation spectra. *J. Magn. Reson.*, **55**, 157–163.

Koskela, H, Heikkinen, O, Kilpeläinen, I, and Heikkinen, S. 2007. Rapid and accurate processing method for amide proton exchange rate measurement in proteins. *J. Biomol. NMR*, **37**, 313–320.

Kövér, KE, and Batta, G. 1999. A general scheme for suppression of ABX strong coupling signals in heteronuclear scalar and dipolar correlation experiments. *J. Magn. Reson.*, **138**, 89–97.

Kövér, KE, Uhrín, D, and Hruby, VJ. 1998. Gradient- and sensitivity-enhanced TOCSY experiments. *J. Magn. Reson.*, **130**, 162–168.

Krishnamurthy, VV, Russell, DJ, Hadden, CE, and Martin, GE. 2000. ^2J,^3J-HMBC: A new long-range heteronuclear shift correlation technique capable of differentiating $^2J_{\mathrm{CH}}$ from $^3J_{\mathrm{CH}}$ correlations to protonated carbons. *J. Magn. Reson.*, **146**, 232–239.

Kumar, A, Ernst, RR, and Wüthrich, K. 1980. A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules. *Biochem Biophys Res Comm*, **95**, 1–6.

Kupče, E. 2001. Applications of adiabatic pulses in biomolecular nuclear magnetic resonance. *Methods Enzymol.*, **338**, 82–111.

Laatikainen, R, Niemitz, M, Malaisse, WH, Biesemans, M, and Willem, R. 1996. A computational strategy for the deconvolution of NMR spectra with multiplet structures and constraints: analysis of overlapping ^{13}C-^{2}H multiplets of 13C enriched metabolites from cell suspensions incubated in deuterated media. *Magn. Reson. Med.*, **36**, 359–366.

Lange, OF, Lakomek, NA, Farès, C, Schröder, GF, Walter, KF, Becker, S, Meiler, J, Grubmüller, H, Griesinger, C, and de Groot, BL. 2008. Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution. *Science*, **320**, 1471–1475.

Lindorff-Larsen, K, Best, RB, Depristo, MA, Dobson, C, and Vendruscolo, M. 2005. Simultaneous determination of protein structure and dynamics. *Nature*, **433**, 128–132.

Maly, T, Debelouchina, GT, Bajaj, VS, Hu, KN, Joo, CG, Mak-Jurkauskas, ML, Sirigiri, JR, van der Wel, PCA, Herzfeld, J, Temkin, RJ, and Griffin, RG. 2008. Dynamic nuclear polarization at high magnetic fields. *J. Chem. Phys.*, **128**, 052211.

Mattinen, ML, Pääkkonen, K, Ikonen, T, Craven, J, Drakenberg, T, Serimaa, R, Waltho, J, and Annila, A. 2002. Quaternary structure built from subunits combining NMR and small-angle X-ray scattering data. *Biophys. J.*, **83**, 1177–1183.

Meissner, A, and Sørensen, OW. 2000. Economizing spectrometer time and broadband excitation in small-molecule heteronuclear NMR correlation spectroscopy. Broadband HMBC. *Magn. Reson. Chem.*, **38**, 981– 984.

Meissner, A, Duus, JØ, and Sørensen, OW. 1997. Spin-state-selective excitation. Application for E.COSY-type measurement of homonuclear coupling constants. *J. Magn. Reson.*, **128**(92–97).

Meyer, B, and Peters, T. 2003. NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew. Chem. Int. Ed. Engl.*, **42**, 864–890.

Mierke, DF, Grdadolnik, SG, and Kessler, H. 1992. Use of one-bond $C^{\alpha}H^{\alpha}$ coupling constants as restraints in MD simulations. *J. Am. Chem. Soc.*, **114**, 8283–8284.

Montelione, GT, Lyons, BA, Emerson, SD, and Tashiro, M. 1992. An efficient triple resonance experiment using carbon-13 isotropic mixing for determining sequence-specific resonance assignments of isotopically-enriched proteins. *J. Am. Chem. Soc.*, **114**, 10974–10975.

Nakai, T, and McDowell, CA. 1994. Product operator theory for ABX spin systems and its application to H-C-C INEPT NMR experiments. *Mol. Physics*, **81**, 337–358.

Neri, D, Szyperski, T, Otting, G, Senn, H, and Wüthrich, K. 1989. Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ^{13}C labeling. *Biochemistry*, **28**, 7510–7516.

Nolis, P, and Parella, T. 2007. Simultaneous α/β spin-state selection for ^{13}C and ^{15}N from a time-shared HSQC-IPAP experiment. *J. Biomol. NMR*, **37**, 65–77.

Nolis, P, Espinosa, JF, and Parella, T. 2006. Optimum spin-state selection for all multiplicities in the acquisition dimension of the HSQC experiment. *J. Magn. Reson.*, **180**, 39–50.

Nyberg, NT, and Sørensen, OW. 2006. Multiplicity-edited broadband HMBC NMR spectra. *Magn. Reson. Chem.*, **44**, 451–454.

Nyberg, NT, Duus, JØ, and Sørensen, OW. 2005. Heteronuclear two-bond correlation: suppressing heteronuclear three-bond or higher NMR correlations even for vanishing $^2J_{\mathrm{CH}}$. *J. Am. Chem. Soc.*, **127**, 6154–6155.

Olejniczak, ET, Xu, RX, and Fesik, SW. 1992. A 4D HCCH-TOCSY experiment for assigning the side chain ^{1}H and ^{13}C resonances of proteins. *J. Biomol. NMR*, **2**, 655–659.

Ottiger, M, Delaglio, F, and Bax, A. 1998. Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.*, **131**, 373–378.

Otting, G, Senn, H, Wagner, G, and Wüthrich, K. 1986. Editing of 2D ^{1}H NMR spectra using X half-filters. Combined use with residue-selective ^{15}N labeling of proteins. *J. Magn. Reson.*, **70**, 500–505.

Palmer, AG. 2004. NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.*, **104**, 3623–3640.

Pellecchia, M, Sem, DS, and Wüthrich, Kurt. 2002. Nmr in drug discovery. *Nature Rev. Drug Discov.*, **1**, 211–219.

Permi, P. 2002a. Intraresidual HNCA: an experiment for correlating only intraresidual backbone resonances. *J. Biomol. NMR*, **23**, 201–209.

Permi, P. 2002b. A spin-state-selective experiment for measuring heteronuclear one-bond and two-bond couplings from an HSQC-type spectrum. *J. Biomol. NMR*, **22**, 27–35.

Permi, P, and Annila, A. 2000. Transverse relaxation optimised spin-state selective NMR experiments for measurement of residual dipolar couplings. *J. Biomol. NMR*, **16**, 221–227.

Permi, P, and Annila, A. 2004. Coherence transfer in proteins. *Progr. NMR Spectrosc.*, **44**, 97–137.

Permi, P, Tossavainen, H, and Hellman, M. 2004. Efficient assignment of methyl resonances: enhanced sensitivity by gradient selection in a DE-MQ(H)CCmHmTOCSY experiment. *J. Biomol. NMR*, **30**, 275–282.

Pervushin, K, Riek, R, Wider, G, and Wüthrich, K. 1997. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. USA*, **94**, 12366–12371.

Petersen, BO, Vinogradov, E, Kay, W, Würtz, P, Nyberg, NT, Duus, JØ, and Sørensen, OW. 2006. H2BC: a new technique for NMR analysis of complex carbohydrates. *Carbohydr. Res.*, **341**, 550–556.

Prestegard, JH, Al-Hashimi, HM, and Tolman, JR. 2000. NMR structures of biomolecules using field oriented media and residual dipolar couplings. *Q. Rev. Biophys.*, **33**, 371–424.

Reif, B, Hennig, M, and Griesinger, C. 1997. Direct measurement of angles between bond vectors in high-resolution NMR. *Science*, **276**, 1230 – 1233.

Salzmann, M, Pervushin, K, Wider, G, Senn, H, and Wüthrich, K. 1998. TROSY in triple-resonance experiments: New perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. USA*, **95**, 13585–13590.

Sattler, M, and Fesik, SW. 1996. Use of deuterium labeling in NMR: overcoming a sizeable problem. *Structure*, **4**, 1245–1249.

Sattler, M, Schmidt, P, Schleucher, J, Schedletzky, O, Glaser, SJ, and Griesinger, C. 1995a. Novel pulse sequences with sensitivity enhancement for in-phase coherence transfer employing pulsed field gradients. *J. Magn. Reson.*, **108**, 235–242.

Sattler, M, Maurer, M, Schleucher, J, and Griesinger, C. 1995b. A simultaneous ^{15}N, ^{1}H- and ^{13}C^{1}H-HSQC with sensitivity enhancement and a heteronuclear gradient echo. *J. Biomol. NMR*, **5**, 97.

Sattler, M, Schleucher, J, and Griesinger, C. 1999. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Progr. NMR Spectrosc.*, **34**, 93–158.

Schleucher, J, Schwendinger, M, Sattler, M, Schmidt, P, Glaser, SJ, Sørensen, OW, and Griesinger, C. 1994. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR*, **4**, 301–306.

Shen, Y, Lange, O, Delaglio, F, Rossi, P, Aramini, JM, Liu, G, Eletsky, A, Wu, Y, Singarapu, KK, Lemak, A, Ignatchenko, A, Arrowsmith, CH, Szyperski, T, Montelione, GT, Baker, D, and Bax, A. 2008. Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. USA*, **105**, 4685–4690.

Shuker, SB, Hajduk, PJ, Meadows, RP, and Fesik, SW. 1996. Discovering high-affinity ligands for proteins: SAR by NMR. *Science*, **274**, 1531–1534.

Skinner, TE, Kobzar, K, Luy, B, Bendall, R, Bermel, W, Khaneja, N, and Glaser, SJ. 2006. Optimal control design of constant amplitude phase-modulated pulses: application to calibration-free broadband excitation. *J. Magn. Reson.*, **179**, 241–249.

Sørensen, OW. 1989. Polarization transfer experiments in high-resolution NMR spectroscopy. *Progr. NMR Spectrosc.*, **21**, 503–570.

Sørensen, OW. 1990. Aspects and prospects of multidimensional time-domain spectroscopy. *J. Magn. Reson.*, **89**, 210–216.

Sørensen, OW, Eich, GW, Levitt, MH, Bodenhausen, G, and Ernst, RR. 1983. Product operator formalism for the description of NMR pulse experiments. *Progr. NMR Spectrosc.*, **16**, 169–192.

Sprang, T, and Bigler, P. 2004. HMBC-RELAY: a combined technique for the differentiation of simultaneously detected ^{2}J(C,H) and ^{n}J(C,H) connectivities. *Magn. Reson. Chem.*, **42**, 55–60.

Sprangers, R, and Kay, LE. 2007a. Probing supramolecular structure from measurement of methyl ^{1}H-^{13}C residual dipolar couplings. *J. Am. Chem. Soc.*, **129**, 12668–12669.

Sprangers, R, and Kay, LE. 2007b. Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature*, **445**, 618–622.

Stuart, AC, Borzilleri, KA, Withka, JM, and Palmer, AG. 1999. Compensating for variations in ^{1}H-^{13}C scalar coupling constants in isotope-filtered NMR experiments. *J. Am. Chem. Soc.*, **121**, 5346–5347.

Takahashi, H, Nakanishi, T, Kami, K, Arata, Y, and Shimada, I. 2000. A novel NMR method for determining the interfaces of large protein-protein complexes. *Nat Struct Biol*, **7**, 220–223.

Takahashi, H, Miyazawa, M, Ina, Y, Fukunishi, Y, Mizukoshi, Y, Nakamura, H, and Shimada, I. 2006. Utilization of methyl proton resonances in cross-saturation measurement for determining the interfaces of large proteinprotein complexes. *J. Biomol. NMR*, **34**, 167–177.

Takeuchi, K, and Wagner, G. 2006. NMR studies of protein interactions. *Curr. Opin. Struct. Biol.*, **16**, 109–117.

Thrippleton, MJ, Edden, RAE, and Keeler, J. 2005. Suppression of strong coupling artefacts in J-spectra. *J. Magn. Reson.*, **174**, 97–109.

Tjandra, N, and Bax, A. 1997. Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science*, **278**, 1111–1112.

Tolman, JR, and Ruan, K. 2006. NMR residual dipolar couplings as probes of biomolecular dynamics. *Chem. Rev.*, **106**, 1720–1736.

Tugarinov, V, and Kay, LE. 2003. Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J. Am. Chem. Soc.*, **125**, 13868–13878.

Tugarinov, V, and Kay, LE. 2005. Methyl groups as probes of structure and dynamics in NMR studies of high-molecular-weight proteins. *ChemBioChem*, **6**, 1567–1577.

Tugarinov, V, Hwang, PM, Ollerenshaw, JE, and Kay, LE. 2003. Cross-correlated relaxation enhanced ^{1}H-^{13}C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.*, **125**, 10420–10428.

Uhrín, D, Uhrínová, S, Leadbeater, C, Nairn, J, Price, NC, and Barlow, PN. 2000a. 3D $HCCH_3$-TOCSY for resonance assignment of methyl-containing side chains in ^{13}C-labeled proteins. *J. Magn. Reson.*, **142**, 288–293.

Uhrín, D, Bramham, J, Winder, SJ, and Barlow, PN. 2000b. Simultaneous CT-^{13}C and VT-^{15}N chemical shift labelling: application to 3D NOESY-CH_3NH and 3D ^{13}C,^{15}N HSQC-NOESY-CH_3NH. *J. Biomol. NMR*, **18**, 253–259.

Vuister, GW, and Bax, A. 1992. Measurement of two-bond $J_{\mathrm{COH}\alpha}$ coupling constants in proteins uniformly enriched with ^{13}C. *J. Biomol. NMR*, **2**, 401–405.

Vuister, GW, and Bax, A. 1993. Quantitative J correlation: A new approach for measuring homonuclear three-bond $J(\mathrm{H^N H^\alpha})$ coupling constants in ^{15}N-enriched proteins. *J. Am. Chem. Soc.*, **115**, 7772–7777.

Williamson, MP, Havel, TF, and Wüthrich, K. 1985. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ^{1}H NMR spectroscopy. *J. Mol. Biol*, **182**, 295–315.

Wishart, DS, and Sykes, BD. 1994. The ^{13}C Chemical-Shift Index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR*, **4**, 171–180.

Xu, Y, Zheng, Y, Fan, JS, and Yang, D. 2006. A new strategy for structure determination of large proteins in solution without deuteration. *Nat Methods*, **3**, 931–937.

Yang, D, and Nagayama, K. 1996. A sensitivity-enhanced method for measuring heteronuclear long-range coupling constants from the displacement of signals in two 1D subspectra. *J. Magn. Reson. A*, **118**, 117–121.

Yang, D, Zheng, Y, Liu, D, and Wyss, DF. 2004. Sequence-specific assignments of methyl groups in high-molecular weight proteins. *J. Am. Chem. Soc.*, **126**, 3710–3711.

Zwahlen, C, Vincent, SJF, Gardner, KH, and Kay, LE. 1998. Significantly improved resolution for NOE correlations from valine and isoleucine ($C^{\gamma 2}$) methyl groups in ^{15}N,^{13}C- and ^{15}N,^{13}C,^{2}H-labeled proteins. *J. Am. Chem. Soc.*, **120**, 4825–4831.