Measurements of Side-Chain 13C-13C Residual Dipolar Couplings in Uniformly Deuterated Proteins

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · APRIL 2004

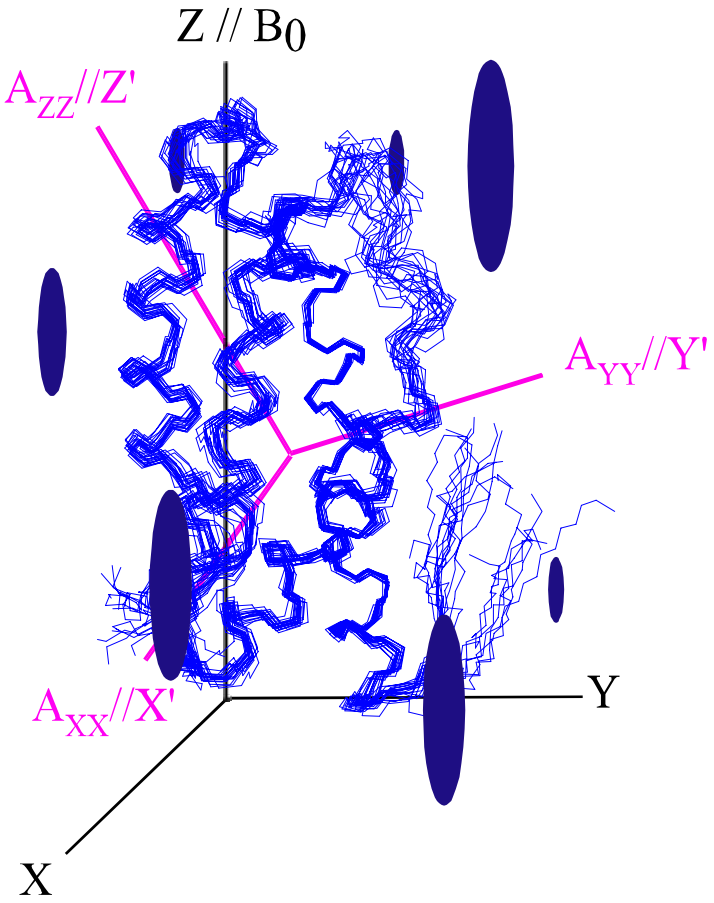
Impact Factor: 12.11 · DOI: 10.1021/ja0381813 · Source: PubMed

3 AUTHORS, INCLUDING:

Beat Vögeli
ETH Zurich
46 PUBLICATIONS **811** CITATIONS
SEE PROFILE

Konstantin Pervushin
Nanyang Technological University
84 PUBLICATIONS **4,512** CITATIONS
SEE PROFILE

Towards Structure and Dynamics of Large and Dynamically Disordered Biomacromolecules: New Methods in Solution NMR Spectroscopy



Beat Vögeli

Diss. ETH No. 15993

Diss. ETH No. 15993

Towards Structure and Dynamics of Large and Dynamically Disordered Biomacromolecules: New Methods in Solution NMR Spectroscopy

A dissertation submitted to the
Swiss Federal Institute of Technology Zürich
for the degree of
Doctor of Natural Sciences

presented by
Beat Vögeli
Dipl. Phys. ETH

born on May 13, 1976
citizen of Bülach (Zürich)

accepted on the recommendation of
Prof. Dr. Konstantin Pervushin, examiner
Prof. Dr. Beat Meier, co-examiner

2005

Acknowledgement

I would like to thank all the people involved in my projects. My collaborators from such a variety of different fields gave me the wonderful possibility to combine different topics in the framework of my thesis. These are: Dr. Alexander Eletsky (LTrosy project); Dr. Helena Kovacs (Bruker BioSpin AG), Dr. Kaifeng Hu and Dr. Donghan Lee (^{13}C spectroscopy); Tim Heinz and Prof. Philippe Hünenberger (Molecular dynamics); Katherina Vamvaca, Polina Anikeeva, Tuomas Knowles, Dr. Osvaldo Moreira, Dr. Peter Kast and Prof. Donald Hilvert (mMjCM project). I want to appreciate especially the intense and fruitful collaboration with Helena, Kaifeng and Katherina.

During my work, I obtained help from the following people: Dr. Elisabeth Enggist and Prof. Linda Thöny-Meier (CcmE samples); Sebastian Meier and Dr. Florence Cordier (Biozentrum Basel, RDC calculations); R. Pulido (Synthesizing TSA); Alexander Kienhöfer (BsCM preparation); Dr. Wolfgang Bermel, Dr. Detlef Moskau and Dr. Gerhard Eber (Bruker BioSpin AG, decoupling); Dr. Fred Damberger, Dr. James Masse and Ken Woycechowsky (Careful reading of parts of the thesis).

I am grateful to the following companies for cooperation: Bruker BioSpin AG (Spectrometer facilities and advice) and Silantes AG (RNA sample and bacterial growth medium for BsCM preparation).

I thank the reviewers of the papers and the thesis for valuable comments.

I thank the ETH Zürich for financial support.

Finally, I would like to thank my thesis supervisor Prof. Konstantin Pervushin. He made all these projects possible by his truly brilliant inspiration and strong belief in the success of new ways.

My Ph.D. time was also a great social experience. I thank all the members of our research group with whom I shared so much pleasure. Besides the ones mentioned above, I thank Veniamin Gallius, Simon Alioth, Prof. Maria Johansson and Reto Walser. Our talks covered topics concerning science as well as life as a whole. Some sayings will probably be cited for a long time by our groups members: „Life is not easy!“, „Saletti“ or „Sie ist perfekt!“ I experienced science as a melting pot of different cultures ranging from South American to Russian and Asian. This irreplacable experience has broadened my views of life!

My family and friends gave me counterweight to science and therefore, I thank them very much!

" Allein, wie grosse Fortschritte auch die *Physik* (im weiten Sinne der Alten verstanden) je machen möge; so wird damit noch nicht der kleinste Schritt zur *Metaphysik* geschehn sein (...) Vielmehr werden die grössten Fortschritte der *Physik* das Bedürfnis einer *Metaphysik* nur immer fühlbarer machen (...) Dies alles wird freilich der einzelne, simple Naturforscher in einem abgesonderten Zweige der Physik nicht sofort deutlich inne: vielmehr schläft er behaglich bei seiner erwählten Magd im Hause des Odysseus, sich aller Gedanken an die Penelopeia entschlagend. "

Arthur Schopenhauer, Die Welt als Wille und Vorstellung, Zweiter Band

Table of Contents

Zusammenfassung

Die in dieser Dissertation zusammengefasste Arbeit präsentiert Beiträge zur Methodik der Kernspinresonanz (NMR), die auf Fortschritten in den Gebieten der Relaxationsoptimierung, dem Gebrauch von ^{13}C-Detektion und der Entwicklung von theoretisch optimalen Experimenten beruht, als auch NMR mit einem ungeordneten Protein.

LTROSY ist eine allgemeine Methode, welche das Signal-zu-Rauschen-Verhältnis in mehrdimensionalen NMR-Experimenten mittels signifikanter Erniedrigung der longitudinalen Protonenrelaxationszeit um einen Faktor von 2 – 2.5 erhöht. Die Methode basiert auf dem Gebrauch eines riesigen Protonenspinreservoirs. Das TROSY-Prinzip wird in einem HNCA-Experiment verwendet, welches für die Messung von den intraresiduellen und sequentiellen H^{α}-C^{α}/H^{N}-N Dipol/Dipol- und H^{α}-C^{α}/N Dipol/CSA-kreuzkorrelierten Relaxationsraten sowie den mittels E.COSY-Prinzip erhaltenen $^{3,4}J_{H\alpha HN}$-Kopplungskonstanten entworfen wurde. Daraus erhält man Einschränkungen für die dihedralen Winkel ψ und ϕ.

Es werden Fortschritte in der ^{13}C-Spektroskopie demonstriert. Die Einführung einer dritten Dimension (^{1}H) in ^{13}C,^{13}C-TOCSY reduziert die Signalüberlappung und erhöht die Sensitivität pro Zeiteinheit, sogar für stark deuterierte (>85%) Proteinproben, was aus dieser Methode ein attraktives Werkzeug macht für die Zuordnung von Seitenketten-H und -C von mittelgrossen Proteinen mit sowohl natürlicher Isotopenhäufigkeit als auch starker Deuteriumanreicherung. Die Experimente werden mit 16 kDa-^{15}N,^{13}C-markiertem nichtdeuteriertem apo-CcmE und 48 kDa-einheitlich ^{15}N,^{13}C-markiertem und teilweise (~90%) deuteriertem dimerischem sFkpA ausgeführt. Es wird vorhergesagt, dass diese Methode geeignet sein sollte für die Zuordnung von den ^{13}C und ^{1}H-chemischen Verschiebungen in Methylgruppen von methylprotonierten und sonst stark deuterierten ^{13}C-markierten Proteinen mit sogar noch höherem molekularem Gewicht. Signifikante Verbesserung in der Auflösung und Vereinfachung der Aufspaltungsmuster in ^{13}C,^{13}C-TOCSY-Spektra von einheitlich deuterierten und ^{13}C,^{15}N-markierten Proteinen und ^{13}C,^{15}N-markierten RNS-Proben wird erreicht durch die Einführung von mehrfach-bandselektiver ^{13}C-Homoentkopplung, angewandt mit gleichzeitiger ^{1}H- oder ^{2}H- und ^{15}N-Entkopplung zu allen Zeitabschnitten in mehrdimensionalen Experimenten einschliesslich der Signalaufnahmeperiode. Des Weiteren wurden mehrfache residuelle ^{13}C-^{13}C-dipolare Kopplungen (RDCs) in einheitlich deuterierten und ^{13}C-markierten Proteinen gemessen. Es wird etabliert, dass unter verschiedenen Breitband-Polarisationstransferschemata die FLOPSY-Familie gebraucht werden kann, um Magnetisierung innerhalb eines *J*-gekoppelten Spinnetzwerkes auszutauschen unter nahezu vollständiger Entkopplung der dipolaren Wechselwirkungen zwischen diesen Spins. Eine exzellente Korrelation zwischen den gemessenen RDCs und der 3D-Struktur von Ubiquitin wurde beobachtet, was eine

mögliche Verwendung von ^{13}C-^{13}C-RDCs in der Strukturbestimmung von deuterierten Proteinen anzeigt.

Die Konstruktion von Pulssequenzen für die Auswahl von individuellen Übergängen in Methylgruppen wird umrissen. Unmittelbare Anwendung beinhaltet quantitative Messung von ^{1}H-^{1}H- und ^{1}H-^{13}C-RDCs und kreuzkorrelierter Relaxation zwischen ^{1}H- und ^{13}C-chemischer Verschiebungsanisotropie und ^{1}H-^{13}C- und ^{1}H-^{1}H Dipol/Dipol-Wechselwirkungen. Die Optimierung von Kohärenztransfer-Pulssequenzelementen erreicht Sensitivität nahe am theoretischen Maximum (unter Vernachlässigung der Relaxation) innerhalb der kürzest möglichen experimentellen Zeit und unterdrückt aktiv ungewünschte Signale. Basierend auf der Anwendung von Molekularer Dynamik im Raum der Pulssequenzvariablen, hat die Methode allgemeinen Charakter.

NMR-Spektroskopie ist eine von wenigen Informationsquellen über ungeordnete Proteine und ihren Faltungsprozesson. mMjCM, eine hoch aktive, monomerische Chorismatmutase, erhalten durch topologisches Redesign eines dimerischen Helixbündel-Proteinenzyms von *Methanococcus jannaschii*, wird untersucht mit TROSY-NMR. Obwohl strukturelle Unordnung im Allgemeinen für inkompatibel mit effizienter Katalyse gehalten wird, besitzt das Monomer, im Gegensatz zu seinem natürlichen Gegenstück, alle Charakteristika eines „Molten Globule"-Proteins. Globales konformationelles Ordnen, beobachtet nach Bindung eines Übergangszustandanalogons (TSA), zeigt, dass Faltung mit Katalyse gekoppelt werden kann unter minimalem Energieaufwand. Eine strukturelle und dynamische Analyse von monomerischem MjCM wird präsentiert. Die Prädominanz von α-helikaler Struktur in monomerischem MjCM bestätigt das durch Homologiemodelierung erhaltene Motif bestehend aus vier α-Helices. Durch den Einbau eines Turns aus sechs Residua in die Helix H1 des Dimers klappt deren obere Hälfte, die zum N-Terminus des Monomers wird, nach unten und bildet die aktive Stelle zusammen mit der zweiten Hälfte. TSA-gebundenes mMjCM unterliegt grossem globalem Austausch. Des Weiteren zeigt die Relaxationsanalyse von ungebundenem mMjCM, dass die räumliche Bindungseinschränkung nieder ist. Obwohl die Bindung von TSA strukturelle Ordnung einführt und die Beweglichkeit der Bindungen einschränkt bis auf das Ausmass eines regulären Proteins, erlaubt das TSA-Bindungsgleichgewicht globalen Austausch auf einer Zeitskala, die so schnell ist, dass sie nicht messbar ist mit NMR-Spektroskopie.

Summary

The work summarized in this thesis represents contributions to Nuclear Magnetic Resonance (NMR) methodology based on the recent progress in the directions of relaxation optimization, use of ^{13}C detection, development of theoretically optimal NMR experiments, as well as NMR with a disordered protein.

LTROSY is a general method yielding a factor 2 - 2.5 increase of the maximal signal-to-noise ratio of multidimensional NMR experiments via significant reduction of the longitudinal proton relaxation times. The method is based on the use of vast pools of "thermal bath" ^{1}H spins. The TROSY principle is introduced into an HNCA experiment, which is designed for measurements of the intraresidual and sequential H^{α}-C^{α}/H^{N}-N dipole/dipole and H^{α}-C^{α}/N dipole/CSA cross-correlated relaxation rates as well as $^{3,4}J_{H\alpha HN}$-coupling constants obtained in an E.COSY manner. Therefrom, conformational restraints are obtained for the ψ and ϕ angles.

Progress in ^{13}C spectroscopy is demonstrated. Introduction of a third dimension (^{1}H) to 2D ^{13}C,^{13}C-TOCSY reduces the peak overlap and increases the sensitivity per unit time, even for highly deuterated (>85%) protein samples, which makes this new method an attractive tool for the side-chain H and C assignment of average sized proteins with natural isotope abundance as well as large partially deuterated proteins. The experiments are demonstrated with a 16 kDa ^{15}N,^{13}C-labeled non-deuterated apo-CcmE and a 48 kDa uniformly ^{15}N,^{13}C-labeled and fractionally (~90%) deuterated dimeric sFkpA. It is predicted that this method should be suitable for the assignment of methyl ^{13}C and ^{1}H chemical shifts of methyl protonated, highly deuterated and ^{13}C-labeled proteins with even higher molecular weight. Significant resolution improvement and simplification of the splitting pattern in ^{13}C,^{13}C-TOCSY spectra of uniformly deuterated and ^{13}C,^{15}N-labeled protein and ^{13}C,^{15}N-labeled RNA samples is achieved by introduction of multiple-band-selective ^{13}C homodecoupling applied simultaneously with ^{1}H or ^{2}H, and ^{15}N decoupling at all stages of multidimensional experiment including signal acquisition period. Furtheron, multiple residual ^{13}C-^{13}C dipolar couplings (RDCs) in uniformly deuterated and ^{13}C-labeled proteins were measured. It is established that, among different broadband polarization transfer schemes, the FLOPSY family can be used to exchange magnetization within a *J*-coupled network of spins while largerly decoupling dipolar interactions between these spins. An excellent correlation between measured RDCs and the 3D structure of Ubiquitin was observed indicating a potential use of the ^{13}C-^{13}C RDCs in structure determination of perdeuterated proteins.

Construction of pulse sequences for selection of individual transitions in methyl groups is outlined. Immediate applications include quantitative measurements of ^{1}H-^{1}H and ^{1}H-^{13}C RDCs and cross-correlated relaxation between ^{1}H and ^{13}C chemical shift anisotropy and ^{1}H-^{13}C and ^{1}H-^{1}H

dipole/dipole interactions. Optimization of coherence-transfer pulse-sequence elements achieves sensitivity close to its theoretical maximum (in the absence of relaxation) in the shortest possible experimental time and features active suppression of undesired signals. Based on the application of molecular dynamics in the space of pulse-sequence variables, the method constitutes general character.

NMR spectroscopy is one of the few sources for information on disordered proteins and their folding processes. mMjCM, a highly active, monomeric chorismate mutase, obtained by topological redesign of a dimeric helical bundle enzyme from *Methanococcus jannaschii*, is investigated by TROSY NMR. Although structural disorder is generally considered to be incompatible with efficient catalysis, the monomer, unlike its natural counterpart, possesses all of the characteristics of a molten globule. Global conformational ordering, observed upon binding of a transition state analog, indicates that folding can be coupled to catalysis with minimal energetic penalty. Structural and dynamical analysis of the monomeric TSA-bound MjCM are presented. The predominance of α-helical structure in monomeric MjCM confirms the four α-helix-bundle motif obtained by homology modeling. By engineering of the six-residue turn into the H1 helix of the dimer, the upper half becoming the N-terminus of the monomer folds down and forms the active site together with the second half. TSA-bound mMjCM undergoes large global exchange. Furthermore, relaxation analysis of unbound mMjCM shows that the spatial bond restriction is low. Although ligation of TSA introduces structural order and restricts spatial bond mobility to a level of regular protein, the TSA binding equilibrium allows for global exchange on a timescale faster than is detectable by NMR spectroscopy.

Abbreviations

1D, 2D, 3D	1-dimensional, 2-dimensional, 3-dimensional
BsCM	*B. Subtilis* chorismate mutase
CcmE	Cytochrome *c* maturation heme chaperone protein E
CEE	Coherence excitation element
COCAINE	CO-CA in- and antiphase spectra with sensitivity enhancement
COSY	Correlated spectroscopy
CRINEPT	Cross relaxation-enhanced polarization transfer
CSA	Chemical shift anisotropy
CSE	Chemical shift encoding element
CSR	Curie spin relaxation
CTE	Coherence transfer element
DD	Dipole/dipole
DHPC	Dihexanoylphosphatidylcholine
DMPC	Dimyristoyl phosphatidylcholine
DQ	Double quantum
DSS	Sodium 2,2-dimethyl-2-silapentane-5-sulfonate
E. Coli	*Escherichia Coli*
E.COSY	Exclusive COSY
EEHT	Exact effective Hamiltonian theory
FID	Free induction decay
FkpA	Heat shock periplasmic peptidyl-prolyl cis/trans isomerase (PPIase) with chaperone activity
FLOPSY	Flip-flop transition spectroscopy
GFT	G-matrix Fourier transform
HACACO	H^{α}-C^{α}-CO correlation spectroscopy
HDDUTY	Homodecoupling duty cycle
HMQC	Heteronuclear multiple quantum coherence
HNCA	H^{N}-N-C^{α} correlation spectroscopy
HNCA	H^{N}-N-C^{α} correlation spectroscopy
HSQC	Heteronuclear single quantum coherence
INEPT	Insensitive nuclei enhanced by polarization transfer
IP	In-phase
ISIS	Increased steady-state magnetization by intermediate storage

L-	Longitudinal relaxation-optimized
MjCM	AroQ chorismate mutase from *Methanococcus jannaschii*
mMjCM	Monomeric MjCM
MRI	Magnetic resonance imaging
MW	Molecular weight
NMR	Nuclear magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	NOE spectroscopy
OCTE	Optimized coherence transfer element
PDB	Protein data bank
PFG	Pulse field gradient
ppm	Parts per million
PTE	Polarization transfer element
RDC	Residual dipolar coupling
rf	Radio frequency
SE	Sensitivity enhanced
SEA	Solvent exposed amides
S/N	Signal-to-noise ratio
SQ	Single quantum
ST2-PT	Single transition-to-single transition polarization transfer
TOCSY	Total correlation spectroscopy
TPPI	Time-proportional phase-incrementation
TROSY	Transverse relaxation-optimized spectroscopy
u-	uniformly
ZQ	Zero quantum

Chapter I

Introduction

1. NMR Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is used by a variety of disciplines, including physics, chemistry, biochemistry, biology and medical science [1-3]. Together with X-ray crystallography, NMR spectroscopy is the only technique for structure determination of macromolecules at high resolution. Moreover, in contrast to X-ray crystallography, NMR methods provide observation of time-dependent phenomena over a large time-scale. This thesis contributes to the methodological areas of biomolecular solution NMR where experimental sensitivity is critical and the quality of the data is limiting, such as high molecular weight structures, membrane proteins [4], high-throughput systems like drug screening [5], in vivo NMR [6], or genomics and proteomics [7].

NMR signals from solid and condensed phase were first detected and theoretically treated by Purcell et al. and Bloch et al. in 1946 [8-11], rapidly followed by detection of NMR in solution. NMR was immediately recognized as a very powerful analytical tool in chemistry and biology. Since then the complexity of molecules amenable to NMR spectroscopy has gradually been pushed up by increasing magnetic fields, more sensitive signal detection schemes, the use of stable isotopes, as well as advancements in NMR methodology. After introduction of pulse techniques and Fourier transform NMR by Ernst and Anderson [12] and of multidimensional NMR by Jeener [13,14], structure and dynamics of molecules of a size up to 10 kDa were routinely analyzed by homonuclear ^{1}H spectroscopy in solution. ^{13}C- and ^{15}N-isotope labeling allowed for heteronuclear NMR, pushing the molecular size limit up to 30 kDa. ^{2}H enrichment reduces relaxation to an extent that spectra of 50 kDa molecules became possible. More sophisticated pulse sequences making use of transverse relaxation-optimized spectroscopy (TROSY) [15] and cross relaxation-enhanced polarization transfer (CRINEPT) [16] boosted the limit to 100 kDa in 1999 [17,18] and 800 kDa in 2002 [19,20], respectively.

In parallel, the interpretation of NMR data in terms of structural constraints and dynamics has undergone significant progress: Upper distance constraints from the nuclear Overhauser effect (NOE) [1], dihedral angle constraints from cross-correlated relaxation [21] or *J*-coupling [1] and orientational restraints from residual dipolar couplings (RDCs) [22] are used for structure calculations. Measurements of time scales and dynamics has become possible mostly by studying relaxation rates [23].

The work summarized in this thesis represents further contribution to NMR methodology based on the recent progress in the directions of relaxation optimization, use of ^{13}C detection, as well as development of theoretically optimal NMR experiments. In the following

paragraphs we briefly describe several areas of NMR spectroscopy of biomolecules where our contributions were made. Some of the developed methods are finally applied to solve an intriguing problem of catalytical activity of a partially folded engineered enzyme.

1.1. Transverse Relaxation-Optimized Spectroscopy (TROSY)

Broadening resonances beyond detection due to fast transverse relaxation of NMR-sensitive spins by interactions modulated by rotational molecular motions has a dominant impact on the size limitation for biomacromolecules studied by liquid NMR spectroscopy. Transverse relaxation can be significantly reduced by TROSY NMR.

The introduction of transverse relaxation-optimized spectroscopy shifted the application limit of liquid-state NMR from a protein size of ~ 50 kDa to > 100 kDa [24,25]. In a single-quantum TROSY experiment, the transverse relaxation rate R_2 is reduced by selection of the slowly relaxing doublet component of a spin I, scalar coupled to a spin S being in the α or β state. At a polarizing magnetic field B_0 of ~ 1 GHz, the cross-correlated relaxation between the H^N-N dipolar interaction (DD) and the H^N or N chemical shift anisotropy (CSA) vanishes almost completely in $^1H^N$-^{15}N bonds [15]. [$^1H^N$,^{15}N]-TROSY can be implemented in most triple-resonance experiments which are extensions of a 2D [$^1H^N$,^{15}N]-correlation basic element [17,26-29]. Complementarily, R_2 of ^{13}C can be significantly suppressed by deuteration [30]. The most prominent examples of experiments which combine TROSY and deuteration are the TROSY-HNCA, TROSY-HNCACB, TROSY-HNCO and TROSY-HN(CO)CA experiments. These experiments are used for backbone-resonance assignment of ^2H,^{13}C,^{15}N-labeled proteins of up to 110 kDa [18]. The same two techniques have been used to reduce R_2 of carbon in aromatic ^1H-^{13}C bonds [31]. However, R_2 of ^1H is hardly decreased due to little CSA. In the zero-quantum [$^1H^N$,^{15}N]-TROSY which is used in a ^{15}N-resolved [^1H,^1H]-NOESY, R_{ZQ} is largely suppressed by cross-correlated relaxation between ^1H CSA and ^{15}N CSA [32]. Besides resonance assignment, the TROSY technique has found a wide range of application in other aspects of NMR spectroscopy, such as for detection of J_{HN} and J_{NN} scalar couplings across Watson-Crick base pairs [33], for measurements of residual dipolar couplings (RDCs) [34] and for detection of solvent exposed amides by SEA-TROSY [35].

1.2. ^{13}C-Detection Spectroscopy in Liquid-State NMR

^{13}C-detection spectroscopy has several advantages over ^{1}H-detection spectroscopy in application to large and paramagnetic biomolecules.

2D [^{1}H,^{13}C]-correlation spectroscopy was introduced for analysis of small molecules such as methyl ethyl ketone, ethanol and carbohydrates as early as in 1977 [36-39]. Applications to proteins were made feasible by uniform ^{13}C enrichment [40,41]. The first 2D [^{13}C,^{13}C]-correlation spectra were used for sequence-specific assignment of backbone resonances [42,43]. Farmer and Venters proposed a ^{13}C-start, ^{1}H^{N}-detection experiment, C(CC)(CO)NH, with ^{13}C-TOCSY mixing for side-chain assignment in deuterated proteins [44]. However, not until recent progress in hardware technology, such as development of NMR probes operating at cryo temperatures and probes optimized for ^{13}C detection did ^{13}C-detection spectroscopy became a competitive alternative to ^{1}H-detection spectroscopy [45-59]. A 3D HACACO experiment has been developed for backbone-resonance assignment [47] and a 2D [^{13}C,^{13}C]-TOCSY experiment for nearly complete side-chain assignment of a 44 kDa protein [48]. Although the sensitivity is decreased by almost an order of magnitude,

$$\frac{S_{\mathrm{H}}/N}{S_{\mathrm{C}}/N} = (\gamma_{\mathrm{H}}/\gamma_{\mathrm{C}})^{3/2} \approx 8, \qquad (1)$$

^{13}C detection offers unique advantages: (i) Due to the fact that water signals are not detected, no water suppression is needed. Therefore, no spectral water artifacts have to be eliminated and ^{13}C$^{\alpha}$ signals can be detected instead of ^{1}H$^{\alpha}$ signals which are otherwise obscured by the strong water signal; (ii) Pulse sequences can be considerably shortened. Replacement of "out-and-back" by "out-and-stay" experiments eliminates INEPT elements; (iii) Fewer radio frequency pulses are used reducing losses caused by B_1 inhomogeneity; (iv) 13-Carbons, which are not *J*-coupled to protons (e.g. carbonyls or carbons in deuterated proteins), become directly detectable; (v) In favorable cases, substitution of fast-relaxing protons with slower-relaxing carbons as detection nuclei increases sensitivity by decreasing the acquisition time; (vi) ^{13}C-detection spectroscopy is superior at identifying connectivities of paramagnetically broadened signals, because all paramagnetic relaxation contributions are decreased by up to a factor of 16 [50-56]:

$$R^{\mathrm{paramH}}/R^{\mathrm{paramC}} \approx (\gamma_{\mathrm{H}}/\gamma_{\mathrm{C}})^{2} \approx 16; \qquad (2)$$

(vii) Aliphatic carbons may exhibit fewer resonances and be more favorably distributed than those of their attached protons; (viii) ^{13}C detection is less sensitive to the ionic strength in the NMR sample than ^{1}H detection. The sensitivity factor L for the probes operating at cryo temperatures is

$$L = \left(1 + A\omega^2 / c\right)^{-1/2}, \quad (3)$$

where A is a function of the probe geometry, temperatures and salt properties, ω is the operational angular frequency, and c is the ionic strength [58]. It has been shown that $c > 200$ mM of NaCl yields 50% loss in ^{1}H detection, but only 20% in ^{13}C detection [59].

1.3. Residual Dipolar Couplings (RDCs)

In addition to ^{1}H-^{1}H NOEs, residual dipolar couplings (RDCs) provide a rich source of structural information [60-62]. Refinement of NMR or X-ray structures with RDC data is currently included in many structure-determination protocols, whereas de novo structure determination, identification of homologous folds, and calculation of internal motion by inconsistency between measured and calculated RDCs, are yet to evolve to their maximal potential [61,62].

Placement of biomolecules in a strong polarizing magnetic field B_0 reintroduces anisotropic interactions, which are observable as quadrupolar splittings or residual dipolar couplings in high-resolution NMR spectra [63-65]. In practice, however, the magnetic susceptibility of macromolecules is mostly too small for extraction of reliable RDCs. Dissolution of a more strongly aligned medium (dilute liquid crystalline phase) in the solvent containing the molecule enhances the degree of alignment of the molecule through steric or electrostatic interaction [66]. The use of media with which the degree of alignment can be precisely adjusted has allowed application of RDCs to protein-structure determination [22,67]. To date, a large array of alignment media have been used, such as bicelles composed of a mixture of dihexanoyl phosphatidylcholine (DHPC) and dimyristoyl phosphatidylcholine (DMPC) [67], rod-shaped viruses [68], filamentous phage [69], crystals of the purple membrane protein bacteriorhodopsin [70] or an anisotropically compressed or stretched hydrogel [71]. Recently, charged gels have been used to align membrane proteins reconstituted in detergent micelles [72].

The angular dependence of the dipolar Hamiltonian in first-order perturbation theory is given by an orientationally averaged Legendre polynomial $P(\cos\psi)$ [73]:

$$H_{ij}^{DD} \sim \langle P(\cos\psi)\rangle = \langle \frac{3(\cos\psi)^2 - 1}{2} \rangle , \quad (4)$$

where ψ is the angle formed by the magnetic field B_0 and the internuclear vector connecting spins *i* and *j*. Therefore, in a rigid molecule the relevant part of the probability distribution can be expressed by the five second-rank spherical harmonics. Specific linear combinations are referred to as the elements of the Saupe matrix [74] or the alignment tensor $\tilde{A}$ of the molecule. For complete description of the tensor, at least five measured RDCs are needed. RDCs can then be expressed as a function of the two angles θ and ϕ of the spherical coordinates orienting the internuclear vector within the principal axis system of $\tilde{A}$ as depicted in Figure 1 [22]:

$$D_{ij} = S(\mu_0/4\pi)\gamma_i\gamma_j /r_{ij}^3(h/2\pi^2)A_a \{(3\cos^2\theta - 1)/2 + {}^3/_4\ \eta\ (\sin^2\theta\cos 2\phi)\}, \quad (5)$$

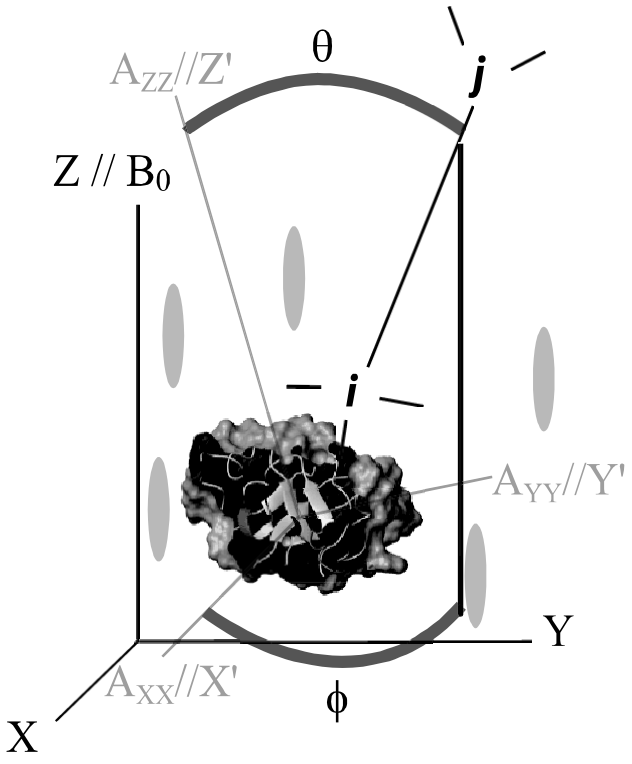


Figure 1. An alignment medium (grey ellipses) introduces slightly anisotropical tumbling of the molecule within the laboratory frame XYZ. The residual dipolar coupling (RDC) between spins *i* and *j* is described by the two angles θ and ϕ of the spherical coordinates within the main axis system X'Y'Z' of the molecule-fixed alignment tensor $\tilde{A}$.

where $A_a = (^3/_2)A_{zz}$ is the axial component and $\eta = {}^2/_3(A_{xx} - A_{yy})/A_{zz}$ the rhombicity of the alignment tensor, with A_{xx}, A_{yy} and A_{zz} being the Cartesian components in the principle axis system of $\tilde{A}$, S^2 is the generalized order parameter, μ_0 is the magnetic permeability of free space, γ_i is the gyromagnetic ratio of the spin *i*, *r* is the *i* - *j* internuclear distance, and h is Planck's constant. Determination of an internuclear vector is restricted to two cones. However, the global character of the constraints makes up for this ambiguity. In contrast, the most widely used structural constraint, upper distance limits stemming from the nuclear Overhauser effect (NOE), is only local [1]. Most RDCs are extracted from protein-backbone bonds, which are less flexible than side-chain bonds and consequently local motions do not rescale the RDCs strongly [22].

1.4. Unfolded and Partially Folded Proteins

Unfolded and partially folded proteins play an important role in cellular processes and signaling events. Many of these become locked into a specific structure upon binding to a partner. Conformationally disordered proteins are difficult to crystallize and may not be representive of the conformational ensemble. Therefore X-ray crystallography is not amenable and NMR spectroscopy is one of the few sources for information on these proteins and their folding processes [75]. Multidimensional experiments together with ^{13}C- and ^{15}N-isotope labeling have overcome the problem of strong ^{1}H-signal overlap [76]. NMR signals are often weak due to the fact that the population of transiently structured forms is too low. Proteins in a state containing nativelike secondary structure but loosened tertiary structure are referred to as "molten globule" proteins [77]. Its NMR analysis is extraordinary difficult because the conformational equilibrium is mostly on a time-scale comparable to the chemical shift time scale.

2. Objective of Thesis

This thesis describes contributions made to various branches of the NMR methodology. Chapter II and III present developments in TROSY NMR. Chapter X demonstrates an application of [$^1H^N$,^{15}N]-TROSY NMR to a low-concentration sample of a molten globule protein. The contributions to ^{13}C spectroscopy of large proteins are discussed in Chapters IV-

VI. Construction of pulse sequences for selection of individual transitions in methyl groups is outlined in Chapters VII and VIII. Chapters VI-VII mainly constitute new approaches to collecting RDC data.

In Chapter II, a general method to enhance sensitivity of multidimensional NMR experiments in a high-polarizing magnetic field by reducing the longitudinal proton-relaxation times is described. The method is based on the use of two vast pools of "thermal bath" ^{1}H spins residing on hydrogens covalently attached to carbon and oxygen atoms in ^{13}C,^{15}N-labeled and fully protonated or fractionally deuterated proteins to uniformly enhance longitudinal relaxation of the $^{1}H^{N}$ spins and concomitantly the sensitivity of multi-pulse NMR experiments. The proposed longitudinal relaxation optimization is implemented in the 2D [^{15}N,^{1}H]-LTROSY, 2D [^{15}N,^{1}H]-LHSQC and 3D LTROSY-HNCA experiments yielding the factor 2 - 2.5 increase of the maximal signal-to-noise ratio per unit time at 600 MHz. At 900 MHz the predicted decrease of the $^{1}H^{N}$ longitudinal relaxation times can be as large as an order of magnitude, making the proposed method an important tool for protein NMR at high magnetic fields.

In Chapter III, the TROSY principle is introduced into an HNCA experiment, which is designed for measurements of the intraresidual and sequential H^{α}-C^{α}/H^{N}-N dipole/dipole and H^{α}-C^{α}/N dipole/CSA cross-correlated relaxation rates. In addition, the new experiment provides values of the $^{3,4}J_{H\alpha HN}$ coupling constants measured in an E.COSY manner. The conformational constraints for the ψ and ϕ angles are obtained through the use of the cross-correlated relaxation rates together with the Karplus-type dependencies of the coupling constants. Improved signal-to-noise ratio is achieved through preservation of all coherence-transfer pathways and application of the TROSY principle. The application of the [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY experiment to the 16 kDa apo-form of the *E.Coli* Heme Chaperon protein CcmE is described. Overall good agreement is achieved between ψ and ϕ angles measured with the new experiment and the average values determined from an ensemble of 20 NMR conformers.

In Chapter IV it is shown that significant resolution improvement in [^{13}C,^{13}C]-TOCSY spectra of uniformly deuterated and ^{13}C,^{15}N-labeled protein and ^{13}C,^{15}N-labeled RNA samples is achieved by introduction of multiple-band-selective ^{13}C homodecoupling applied simultaneously with ^{1}H or ^{2}H, and ^{15}N decoupling at all stages of multidimensional experiments including the signal-acquisition period. The application of single, double or triple band-selective ^{13}C decoupling in 2D [^{13}C,^{13}C]-TOCSY experiments during acquisition strongly simplifies the homonuclear splitting pattern. The technical aspects of complex

multiple-band homonuclear decoupling and hardware requirements are discussed. The use of this technique (i) facilitates the resonance-assignment process as it reduces signal overlap in homonuclear ^{13}C spectra and (ii) possibly improves the signal-to-noise ratio through multiplet collapse. It can be applied in any ^{13}C-detected experiment.

Chapter V proposes the use of ^{13}C-detected 3D HCC-TOCSY experiments for assignment of ^{1}H and ^{13}C resonances in protonated and partially deuterated proteins. The experiments extend 2D ^{13}C-start and ^{13}C-observe TOCSY-type experiments proposed earlier. Introduction of the third ^{1}H dimension to 2D TOCSY (i) reduces the peak overlap and (ii) increases the sensitivity per unit time, even for highly deuterated (> 85%) protein samples, which makes this new method an attractive tool for the side-chain H and C assignment of average sized proteins with natural isotope abundance as well as large partially deuterated proteins. The experiments are demonstrated with a 16 kDa ^{15}N,^{13}C-labeled non-deuterated apo-CcmE and a 48 kDa uniformly ^{15}N,^{13}C-labeled and fractionally (~90%) deuterated dimeric sFkpA. It is predicted that this method should be suitable for the assignment of methyl ^{13}C and ^{1}H chemical shifts of methyl protonated, highly deuterated and ^{13}C-labeled proteins with even higher molecular weight.

^{13}C-only spectroscopy was used in Chapter VI to measure multiple residual ^{13}C-^{13}C dipolar couplings (RDCs) in uniformly deuterated and ^{13}C-labeled proteins. We demonstrate that ^{13}C-start and ^{13}C-observe spectra can be routinely used to measure an extensive set of the side-chain residual ^{13}C-^{13}C dipolar couplings upon partial alignment of human ubiquitin in the presence of bacteriophage Pf1. We establish that, among different broadband polarization-transfer schemes, the FLOPSY family can be used to exchange magnetization between a *J*-coupled network of spins while largerly decoupling dipolar interactions between these spins. An excellent correlation between measured RDCs and the 3D structure of the protein was observed indicating a potential use of the ^{13}C-^{13}C RDCs in structure determination of perdeuterated proteins.

Individual transitions of magnetically equivalent spin systems such as methyl groups residing on isotropically tumbling molecules in solution usually can not be observed as multiplet-split NMR lines. In Chapter VII, we propose a pair of NMR experiments, 2D $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$ and $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$ HSQC, to overcome this limitation and enable direct and selective observation of individual ^{1}H transitions in ^{13}C-labeled methyl spin systems. Immediate applications include quantitative measurements of ^{1}H-^{1}H residual dipolar couplings (RDC) and cross-correlated relaxation between ^{1}H chemical shift anisotropy and ^{1}H-^{1}H dipole-dipole interactions. The use of these experiments for the measurement of RDCs is

demonstrated with two proteins, one weakly aligned in a Pf1 phage containing medium and the other using a native paramagnetic heme group.

Chapter VIII goes further to describe the computer optimization of coherence-transfer pulse-sequence elements (CTEs). Through this most challenging step in the construction of heteronuclear correlation NMR experiments, we achieved a sensitivity close to the theoretical maximum (in the absence of relaxation) in the shortest possible experimental time and featuring active suppression of undesired signals. As reported in this thesis, this complex optimization problem in a space of high dimensionality turns out to be numerically tractable. Based on the application of molecular dynamics in the space of pulse-sequence variables, a general method is proposed for constructing optimized CTEs capable of transferring an arbitrary (generally non-Hermitian) spin operator encoding the chemical shift of heteronuclear spins to an arbitrary spin operator suitable for signal detection. The CTEs constructed in this way are evaluated against benchmarks provided by the theoretical unitary bound for coherence transfer and the minimal required transfer time (when available). This approach is used to design a set of NMR experiments enabling direct and selective observation of individual ^{1}H-transitions in ^{13}C-labeled methyl spin systems close to optimal sensitivity and using a minimal number of spectra. As an illustration of the method, optimized CTEs are used to quantitatively measure ^{1}H-^{1}H and ^{1}H-^{13}C residual dipolar couplings (RDCs) in a 17 kDa protein weakly aligned by means of Pf1 phage.

The methods presented in the Chapters II to VIII are discussed in Chapter IX.

In Chapters X and XI, a dynamically disordered protein is characterized by means of NMR.

In Chapter X, a highly active, monomeric chorismate mutase, obtained by topological redesign of a dimeric helical bundle enzyme from *Methanococcus jannaschii*, is investigated by 2D [^{15}N,^{1}H]-TROSY NMR. Although structural disorder is generally considered to be incompatible with efficient catalysis, the monomer, unlike its natural counterpart, unexpectedly possesses all of the characteristics of a molten globule. Global conformational ordering, observed upon binding of a transition state analog, indicates that folding can be coupled to catalysis with minimal energetic penalty. These results support the suggestion that many modern enzymes might have evolved from molten globule precursors. Insofar as their structural plasticity confers relaxed substrate specificity and/or catalytic promiscuity, molten globules may also be attractive starting points for the evolution of new catalysts in the laboratory.

Chapter XI presents a structural and dynamical analysis of the monomeric TSA-bound MjCM. The predominance of α-helical structure in monomeric MjCM confirms the four α-

helix-bundle motif obtained by homology modeling. By introduction of the six-residue turn into H1 of the dimer, the upper half becoming the N-terminus of the monomer folds down and forms the active site together with the second half. Relaxation analysis shows that TSA-bound mMjCM undergoes large global exchange. Furthermore, relaxation analysis of unbound mMjCM shows that the spatial bond restriction is low. Although ligation of TSA introduces structural order and restricts spatial bond mobility to a level of regular protein, the TSA binding equilibrium allows for global exchange on a timescale faster than is detectable by NMR spectroscopy.

3. References

[1] Wüthrich, K. *NMR of Proteins and Nucleic Acid.* **1986**, Wiley, New York.
[2] Ernst, R.R.; Bodenhausen, G.; Wokaun, A. *The Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. **1987**, Oxford, Clarendon.
[3] Cavanagh, J.; Fairbrother, W.J.; Palmer, A.G.; Skleton, N.J. . *Protein NMR Spectroscopy. Principles and Pratice*. **1996**, Academic Press, San Diego.
[4] Nielsen, N.C.; Malmendal, A.; Vosegaard, T. *Mol. Membr. Biol.* **2004**, *21*, 129-141.
[5] Johnson, E. C.; Feher, V. A.; Peng, J. W.; Moore, J. M.; Williamson, J. R.; *Biochemistry* **2003**, *125*, 15724-15725.
[6] Serber, Z.; Dötsch, V. *Biochemistry* **2001**, *40*, 14317-14323.
[7] Staunton, D.; Owen, J.; Campbell, I.D. *Acc. Chem. Res.* **2003**, *36*, 207-214.
[8] Purcell, E.M.; Torrey, H.C.; Pound, R.V. *Proc. Rev.* **1946**, *69*, 37-38.
[9] Bloch, F.; Hansen, W.W.; Packard, M. *Phys. Rev.* **1946**, *69*, 127.
[10] Bloch, F. *Phys. Rev.* **1946**, *70*, 460-474.
[11] Bloch, F.; Hansen, W.W.; Packard, M. *Phys. Rev.* **1946**, *70*, 474-485.
[12] Ernst, R.R.; Anderson, W.A. *Rev. Sci. Instrum. Sci.* **1966**, *37*, 93-102.
[13] Jeener, J. (Lecture , Ampère International Summer School II, Basko Polje, Yugoslavia, 1971).
[14] Aue, W.P.; Bartholdi, E.; Ernst, R.R. *J. Chem. Phys.* **1976**, *64*, 2229-2246.
[15] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366-12371.
[16] Riek, R.; Wider, G.; Pervushin, K.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 4918-4923.
[17] Salzmann, M.; Wider, G.; Pervushin, K.; Wüthrich, K. *J. Biomol. NMR* **1999**, *15*, 181-184.
[18] Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *J. Am. Chem. Soc.* **2000**, *122*, 7543-7548.
[19] Riek, R.; Fiaux, J.; Bertelsen, E.B.; Horwich, A.L.; Wüthrich, K. *J. Am. Chem. Soc.* **2002**, *124*, 12144-12153.
[20] Fiaux, J.; Bertelsen, E.B.; Horwich, A.L.; Wüthrich, K. *Nature* **2002**, *418*, 207-211.
[21] Goldman, M. *J. Magn. Reson.* **1984**, *60*, 437-452.
[22] Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111-1114.
[23] Korzhnev, D. M.; Billeter, M.; Arseniev, A. S.; Orekhov, V.Y. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *38*, 197-266.
[24] Pervushin, K. *Q. Rev. Biophys.* **2000**, *33*, 161-197.
[25] Riek, R.; Pervushin, K.; Wüthrich, K *Trends Biochem. Sci.* **2000**, *25*, 462-468.
[26] Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 13585-13590.
[27] Yang, D; Kay, L.E. *J. Biomol. NMR* **1999**, *13*, 3-10.
[28] Salzmann, M.; Wider, G.; Pervushin, K.; Senn, H.; Wüthrich, K. *J. Am. Chem. Soc.* **1998**, *121*, 844-848.
[29] Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *J. Biomol. NMR* **1999**, *14*, 85-88.
[30] Venters, R.A.; Farmer, B.T.; Fierke, C.A.; Spicer, L.D. *J. Mol. Biol.* **1996**, *264*, 1101-1116.
[31] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *J. Am. Chem. Soc.* **1998**, *120*, 6394-6400.
[32] Pervushin, K.; Wider, G.; Riek, R.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9607-9612.
[33] Pervushin, K.; Ono, A.; Fernandez, C.; Szyperski, T.; Kainosho, M.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 14147-14151.

[34] Yang, D; Venters, R.A.; Mueller, G.A.; Choy, W.Y.; Kay, L.E. *J. Biomol. NMR* **1999**, *14*, 333-343.
[35] Pellecchia, M.; Meininger, D.; Shen, A.L.; Jack, R.; Kasper, C.B.; Sem, D.S. *J. Am. Chem. Soc.* **2001**, *123*, 4633-4634.
[36] Maudsley, A.A.; Müller, L.; Ernst, R.R. *J. Magn. Reson.* **1977**, *28*, 463-469.
[37] Bodenhausen, G.; Freeman, R. *J. Magn. Reson.* **1977**, *28*, 471-476.
[38] Freeman, R.; Morris, G.A. *Bull. Magn. Reson.* **1979**, *1*, 5-26.
[39] Morris, G.A.; Hall, L.D. *J. Am. Chem. Soc.* **1981**, *103*, 4703-4711.
[40] Chan, T.M.; Markley, J.L. *J. Am. Chem. Soc.* **1982**, *104*, 4010-4011.
[41] Kojiro, L.; Markley, J.L. *FEBS Lett.* **1983**, *162*, 52-56.
[42] Oh, B.H.; Westler, W.M.; Darba, P.; Markley, J.L. *Science* **1988**, *240*, 908-911.
[43] Westler, W.M.; Kainosho, M.; Nagao, H.; Tomonaga, N.; Markley, J.L. *J. Am. Chem. Soc.* **1988**, *110*, 4093-4095.
[44] Farmer, B.T.; Venters, R.A. *J. Am. Chem. Soc.* **1995**, *117*, 4187-4188.
[45] Serber, Z.; Richter, C.; Moskau, D.; Böhlen, J.M.; Gerfin, T.; Marek, D.; Häberli, M.; Baselgia, L.; Laukien, F.; Stern, A.S.; Hoch, J.C.; Dötsch, V. *J. Am. Chem. Soc.* **2000**, *122*, 3554-3555.
[46] Serber, Z.; Richter, C.; Dötsch, V. *Chembiochem* **2001**, *2*, 247-251.
[47] Pervushin, K.; Eletsky, A. *J. Biomol. NMR* **2003**, *25*, 147-152.
[48] Eletsky, A.; Moreira, O.; Kovacs, H.; Pervushin, K. *J. Biomol. NMR* **2003**, *26*, 167-179.
[49] Bertini, I.; Felli, I.C.; Kümmerle, R.; Moskau, D.; Pierattelli, R. *J. Am. Chem. Soc.* **2003**, *126*, 464-465.
[50] Bermel, W.; Bertini, I; Felli, I.C.; Kümmerle, R.; Pierattelli, R. *J. Am. Chem. Soc.* **2003**, *125*, 16423-16429.
[51] Bertini, I; Lee, Y.M.; Luchinat, C.; Piccioli, M.; Poggi, L. *Chembiochem* **2001**, *2*, 550-558.
[52] Machonkin, T.E.; Westler, W.M.; Markley, J.L. *J. Am. Chem. Soc.* **2002**, *124*, 3204-3205.
[53] Kostic, K.; Pochapsky, S.S.; Pochapsky, T.C. *J. Am. Chem. Soc.* **2002**, *124*, 9054-9055.
[54] Arnesano, F.; Banci, L.; Bertini, I.; Felli, I.C.; Luchinat, C.; Thompsett, A.R. *J. Am. Chem. Soc.* **2003**, *125*, 7200-7208.
[55] Machonkin, T.E.; Westler, W.M.; Markley, J.L. *J. Am. Chem. Soc.* **2004**, *126*, 5413-5426.
[56] Babini, E.; Bertini, I.; Capozzi, F.; Felli, I.C.; Lelli, M.; Luchinat, C. *J. Am. Chem. Soc.* **2004**, *126*, 10496-10497.
[57] Bertini, I.; Felli, I.C.; Kümmerle, R.; Luchinat, C.; Pierattelli, R. *J. Biomol. NMR* **2004**, *30*, 245-251.
[58] Kelly, A.E.; Ou, H.D.; Withers, R.; Dötsch, V. *J. Am. Chem. Soc.* **2002**, *124*, 12013-12019.
[59] Shimba, N.; Kovacs, H.; Stern, A.S.; Nomura, A.M.; Shimada, I.; Hoch, J.C.; Craik,C.S.; Dötsch, V. *J. Biomol. NMR* **2004**, *30*, 175-179.
[60] Prestegard, J.H.; Al-Hashimi, H.M.; Tolman, J.R. *Q. Rev. Biophys.* **2000**, *33*, 371-424.
[61] De Alba, E.; Tjandra, N. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *40*, 175-197.
[62] Bax, A. *Protein Sci.* **2003**, *12*, 1-16.
[63] Lohman, J.A.B.; MacLean, C. *Chem. Phys. Lett.* **1978**, *58*, 483-486.
[64] Bothner-By, A.A.; Domaille, P.J.; Gayathri, C. *J. Am. Chem. Soc.* **1981**, *103*, 5602-5603.
[65] Tolman, J.R.; Flanagan, J.M.; Kennedy, M.A.; Prestegard, J.H. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 9279-9283.
[66] Saupe, A.Z.; Englert, G. *Phys. Rev. Lett.* **1963**, *11*, 462-464.
[67] Bax, A.; Tjandra, N. *J. Biomol. NMR* **1997**, *10*, 289-292.
[68] Clore, G.M.; Starich, M.R.; Gronenborn, A.M. *J. Am. Chem. Soc.* **1998**, *120*, 10571-10572.
[69] Hansen, M.R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065-1074.
[70] Sass, J.; Cordier, F.; Hoffmann, A.; Cousin, A.; Omichinski, J.G.; Lowen, H.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 2047-2055.
[71] Tycko, R.; Blanco, F.J.; Ishii, Y. *J. Am. Chem. Soc.* **2000**, *122*, 9340-9341.
[72] Cierpicki, T.; Bushweller, J.H. *J. Am. Chem. Soc.* **2004**, *126*, 16259-16266.
[73] Abragam, A. *The Principles of Nuclear Magnetism.* **1961**, *17*, Oxford, Clarendon.
[74] Saupe, A. Z. *Naturforsch.* **1964**, *19A*, 161-171.
[75] Dyson, H.J.; Wright, P.E.; *Chem. Rev.* **2004**, *104*, 3607-3622.
[76] Yao, J.; Dyson, H.J.; Wright, P.E.; *FEBS Lett.* **1997**, *419*, 285-289.
[77] Dolgikh, D.A.; Gilmanshin, R.I.; Brazhnikov, E.V.; Bychkova, V.E.; Semisotnov, G.V.; Venyaminov, S.Y.; Ptitsyn, O.B. *FEBS Lett.* **1981**, *136*, 311-315.

Chapter II

Longitudinal ^{1}H-Relaxation Optimization in TROSY NMR Spectroscopy

Konstantin Pervushin[1], Beat Vögeli[1] and Alexander Eletsky[2]*

[1]Laboratorium für Physikalische Chemie, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

[2]Departments of Chemistry and Structural Biology, State University of New York, Northeast Structural Genomics Consortium, Buffalo, NY 14260, USA.

* Spectroscopy together with A.E.

Journal of the American Chemical Society, 2002, 124, 12898-12902

1. Introduction

The application of NMR spectroscopy for structure determination of proteins and nucleic acids [1] with molecular mass exceeding 30 kDa is largely constrained by two factors, fast transverse relaxation of spins of interest and complexity of NMR spectra, both of which increase with increasing molecular size [2-5]. The TROSY technique offers a simple and robust bypass solution to both of these problems when NMR spectroscopy is conducted at sufficiently high magnetic fields [6]. What is less appreciated as a factor limiting the sensitivity is the long longitudinal relaxation of the ^{1}H spins at high magnetic fields and larger molecular weights [1]. Long longitudinal relaxation times, which for proteins with MWs above 30 kDa are typically found to be in the range of many seconds [7], result in a substantial loss of the signal-to-noise ratio of the multipulse NMR experiments per unit time of the data acquisition [8]. Thus, any enhancement of longitudinal relaxation increases the sensitivity of NMR experiments.

Thus far, methods used to enhance longitudinal relaxation include the use of the paramagnetic compounds such as MRI contrasting reagents [9,10], naturally occurring paramagnetic metal centers [11-13], fast exchange of the labile protons with water in the SEA-TROSY experiment [14,15]. The limitations of these methods stem from their abilities to enhance sensitivity only of the very limited subset of the ^{1}H spins. For the special HMQC class of heteronuclear correlation experiments, implementation of the Ernst angle [8] might yield signal enhancement in the order of 25% by accounting for the finite rate of recovery to the thermal equilibrium [16]. A sensitivity gain in the range of 10-20% in TROSY spectroscopy can be obtained by an alternative approach that utilizes the intermediate storage of a part of the steady-state magnetization of spin *I* on insensitive spin *S* with a pulse sequence element ISIS during the interscan delay [17]. Although the later represents a general spectroscopic technique, the achieved gain is rather modest.

Here we propose the use of two vast pools of ^{1}H spins residing on hydrogens covalently attached to carbon and oxygen atoms in ^{13}C,^{15}N-, or ^{15}N-labeled and fully protonated or fractionally deuterated proteins to significantly and uniformly enhance longitudinal relaxation of the $^{1}H^{N}$ spins and concomitantly the sensitivity of multipulse NMR experiments. The NMR experiments with the excitation/detection performed with the $^{1}H^{N}$ spins covalently attached to the ^{15}N spins are designed for the assignment of the backbone resonances and collect structural constraints with larger proteins at high magnetic fields and frequently utilize the TROSY principle [6]. The proposed longitudinal relaxation optimization is implemented in the 2D [^{15}N,^{1}H]-LTROSY, 2D [^{15}N,^{1}H]-LHSQC and 3D LTROSY-HNCA experiments yielding the factor 2 - 2.5 increase of the maximal signal-to-noise ratio per unit time measured with two proteins, 7 kDa ubiquitin (PDB code 1UBI) and 44 kDa trimeric *B. Subtilis* chorismate mutase (PDB code 1COM). Calculations predict that optimal interscan d_1^{opt} delays in the longitudinal relaxation-optimized NMR experiments can be an order of

magnitude shorter than those in the conventional experiments. Thus, the longitudinal relaxation optimization might become an important tool for NMR of proteins at 900 MHz.

2. Methods

Signal-to-Noise Ratio per Unit Time and Longitudinal ^{1}H Relaxation. The repetition rate of an NMR experiment is a function of the interscan delay d_1 performed between the last pulse of one scan to the first pulse of the next scan. The signal-to-noise ratio per measurement time, referred to as the sensitivity of the experiment, is given by Equation 1 [8].

$$S/N_t = c\frac{(1-\exp(-Rd_1))}{\sqrt{d_0+d_1}}, \quad (1)$$

where R is the longitudinal relaxation rate of the excited spins, d_0 is the duration of the pulse sequence excluding data acquisition, d_1 is the interscan delay, and c is a proportionality coefficient. For monoexponential relaxation, the maximal sensitivity $S/N_{tmax} = c0.71/\sqrt{d_1^{opt}} \approx c0.64\sqrt{R}$ is achieved at the optimal interscan delay $\sqrt{d_1^{opt}} = 1.25/R - d_0$. Thus, the efficiency of longitudinal relaxation of the excited spins defines the sensitivity of NMR experiments in the case, where many repetitions of the excitation/detection cycles are performed.

Longitudinal relaxation of the ^{1}H spins, I_z, to the thermal equilibrium, I_z^{eq}, can be modeled by the system of linear differential equations given by Equations 2, which are inhomogeneous with respect to exchangeable protons in OH, NH_3^+, COOH, guanidinium groups, and structural water molecules [18].

$$d\boldsymbol{I}_z/dt = \mathbf{R}\,(\boldsymbol{I}_z - \boldsymbol{I}_z^{eq}) + \mathbf{K}\,\boldsymbol{I}_z \quad (2)$$

where **R** is the relaxation matrix [8], **K** is the kinetic matrix collecting chemical exchange constants between the individual ^{1}H spins and water [1,19], $\boldsymbol{I}_z$ is a column vector consisting of the four groups of spins, I_z^N, I_z^C, I_z^O, and I_z^W, which represent the ^{1}H spins covalently attached to the ^{15}N, ^{13}C, and ^{16}O atoms, and residing on the water molecules, respectively. During the numerical integration of Equation 2, the boundary condition $I_z^W(t) - I_z^W(0) = 0$ is maintained assuming that (i) the longitudinal relaxation time of water protons is significantly longer than that of the protein ^{1}H spins and (ii) the protein solution is sufficiently diluted. The initial values of the I_z^N, I_z^C, I_z^O, and I_z^W spins are set according to the condition of the corresponding magnetization after signal acquisition in analyzed NMR experiments. For each individual amide proton, $\sqrt{d_1^{opt}}$ is reported as a time point, at which the global maximum of the function $I_z^N(t)/\sqrt{t}$ occurs for $t > 0$, thus avoiding the problem of the deconvolution of the multiple exponential time course of the function $I_z^N(t)$.

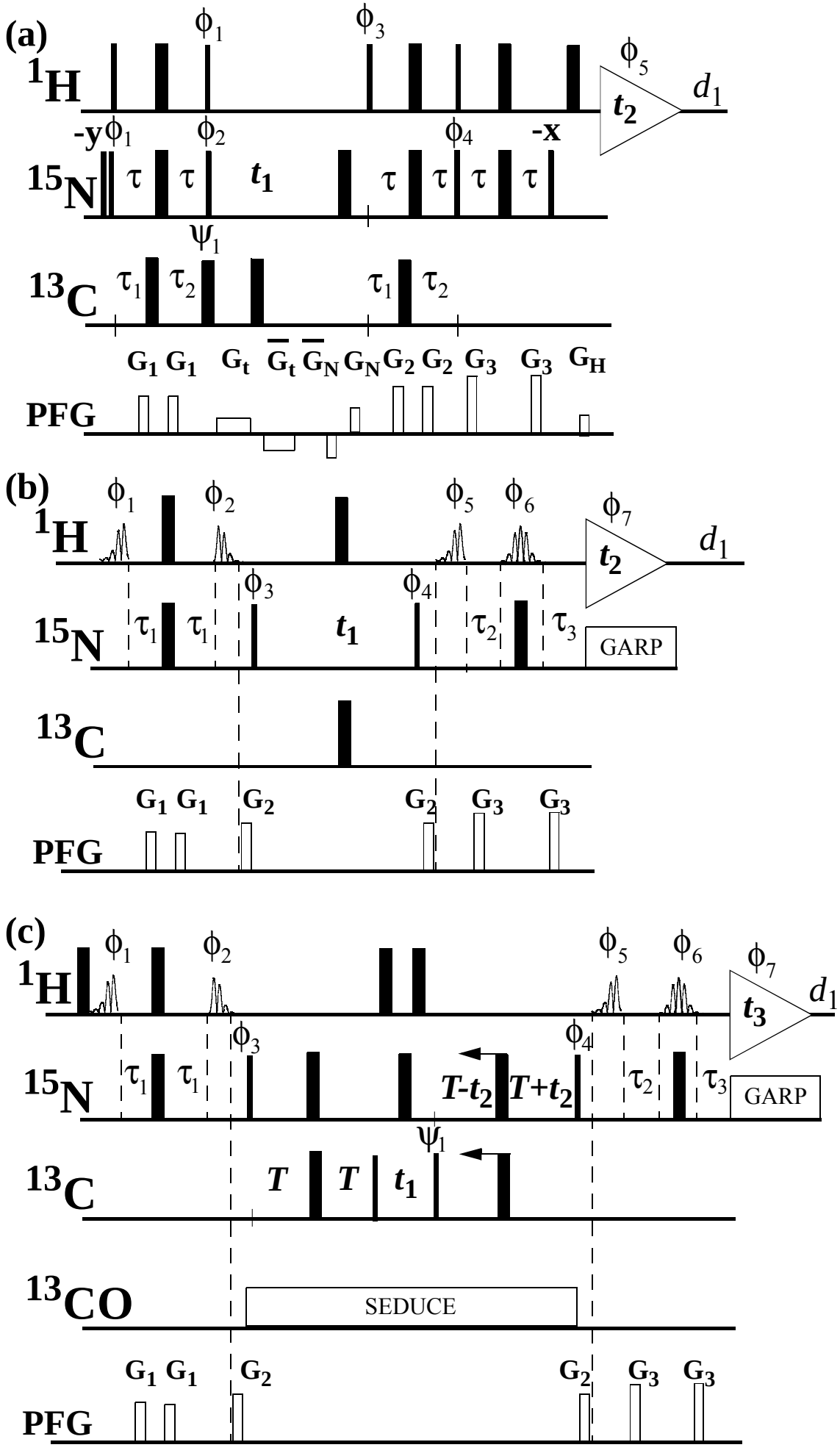

(a)
^{1}H
^{15}N
^{13}C
PFG
ϕ_1
ϕ_2
ϕ_3
ϕ_4
ϕ_5
ψ_1
-y
-x
τ
t_1
t_2
d_1
τ_1
τ_2
G_1 G_1 G_t $\overline{G_t}$ $\overline{G_N}$ G_N G_2 G_2 G_3 G_3 G_H
(b)
ϕ_6
ϕ_7
τ_3
GARP
G_2
(c)
t_3
T-t_2
T+t_2
T
SEDUCE
^{13}CO

Figure 1 (previous page). Experimental schemes for the longitudinal ^{1}H relaxation optimized (a) 2D [^{15}N,^{1}H]-LTROSY, (b) 2D [^{15}N,^{1}H]-LHSQC, and (c) 3D [^{15}N,^{1}H]-LTROSY-HNCA experiments. The radio-frequency pulses on ^{1}H, ^{15}N, ^{13}C, ^{13}CO, and $^1H^{\alpha}$ are applied at 4.7, 118, 55, 174, and 4.7 ppm, respectively. Narrow and wide black bars indicate non-selective 90° and 180° pulses, respectively. Complex shapes on the line marked ^{1}H indicate the $^1H^N$ band-selective 1.5 ms excitation E-Burp2 pulses [31] with the phases ϕ_1 and ϕ_5 and γB_1 = 2733 Hz, the identical to the pulses ϕ_1 and ϕ_5 but time-reversed excitation E-Burp2 pulse with the phase ϕ_2 and the 1.8 ms refocusing Re-Burp pulse [31] with the phase ϕ_6 and γB_1 = 3050 Hz. The center of the excitation of all $^1H^N$ band-selective pulses is placed at 9.7 ppm. The line marked PFG indicates the duration and strength of pulsed magnetic field gradients applied along the z-axis. The variable interscan ^{1}H relaxation delay d_1 is performed between repeated applications of the experimental pulse sequences. In (a) the delays are τ = 2.7 ms, τ_1 = 1.7 ms and τ_2 = 2.0 ms. The PFGs are G_1: 800 μs, 50 G/cm; G_2: 800 μs, 60 G/cm; G_3: 800 μs, 80 G/cm; G_t: t_1/2, 10 G/cm; $\overline{G}t$: t_1/ 2, -10 G/cm; G_N: 800 μs, 50 G/cm; $\overline{G}_N$: 800 μs, -50 G/cm; G_H: 800 μs, 10.13 G/cm. The phases are ϕ_1 = {y}; ϕ_2 = { y,-y, x,-x}; ϕ_3 = {y}; ϕ_4 = {-y}; ϕ_5 = { y,-y,-x, x}; ψ_1 = {x}; {x} for all other pulses. A phase-sensitive spectrum in the ^{15}N(t_1) dimension is obtained by recording a second FID for each t_1 value, with ϕ_1 = {-y}; ϕ_2 = {y,-y,-x, x}; ϕ_3 = {-y}; ϕ_4 = {y}; ϕ_5 = {-y, y, x,-x}, G_N: 800 μs, -50 G/cm; $\overline{G}_N$: 800 μs, 50 G/cm and data processing as described by [27]. The broad band inversion of the ^{13}C spins is achieved with the 90^o_x-225^o_y-90^o_x composite pulses [28]. To eliminate the effect of the longitudinal ^{1}H relaxation optimization the ^{13}C composite pulse ψ_1 is not performed. In (b), the delays are τ_1 = 2.4 ms, τ_2 = 2.34 ms and τ_3 = 2.4 ms. The PFGs are G_1: 800 μs, 50 G/cm; G_2: 800 μs, 70 G/cm; G_3: 800 μs, 80 G/cm. The phases are ϕ_1 = {x}; ϕ_2 = {y}; ϕ_3 = {x,-x}; ϕ_4 = {x, x,-x,-x}; ϕ_5 = {x}; ϕ_6 = {x}; ϕ_7 = {x,-x,-x, x}; {x} for all other pulses. Quadrature detection in the ^{1}H(t_2) dimension is achieved by the States-TPPI method [37] applied to the phase ϕ_3. To eliminate the effect of the longitudinal ^{1}H relaxation optimization, a 2 ms ^{1}H spin lock pulse at a field strength of γB_1 = 20 kHz is employed to dephase the ^{1}H magnetization immediately after data acquisition. In (c), the delays are τ_1 = 2.4 ms, τ_2 = 2.34 ms, τ_3 = 2.4 ms and T = 15 ms. The PFGs are G_1: 800 μs, 50 G/cm; G_2: 800 μs, 70 G/cm; G_3: 800 μs, 80 G/cm. The phases are ϕ_1 = {x}; ϕ_2 = {y}; ϕ_3 = {x,-x}; ϕ_4 = {x, x,-x,-x}; ϕ_5 = {x}; ϕ_6 = {x}; ϕ_7 = {x,-x,-x, x,-x, x, x,-x}; ψ_1 = {x,-x}; {x} for all other pulses. Quadrature detection in the $^{13}C^{\alpha}$(t_1) and ^{1}H(t_2) dimensions is achieved by the States-TPPI method [37] applied to the phases ψ_1 and ϕ_3, respectively. The ^{13}CO spins are decoupled by the SEDUCE composite pulse decoupling [28]. For the fractionally deuterated proteins, the experiment can be performed with ^{2}H decoupling during t_1 achieved with WALTZ-16 [38] at a field strength of γB_1 = 0.7 kHz. In (b) and (c), ^{15}N spins are decoupled with GARP [28] at a field strength of γB_1 = 2.5 kHz.

Longitudinal Relaxation-Optimized NMR Experiments. TROSY-type heteronuclear correlation experiments [6,20-26] select the two coherence transfer pathways obtained in alternate scans connecting the single transition of the heteronuclear N spin with the single transition of spin I.

$$I_z^N \rightarrow N_{34}^{+} \rightarrow I_{24}^{N-} \tag{3.1}$$

$$I_z^N \rightarrow N_{34}^{-} \rightarrow I_{24}^{N-} \tag{3.2}$$

The single-transition basis operators refer to the transitions $1 \rightarrow 2$ and $3 \rightarrow 4$ of spin N and $1 \rightarrow 3$ and $2 \rightarrow 4$ of spin I in the standard energy-level diagram for a system of two spins ½. The sensitivity of these experiments is not compromised when a heteronuclear gradient echo is applied [20,27] providing efficient elimination of the water resonance from the resulting spectrum.

In addition to the magnetization transfer pathways given by Equations 3, the experimental scheme of Figure 1a controls two additional magnetization flows given by Equations 4 and 5 outlining the evolution of the density operators of the I_z^C and I_z^W spin groups, respectively,

$$I_z^C \rightarrow C_z I_z^C \rightarrow \pm I_z^C, \quad (4)$$

$$I_z^W \rightarrow I_x^W \rightarrow \pm I_z^W, \quad (5)$$

where the "plus" or "minus" signs can be arbitrarily controlled. For example, the "minus" sign in Equation 4 can be selected by removing the ^{13}C 180º pulse ψ_1 from the experimental scheme of Figure 1a, otherwise the "plus" sign is obtained. The longitudinal ^{1}H relaxation optimization occurs when both positive signs for the magnetization transfer pathways of Equations 4 and 5 are selected.

Similar considerations allow the control of the parallel magnetization transfer pathways in the experimental schemes of Figures 1b and c. In these cases, the separation between I^N and I^C transfer pathways is achieved by applying $^1H^N$ band selective 90º excitation pulses using the distinct difference between chemical shifts of the amide and aliphatic protons. In the case of the 3D LTROSY-HNCA experiment, the TROSY-type transverse relaxation optimization is employed during the long $N \rightarrow C$ and $C \rightarrow N$ polarization transfer periods.

3. Results and Discussion

To enhance longitudinal $^1H^N$ relaxation in the experimental schemes of Figure 1 three independent magnetization transfer pathways are constructed. The main pathway transfers the steady-state $^1H^N$ magnetization to the directly attached ^{15}N spin and back for the signal acquisition. This pathway defines the type of the NMR experiment performed. In the case of TROSY of Figure 1a, the main pathway is outlined by Equations 3 [20]. The $^1H^C$ pathway represented by Equation 4 returns the steady-state $^1H^C$ magnetization to the z direction after the last 90º ^{1}H pulse of the pulse sequence. During the t_1 chemical shift evolution period, the $^1H^C$ magnetization is stored as a slowly relaxing two spin order operator $C_z I_z^C$ [28]. The steady state of the $^1H^C$ magnetization can be controlled by executing or omitting the 180º ^{13}C pulse ψ_1 in Figure 1a. The inversion of the I_z^C at the beginning of the signal acquisition results in the rapid saturation of the $^1H^C$ magnetization and as a consequence retarded return of the $^1H^N$ spins to the steady-state. The evolution pathway of water magnetization is represented by Equation 5. During the t_1 chemical shift evolution period transverse water magnetization is dephased and rephased by a pair of weak PFG pulses to prevent radiation damping. In the experimental scheme of Figure 1a, almost 90% of the initial water magnetization is recovered. Due to the fast chemical exchange between ^{1}H of water and majority of the O-bound ^{1}H spins the I_z^O magnetization quickly equilibrates with the I_z^W magnetization with k_{ex} of the order of 10^3 s^{-1} [1].

As an application, the [^{15}N,^{1}H]-correlation spectrum of the 44 kDa trimeric *B. Subtilis* chorismate mutase (BsCM, PDB code 1COM) (Figure 2) was recorded with the LTROSY experiment of Figure 1a with (Figure 2a) and without (Figure 2b) the longitudinal relaxation optimization of the $^{1}H^{N}$ spins. In all of the spectra of Figure 2 water magnetization was flipped to the +z direction before signal acquisition. The protein was uniformly ^{15}N,^{13}C-labeled and 35% randomly deuterated, dissolved in 95%/5% H_2O/D_2O at a concentration of 2 mM and pH = 7.5 [29]. As expected, a uniform and significant improvement of sensitivity was observed. Figure 2c shows the spectrum measured with the conventional 2D [^{15}N,^{1}H]-TROSY [20] where the water-flip-back pulses [30] were replaced with $^{1}H^{C}$ band-selective 1 ms E-Burp2 pulses [31] centered at 1.4 ppm with all ^{1}H spins resonating between 4.9 ppm and -2 ppm (including water) flipped to the +z direction before data acquisition. In that case, a substantial improvement of sensitivity is also observed (Figure 2c).

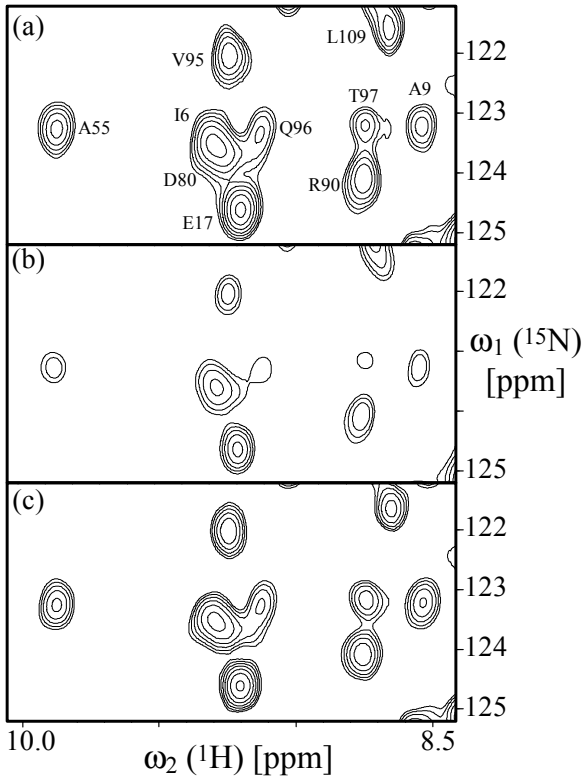


Figure 2. Region of the 2D [^{15}N,^{1}H]-LTROSY spectra of the u-^{13}C,^{15}N-labeled 44 kDa BsCM measured with the experimental scheme of Figure 1a, (a), with the ^{13}C 180° ψ_1 pulse applied and, (b), the ^{13}C 180° pulse ψ_1 not performed. (c), the same spectrum measured with the conventional 2D [^{15}N,^{1}H]-TROSY [20] where the water selective pulses were replaced with $^{1}H^{C}$ band-selective 1 ms E-Burp2 pulses [31] with γB_1 = 4100 Hz centered at 1.4 ppm. In (c), all ^{1}H spins resonating between 4.9 ppm and -2 ppm (including water) are flipped to the +z direction before data acquisition. For all spectra 140(t_1) x 1024(t_2) complex points were accumulated yielding t_{1max} = 20 ms and t_{2max} = 51.2 ms, respectively, with the interscan ^{1}H relaxation delay d_1 = 340 ms and 4 scans per increment. The sequence specific resonance assignments of the cross-peaks are shown.

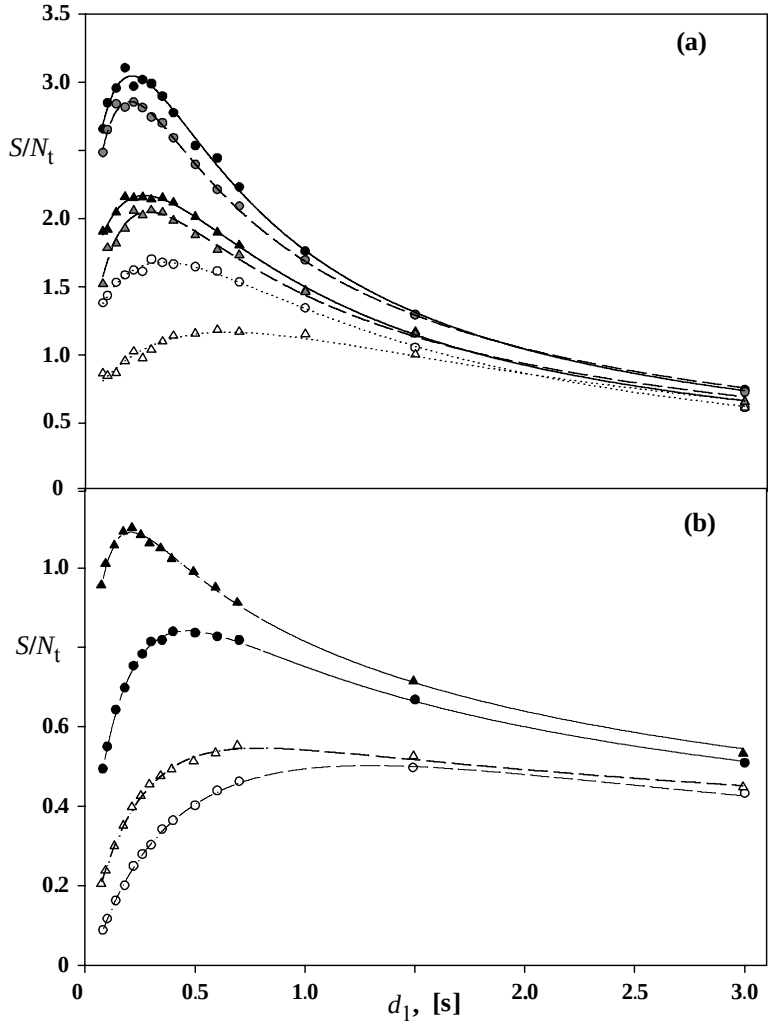


Figure 3. Signal-to-noise ratios per unit time, S/N_t, in a series of the 2D [^{15}N,^{1}H]-correlation spectra versus the interscan ^{1}H relaxation delay d_1 measured with (a) the u-^{13}C,^{15}N-labeled 7 kDa ubiquitin and (b) the u-^{13}C,^{15}N-labeled 44 kDa BsCM. In (a), circles and triangles represent S/N_t measured for Gln2 and Ile13, respectively. The filled and open symbols indicate S/N_t obtained using the experimental scheme of Figure 1a with and without the ^{13}C 180^{o} pulse ψ_1, respectively. The grey symbols represent S/N_t in a series of the 2D [^{15}N,^{1}H]-TROSY [20] experiments with the ^{1}H^{C} flip-back (see Caption to Figure 2c). In (b), circles and triangles represent S/N_t measured for Ala55 and Glu17, respectively. The open and filled symbols indicate S/N_t obtained using the experimental scheme of Figure 1b performed with and without a 2 ms ^{1}H spin lock pulse, respectively. The fit of the experimental relaxation rates using Equation 1 are shown by solid and dashed lines.

Systematic measurements of S/N_t as a function of the interscan delay d_1 were performed using the experimental schemes of Figure 1a and b, with two proteins, 7 kDa uniformly ^{15}N,^{13}C- labeled and uniformly ^{2}H,^{15}N,^{13}C-labeled ubiquitin and 44 kDa uniformly ^{15}N,^{13}C-labeled and 35% randomly deuterated and uniformly ^{2}H,^{15}N,^{13}C-labeled BsCM. Figure 3 shows a 2 - 2.5 times improvement of the maximal S/N_t achieved for both proteins in the protonated and fractionally deuterated states. An excellent fit between Equation 1 and the experimental data is observed for both proteins (Figure 3). For the spectra measured with the longitudinal relaxation optimization the optimal d_1^{opt} delays

corresponding to the maximal values of the S/N_t function are reduced by a factor of 4 on average as it is predicted from Equation 1 and distributed in much narrower range than the optimal d_1^{opt} delays measured without optimization (Figure 4). On the other hand, no improvement in S/N_t was observed in the case where all nonexchangeable protons were replaced with deuterons (Figure 7 in the Appendix) establishing that the sensitivity enhancement mainly results from the construction of the $^1H^C$-magnetization transfer pathway of Equation 4.

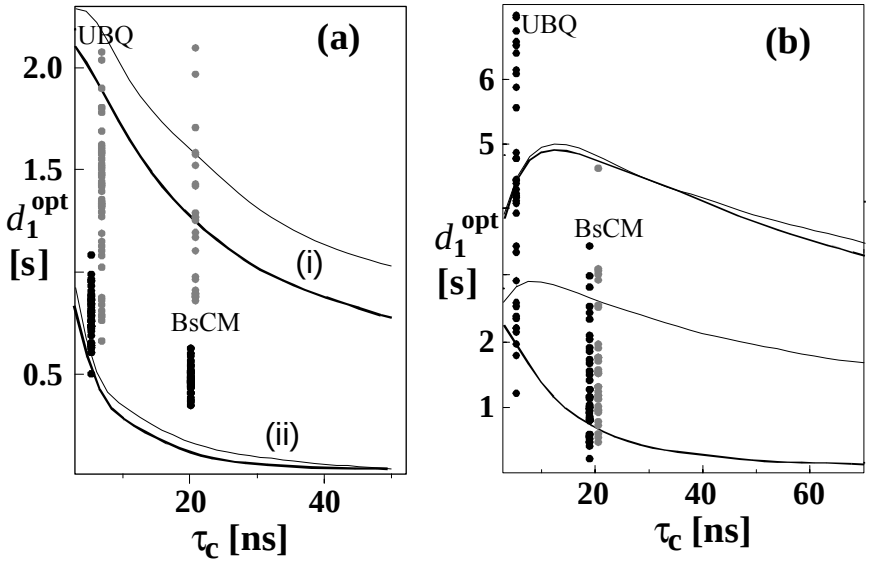


Figure 4. The optimal interscan ^{1}H relaxation delays, d_1^{opt}, versus the protein rotational correlation time of a model protein (ubiquitin, PDB code 1UBI) calculated using the full ^{1}H relaxation matrix approach [32,33] coupled with the kinetic exchange of the labile protons with water (Equation 2) (see the Methods section). The thin and thick lines represent two sets of calculations with the initial water magnetization fully saturated ($I_z^W(0) = 0$) or in the thermal equilibrium with the lattice ($I_z^W(0) = I_z^W$eq), respectively. In (a), calculations are performed for the u-^{15}N,^{13}C-labeled ubiquitin for two sets of initial conditions, (i), the initial magnetization of all ^{1}H spins is zero, (ii), the initial magnetization of $^1H^N$ spins is set to zero (upper curves) and the $^1H^C$ spins are in the thermal equilibrium with the lattice (lower curves). The curves represent the averaged recovery of all $^1H^N$ spins of ubiquitin. For calculations, the Lipari and Szabo spectral density function with $S^2 = 0.8$ and $\tau_e = 100$ ps are assumed. The experimentally determined optimal interscan ^{1}H relaxation delays for each individual $^1H^N$ spin of u-^{15}N,^{13}C-labeled ubiquitin and u-^{15}N,^{13}C-labeled BsCM are shown with grey and filled circles for the experiment of Figure 1b with and without ^{1}H saturation shortly after the $^1H^N$ acquisition. In (b), calculations are performed for the u-^{2}H,^{15}N,^{13}C-labeled ubiquitin in H_2O. The upper and lower curves correspond to the range of the calculated responses for each individual $^1H^N$ spin. The experimentally determined optimal interscan ^{1}H relaxation delays for each individual $^1H^N$ spin of u-^{2}H,^{15}N,^{13}C-labeled ubiquitin and u-^{2}H,^{15}N,^{13}C-labeled BsCM are shown with grey and filled circles for the experiment of Figure 1b with and without ^{1}H saturation short after the $^1H^N$ acquisition.

Full relaxation-matrix analysis [32,33] coupled with the kinetic exchange of the labile protons with water was performed by the numerical integration of Equation 2. The calculated recovery of the $^1H^N$ magnetization indicates that the significantly shorter optimal interscan d_1^{opt} delays are achieved in the

longitudinal relaxation-optimized experiments as it is compared to the standard experiments. In addition, a reduced dependence on the saturation level of water magnetization (thin and thick curves in Figures 4a and b) is observed. A satisfactory correspondence between the measured optimal d_1^{opt} delays and the theoretical d_1 delays is obtained. Some deviation of the optimal interscan d_1^{opt} delays measured with ^{13}C,^{15}N-labeled and fractionally deuterated BsCM from the theoretical curve might originate from (i) the partial dilution of the $^{1}H^{C}$ spins by the ^{2}H spins, (ii) saturation of the $^{1}H^{C}$ spins of aromatic groups resonating close to the $^{1}H^{N}$ spins, and (iii) failure of the full relaxation-matrix analysis to adequately represent spin dynamics in fractionally labeled proteins [34]. For the uniformly deuterated ubiquitin and BsCM, a wide distribution of the optimal interscan delays throughout the backbone $^{1}H^{N}$ spins is predicted and actually observed in the experiment (Figure 4b) resulting in a nonuniform appearance of the signal intensities in NMR spectra. Within the scope of the proposed approach this situation can not be improved.

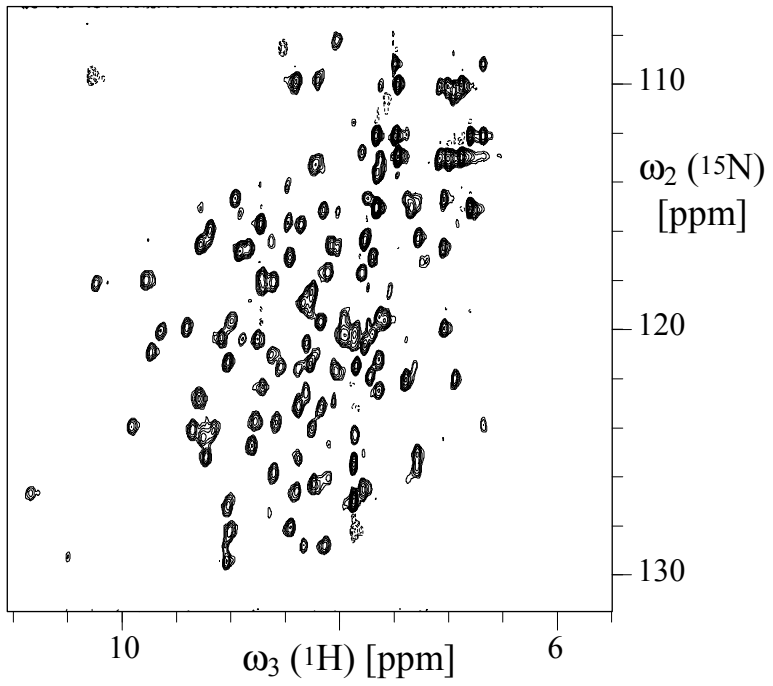


Figure 5. 2D [^{15}N,^{1}H]-projection of the 3D LTROSY-HNCA experiment of Figure 1c with the ^{1}H relaxation optimization applied to the u-^{15}N,^{13}C-labeled BsCM. The interscan ^{1}H-relaxation delay d_1 = 340 ms. 12(t_1) x 36(t_2) x 1024(t_3) complex points were accumulated yielding t_{1max} = 5 ms, t_{2max} = 13 ms and t_{3max} = 51.2 ms, respectively, and a total measurement time of 6 h (32 scans).

At high magnetic fields a viable solution to the longitudinal relaxation optimization can be obtained by the use of the $^{1}H^{N}$ band-selective excitation pulses. Figure 5 shows a 2D [^{15}N,^{1}H]-projection of the 3D LTROSY-HNCA spectrum of ^{13}C,^{15}N-labeled and 35% deuterated BsCM measured with the experimental scheme of Figure 1c. The uniform excitation via the $^{1}H^{N}$ region was obtained with the E-Burp2 and time reversed E-Burp2 pulses [31]. No active selection of the

TROSY pathway was employed which would require implementation of additional $^1H^N$ selective 90º pulses. For the majority of the backbone amide groups of BsCM, the anti-TROSY cross-peaks were not detected due to strong transverse relaxation during the N-C magnetization transfer periods (Figure 5).

The source of enhanced longitudinal relaxation of $^1H^N$ spins to the steady-state level of the magnetization is the dipole-dipole interactions between the $^1H^N$ spins and all other ^{1}H spins in protein [8] and in bulk water [18,35]. The maximum effect is achieved when the magnetization of the $^1H^N$ group of spins is selectively perturbed from the corresponding thermal equilibrium. Frequently, this regime is called selective ^{1}H T_1 relaxation, which proceeds much faster and exhibits only small polarizing magnetic field strength dependence in contrast to nonselective ^{1}H T_1 relaxation, where the effects of dipole-dipole auto- and cross-relaxation tend to compensate each other [36]. Figure 6 compares the optimal interscan d_1^{opt} delays calculated as a function of the polarizing magnetic field strength B_0 for the longitudinal relaxation-optimized experiments and conventional experiments. Due to the increase of the nonselective longitudinal relaxation times of the ^{1}H spins in proteins with the magnetic field strength (Figure 6), the optimal interscan delays in the longitudinal relaxation-optimized NMR experiments are of an order of magnitude shorter than that in the conventional experiments.

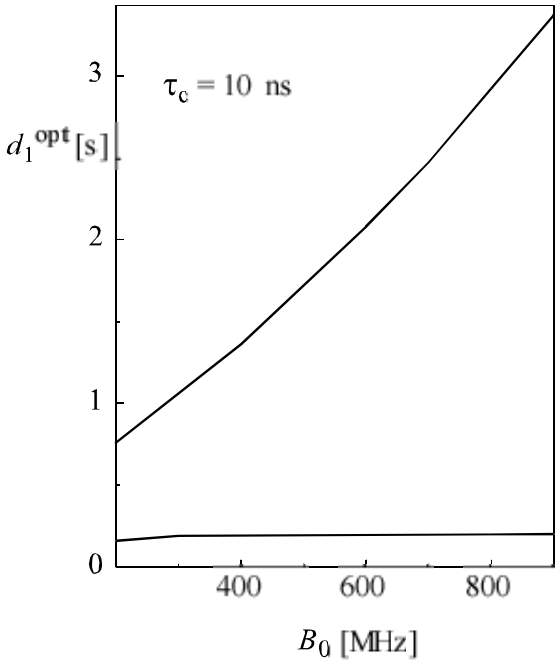


Figure 6. Optimal interscan $^1H^N$ relaxation delays, d_1^{opt}, of a model protein (ubiquitin, PDB code 1UBI) versus polarizing magnetic field strength B_0 calculated using the full ^{1}H relaxation matrix approach (Equation 2). The lower and upper curves correspond to the $^1H^N$ relaxation in the LTROSY experiment of Figure 1a and conventional [^{15}N,^{1}H]-correlation techniques, respectively.

The concentration of the saturated protons is almost 5 times smaller than that of the protons with the corresponding magnetization returned back to the thermal equilibrium. Thus, in the spin diffusion regime the dipole-dipole interactions between diluted $^1H^N$ spins and abundant "thermal bath" spins are not expected to perturb significantly the equilibrium state of the abundant spins. The effective longitudinal relaxation enhancement of the $^1H^N$ spins throughout entire multidimentional NMR experiments can be obtained provided that the magnetization of the thermal bath spins is maintained close to its thermal equilibrium or returned to the thermal equilibrium state before the signal acquisition as it is implemented in the experimental scheme of Figure 1, parts b and a, respectively. The dilution of the thermal bath ^{1}H spins with ^{2}H spins reduces the efficiency of the longitudinal relaxation optimization, so that the uniform deuteration of proteins should be undertaken with care. At 900 MHz, the predicted decrease of the $^1H^N$ longitudinal relaxation times can be as large as an order of magnitude making the proposed method a valuable tool for protein NMR at high magnetic fields.

4. Appendix

Figure 7 illustrates that no improvement in S/N_t was observed in the case where all nonexchangeable protons were replaced with deuterons.

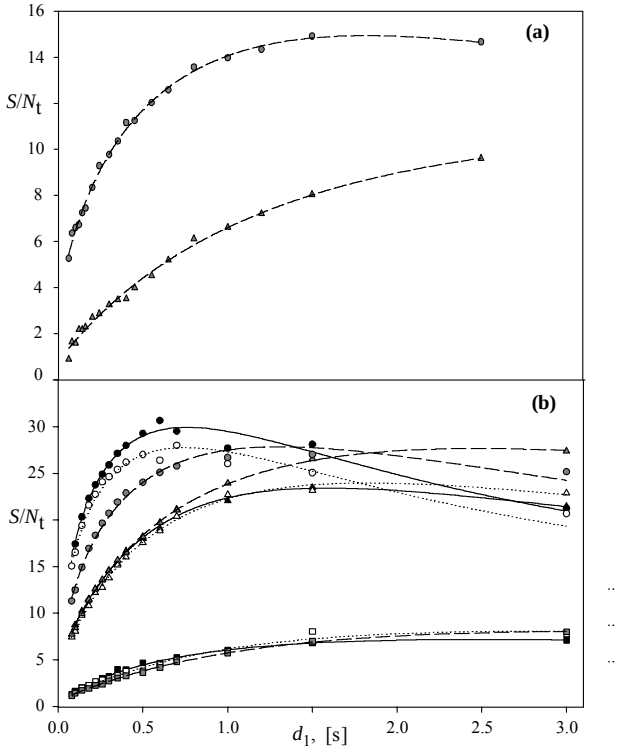


Figure 7 (previous page). The signal-to-noise ratios per unit time, S/N_t, in a series of the 2D [^{15}N,^{1}H]-correlation spectra versus the interscan ^{1}H relaxation delay d_1 measured with (a) the u-^{2}H,^{13}C,^{15}N-labeled 7 kDa Ubiquitin and (b) the u-^{2}H,^{13}C,^{15}N-labeled 44 kDa BsCM. In (a) circles and triangles represent S/N_t measured for Gln2 and Ile13, respectively. In (b), circles, triangles and rectangles represent S/N_t measured for Met2, Glu110 and Ala55, respectively. The filled and open symbols indicate S/N_t obtained using the experimental scheme of Figure 1a with and without the ^{13}C 180^o pulse ψ_1, respectively. The grey symbols represent S/N_t in a series of the 2D [^{15}N,^{1}H]-TROSY [20] experiments with the $^1H^C$ flip-back. The fits of the experimental relaxation rates using Equation 1 are shown by solid and dashed lines.

5. References

[1] Wüthrich, K., *NMR of Proteins and Nucleic Acids*. **1986**, Wiley, New York.
[2] Wagner, G. *J. Biomol. NMR* **1993**, *3*, 375-385.
[3] Kay, L.E.; Gardner, K.H. *Curr. Opin. Struct. Biol.* **1997**, *7*, 722-731.
[4] Clore, G.M.; Gronenborn, A.M. *Nat. Struct. Biol.* **1997**, *4 Suppl*, 849-853.
[5] Clore, G.M.; Gronenborn, A.M. *Curr. Opin. Chem. Biol.* **1998**, *2*, 564-70.
[6] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366-12371.
[7] Wang, Y.X.; Jacob, J.; Cordier, F.; Wingfield, P.; Stahl, S.J.; Lee-Huang, S.; Torchia, D.; Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1999**, *14*, 181-184.
[8] Ernst, R.R.; Bodenhausen, G.; Wokaun, A. *The Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. **1987**, Oxford, Clarendon.
[9] Pintacuda, G.; Otting, G. *J. Am. Chem. Soc.* **2002**, *124*, 372-373.
[11] Liepinsh, E.; Baryshev, M.; Sharipo, A.; Ingelman-Sundberg, M.; Otting, G.; Mkrtchian, S. *Structure* **2001**, *9*, 457-471.
[11] Machonkin, T.E.; Westler, W.M.; Markley, J.L. *J. Am. Chem. Soc.* **2002**, *124*, 3204- 3205.
[12] Chen, Z.G.; Deropp, J.S.; Hernandez, G.; Lamar, G.N. *J. Am. Chem. Soc.* **1994**, *116*, 8772-8783.
[13] Krugh, T.R.; Schaefer, W.C. *J. Magn. Reson.* **1975**, *19*, 99-107.
[14] Pellecchia, M.; Meininger, D.; Shen, A.L.; Jack, R.; Kasper, C.B.; Sem, D.S. *J. Am. Chem. Soc.* **2001**, *123*, 4633-4634.
[15] Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D.S. *J. Biomol. NMR* **2002**, *22*, 165-173.
[16] Ross, A.; Salzmann, M.; Senn, H. *J. Biomol. NMR* **1997**, *10*, 389-396.
[17] Riek, R. *J. Biomol. NMR* **2001**, *21*, 99-105.
[18] Otting, G.; Liepinsh, E.; Wuthrich, K. *Science* **1991**, *254*, 974-980.
[19] Englander, S.W.; Mayne, L.; Bai, Y.; Sosnick, T.R. *Protein Sci.* **1997**, *6*, 1101-1109.
[20] Pervushin, K.; Wider, G.; Wüthrich, K. *J. Biomol. NMR* **1998**, *12*, 345-348.
[21] Czisch, M.; Boelens, R. *J. Magn. Reson.* **1998**, *134*, 158-160.
[22] Weigelt, J. *J. Am. Chem. Soc.* **1998**, *120*, 10778-10779.
[23] Brutscher, B.; Boisbouvier, J.; Pardi, A.; Marion, D.; Simorre, J.P. *J. Am. Chem. Soc.* **1998**, *120*, 11845-11851.
[24] Andersson, P.; Annila, A.; Otting, G. *J. Magn. Reson.* **1998**, *133*, 364-367.
[25] Meissner, A.; Schulte-Herbrüggen, T.; Briand, J.; Sørensen, O.W. *Mol. Phys.* **1998**, *95*, 1137-1142.
[26] Rance, M.; Loria, J.P.; Palmer, A.G. *J. Magn. Reson.* **1999**, *136*, 92-101.
[27] Kay, L.E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663-10665.
[28] Cavanagh, J.; Fairbrother, W.J.; Palmer, A.G.; Skelton, N.J., *Protein NMR Spectroscopy: Principles and Practice*. **1996**, Academic Press, New York.
[29] Eletsky, A.; Kienhöfer, A.; Pervushin, K. *J. Biomol. NMR* **2001**, *20*, 177-180.
[30] Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 12593-12594.
[31] Geen, H.; Freeman, R. *J. Magn. Reson.* **1991**, *93*, 93-141.
[32] Olejniczak, E.T.; Gampe, R.T.; Fesik, S.W. *J. Magn. Reson.* **1986**, *67*, 28-38.
[33] Boelens, R.; Koning, T.M.G.; Kaptein, R. *J. Mol. Struct.* **1988**, *173*, 299-309.
[34] Zolnai, Z.; Juranic, N.; Macura, S. *J. Biomol. NMR* **1998**, *12*, 333-337.
[35] Brüschweiler, R.; Wright, P.E. *Chem. Phys. Lett.* **1994**, *229*, 75-81.
[36] Rossi, C., Selective Relaxation Techniques in Biological NMR, in *Encyclopedia of nuclear magnetic resonance*, Grant, D.M.; Harris, R.K., Eds.; **1996**, John Wiley & Sons, New York, p. 4237-4246.
[37] Marion, D.; Ikura, M.; Tschudin, R.; Bax, A. *J. Magn. Reson.* **1989**, *85*, 393-399.
[38] Shaka, A.J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335-338.

Chapter III

TROSY Experiment for Refinement of Backbone ψ and ϕ by Simultaneous Measurements of Cross-Correlated Relaxation Rates and $^{3,4}J_{H\alpha HN}$ Coupling Constants

Beat Vögeli and Konstantin Pervushin

Laboratorium für Physikalische Chemie, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

Journal of Biomolecular NMR, 2002, 24, 291-300

1. Introduction

In the pursuit of more precise solution structures of proteins the usual sources of structural constraints based on proton-proton NOE cross-relaxation and scalar coupling constants [1] are effectively complemented by dihedral angle estimates derived from the effects of cross- correlated cross relaxation [2,3]. The main focus of newly developed NMR experiments is the determination of various dihedral angles in proteins by the use of quantum interference between spin Hamiltonians, which describe chemical shift anisotropy and dipolar interactions [4-15]. On the other hand, preservation of the spin state selectivity in NMR experiments enables effective determination of the various scalar coupling constants in an E.COSY manner [16-18]. The quality of spectra obtained with proteins of larger molecular weight can be significantly improved by the use of TROSY [19]. TROSY is a spectroscopic technique based on the constructive use of interference effects without mixing of the relevant spin states. Thus, all these experiments share the same physical basis and consequently, it is instructive to design a highly optimized experiment unifying these approaches, which enables direct determination of the ψ and ϕ angles in larger proteins by the use of the cross-correlated relaxation rates together with the Karplus-type dependencies of the coupling constants.

In this paper we present a new TROSY-type experiment where backbone ψ and ϕ angles are determined using cross-correlated cross relaxation rates between H^{α}-C^{α} and H^{N}-N dipolar interactions and between H^{α}-C^{α} dipolar interaction and N chemical shift anisotropy (CSA).

Simultaneously, the values of the $^{3,4}J_{H\alpha HN}$ scalar couplings are measured in an E.COSY manner.

In the case of the ϕ angle, the Karplus curve for $^{3}J_{H\alpha HN}$ yields an independent angle determination [20], whereas $^{4}J_{H\alpha HN}$ might provide additional yet not parametrized constraints on the ψ angle [21].

2. Methods

We introduce two basic variants of the experimental scheme of Figure 1, which are optimized for applications at lower (up to 600 MHz) and higher polarizing magnetic field strength B_0, respectively. In the following, the coherence transfer pathways in each of the two experiments are analyzed using the product operator formalism [22].

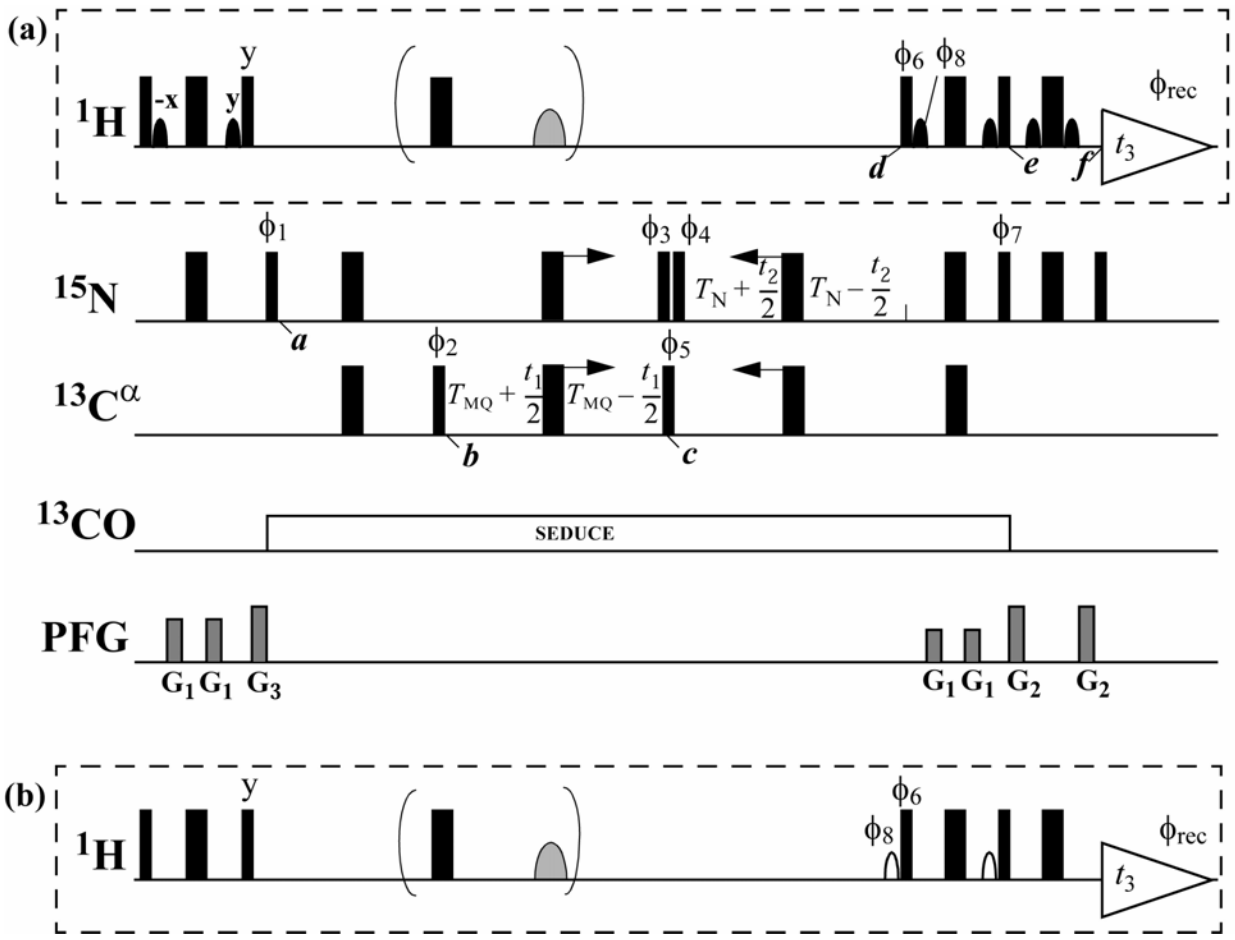


Figure 1. The experimental scheme of the (a) [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY and (b) [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY. A variant of each experiment designed for use at polarizing magnetic field strengths B_0 higher than 600 MHz (see Methods) is obtained by the insertion of two $^1H^N$ inversion pulses indicated in brackets. Narrow (wide) black bars represent non-selective 90^o (180^o) pulses on ^{1}H, $^{13}C^\alpha$, ^{15}N and ^{13}C' with carrier offsets placed at 4.7 ppm, 55.0 ppm, 118.7 ppm and 175.5 ppm, respectively. Black and open curved shapes indicate selective 90^o Gaussian pulses on the water ^{1}H resonance with a duration of 1 ms and selective 90^o E-Burp pulses on $^1H^\alpha$ and water ^{1}H with a duration of 1.5 ms , respectively [23]. The grey curved shapes represent $^1H^N$-selective 180^o I-Burp pulses. The boxes on the line labeled PFG indicate sine-shaped pulsed magnetic field gradients along the z-axis with 1 ms duration except for G_2 in (b), which has a duration of 950 μs: G_1, 30 G/cm; G_2 80 G/cm; 60 G/cm. The delays are: τ_1 = 2.7 ms, τ_2 = 18 ms and 34 ms for the determination of the ϕ and ψ angles, respectively, T_{MQ} = 10.8 ms, $T_N = \tau_2$ in (a) and $T_N = \tau_2 - \tau_1$ in (b). The 16 separately stored interferograms c_i (i = 1..16) are obtained by the following phases: $\phi_1 = \phi_2$ = {x, x, x, x, x, x, x, x, y, y, y, y, y, y, y, y}; ϕ_3 = {x}; ϕ_4 = {x, x, x, x, -x, -x, -x, -x}; ϕ_5 = {x, y, x, y, x, y, x, y, y, -x, y, - x, y, -x, y, -x}; $\phi_6 = \phi_7$ = {y, y, -y, -y}; ϕ_8 = {-y, -y, y, y}; ϕ_{rec} = {x, y, x, -y, x, y, x, -y, -y, x, -y, - x, -y, x, -y, -x}; x on all pulses without phase specification. The basic 4-step phase cycle is carried out for each of the 16 interferograms in the following manner: ϕ_1 = {u, -u, -u, u}; ϕ_2 = {u, -u, u, - u}; ϕ_{rec} = {v, v, -v, -v}, where u and v stand for x or y. ^{13}C' decoupling is performed using the SEDUCE phase modulation [24]. Spectra are processed by a linear combination of the obtained set of the interferograms as described in Methods.

The [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY designed for use at lower B_0 omits the double inversion of the $^1H^N$ spins during MQ evolution making full use of TROSY-type relaxation

compensation. To emphasize this fact, the symbol $\bar{\mathrm{H}}$ (= inversion) is added to the corresponding experiment names. The four relevant coherence tranfer pathways are given by Equation 1.

$$\begin{aligned} &H_z \rightarrow MQ\,(E/2 - H_z^{\mathrm{N}})(E/2 \mp H_z^{\alpha})\exp[-2R_{\beta/\alpha}{}^{\mathrm{MQ}}T_{\mathrm{MQ}} - i(\omega^{\mathrm{MQ}} \mp \pi J_{\mathrm{HC}} - \pi J_{\mathrm{HN}})t_1] \\ &\rightarrow N_+^{34}(E/2 \mp H_z^{\alpha})\exp[-2R^{34}T_{\mathrm{N}} - i\omega^{34}t_2] \rightarrow H_-^{24}(E/2 \mp H_z^{\alpha})\exp[-2R^{24}t_3 - i\omega^{24}t_3], \end{aligned} \tag{1}$$

where *MQ* represents either one of the two double-quantum operators (*DQ*), $C_{\pm}N_{\pm}$, or one of the two zero-quantum operators (*ZQ*), $C_{\mp}N_{\pm}$, and N_+^{34} and H_-^{24} are the single-transition basis operators $N_+(E/2 - H_z)$ and $N_-(E/2 + H_z)$, respectively. The chemical shifts relative to the carrier frequencies are $\omega^{\mathrm{DQ}} = \omega^{\mathrm{N}} + \omega^{\mathrm{C}}$, $\omega^{\mathrm{ZQ}} = \omega^{\mathrm{N}} - \omega^{\mathrm{C}}$, $\omega^{34} = \omega^{\mathrm{N}} - \pi J_{\mathrm{HN}}$ and $\omega^{24} = \omega^{\mathrm{H}} + \pi J_{\mathrm{HN}}$. J_{HN} and J_{HC} are the scalar coupling constants between $^1\mathrm{H}^{\mathrm{N}}$ and $^{15}\mathrm{N}$ spins and between $^1\mathrm{H}^{\alpha}$ and $^{13}\mathrm{C}^{\alpha}$ spins, respectively. R^{34} and R^{24} are the relaxation rates of the corresponding TROSY operators [19]. The MQ relaxation rates $2R_{\beta/\alpha}{}^{\mathrm{MQ}}$ are given by Equations 2-5, where the indices α and β correspond to the spin state operators $E/2 + H_z^{\alpha}$ and $E/2 - H_z^{\alpha}$, respectively [25]:

$$\begin{pmatrix} R_{\beta}^{\mathrm{DQ}} \\ R_{\alpha}^{\mathrm{DQ}} \\ R_{\beta}^{\mathrm{ZQ}} \\ R_{\alpha}^{\mathrm{ZQ}} \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 & -1 & -1 & -1 & 1 & 0 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & -1 & -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} [\tfrac{8}{3}(A_{\mathrm{CSA(C)}}{}^2 + A_{\mathrm{CSA(N)}}{}^2)B_0{}^2 + \tfrac{2}{3}(A_{\mathrm{D(HC)}}{}^2 + A_{\mathrm{D(HN)}}{}^2)]J(0) \\ \tfrac{16}{3}A_{\mathrm{CSA(C)}}A_{\mathrm{CSA(N)}}Q(\Theta_{\mathrm{CSA(C)CSA(N)}})B_0{}^2J(0) \\ \tfrac{8}{3}A_{\mathrm{D(HC)}}A_{\mathrm{CSA(C)}}Q(\Theta_{\mathrm{D(HC)CSA(C)}})B_0J(0) \\ \tfrac{8}{3}A_{\mathrm{D(HN)}}A_{\mathrm{CSA(N)}}Q(\Theta_{\mathrm{D(HN)CSA(N)}})B_0J(0) \\ \tfrac{8}{3}A_{\mathrm{D(HC)}}A_{\mathrm{CSA(N)}}Q(\Theta_{\mathrm{D(HC)CSA(N)}})B_0J(0) \\ \tfrac{8}{3}A_{\mathrm{D(HN)}}A_{\mathrm{CSA(C)}}Q(\Theta_{\mathrm{D(HN)CSA(C)}})B_0J(0) \\ \tfrac{4}{3}A_{\mathrm{D(HC)}}A_{\mathrm{D(HN)}}Q(\Theta_{\mathrm{D(HC)D(HN)}})J(0) \\ \tfrac{2}{3}A_{\mathrm{D(H}\alpha\mathrm{HN)}}{}^2J(\omega^{\mathrm{H}\alpha} - \omega^{\mathrm{HN}}) \end{pmatrix} \tag{2-5}$$

The spectral density functions $J(\omega)$ of the spectral frequency ω in an approximation of fast internal motions and interaction constants for CSA of nucleus *k* ($A_{\mathrm{CSA}(k)}$) and for dipole-dipole interaction of a given $^1\mathrm{H}$ and a nucleus *k* ($A_{\mathrm{D(H}k)}$) are given by the Equations 6-9. In addition, we assume that various auto- and cross-correlated relaxation spectral densities encountered in the same spin system can be reasonably represented by the same functional form given by Equation 6.

$$J(\omega) = \frac{\tau_{\mathrm{c}}S^2}{1 + \omega^2\tau_{\mathrm{c}}{}^2}, \tag{6}$$

$$A_{\mathrm{CSA}(k)} = -\sqrt{\tfrac{3}{10}}\,\tfrac{1}{2}\gamma_k\delta_{\mathrm{z}}^k, \tag{7}$$

$$A_{\mathrm{D(H}k)} = \sqrt{\tfrac{3}{10}}\,\frac{\mu_0}{8\pi^2}\gamma_{\mathrm{k}}\gamma_{\mathrm{H}}\mathrm{h}\left\langle\frac{1}{r_{\mathrm{H}k}^3}\right\rangle, \tag{8}$$

$$Q(\Theta) = P(\cos\Theta) = \frac{3(\cos\Theta)^2 - 1}{2}, \tag{9}$$

where τ_c is the correlation time, S^2 the generalized order parameter, γ_k the gyromagnetic ratio of the nucleus k, δ_z^k the axial principal component of the anti-symmetrical part of the chemical shift tensor for nucleus k, which is assumed to be axially symmetric, μ_0 the permeability of free space, h Planck's constant, r_{Hk} the distance between the spins ^{1}H and k, and Θ the projection angle between the indicated interactions (Figure 2).

For glycine residues, a set of equations similar to Equations 2-5 have been derived in closed form. In brief, each of the Equations 2-5 is replaced by two equations corresponding to the up or down state of the second proton directly attached to the C^α atom. In the right hand expression of each modified equation, the terms containing $A_{D(HC)}$ are replaced by two terms corresponding to the spin states of the $^1H^{\alpha 1}$ and $^1H^{\alpha 2}$ spins. In addition, a term describing the $H^{\alpha 1}$-C^α/$H^{\alpha 2}$-C^α dipole/dipole interaction is added.

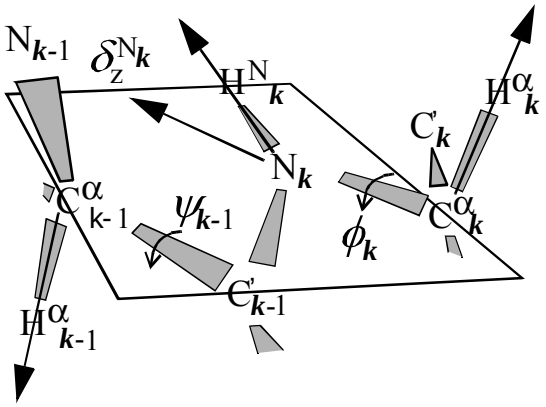


Figure 2. Schematic representation of the protein backbone geometry. The arrows indicate the vectors on which the measured relaxation rates depend. δ_z^N is the axial principal component of the anti-symmetrical part of the ^{15}N chemical shift tensor, ψ and ϕ are the dihedral angles defining the backbone conformation. Small variations of the ^{15}N chemical shift tensor principal components and their directions throughout the polypeptide chain can be expected [26].

The initial ^{1}H polarization is transfered by an INEPT step to N_+^{34} at time point *a*, followed by a pulse segment which creates double- and zero-quantum coherence between ^{15}N and ^{13}C at time point *b* (Figure 1a). The MQ operators are represented as a linear combination of two single-transition operators with respect to the single transitions of the $^1H^N$ and $^1H^\alpha$ spins, $MQ(E/2 - H_z^N)(E/2 - H_z^\alpha)$ and $MQ(E/2 - H_z^N)(E/2 + H_z^\alpha)$. During the constant time period T_{MQ} between the time points *b* and *c*, these operators evolve due to the MQ chemical shifts Ω^{DQ} and Ω^{ZQ} and the scalar couplings of $^1J_{HN}$ and $^1J_{HC}$. Between time points *c* and *e*, the MQ coherence is transfered back to the ^{15}N spins via the single-transition TROSY operator N_+^{34} followed by the standard ST2-PT element, which transfers N_+^{34} to H_-^{24} [27]. For the experiment of Figure 1b, the $(E/2 - H_z^\alpha)$ and $(E/2 + H_z^\alpha)$ spin

states are not mixed during the ST2-PT element, so that the peaks of the resulting doublets (or triplets in the case of glycines) shown in Figure 4 are shifted by $^{3,4}J_{H\alpha HN}$ relative to each other in the Ω^{H}-dimension. The inclusion of a water-flip-back as implemented in the experimental scheme of Figure 1a, results in partial mixing of the ($E/2 - H_z^{\alpha}$) and ($E/2 + H_z^{\alpha}$) spin states for $^1H^{\alpha}$ excited by the selective pulses, which prevents quantitative measurements of the $^{3,4}J_{H\alpha HN}$ coupling constants, but not cross-correlated relaxation rates. Thus, if sensitivity losses due to the water saturation is not a question, the water flip back trick should be avoided.

In order to maximize the spectral sensitivity, all coherence transfer pathways are retained. The individual phase sensitive subspectra such as the 3D TROSY-ZQ-HNCA and the 3D TROSY-DQ-HNCA are then obtained from linear combinations of 16 separately stored interferograms measured for each (t_1, t_2)-pair (Figure 1). Table 1 summarizes the linear combinations employed to obtain each subspectrum. In Table 1 the column c_i represents the sign of the contribution of the i-th interferogram given in Figure 1 to the j-th subspectrum Σ_j, where both i and j range from 1 to 16. The complex exponents $\exp(-\Omega^{MQ}t_1)\exp(-\Omega^{12/34}t_2)$ required for the complex Fourier transform are obtained by combining four of the cosine and sine modulated terms. For example, $\exp(-\Omega^{DQ}t_1)\exp(-\Omega^{34}t_2)$ is built up with the subspectra Σ_5 - Σ_8. Finally, only spectra showing the slowly relaxing TROSY components in the [^{1}H,^{15}N]-plane [19] are retained for further analysis. Neglecting transverse relaxation, one quarter of the initially combined $^1H^N$ and ^{15}N Boltzmann steady state magnetizations contribute to the individual TROSY-DQ-HNCA or TROSY-ZQ-HNCA spectra. The losses stem from the TROSY selection and the States-TPPI quadrature detection of the MQ frequencies.

In order to obtain structural information, the four relaxation rates of the doublet peaks in the DQ- and ZQ-spectra are linearly combined resulting in cancelation of all terms except for the dipole(HC)/dipole(HN)- and dipole(HC)/CSA(N)-cross-correlated terms given by Equation 10.

$$\tfrac{4}{3}[A_{D(HC)}A_{D(HN)}Q(\Theta_{D(HC)D(HN)}) - 2A_{D(HC)}A_{CSA(N)}B_0Q(\Theta_{D(HC)CSA(N)})]J(0) = \tfrac{1}{4}(R_\beta^{DQ} - R_\alpha^{DQ} - R_\beta^{ZQ} + R_\alpha^{ZQ}) \tag{10}$$

To be able to derive the expressions for the ψ and ϕ angles in a closed form, we further assume that the difference in contribution of the passive scalar couplings to ZQ and DQ coherences is negligible. In that case, the linear combination of the relaxation rates on the right side of Equation 10 can be expressed via the experimentally measured intensities of the 3D cross-peaks in the DQ and ZQ spectra:

$$\tfrac{1}{4}(R_\beta^{DQ} - R_\alpha^{DQ} - R_\beta^{ZQ} + R_\alpha^{ZQ}) = \frac{1}{T_{MQ}}\ln\left(\frac{I_\alpha^{DQ}I_\beta^{ZQ}}{I_\beta^{DQ}I_\alpha^{ZQ}}\right) \tag{11}$$

A combination of the Equations 10 and 11 yields:

$$\tfrac{4}{3}[A_{\mathrm{D(HC)}}A_{\mathrm{D(HN)}}Q(\Theta_{\mathrm{D(HC)D(HN)}})-2A_{\mathrm{D(HC)}}A_{\mathrm{CSA(N)}}B_0Q(\Theta_{\mathrm{D(HC)CSA(N)}})]J(0)=\frac{1}{T_{\mathrm{MQ}}}\ln\left(\frac{I_\alpha^{\mathrm{DQ}}I_\beta^{\mathrm{ZQ}}}{I_\beta^{\mathrm{DQ}}I_\alpha^{\mathrm{ZQ}}}\right) \qquad (12)$$

Table 1. The subspectra Σ_1 - Σ_{16} obtained by linear combination of the interferograms c_1 - c_{16} and the resulting spectral terms used to reconstruct phase sensitive 3D [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]- TROSY-E.COSY subspectra.

	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	c_{16}	Spectral terms
Σ_1	–	–	–	–	–	+	–	+	–	–	+	+	+	–	–	+	$\cos(\omega^{\mathrm{ZQ}}t_1)\sin(\omega^{12}t_2)H_y{}^{24}$
Σ_2	+	+	+	+	–	+	–	+	+	+	–	–	+	–	–	+	$\sin(\omega^{\mathrm{ZQ}}t_1)\cos(\omega^{12}t_2)H_y{}^{24}$
Σ_3	–	–	+	+	+	–	–	–	+	+	+	+	+	–	+	–	$\sin(\omega^{\mathrm{ZQ}}t_1)\sin(\omega^{12}t_2)H_x{}^{24}$
Σ_4	–	–	+	+	–	+	+	–	–	–	–	–	+	–	+	–	$\cos(\omega^{\mathrm{ZQ}}t_1)\cos(\omega^{12}t_2)H_x{}^{24}$
Σ_5	–	–	–	–	–	+	–	+	+	+	–	–	–	+	+	–	$\cos(\omega^{\mathrm{DQ}}t_1)\sin(\omega^{34}t_2)H_y{}^{13}$
Σ_6	+	+	+	+	–	+	–	+	–	–	+	+	–	+	+	–	$\sin(\omega^{\mathrm{DQ}}t_1)\cos(\omega^{34}t_2)H_y{}^{13}$
Σ_7	+	+	–	–	–	+	+	–	–	–	–	–	–	+	–	+	$\sin(\omega^{\mathrm{DQ}}t_1)\sin(\omega^{34}t_2)H_x{}^{13}$
Σ_8	+	+	–	–	+	–	–	+	–	–	–	–	+	–	+	–	$\cos(\omega^{\mathrm{DQ}}t_1)\cos(\omega^{34}t_2)H_x{}^{13}$
Σ_9	–	+	–	+	–	–	–	–	–	+	+	–	+	+	–	–	$\cos(\omega^{\mathrm{DQ}}t_1)\sin(\omega^{12}t_2)H_y{}^{24}$
Σ_{10}	+	–	+	–	–	–	–	–	+	–	–	+	+	+	–	–	$\sin(\omega^{\mathrm{DQ}}t_1)\cos(\omega^{12}t_2)H_y{}^{24}$
Σ_{11}	–	+	+	–	+	+	–	–	–	+	–	+	–	–	–	–	$\sin(\omega^{\mathrm{DQ}}t_1)\sin(\omega^{12}t_2)H_x{}^{24}$
Σ_{12}	–	+	+	–	–	–	+	+	–	+	–	+	+	+	+	+	$\cos(\omega^{\mathrm{DQ}}t_1)\cos(\omega^{12}t_2)H_x{}^{24}$
Σ_{13}	–	+	–	+	–	–	–	–	+	–	–	+	–	–	+	+	$\cos(\omega^{\mathrm{ZQ}}t_1)\sin(\omega^{34}t_2)H_y{}^{13}$
Σ_{14}	+	–	+	–	–	–	–	–	–	+	+	–	–	–	+	+	$\sin(\omega^{\mathrm{ZQ}}t_1)\cos(\omega^{34}t_2)H_y{}^{13}$
Σ_{15}	–	+	+	–	+	+	–	–	+	–	+	–	+	+	+	+	$\sin(\omega^{\mathrm{ZQ}}t_1)\sin(\omega^{34}t_2)H_x{}^{13}$
Σ_{16}	+	–	–	+	+	+	–	–	–	+	–	+	+	+	+	+	$\cos(\omega^{\mathrm{ZQ}}t_1)\cos(\omega^{34}t_2)H_x{}^{13}$

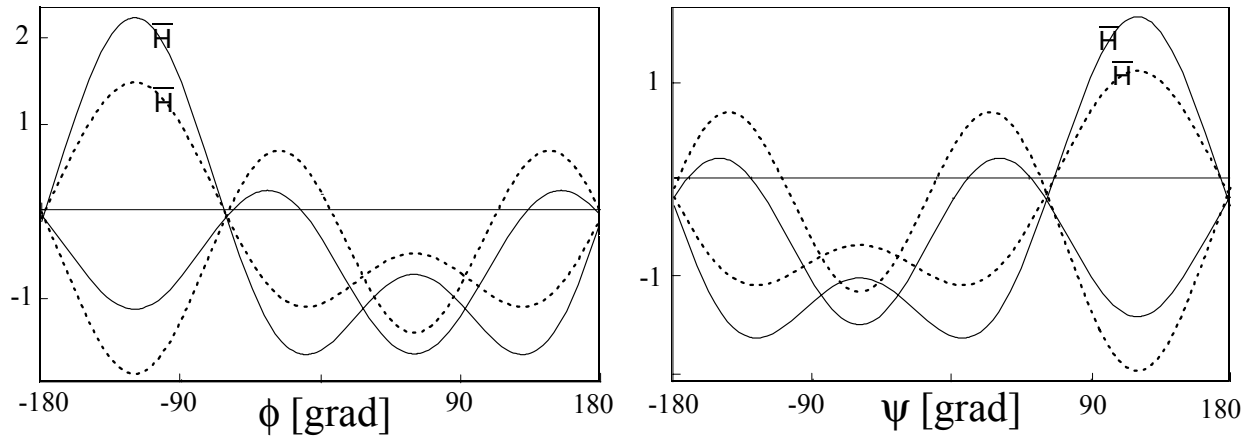


Figure 3. The left-hand sides of the Equations 12 and 14 as functions *F* of the dihedral ϕ and ψ angles at 900 MHz (solid curves) and 600 MHz (broken curves) fields. The assumptions mentioned in the Results part are made. The modification $\bar{H}$ yields a doubled variation of the relaxation rates.

which indicates the relationship between the intensities of the MQ peak components and the projection angles resulting from cross-correlated relaxation of both the dipole-dipole and dipole-CSA interactions involving C^{α} and N spins.

A similar relationship for Gly residues is provided by Equation 13, where only the outer components of the MQ triplets corresponding to both $^1H^{\alpha 1}$ and $^1H^{\alpha 2}$ spins in $\alpha\alpha$ or $\beta\beta$ state are used.

$$\begin{aligned}&\tfrac{4}{3}[A_{\mathrm{D(HC)}}A_{\mathrm{D(HN)}}(Q(\Theta^{1}_{\mathrm{D(HC)D(HN)}})+Q(\Theta^{2}_{\mathrm{D(HC)D(HN)}}))\\&-2A_{\mathrm{D(HC)}}A_{\mathrm{CSA(N)}}B_0(Q(\Theta^{1}_{\mathrm{D(HC)CSA(N)}})+Q(\Theta^{2}_{\mathrm{D(HC)CSA(N)}}))]J(0)=\frac{1}{T_{\mathrm{MQ}}}\ln\left(\frac{I^{\mathrm{DQ}}_{\alpha\alpha}I^{\mathrm{ZQ}}_{\beta\beta}}{I^{\mathrm{DQ}}_{\beta\beta}I^{\mathrm{ZQ}}_{\alpha\alpha}}\right)\end{aligned} \tag{13}$$

All elements of the described experiment are optimized for TROSY relaxation. A graphical representation of Equations 12 and 13 given by Figure 3 shows that the most effective use of the experiment can be achieved at polarizing magnetic fields up to 600 MHz. Since the second terms on the left side of both Equations 12 and 13 depend linearly on B_0, the right side becomes smaller at higher fields. To circumvent this reduction, we propose the use of the same experimental scheme supplemented with $\bar{\mathrm{H}}$. Due to the interchange of the single-transition operators in the middle of the MQ-evolution period, the relaxation term describing the dipole(HC)/dipole(HN) cross-correlation is canceled out, whereas the term stemming from the dipole(HC)/CSA(N) cross-correlation remains unchanged as it is described by Equations 14 and 15.

$$\tfrac{4}{3}[-2A_{\mathrm{D(HC)}}A_{\mathrm{CSA(N)}}B_0Q(\Theta_{\mathrm{D(HC)CSA(N)}})]J(0)=\frac{1}{T_{\mathrm{MQ}}}\ln\left(\frac{I^{\mathrm{DQ}}_{\alpha}I^{\mathrm{ZQ}}_{\beta}}{I^{\mathrm{DQ}}_{\beta}I^{\mathrm{ZQ}}_{\alpha}}\right)\quad , \tag{14}$$

$$\tfrac{4}{3}[-2A_{\mathrm{D(HC)}}A_{\mathrm{CSA(N)}}B_0(Q(\Theta^{1}_{\mathrm{D(HC)CSA(N)}})+Q(\Theta^{2}_{\mathrm{D(HC)CSA(N)}}))]J(0)=\frac{1}{T_{\mathrm{MQ}}}\ln\left(\frac{I^{\mathrm{DQ}}_{\alpha\alpha}I^{\mathrm{ZQ}}_{\beta\beta}}{I^{\mathrm{DQ}}_{\beta\beta}I^{\mathrm{ZQ}}_{\alpha\alpha}}\right). \tag{15}$$

The right-hand sides of the Equations 12 and 14 are compared as functions F of the dihedral angles ϕ and ψ at 600 MHz (dotted curves) and 900 MHz ^{1}H fields (solid curves) in Figure 3. The double inversion of $^1H^N$ at 900 MHz achieved in the experimental scheme of Figure 1 results in a twofold increase in the range of relaxation rates observed for different dihedral angles in comparison with the full TROSY version. In addition, the redundancy of possible dihedral angles deduced from the relaxation rates is reduced to two from four possible angles. Pure dipole(HC)/dipole(HN) cross-correlated relaxation terms can be obtained by subtracting Equation 14 from Equation 12 and Equation 15 from Equation 13, respectively, whereas $\bar{\mathrm{H}}$ increases the amplitude of the variation of the relaxation rates measured at 900 MHz fields. No such benefit is derived at 600 MHz fields so that $\bar{\mathrm{H}}$ can be omitted.

3. Results and Discussion

The experimental schemes of Figure 1 were applied to the 16 kDa uniformly ^{13}C, ^{15}N labeled cytochrome *c* maturation heme chaperone protein E from *E.Coli* (CcmE). The NMR sample contained 350 µl of 1 mM protein solution in 20 mM sodium phosphate buffer at pH = 6.0 containing in addition 300 mM NaCl. The CcmE family of proteins exhibits no significant amino acid sequence homology to any protein of a known 3D structure [28]. The high resolution 3D structure of CcmE determined by solution NMR techniques (PDB code 1LIZ) serves as a representative for this class of proteins. ^{15}N relaxation rates were used to estimate the product of the averaged generalized order parameter and the rotational correlation time of the protein, $\tau_c S^2 = 10 \pm 1$ ns [28].

In order to maximize the spectral sensitivity and facilitate data evaluation, the experiments of Figure 1 were run separately for the ϕ_k and ψ_{k-1} angle estimations using either $\tau_2 = 18$ ms or $\tau_2 = 34$ ms yielding mostly transfer from $^{15}N_k$ to $^{13}C^{\alpha}_k$ or $^{15}N_k$ to $^{13}C^{\alpha}_{k-1}$, respectively. Figure 2 shows the geometry of a fragment of the polypeptide backbone with the nomenclature indicated for atoms and angles involved in the analysis. We assume that for any given polypeptide fragment relaxation of the corresponding multiple quantum coherences excited between the $^{15}N_k$ and $^{13}C^{\alpha}_k$ and $^{15}N_k$ and $^{13}C^{\alpha}_{k-1}$ spins depends only on the values of the ϕ_k and ψ_{k-1} angles. The following projection angles and distances were assumed to be identical for all peptide groups: H^{α}_k-C^{α}_k-N_k = 108.4º and $H^{\alpha 1/2}$k-C^{α}k-N_k = 111º for glycines, H^N_k-N_k-C^{α}_k = 116º, C^{α}_{k-1}-C'_{k-1}-N_k = 116º, C'_{k-1}-N_k-$H^N k$ = 116.5º, H^{α}_k-C^{α}_k-C'_k = 107.2º, r_{HN} = 1.04 x 10^{-10} m [29], r_{HC} = 1.09 x 10^{-10} m [30]. The main axis of the ^{15}N CSA tensor is located in the peptide plane deviating from the HN bond by 20º [31] with the main eigenvalue $\Delta\delta(^{15}N\ CSA) = 170$ ppm [32]. The dihedral angles ϕ_k and ψ_{k-1} are related to the projection angles between the interactions by geometrical considerations. The ratios of the MQ relaxation rates as functions of the ϕ and ψ angles are listed in the Tables 2a-c. The functions are given in units of $S^2\tau_c$ x 10^8/s and therefore are independent of τ_c. The Tables 2a and b show the functions of the Equations 12-15 calculated for B_0 = 600 MHz without and with $\bar{H}$. Table 2c reports the corresponding functions for B_0 = 900 MHz, where $\bar{H}$ is important due to the significantly reduced strength of the combined HC/HN dipole/dipole and HC/N dipole/CSA interactions observed in TROSY spectra recorded without $\bar{H}$.

The logarithms of the ratio of the measured MQ relaxation rates, $F = \ln\left(\frac{I^{DQ}_{\alpha(\alpha)} I^{ZQ}_{\beta(\beta)}}{I^{DQ}_{\beta(\beta)} I^{ZQ}_{\alpha(\alpha)}}\right)$, for the ϕ and ψ angles were obtained from the 3D [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY spectra. In the $\bar{H}$ experiment, both the ratio of the relaxation rates and the experimental $^3J_{H\alpha HN}$ scalar coupling

constants are extracted from the same data set measured with the experimental scheme of Figure 2b. Figures 4a and b show the 2D doublet corresponding to the backbone moiety of Gln72 measured with the experiment [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY-$\bar{\mathrm{H}}$ in the ZQ- and DQ-spectra, respectively. An example of a glycine 2D triplet is given in Figures 4c and d. The 2D triplet of Gly113 is obtained from the [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY experiment.

Table 2. Logarithm of the ratio of peak amplitudes in 3D [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY subspectra as a function of the ϕ and ψ dihedral angles measured at B_0 = 600 MHz (a and b) and B_0 = 900 MHz (c). The functions in (a) correspond to the experimental scheme of Figure 1 applied without $\bar{\mathrm{H}}$ and the functions in (b) and (c) correspond to the scheme applied with $\bar{\mathrm{H}}$ during DQ and ZQ evolution.

(a) $F = \ln\left(I^{DQ}_{\alpha(\alpha)} I^{ZQ}_{\beta(\beta)} / I^{DQ}_{\beta(\beta)} I^{ZQ}_{\alpha(\alpha)}\right) [\tau_c \text{ x } 10^8/\text{s}]$

ϕ	$0.696 + 0.234 \cos(\phi - 60) - 2.327 \cos^2(\phi - 60)$
ψ	$0.677 - 0.410 \cos(\psi - 120) - 2.252 \cos^2(\psi - 120)$
ϕ_{Gly}	$1.416 + 0.251 \cos(\phi - 60) - 2.283 \cos^2(\phi - 60) + 0.251 \cos(\phi + 60) - 2.283 \cos^2(\phi + 60)$
ψ_{Gly}	$1.354 - 0.410 \cos(\psi - 120) - 2.252 \cos^2(\psi - 120) - 0.410 \cos(\psi + 120) - 2.252 \cos^2(\psi + 120)$

(b)

ϕ	$-0.927 - 0.985 \cos(\phi - 60) + 1.429 \cos^2(\phi - 60)$
ψ	$-0.918 + 0.902 \cos(\psi - 120) + 1.138 \cos^2(\psi - 120)$
ϕ_{Gly}	$-1.795 - 1.057 \cos(\phi - 60) + 1.402 \cos^2(\phi - 60) - 1.057 \cos(\phi + 60) + 1.402 \cos^2(\phi + 60)$
Ψ_{Gly}	$-1.836 + 0.902 \cos(\psi - 120) + 1.138 \cos^2(\psi - 120) + 0.902 \cos(\psi + 120) + 1.138 \cos^2(\psi + 120)$

(c)

ϕ	$-1.390 - 1.477 \cos(\phi - 60) + 2.144 \cos^2(\phi - 60)$
ψ	$-1.377 + 1.352 \cos(\psi - 120) + 1.707 \cos^2(\psi - 120)$
ϕ_{Gly}	$-2.692 - 1.585 \cos(\phi - 60) + 2.103 \cos^2(\phi - 60) - 1.585 \cos(\phi + 60) + 2.103 \cos^2(\phi + 60)$
ψ_{Gly}	$-2.754 + 1.352 \cos(\psi - 120) + 1.707 \cos^2(\psi - 120) + 1.352 \cos(\psi + 120) + 1.707 \cos^2(\psi + 120)$

Figures 5a-e show the logarithms of the ratio of the measured MQ relaxation rates, F, versus the values of the ϕ angles averaged over 20 NMR conformers of the structurally well defined core of CcmE (residues 34-132) along with the theoretical curves calculated with Equations 12-15.

In general, a satisfactory correlation between the experimental relaxation data and the angle values calculated from the NMR structure of CcmE is observed. The error propagation of the signal-to-

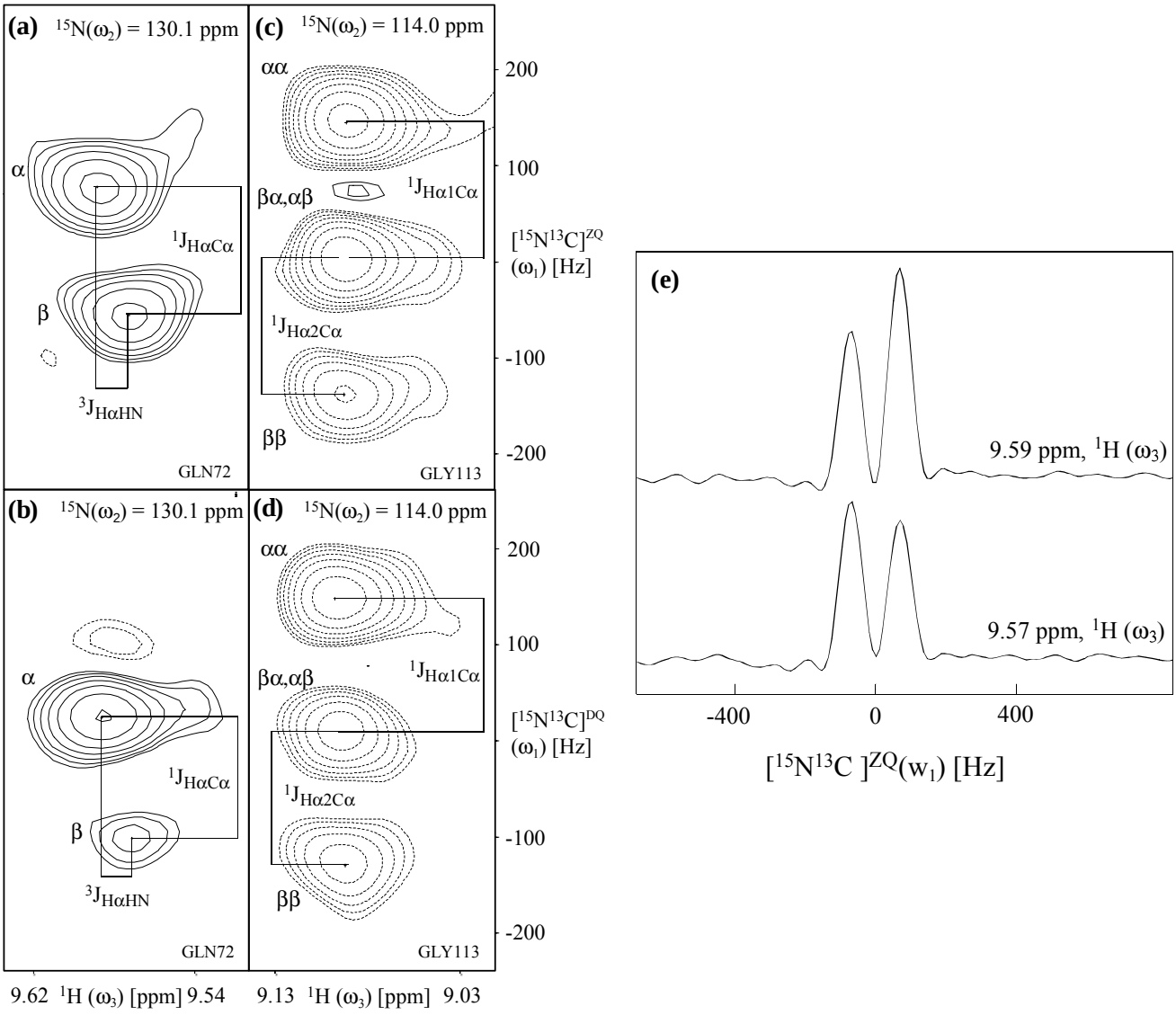


Figure 4. [^{1}H,^{15}N/^{13}C-MQ] planes of selected cross-peaks from the 3D [^{15}N,^{13}C]-DQ/ZQ- [^{15}N,^{1}H]-TROSY-E.COSY spectra of a sample of ^{15}N/^{13}C-labeled apo-CcmE$^{30\text{-}159}$-H$_6$ [28] used for the ϕ angle determination. Spectra of the experiment employing selective 90° E-Burp pulses with $\overline{H}$ are shown in (a) and (b). The peak doublets of Gln72 result from the $J_{H\alpha C\alpha}$ splitting in the ω_{MQ} dimension. Due to the E.COSY effect the peaks are shifted in the ω_H dimension by $^3J_{H\alpha HN}$ = 9.3 Hz and $^3J_{H\alpha HN}$ = 8.2 Hz in (a) and (b), respectively. Spectra of the experiment using selective 90° Gaussian pulses without $\overline{H}$ are shown in (c) and (d). The peak triplet of Gly113 results from the $J_{H\alpha 1C\alpha}$ and $J_{H\alpha 2C\alpha}$ splittings in the ω_{MQ} dimension. For the relaxation calculations, the intensities of the outer peaks corresponding to the αα and ββ states are used. The inner peaks are superpositions of the peaks corresponding to the βα and αβ states. The experiments were performed on a Bruker *Avance* 600 MHz spectrometer. The 16 complex interferograms consisting of 60 x 30 x 512 data points along the ω_1, ω_2 and ω_3 dimensions are collected and processed with PROSA [33] as described in Methods. The experiments using selective Gaussian pulses (E-Burp pulses) were acquired with an interscan delay of 850 ms (800 ms) and 16 (4) scans per datapoint resulting in a total acquisition time of 130 (31) hours. (e) shows 1D [^{15}N,^{13}C]-ZQ (ω_1) slices taken at the positions of the maximum intensity of the cross-peaks in (a) to demonstrate the quality of the acquired spectra.

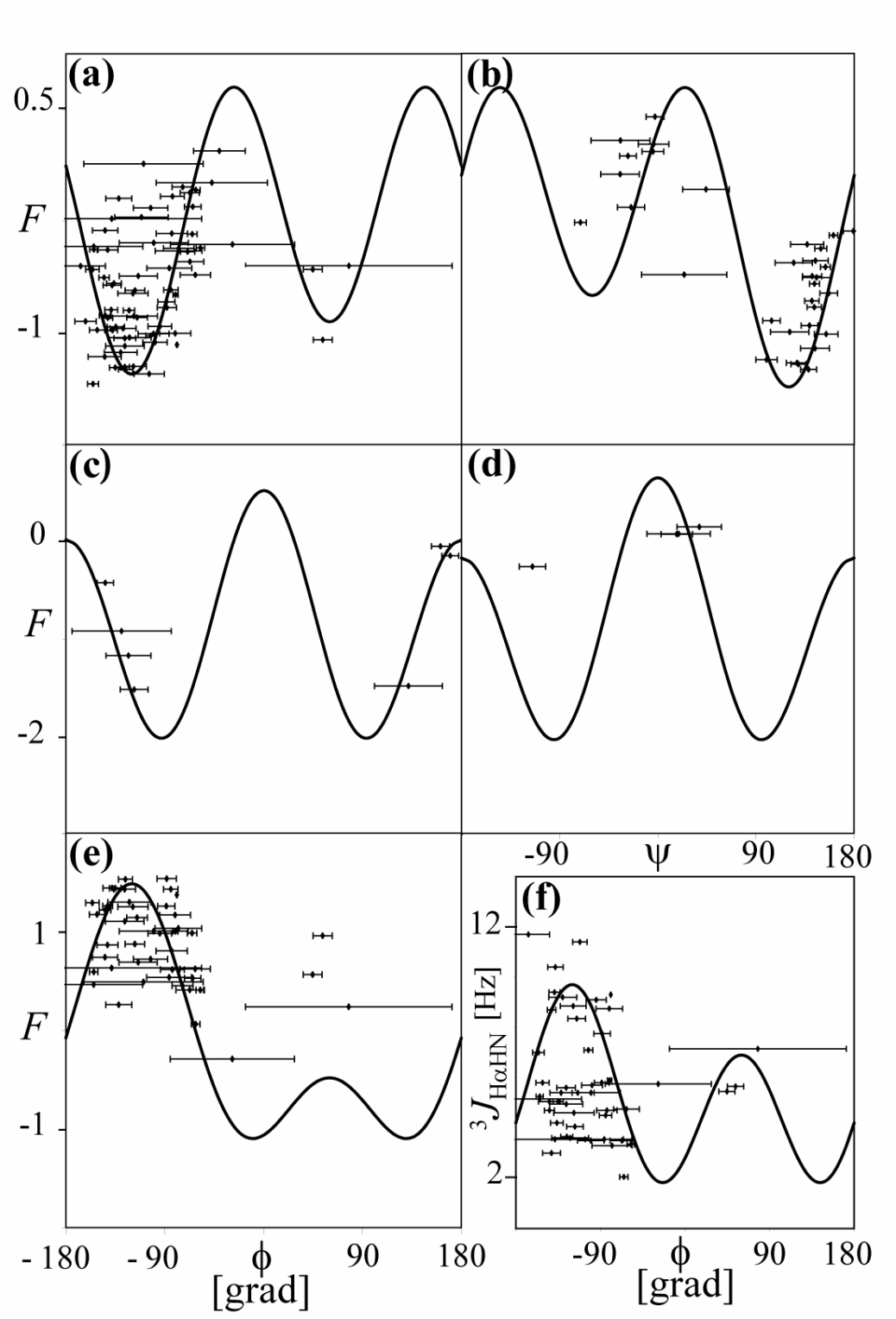


Figure 5. Experimental values of $F = \ln\left(I^{DQ}_{\alpha(\alpha)} I^{ZQ}_{\beta(\beta)} / I^{DQ}_{\beta(\beta)} I^{ZQ}_{\alpha(\alpha)}\right)$ measured at B_0 = 600 MHz vs. mean angles derived from the set of the 20 conformers representing the solution structure of CcmE [28]. (a) F vs. ϕ for all amino acids except for Gly, (b) F vs. ψ for all amino acids except for Gly, (c) F vs. ϕ for Gly, (d) F vs. ψ for Gly, (e) F vs. ϕ for all amino acids except for Gly with $\overline{\overline{H}}$ (see Methods section), (f) $^3J_{H\alpha HN}$ vs. ϕ .The solid curves represent theoretical values. The error bars show the distribution of the angle within the 20 NMR conformers.

noise ratio in the [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY spectra used to estimate the ϕ angles typically yields an error of 6% for the F values. Statistically significant deviations of the experimental values of F from the theoretical curves of Equations 12-15 are attributed to an absence of traditional structural constraints which restrain the corresponding ϕ and ψ angles in the process of NMR structure calculations. This situation was detected for the ϕ angle of the residues 40, 44, 56, 58, 70, 76, 78, 88, 109, 120, 128, 134, 135 and 142 and for the ψ angle of the residues 45, 55, 61, 70, 81, 88, 102, 110, 113, 119 and 126. Structural refinement using cross-correlated relaxation data serves as a useful complementary tool to the conventional NOEs based constraints in loop regions of the protein structure [12].

The values of the $^{3,4}J_{\mathrm{H}\alpha\mathrm{HN}}$ can be independently obtained from both ZQ- and DQ-subspectra of the [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY experiment. This fact can be used to reduce the effects of spectral artifacts or peak overlap and in order to calculate statistical uncertainties. In the case of glycine, only the sum of $^{3,4}J_{\mathrm{H}\alpha1\mathrm{HN}}$ and $^{3,4}J_{\mathrm{H}\alpha2\mathrm{HN}}$ can be measured, since the inner peak of the triplet is a superposition of the peaks corresponding to the αβ and βα states slightly shifted relatively to each one another by 2 x ($^{3,4}J_{\mathrm{H}\alpha1\mathrm{HN}}$ – $^{3,4}J_{\mathrm{H}\alpha2\mathrm{HN}}$). The values of the $^{3}J_{\mathrm{H}\alpha\mathrm{HN}}$ coupling constants are usually analysed solely as a function of the ϕ angle using the Karplus-type dependencies while a structural parametrization for the $^{4}J_{\mathrm{H}\alpha\mathrm{HN}}$ coupling constants is not yet available [21]. Figure 5f correlates measured $^{3}J_{\mathrm{H}\alpha\mathrm{HN}}$ scalar coupling constants with the theoretical Karplus-type curve [1]. Due to transverse relaxation of the $^{1}\mathrm{H}^{\alpha}$ spins in the ST2-PT element, the apparant values of the $^{3,4}J_{\mathrm{H}\alpha\mathrm{HN}}$ scalar couplings smaller than 2 Hz are significantly biased by cross-talk between the α and β states of the $^{1}\mathrm{H}^{\alpha}$ spin [34]. As a result only $^{3}J_{\mathrm{H}\alpha\mathrm{HN}}$ scalar couplings stronger than 3 Hz match well with the angular constraints derived from the relaxation data. We attempted to minimize mixing of the α and β states of the $^{1}\mathrm{H}^{\alpha}$ spin in the ST2-PT by introducing $^{1}\mathrm{H}^{\alpha}$ band selective 90º pulses, which in combination with the subsequent non-selective 90º ^{1}H hard pulses quantitatively flip the $^{1}\mathrm{H}^{\alpha}$ magnetization to the +z axis (Figure 1b). Nonetheless, the experimental values of the scalar coupling constants smaller than 2-4 Hz tend to be systematically underestimated. With the increased transverse relaxation of $^{1}\mathrm{H}^{\alpha}$ spins usually found in larger proteins, the systematic errors stemming from the cross-talk problem result in an underestimation of even larger scalar couplings. The positive and negative values of the ϕ angles can be most effectively discriminated by combining cross-correlated relaxation rates obtained from the the [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY experiment run with and without the $\bar{\mathrm{H}}$ element (see Figures 3 and 5e). The obtained values of the ϕ angles can be then cross-validated using the $^{3}J_{\mathrm{H}\alpha\mathrm{HN}}$ scalar coupling constants.

Thus, the TROSY principle has been introduced into a modified MQ-E.COSY-HNCA experiment designed for measurements of the intraresidual and sequential H^{α}-C^{α}/H^{N}-N dipole/dipole and H^{α}-

C^{α}/N dipole/CSA cross-correlated relaxation rates, which in addition provides values of the $^{3,4}J_{H\alpha HN}$-coupling constants measured in an E.COSY manner. The restraints for backbone ψ and ϕ angles are derived from a ratio of the relaxation rates of ZQ and DQ coherencies excited between the $^{13}C^{\alpha}$ and ^{15}N spins and the Karplus-type dependencies of the coupling constants measured from the corresponding $^{1}H^{\alpha}$ spin state selective [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY spectra reconstructed from a single data set. A comparison of the measured values of the ψ and ϕ angles of the 16 kDa CcmE protein with the average values determined from the NMR solution structure by traditional methods reveals an overall match between directly measured ψ and ϕ angles with the corresponding angles found in the set of 20 NMR conformers.

We conclude that the values of dihedral angles calculated from the relaxation data are more reliable than the ones derived from the scalar coupling alone. In addition, the comparison indicates a necessity to refine NMR structures by including cross-correlated relaxation-based constraints in the structure determination protocol.

4. Appendix

The relaxation rate during t_1 is calculated for the four-spin system H^{C}-C-N-H^{N} using the following basis in the master equation [35]:

$$\begin{pmatrix} ZQ_1 \\ ZQ_2 \\ ZQ_3 \\ ZQ_4 \\ ZQ_5 \\ ZQ_6 \\ ZQ_7 \\ ZQ_8 \\ DQ_1 \\ DQ_2 \\ DQ_3 \\ DQ_4 \\ DQ_5 \\ DQ_6 \\ DQ_7 \\ DQ_8 \end{pmatrix} = \begin{pmatrix} C_{-}N_{+}(E-2H_Z^C)(E-2H_Z^N) \\ C_{+}N_{-}(E-2H_Z^C)(E-2H_Z^N) \\ C_{-}N_{+}(E+2H_Z^C)(E-2H_Z^N) \\ C_{+}N_{-}(E+2H_Z^C)(E-2H_Z^N) \\ C_{-}N_{+}(E-2H_Z^C)(E+2H_Z^N) \\ C_{+}N_{-}(E-2H_Z^C)(E+2H_Z^N) \\ C_{-}N_{+}(E+2H_Z^C)(E+2H_Z^N) \\ C_{+}N_{-}(E+2H_Z^C)(E+2H_Z^N) \\ C_{+}N_{+}(E-2H_Z^C)(E-2H_Z^N) \\ C_{-}N_{-}(E-2H_Z^C)(E-2H_Z^N) \\ C_{+}N_{+}(E+2H_Z^C)(E-2H_Z^N) \\ C_{-}N_{-}(E+2H_Z^C)(E-2H_Z^N) \\ C_{+}N_{+}(E-2H_Z^C)(E+2H_Z^N) \\ C_{-}N_{-}(E-2H_Z^C)(E+2H_Z^N) \\ C_{+}N_{+}(E+2H_Z^C)(E+2H_Z^N) \\ C_{-}N_{-}(E+2H_Z^C)(E+2H_Z^N) \end{pmatrix}. \quad (16)$$

The relaxation matrix has a pure diagonal form. The DD interactions of the H^{α}-C^{α}, H^{N}-N, N-C^{α} and H^{N}-H^{α} spin pairs, the CSA of the H^{α}, C, N and H^{N} spins, and all cross-correlation terms are

taken into account. In the slow tumbling regime we retain terms in $J(0)$ only, the exception are terms in the difference and sum of two proton frequencies.

5. References

[1] Wüthrich, K. *NMR of Proteins and Nucleic Acid.* **1986**, Wiley, New York.
[2] Ernst, M.; Ernst, R.R. *J. Magn. Reson. Ser. A* **1994**, *110*, 202-213.
[3] Reif, B.; Hennig, M.; Griesinger, C. *Science* **1997**, *276*, 1230-1233.
[4] Yang, D.W.; Konrat, R.; Kay, L.E. *J. Am. Chem. Soc.* **1997**, *119*, 11938-11940.
[5] Yang, D.W.; Gardner, K.H.; Kay, L.E. *J. Biomol. NMR* **1998**, *11*, 213-220.
[6] Yang, D.W.; Kay, L.E. *J. Am. Chem. Soc.* **1998**, 1*20*, 9880-9887.
[7] Pelupessy, P.; Chiarparin, E.; Ghose, R.; Bodenhausen, G. *J. Biomol. NMR* **1999**, *13*, 375-380.
[8] Chiarparin, E.; Pelupessy, P.; Ghose, R.; Bodenhausen, G. *J. Am. Chem. Soc.* **1999**, *121*, 6876-6883.
[9] Pelupessy, P.; Chiarparin, E.; Ghose, R.; Bodenhausen, G. *J. Biomol. NMR* **2002**, *14*, 277-280.
[10] Pang, Y.X.; Wang, L.C.; Pellecchia, M.; Kurochkin, A.V.; Zuiderweg, E.R.P. *J. Biomol. NMR.* **1999**, *14*, 297-306.
[11] Reif, B.; Diener, A.; Hennig, M.; Maurer, M.; Griesinger, C. *J. Magn. Reson.* **2000**, *143*, 45-68.
[12] Sprangers, R.; Bottomley, M.J.; Linge, J.P.; Schultz, J.; Nilges, M.; Sattler, M. *J. Biomol. NMR* **2000**, *16*, 47-58.
[13] Kloiber, K.; Konrat, R. *J. Biomol. NMR* **2000**, *17*, 265-268.
[14] Kloiber, K. and Konrat, R. *J. Am. Chem. Soc.* **2000**, *122*, 12033-12034.
[15] Kloiber, K.; Schuler, W.; Konrat, R. *J. Biomol. NMR* **2002**, *22*, 349-363.
[16] Griesinger, C.; Sørensen, O.W.; Ernst, R.R. *J. Magn. Reson.* **1987**, *75*, 474-492.
[17] Wagner, G.; Schmieder, P.; Thanabal, V. *J. Magn. Reson.* **2000**, *93*, 436-440.
[18] Weisemann, R.; Ruterjans, H.; Schwalbe, H.; Schleucher, J.; Bermel, W.; Griesinger, C. *J. Biomol. NMR* **1994**, *4*, 231-240.
[19] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A* . **1997**, *94*, 12366-12371.
[20] Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11-15.
[21] Vuister, G.W.; Bax, A. *J. Biomol. NMR* **1994**, *4*, 193-200.
[22] Sørensen, O.W.; Eich, G.W.; Levitt, M.H.; Bodenhausen, G.; Ernst, R.R. *Prog. Nucl. Magn. Reson. Spectrosc.* **1983**, *16*, 163-192.
[23] Geen, H.; Freeman, R. *J. Magn. Reson.* **1991**, *93*, 93-141.
[24] McCoy, M.A.; Mueller, L. *J. Am. Chem. Soc.* **1992**, *114*, 2108-2112.
[25] Korzhnev, D. M.; Billeter, M.; Arseniev, A. S.; Orekhov, V.Y. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *38*, 197-266.
[26] Pervushin, K. *Quart. Rev. Biophys.* **2000**, *33*, 161-197.
[27] Pervushin, K.; Wider, G.; Wüthrich, K. *J. Biomol. NMR* **1998**, *12*, 345-348.
[28] Enggist, E.; Thöny-Meyer, L.; Güntert, P.; Pervushin, K. *Structure* **2002**, *10*, 1551-1557.
[29] Roberts, J.E.; Harbison, G.S.; Munowitz, M.G.; Herzfeld, J.; Griffin, R.G. *J. Am. Chem. Soc.* **1987**, *109*, 4163-4169.
[30] Case, D.A. *J. Biomol. NMR* **1999**, *15*, 95-102.
[31] Oas, T.G.; Hartzell, C.J.; Dahlquist, F.W.; Drobny, G.P. *J. Am. Chem. Soc.* **1987**, *109*, 5962- 5966.
[32] Tjandra, N.; Szabo, A.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6986-6991.
[33] Güntert, P., Dotsch, V., Wider, G. and Wüthrich, K. *J. Biomol. NMR* **1992**, *2*, 619-629.
[34] Wang, A.C.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 2483-2494.
[35] Abragam, A. *The Principles of Nuclear Magnetism.* **1961**, *17*, Oxford, Clarendon.

Chapter IV

Simultaneous ^{1}H or ^{2}H, ^{15}N and Multiple-Band-Selective ^{13}C Decoupling During Acquisition in ^{13}C-Detected Experiments with Proteins and Oligonucleotides

Beat Vögeli[1], Helena Kovacs[2] and Konstantin Pervushin[1]

[1]Laboratorium für Physikalische Chemie, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

[2]Bruker BioSpin AG, Industriestrasse 26, CH-8117 Fällanden, Switzerland.

Journal of Biomolecular NMR, 2005, 31, 1-9

1. Introduction

Novel strategies based on ^{13}C detection of the backbone and side-chain resonances hold promise for expanding NMR structural analysis to large uniformly deuterated proteins [1,2]. We have combined ^{13}C- and ^{1}H-detected experiments to obtain the backbone resonance assignment of a 44 kDa protein [2,3] and designed an approach based on ^{13}C-start, ^{13}C-observe experiments to get a nearly complete assignment of the side-chain ^{13}C resonances of this protein [1]. We have also demonstrated that the ^{13}C-start, ^{13}C-observe experiments can be used to determine multiple side-chain residual ^{13}C-^{13}C dipolar couplings upon partial alignment of a protein [4] (See Chapter VI). For these studies, a cryogenic probe optimized for ^{13}C detection was used at 500 MHz. The applicability of these methods is, however, limited by spectral overlap due to the complex splitting patterns of the cross-peaks, which leads to difficulties in assigning the peaks and in extracting single $J_{CC} + D_{CC}$ couplings from them, and to a poor signal-to-noise ratio for resonances at large offsets. Similarly, the application of ^{13}C-detected methods to oligonucleotides is compromised by homonuclear couplings particularly in the crowded ribose region. Homodecoupled ^{13}C-detected methods are potentially a useful complement for the resonance assignment of labeled RNA and DNA, particularly in the sugar region where the ^{1}H chemical shift dispersion is low. Bertini and coworkers recently demonstrated band-selective homodecoupling of ^{13}CO from $^{13}C^{\alpha}$ and vice versa in the COCAMQ and CT-COSY experiments at 500 MHz [5]. Earlier Matsuo et al. have implemented band-selective ^{13}C decoupling of ^{13}CO and $^{13}C^{\beta}$ during ^{13}C evolution in the Cbd-HCCH-TOCSY experiment for proteins [6]. Brutscher et al. have applied band-selective ^{13}C decoupling on C2', C5' and low-field aromatic-^{13}C during the ^{13}C evolution in 2D [^{1}H,^{13}C]-TROSY and 3D [^{13}C,^{13}C]-NOESY-[^{1}H,^{13}C]-TROSY experiments on an RNA sample [7]. In this paper, we show that the above mentioned obstacles can be alleviated by employing band-selective ^{13}C homodecoupling, simultaneously with ^{1}H or ^{2}H, and ^{15}N decoupling during all stages of multidimensional experiments including the signal acquisition period. Thus, the approach of ^{13}C detection becomes attractive for biomolecular NMR applications, particularly, if it can be combined with the gains in sensitivity and resolution substantiated by cryogenic probes and high magnetic fields as is shown by the present work.

2. Results

Performance of the band-selective adiabatic inversion pulses was initially tested in 1D ^{1}H- or ^{2}H-decoupled experiments. The experimental inversion profiles are shown in Figures 1 and 2 for the

protein and RNA spectra, respectively. In each case, a 1D spectrum lacking a ^{13}C inversion pulse is shown on the top. The calculated excitation profiles for the double- and triple-band pulses are also presented in Figures 1c and d, respectively. The single-band carbonyl ^{13}C-, double-band $^{13}C^{\alpha}/^{13}C^{\gamma}$- and triple-band $^{13}CO/^{13}C^{\beta}$-inversion pulses for the protein, displayed in Figures 1b, 1c and 1d, and the single-band C2′/C3′- and C4′-selective pulses for the RNA-riboses, displayed in Figures 2b and 2c, all feature high resonance selectivity. These band-selective inversion pulses were subsequently employed in the homodecoupling trains during acquisition [8] in the 2D [^{13}C,^{13}C]-TOCSY experiment [1], depicted in Figure 3. This is, in fact, the most sensitive [^{13}C,^{13}C]-correlation experiment for molecules of current size [9]. The p5m4 supercycle [10,11] with a homodecoupling duty cycle of 20% was used in all cases, and the rf field strength of the adiabatic pulses was increased accordingly.

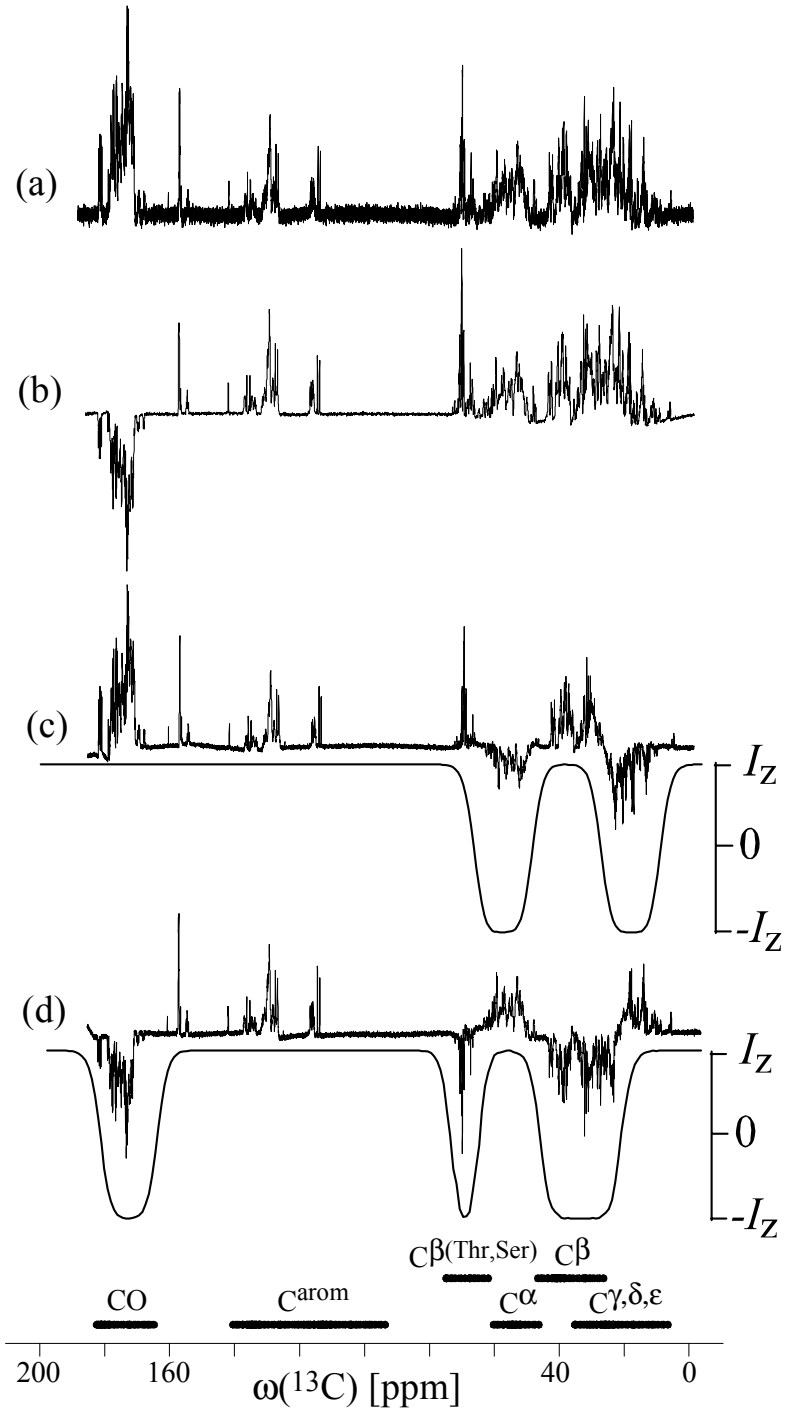


Figure 1. 1D ^{13}C spectra of a 0.4 mM uniformly ^{2}H,^{15}N,^{13}C-labeled ubiquitin sample measured using a Bruker *Avance* 900 spectrometer operating at $\gamma B_0(^{1}\mathrm{H})$ = 900 MHz. The spectra were recorded using WALTZ16 ^{2}H decoupling at $\gamma B_1(^{2}\mathrm{H})$ = 0.83 kHz during data acquisition. The spectra are superimposed with the inversion profiles of I_z magnetization calculated with the program NMRSIM (Bruker BioSpin). (a) Reference spectrum recorded using a single broadband ^{13}C 90^{o} excitation pulse. (b) 4 ms CO-band-selective inversion CHIRP [12] pulse with $\gamma B_1(^{13}\mathrm{C})$ = 1.26 kHz inverting resonances in the CO region, (c) double-band-selective CHIRP pulse of 4 ms length with $\gamma B_1(^{13}\mathrm{C})$ = 1.75 kHz inverting the C^{α} and C^{γ} regions and, (d) triple-band-selective CHIRP pulse of 4 ms length with $\gamma B_1(^{13}\mathrm{C})$ = 3.51 kHz inverting the CO and two C^{β} regions, were applied prior to the 90^{o} excitation pulse. The spectra (a) and (b)-(d) were recorded using 132 and 1024 scans, respectively, with 1s interscan delay.

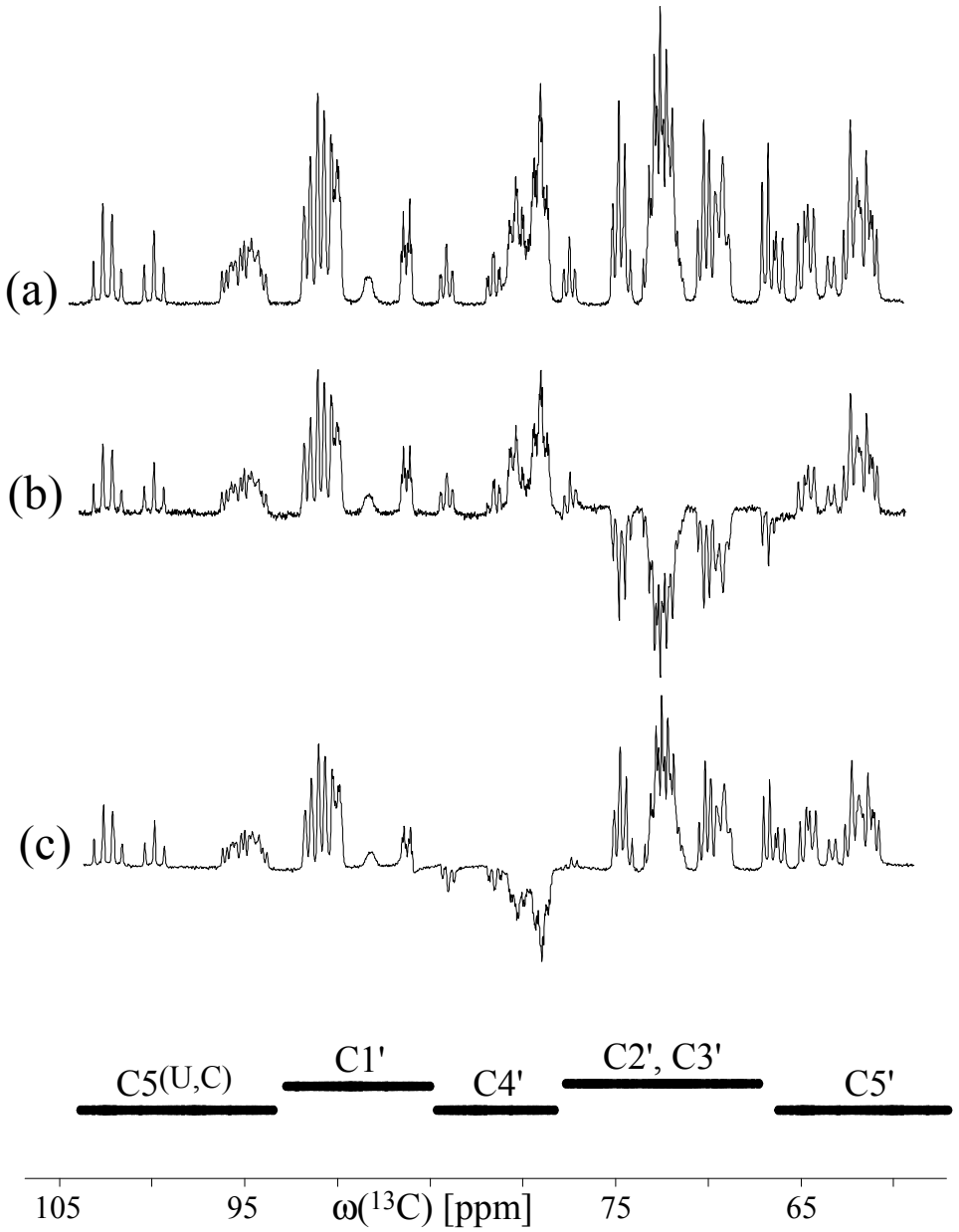


Figure 2. 1D ^{13}C spectra of a ^{15}N,^{13}C-labeled RNA 14-mer, 5′ GGCACUUCGGUGCC 3′, measured using a Bruker *Avance* 500 spectrometer operating at $\gamma B_0(^1\text{H})$ = 500 MHz. The spectra were recorded using WALTZ-16 ^{1}H decoupling at $\gamma B_1(^1\text{H})$ = 2.50 kHz during data acquisition. (a) Reference spectrum recorded using a single broadband ^{13}C 90^{o} excitation pulse. (b) 15 ms band-selective inversion CHIRP pulse with $\gamma B_1(^{13}\text{C})$ = 0.28 kHz inverting resonances in the C2' and C3' region, and, (c) band-selective CHIRP pulse of 20 ms length with $\gamma B_1(^{13}\text{C})$ = 0.21 kHz inverting the C4' region, were applied prior to the 90^{o} excitation pulse. The spectra (a), (b) and (c) were recorded using 512, 322 and 242 scans and 2s, 2s and 1s interscan delay, respectively.

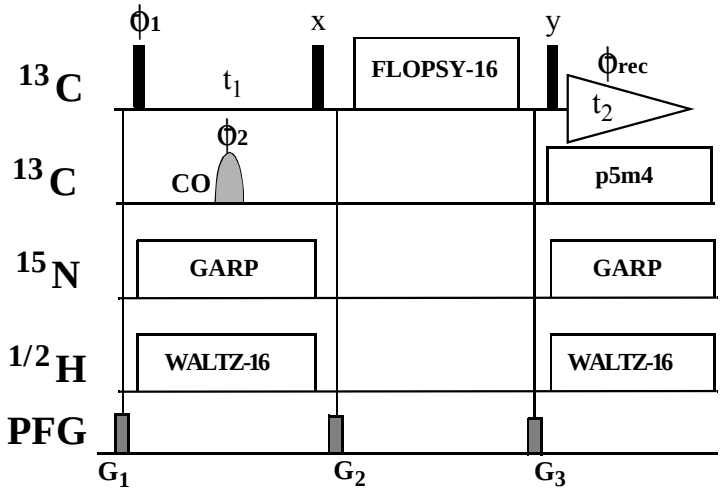
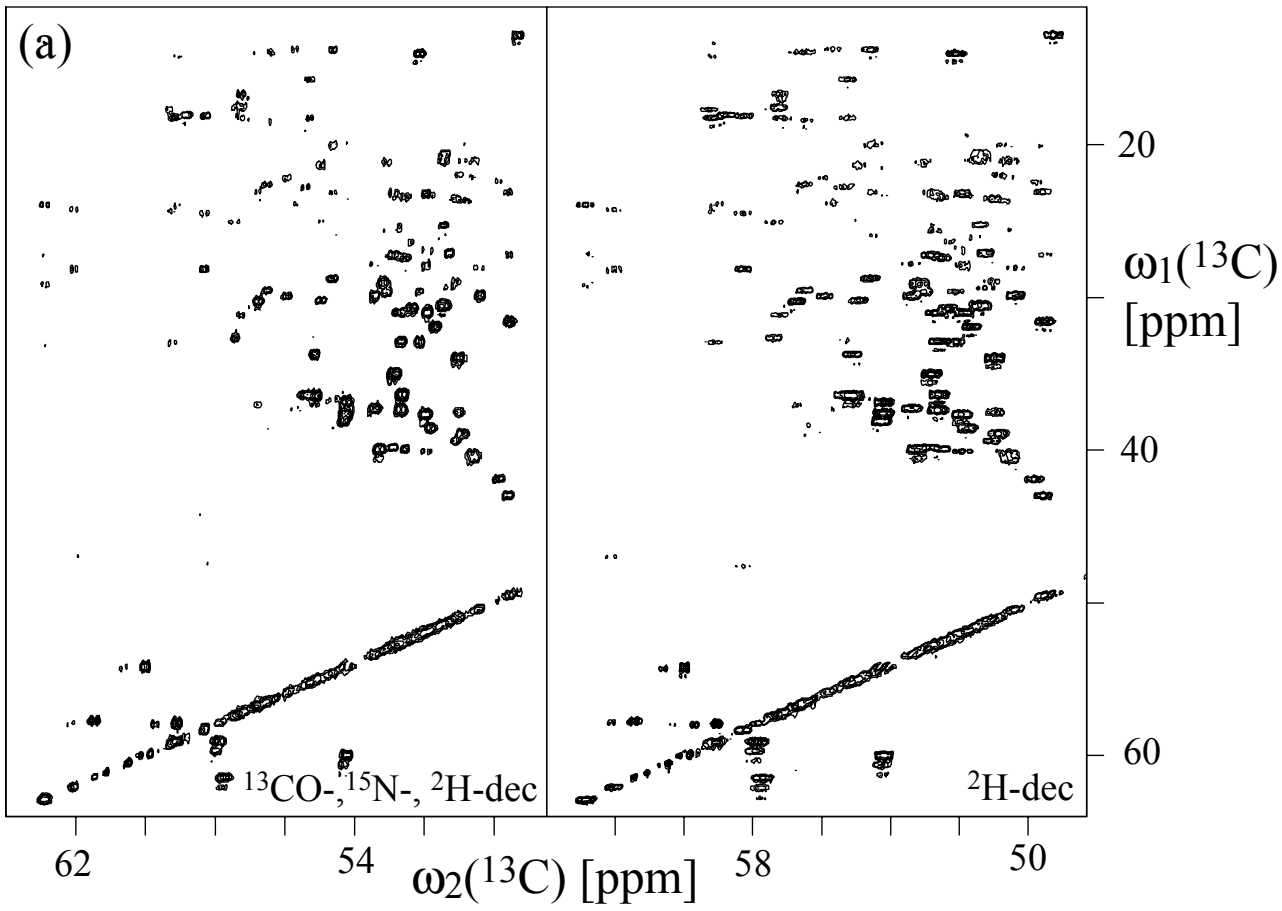
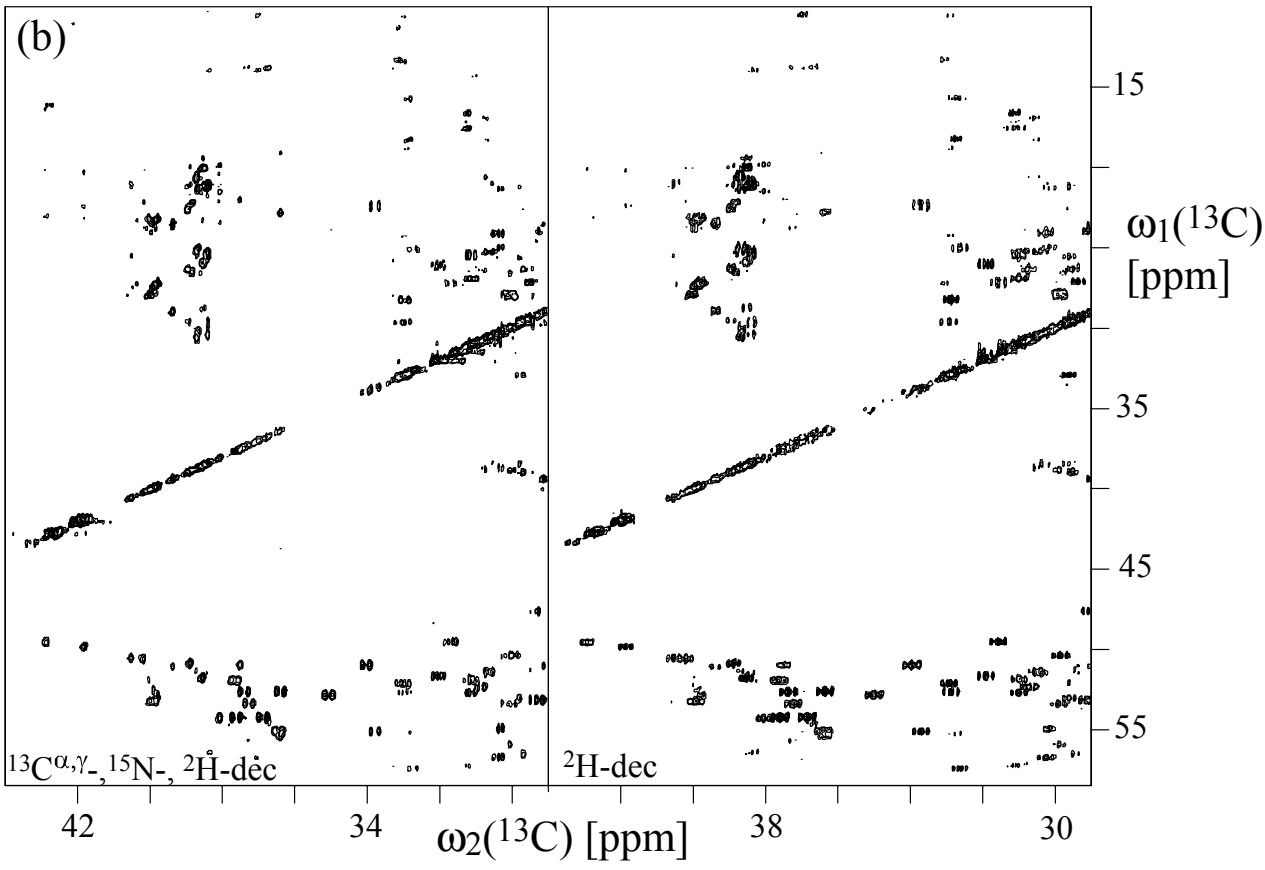


Figure 3. Experimental scheme of the ^{2}H,^{15}N,^{13}C(selective)-decoupled [^{13}C,^{13}C]-TOCSY. Black bars represent broadband ^{13}C 90^{o} pulses. The curved shape indicates a selective, 5%-truncated Gaussian 180^{o} pulse on ^{13}CO with a duration of 100 μsec. The FLOPSY-16 mixing was applied with $\gamma B_1(^{13}C)$ = 8.3 kHz with τ_{mix} = 16.96 ms at 900 MHz and with $\gamma B_1(^{13}C)$ = 10.0 kHz with τ_{mix} = 18.84 ms at 500 MHz. The boxes on the line labeled PFG indicate sine-shaped pulsed magnetic field gradients along the z-axis : G1 and G3, 25 G cm^{-1}, 1 ms; G2, 30 G cm^{-1}, 0.6 ms. The phase cycle is ϕ_1 = {x, -x, x, -x}; ϕ_2 = { x, x, -x, -x}; ϕ_{rec} = {x, -x, x, -x}. ^{2}H- and ^{15}N-decoupling is performed with WALTZ-16 [13] and GARP4 [14] pulse trains, respectively. The homonuclear decoupling employs a p5m4 supercycle [10,11] consisting of single ^{13}CO-, double $^{13}C^{\alpha}/^{13}C^{\gamma}$- or triple $^{13}CO/^{13}C^{\beta}$-band-selective pulses using CHIRP inversion pulses as described in Figures 1 and 2, with the exception that all rf-power settings were increased by 14 dB to account for the 20% homodecoupling duty cycle.

Figure 4 shows relevant expansions of 2D [^{13}C,^{13}C]-TOCSY spectra of human ubiquitin [15]. The different options of band-selective homonuclear ^{13}C decoupling, all of which were applied together with ^{2}H and ^{15}N decoupling at 900 MHz, are compared to a reference spectrum recorded with ^{2}H decoupling only. The single-band-selective ^{13}CO decoupling collapses the carbonyl-coupling for all $^{13}C^{\alpha}$ cross-peaks (Figure 4a) and the sidechain-carbonyl coupling in asparagines and glutamines. Figures 5c, 5d and 5e show representative slices for C^{α}/C^{β} of ASN60, C^{β}/C^{α} of GLN62 and C^{β}/C^{α} of THR22, respectively. The performance of the double-band-selective $^{13}C^{\alpha}/^{13}C^{\gamma}$-decoupling is slighty less clean (Figure 4b). Most $^{13}C^{\beta}$ and $^{13}C^{\delta}$ cross-peaks turn into singlets, as illustrated by Figures 5a and 5b displaying C^{β}/C^{δ} of ARG54 and C^{α}/C^{β} of LEU50, respectively. Figure 5a also depicts the collapse of the coupling from the side-chain ^{15}N in arginine. A minor fraction reveals reduced doublet structure although the coupled spins resonate within the decoupled frequency range, for instance the cross-peak C^{γ}/C^{β} of LEU50. This is even more the case for the triple-band-selective decoupling of $^{13}CO/^{13}C^{\beta}$ (Figure 4c). For instance, the cross-peak C^{β}/C^{α} of THR22, which becomes a singlet, in contrast to C^{γ}/C^{α} of GLN62, which remains a reduced doublet, as shown in Figures 5e and 5d, respectively. Most rewardingly, there are also some cross-peaks such as C^{γ}/C^{β} of LEU50, which become visible only upon $^{13}C^{\alpha}/^{13}C^{\gamma}/^{15}$N-decoupling (Figure 5f).


(a)
^{13}CO-,^{15}N-, ^{2}H-dec
^{2}H-dec
20
40
60
$\omega_1(^{13}C)$ [ppm]
62
54
58
50
$\omega_2(^{13}C)$ [ppm]


(b)
$^{13}C^{\alpha,\gamma}$-, ^{15}N-, ^{2}H-dec
^{2}H-dec
15
35
45
55
$\omega_1(^{13}C)$ [ppm]
42
34
38
30
$\omega_2(^{13}C)$ [ppm]

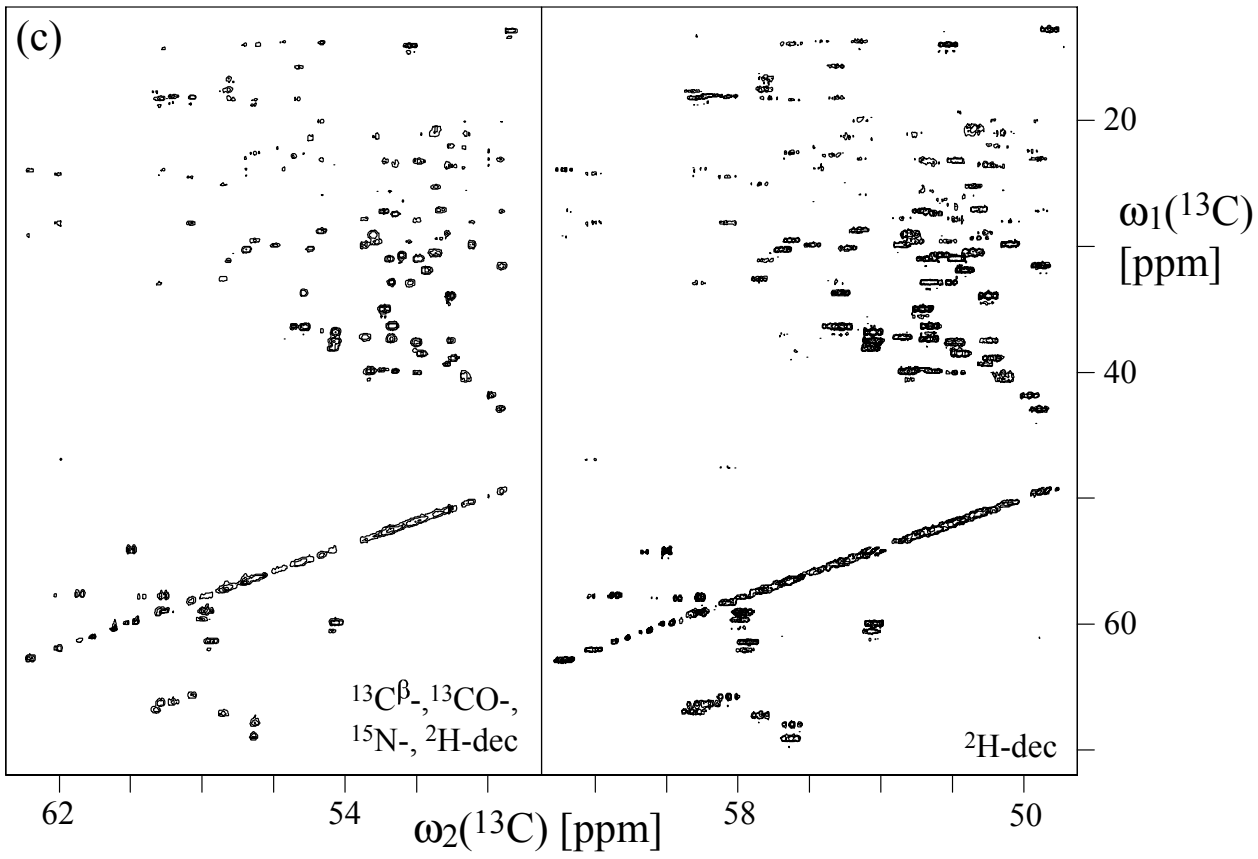


Figure 4 (previous and present pages). Expansions of 2D [^{13}C,^{13}C]-TOCSY spectra of uniformly ^{2}H,^{15}N,^{13}C-labeled ubiquitin samples measured at a Bruker *Avance* spectrometer operating at $\gamma B_0(^1\mathrm{H})$ = 900 MHz using ^{2}H,^{15}N,^{13}C-tunable ^{13}C-detection probe (DUX). The right panels show expansions of a reference spectrum with only ^{2}H decoupling using WALTZ-16 pulse trains at γB_1 = 0.96 kHz, $t_{1\mathrm{max}}$ = 11.0 ms and $t_{2\mathrm{max}}$ = 147.5 ms and 400 x 4096 complex points resulting in an acquisition time of 20 hours. The left panels show expansions of the multiple-band decoupled spectra: (a) ^{2}H decoupling using WALTZ-16 pulse trains at γB_1 = 0.83 kHz, ^{15}N decoupling using GARP pulse trains at γB_1 = 1.3 kHz, ^{13}C-homonuclear decoupling of CO using a p5m4 supercycle with an adiabatic CHIRP pulse of 4 ms length and 20% smoothing according to Figure 1b, $t_{1\mathrm{max}}$ = 7.1 ms and $t_{2\mathrm{max}}$ = 147.5 ms, 256 x 8192 complex points were collected resulting in an acquisition time of 15 hours. (b) ^{2}H decoupling using WALTZ-16 pulse trains at γB_1 = 0.96 kHz, ^{15}N decoupling using GARP pulse trains at γB_1 = 1.3 kHz, ^{13}C-homonuclear decoupling of C^{α} and C^{γ} using a p5m4 supercycle with a double-band-selective CHIRP pulse of 4 ms length and 20% smoothing according to Figure 1c, $t_{1\mathrm{max}}$ = 11.0 ms and $t_{2\mathrm{max}}$ = 147.5 ms, 400 x 8192 complex points were recorded resulting in an acquisition time of 24 hours. (c) ^{2}H decoupling using WALTZ-16 pulse trains at γB_1 = 0.96 kHz, ^{15}N decoupling using GARP pulse trains at γB_1 = 1.3 kHz, ^{13}C-homonuclear decoupling of CO and C^{β} using a p5m4 supercycle with a triple-band-selective adiabatic CHIRP pulse of 4 ms length and 20% smoothing according to Figure 1d, $t_{1\mathrm{max}}$ = 11.0 ms and $t_{2\mathrm{max}}$ = 147.5 ms, 400 x 8192 complex points were collected resulting in an acquisition time of 35 hours. The ^{13}C homodecoupling was applied with 20% duty cycle in all cases. In all experiments the interscan delay was 1.5 s, the radio-frequency carrier offsets were placed at 40 ppm (^{13}C), 174 ppm (^{13}CO), 120.0 ppm (^{15}N) and 3.0 ppm (^{2}H) and the time domain data were multiplied with a cosine function in the t_1 and t_2 dimensions and zero-filled to 1024 and 16384 points, respectively. 1D slices of six representative cross-peaks are shown in Figure 5.

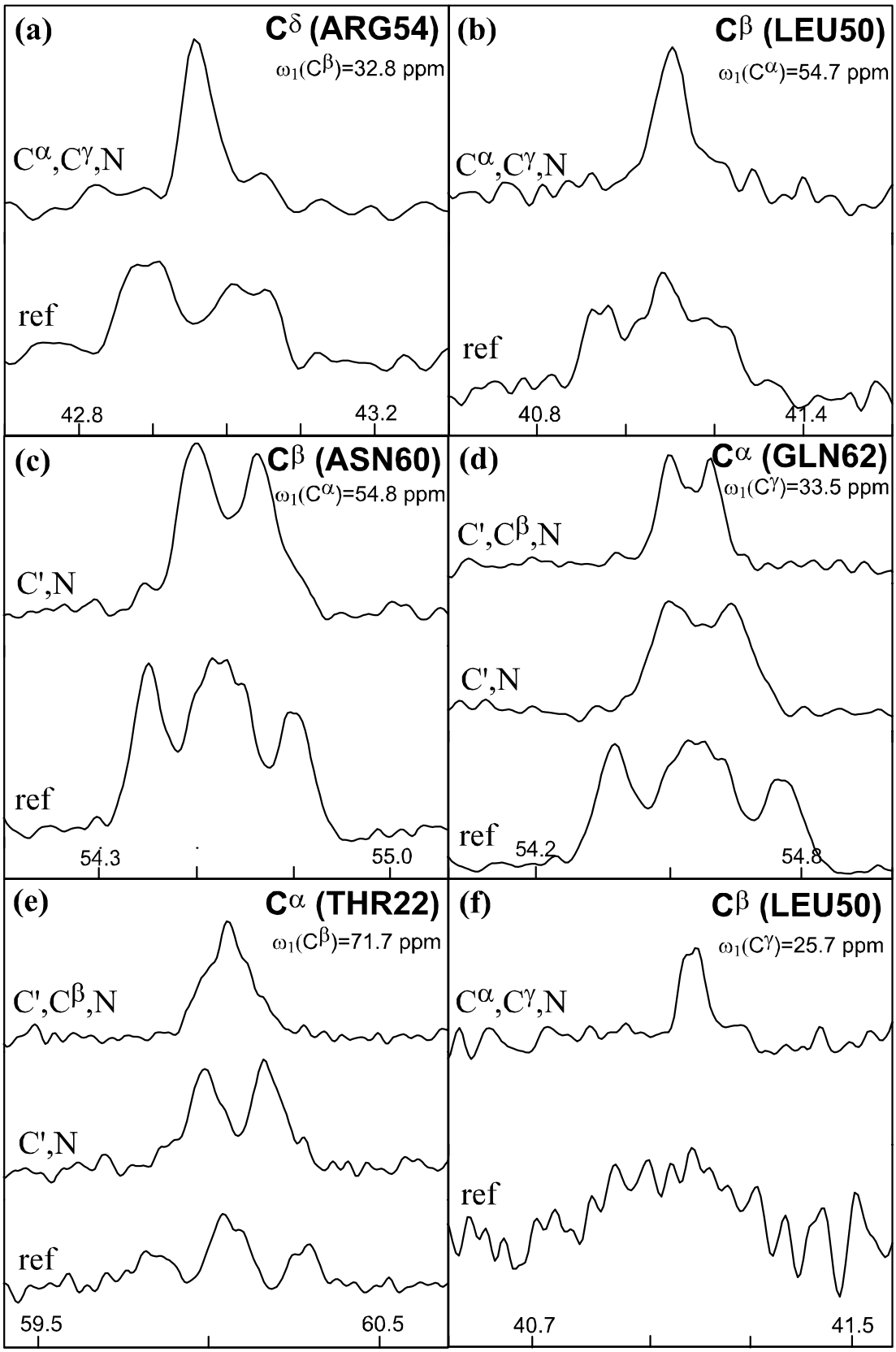
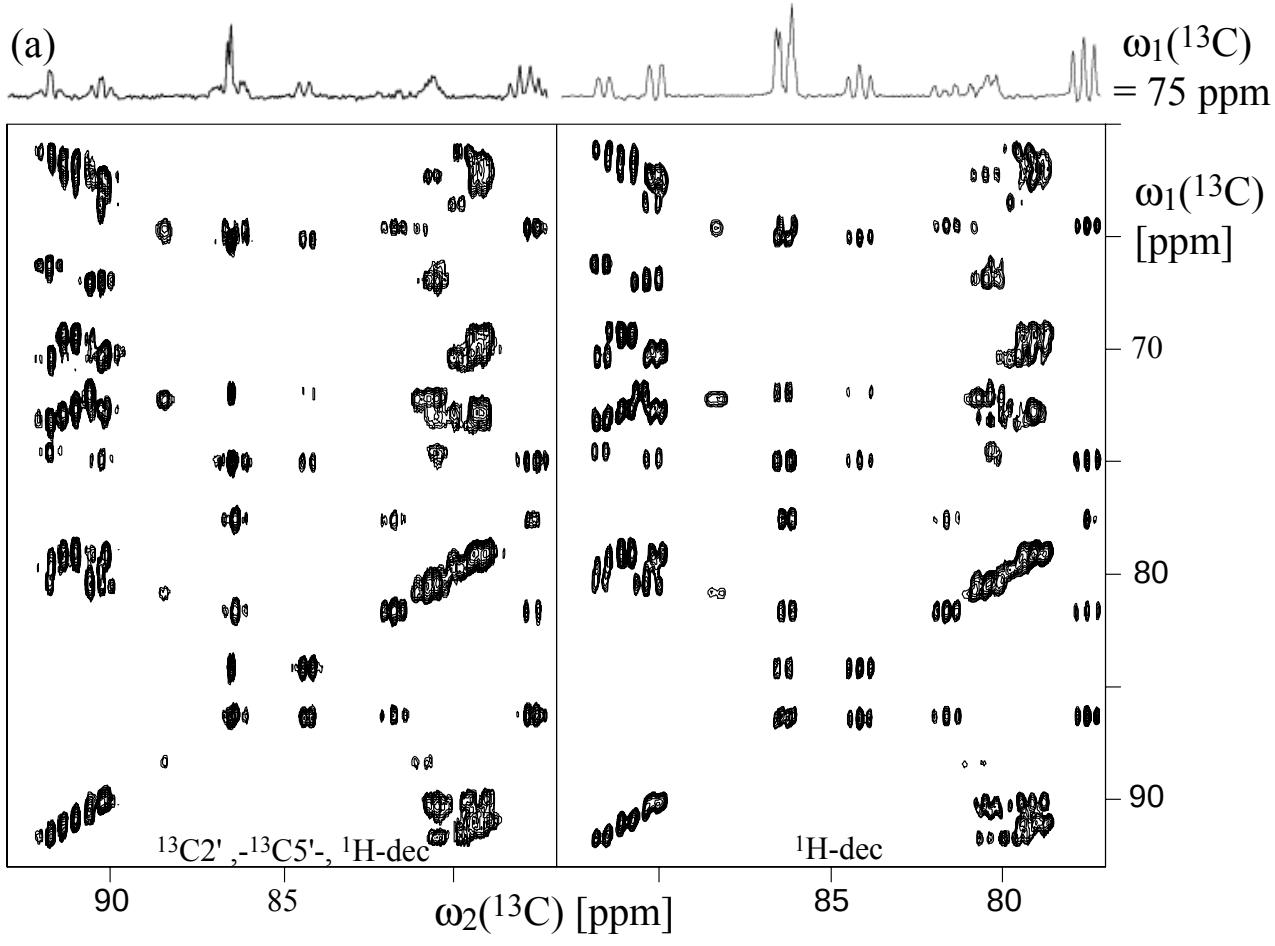


Figure 5. 1D slices of six representative cross-peaks taken along the direct $^{13}C(\omega_2)$ dimension from the spectra shown in Figure 4. Cross-peak assignments are given in the upper right corners. The top slices are taken from [^{13}C,^{13}C]-TOCSY spectra acquired with ^{2}H, ^{15}N and band-selective ^{13}C decoupling as specified on each slice. The bottom slices were recorded with ^{2}H decoupling only.

Figure 6 shows relevant expansions of 2D [^{13}C,^{13}C]-TOCSY spectra of the RNA 14-mer. The two different band-selective homonuclear ^{13}C-decoupling options are compared to a coupled reference spectrum. All these spectra were recorded with ^{1}H decoupling at 500 MHz using a ^{13}C-observe CryoProbe. The band-selective ^{13}C decoupling of C2' and C3' collapses the C2' coupling for all C1' cross-peaks and C3' couplings for C4' cross-peaks, respectively (Figure 6a). The ambiguities due to overlap in the C1′ region between 90-92 ppm and in the C4′ region between 78-81 ppm in the coupled spectrum could be resolved. The band-selective ^{13}C decoupling of C4' collapses the C4' couplings for all C3' and C5'cross-peaks (Figure 6b), resolving particularly the overlap in the C5′ region between 61-63 ppm and in the C2' region between 71-73 ppm. Hence these spectra enabled us to confirm the heteronuclear resonance assignment of the complete spin systems of the riboses [16]. On the top of each spectrum, slices taken along the direct dimension (at 75 ppm in the indirect dimension) offer a comparison between the decoupled and the reference spectra.



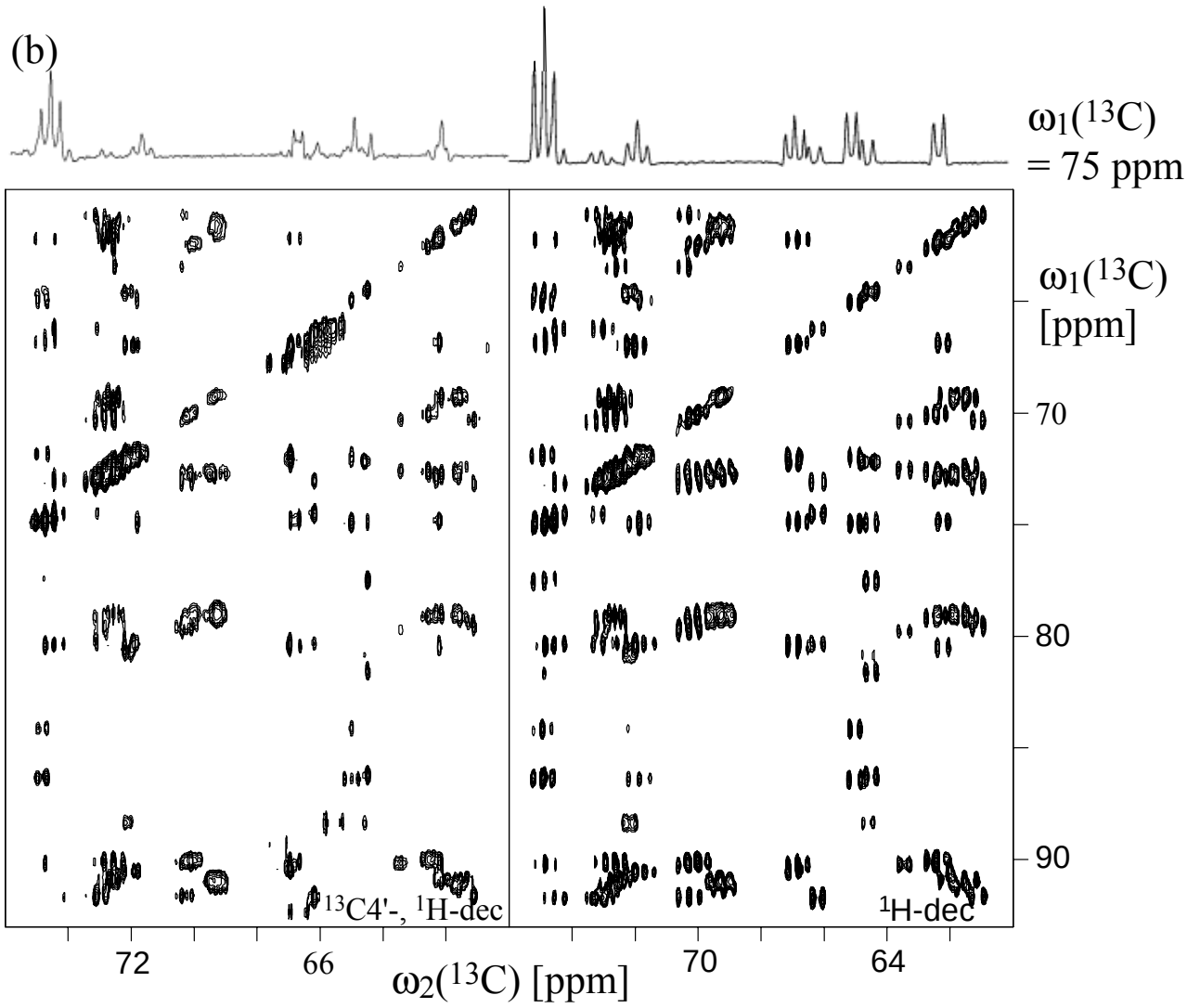


Figure 6 (previous and present pages). Expansions of 2D [^{13}C,^{13}C]-TOCSY spectra measured with a ^{15}N,^{13}C-labeled RNA 14-mer, 5′ GGCACUUCGGUGCC 3′, recorded on a Bruker *Avance* spectrometer operating at a proton frequency of 500 MHz and using the 5mm ^{13}C-observe CryoProbe (DUL). Reference experiments (right): ^{1}H decoupling using WALTZ-16 pulse trains at γB_1 = 3.33 kHz, radio-frequency carrier offsets placed at 110 ppm (^{13}C) and 4.0 ppm (^{1}H), t_{1max} = 8.0 ms and t_{2max} = 116.3 ms, interscan delay 1.0 s and 284 x 4096 complex points resulting in an acquisition time of 36 hours. Homodecoupled spectra (left): (a) ^{13}C-homonuclear decoupling of C2' and C3' using a p5m4 supercycle with an adiabatic CHIRP pulse of 15 ms length and 20% smoothing according to Figure 1b, the radio-frequency carrier offsets were placed at 82 ppm (^{13}C), 71 ppm (^{13}C-decoupling channel) and 4.0 ppm (^{1}H). (b) ^{13}C-homonuclear decoupling of C4' using a p5m4 supercycle with an adiabatic CHIRP pulse of 20 ms length and 20% smoothing according to Figure 1c, the radio-frequency carrier offsets were placed at 79 ppm (^{13}C), 81.5 ppm (^{13}C-decoupling channel) and 4.0 ppm (^{1}H). In both homodecoupled spectra ^{1}H decoupling was applied using WALTZ-16 pulse trains at γB_1 = 2.50 kHz. Both spectra were recorded with t_{1max} = 8.0 ms, t_{2max} = 163.8 ms, 100 x 2048 complex points and an interscan delay of 2.0 s. The acquisition time was 31 hours (a) and 36 hours (b), respectively. The time domain data of all spectra was multiplied with a cosine function in the t_1 and t_2 dimensions and zero-filled to 1024 and 4096 points, respectively.

3. Discussion

For the homodecoupled 2D [^{13}C-^{13}C]-TOCSY spectra presented in Figure 4, the theoretical S/N improvement of 2 x 0.894 = 1.789 per collapsed ^{13}C doublet is not always achieved experimentally. The factor of 0.894 is the theoretical attenuation due to homonuclear decoupling during acquisition with a 20% duty cycle [5] (See also Materials and Methods). Although the comparison to the reference spectrum may not hold exactly due to slightly different experimental settings and the mostly not perfectly resolved single components in the splitting pattern, we estimated the gain factor of the S/N ratio for several peaks. Values ranging roughly from 1.0 to 1.8 per decoupled doublet indicate a rather large distribution in decoupling efficiency and S/N loss.

Initial tests on a ^{13}C,^{15}N-labeled glycine sample indicated clean decoupling of the ^{13}CO and ^{15}N spins under simultaneous irradiation of ^{13}C, ^{15}N and ^{2}H (spectra not shown). The collapse of the $^{13}C^{\alpha}$ quadruplet in glycine due to $J_{C\alpha CO}$ and $J_{C\alpha N}$ couplings results, however, in only a twofold gain in signal intensity instead of the expected value of 4 x 0.894 = 3.578. An experiment with only ^{15}N and ^{2}H decoupling suffers from almost the same loss of intensity, whereas additional ^{13}C decoupling does not scale the signal intensity by more than the factor of 0.894 (2 x 0.894 = 1.789 per collapsed doublet). Because no spurious noise stemming from decoupling is added, we attribute this loss of signal intensity to heteronuclear interference. In order to minimize heteronuclear interference, we employed the latest type of band-pass frequency filters. Nevertheless, because the relations between the gyromagnetic ratios are $\gamma_C : \gamma_N : \gamma_D = 1.00 : 0.40 : 0.61$, the sum of ^{15}N and ^{2}H frequencies can interfere with ^{13}C frequencies. Subsequent 1-dimensional test experiments indicated that the S/N loss can be reduced by using less decoupling power on both ^{2}H and ^{15}N. Certainly, this is only applicable in the cases where a weaker decoupling field is acceptable.

For the RNA ribose spectra presented in Figure 6 neither ^{15}N nor ^{2}H decoupling was needed. The S/N ratio of the decoupled spectrum is comparable to that of the reference spectrum. The recording times were similar, the interscan delay, however, was 1 s in the reference spectrum, whereas 2 s was used in the decoupled spectra. The average longitudinal relaxation rate R_1 of C1' in this molecule is 3.4 Hz at 500 MHz (Radovan Fiala, unpublished results). Based on this value, the S/N is estimated to increase by 36% upon shortening the interscan delay from 2 s to 1 s [17] (See Chapter II). Therefore, for the reference spectrum with an interscan delay of 1s and the decoupled spectra with an interscan delay of 2 s, if recorded equally long, the S/N ratios become comparable. This corresponds to the observation in the quantitative 1D tests on the glycine sample, where the S/N ratio was not further downscaled than by the theoretical factor upon switching on homonuclear ^{13}C decoupling (but no simultaneous ^{2}H and ^{15}N decoupling, see above).

In fact, the ^{13}C homodecoupling could also be used to simplify extraction of scalar and residual dipolar couplings, J_{CC} and D_{CC}, in complex spin systems. However, if the frequency of one of the spins involved in the nondecoupled splitting lies in the transition region of the excited range, rescaling of the apparent splitting is expected [18]. More investigation is required to evaluate the reliability of extracted J_{CC} and D_{CC} values.

4. Conclusions

The current homodecoupling technique proves to be a useful complement to ^{13}C detection based experiments. The use of the technique (i) facilitates the resonance assignment process. Signal overlap is reduced due to line narrowing and collapsing of multiplet pattern in homonuclear spectra. (ii) Because application of ^{13}C decoupling solely (or together with ^{2}H) yields S/N comparable to S/N in the reference spectra, collapsed multiplets result in higher signal intensity. However, simultaneous decoupling of ^{2}H and ^{15}N sacrifices some of the S/N. The homodecoupling technique can be applied in any ^{13}C-detected experiment. In order to further reduce the spectral overlap in ^{13}C–detected experiments, we currently develop experiments which utilize several magnetization transfer steps and make use of spin-state editing (See Chapter V). Thus, direct ^{13}C detection in perdeuterated proteins for resonance assignment and extraction of a large number of side-chain ^{13}C-^{13}C RDC constraints is becoming a useful complement to high quality structure determination. We also pursue ^{13}C detection as an alternative means to obtain signal assignment and structural constraints for oligonucleotides.

5. Materials and Methods

The experiments at 500 MHz were performed with a Bruker *Avance* spectrometer equipped with a cryogenic Z-gradient DUL ^{13}C{^{1}H, ^{2}H} probe, which is optimized for ^{13}C detection. With this probe, however, ^{15}N decoupling is not possible. The sample was 2mM ^{15}N,^{13}C-labeled RNA 14-mer, 5′ GGCACUUCGGUGCC 3′ (produced by Silantes) [16], dissolved in 90%/10% H_2O/D_2O at 295 K. The experiments at 900 MHz were performed with a Bruker *Avance* spectrometer equipped with a conventional Z-gradient DUX ^{13}C{^{15}N, ^{2}H} probe, which is optimized for ^{13}C detection. With this probe, ^{1}H decoupling is not possible. The sample was 1.6mM ^{2}H,^{13}C,^{15}N-labeled (98% isotope enrichment) non-tagged human ubiquitin sample dissolved in 88%/12% H_2O/D_2O at pH 7.3 and 295 K.

The multiple band-selective adiabatic pulses were created through i) integration of the corresponding frequency bands in a 1D spectrum to define the adiabatic sweep width and ii) determination of the pulse length and power by using the ShapeTool in the software package XWIN-NMR 3.5 (Bruker BioSpin). In all cases, the adiabatic pulse shape CHIRP [12] was used, with 20% smoothing and a sweep from low to high field. If created through integration the resulting pulse directly becomes a sum of the band-selective inversion pulses with phase modulations according to the appropriate offset frequencies.

Figure 7 shows schematically the implementation of the complex multi-band homodecoupling during signal acquisition. The homodecoupling duty cycle (HDDUTY) is calculated as

$$\text{HDDUTY} = p/(r + p) \text{ x } 100\%, \quad (1)$$

where p is the irradiation time and r = (dwell time) - p. While switching on the decoupling pulse, 1/3 of the delay between points (v_1) is lost after the last measured point pair. After the last unmeasured point pairs, the pulse is still applied for 1/6 of the delay between points (v_2). v_1 and v_2 are indicated by grey boxes. Thus, the number of lost points (open dots) amounts to all the points in the time window $p + v_1 + v_2$ and the points in the time window $r - v_1 - v_2$ contribute effectively to the signal (black dots). The signal-to-noise ratio S/N as a function of the signal-to-noise ratio in an experiment without decoupling $(S/N)_C$ is

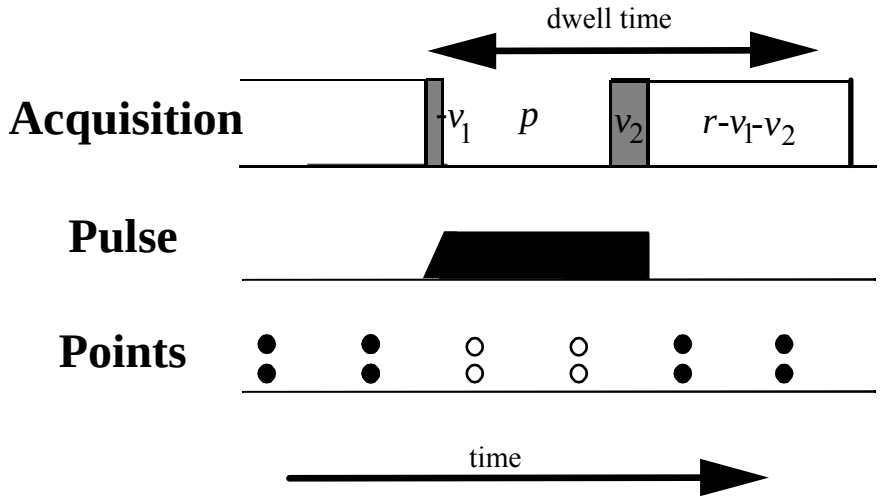


Figure 7. Dwell cycle showing the acquisition status, the pulse p and the dwell window on the time axis. Filled and empty circles represent data acquisition points, respectively, contributing and not contributing to the signal. Time periods v_i, i =1,2 are discussed in the text.

$$S/N = \sqrt{\frac{\text{W}(r - v_1 - v_2)}{\text{W(dwell time)}}} \cdot (S/N)_C \approx \sqrt{\frac{100\%\text{-HDDUTY}}{100\%}} \cdot (S/N)_C \quad (2)$$

where W(t) is the number of points within the time window of length t.

The dwell time was 18 μs in the protein experiments and 80 μs in the RNA experiments, and the total number of points within a dwell time was 6 and 32, respectively. 20% HDDUTY yields a downscale factor of 0.894, which is a gain of 1.789 for a decoupled doublet. 20% is a reasonable compromise between the loss in S/N and decoupling power. According to the convolution theorem in Fourier transform, an irradiation period of total length p spaced by r and with $p \ll r$ in the time domain, produces signals with sidebands spaced by n/r (n = integer) in the frequency domain [19]. Because the dwell time is less than 100 μs in all experiments, the first sidebands on both sides of excitation region appear more than 10 kHz away from the irradiation frequency.

6. Appendix

Additional Figures showing a ^{2}H-, ^{15}N-, and ^{13}C'-decoupled 2D [^{13}C,^{13}C]-TOCSY spectrum recorded on a Bruker *Avance* spectrometer operating at the proton frequency of 900 MHz are presented in the Appendix of Chapter VI.

7. References

[1] Eletsky, A.; Moreira, O.; Kovacs, H.; Pervushin, K. *J. Biomol. NMR* **2003**, *26*, 167-179.
[2] Pervushin, K.; Eletsky, A. *J. Biomol. NMR* **2003**, *25*, 1476-152.
[3] Hu, K.; Eletsky, A.; Pervushin, K. *J. Biomol. NMR* **2003**, *26*, 69-77.
[4] Vögeli, B.; Kovacs, H.; Pervushin, K. *J. Am. Chem. Soc.* **2004**, *126*, 2414-2420.
[5] Bermel, W.; Bertini, I.; Felli, I. C.; Kümmerle, R.; Pierattelli, R. *J. Am. Chem. Soc.* **2003**, *125*, 16423-16429.
[6] Matsuo, H.; Kupce, E.; Wagner, G. *J. Magn. Reson.B* **1996**, *113*, 190-194.
[7] Brutscher, B.; Boisbouvier, J.; Kupce, E.; Tisne, C.; Dardel, F.; Marion, D.; Simorre, J.P. *J. Biomol. NMR* **2001**, *19*, 141-151.
[8] Hammarström, A.; Otting, G. *J. Am. Chem. Soc.* **1994**, *116*, 8847-8848.
[9] Fischer, M.W.F.; Zeng, L.; Zuiderweg, E.R.P. *J. Am. Chem. Soc.* **1996**, *118*, 12457-12458.
[10] Tycko, R.; Pines, A.; Guckenheimer, J. *J. Chem. Phys.* **1985**, *83*, 2775-2802.
[11] Fujiwara, T.; Nagayama, K. *J. Magn. Reson.* **1988**, *77*, 53-63.
[12] Böhlen, J. M.; Bodenhausen, G. *J. Magn. Reson. Ser. A* **1993**, *102*, 293-301.
[13] Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335-338.
[14] Shaka, A. J.; Barker, P.B.; Freeman, R. *J. Magn. Reson.* **1985**, *194*, 547-552.
[15] Vijaykumar, S.; Bugg, C.E.; Cook, W.J. *J. Mol. Biol.* **1987**, *125*, 531-544.
[16] Fürtig, B.; Richter, C.; Bermel, W.; Schwalbe, H. *J. Biomol. NMR* **2004**, *28*, 69-79.
[17] Pervushin, K.; Vögeli, B.; Eletsky, A. *J. Am. Chem. Soc.* **2002**, *124*, 12898-12902.
[18] Freeman, R. *Handbook of Nuclear Magnetic Resonance.* **1986**, Scientific & Technical, Essex.
[19] Morris, G.A.; Freeman, R. *J. Magn. Reson.* **1978**, *29*, 433-462.

Chapter V

Side-Chain H and C Resonance Assignment in Protonated / Partially Deuterated Proteins using a New 3D ^{13}C-Detected HCC-TOCSY

Kaifeng Hu, Beat Vögeli and Konstantin Pervushin*

Laboratorium für Physikalische Chemie, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

* Spectroscopy and Calculations together with K.H.

Journal of Magnetic Resonance, in press

1. Introduction

Recently, ^{13}C-detected NMR spectroscopy [1-4] has been proposed as an attractive alternative for studying large macromolecules [5,6]. Thus, two-dimensional ^{13}C-start and ^{13}C-observe TOCSY NMR experiments were successfully used for the assignment of side-chain aliphatic ^{13}C resonances in a completely deuterated protein with a molecular weight of 44 kDa [5] and measurement of a number of ^{13}C-^{13}C residual dipolar couplings [7] (see Chapter VI). As the protein size increases, the peak overlap becomes a significant problem, especially for monomeric proteins, which have a large number of residues. Simultaneous multiple-band-selective homonuclear ^{13}C and ^{15}N decoupling during both chemical shift evolution periods and signal acquisition can alleviate this problem to some extent by simplifying the multiplet patterns [8] (see Chapter IV). The introduction of a third chemical shift dimension in combination with extensive spin decoupling is expected to offer further reduction in peak overlap.

Here, we propose a pair of ^{1}H-start and ^{13}C-observe 3D HCC-TOCSY to facilitate assignment of the side-chain ^{1}H and ^{13}C resonances. Previously, 3D HCCH-TOCSY experiments were designed for protonated proteins in order to correlate the chemical shifts of H^i, C^i and H^j, which is covalently bound to the adjacent nonfrequency-labeled C^j [9,10]. Apparently, for partially deuterated proteins this experiment fails due to the low probability to find two protons simultaneously in a HCCH moiety. On the other hand, ^{13}C detection relieves this requirement, thus enabling it to work at a very high level of proton dilutions. Comparing to the HMCM[CG]CBCA and amino-acid specific HMCM(CGCBCA)CO "out-and-back" experiments proposed by Kay and coworkers for assignment of methyl groups in a 723 residue protein [11], we expect that the ^{13}C-detected "out-and-stay" HCC is better-suited and can be regarded as a general route for the assignment of methyl ^{13}C and ^{1}H chemical shifts for methyl protonated, highly deuterated and ^{13}C-labeled proteins with high molecular weights [12,13]. A comparison to the original single-quantum ^{13}C-detected 3D HCC-TOCSY reported on a protonated 14 kDa protein alluded to potential application of ^{13}C-detected spectra [1,14]. The use of the ^{1}H–^{13}C multiple-quantum coherence during the chemical shift labeling period for the indirect ^{1}H and ^{13}C dimensions offers more favorable relaxation properties compared to the ^{13}C single-quantum coherence, as demonstrated for the ^{13}C'-detected 3D multiple-quantum-HACACO, 3D TROSY multiple-quantum-HN(CA)HA and 4D TROSY multiple-quantum-HACANH experiments [15]. The HCC version proposed earlier [14] uses a single quantum non-constant time period for ^{1}H labeling and ^{1}H $\rightarrow$ ^{13}C transfer. Subsequently, a constant time period labels ^{13}C. The new version employs the single quantum period only for 4 ms. Single quantum magnetization relaxes with a contribution proportional to the spectral density function at zero frequency, whereas during the multiple quantum period the magnetization relaxes without such contribution [16]. Even though the new element is longer, a clear advantage can be expected,

especially for large proteins. Another advantage of the proposed experiment is the possibility to use the resolved ^{13}C multiplet pattern along the directly acquired dimension in order to match resonances from different strips. In the construction of HCC-TOCSY, we pursued two alternative goals resulting in two complementary experiments. In the first experiment, referred to further on as IP-HCC-TOCSY, a clean, in-phase multiplet pattern in the directly acquired dimension is produced, facilitating sequential matching of 2D strips in the process of resonance assignment. In the second experiment, referred to as SE-HCC-TOCSY (SE stands for "sensitivity-enhanced"), the spectral sensitivity is maximized, which is achieved by relaxing the requirement of "in-phase" appearance of the resulting 3D spectra in the directly acquired dimension.

2. Materials and Methods

CcmE (Cytochrome *c* maturation heme chaperone protein E) is a heme chaperone active in the cytochrome c maturation pathway of *Escherichia coli*, protecting the cell from premature activities of the highly reactive metalloorganic cofactor, which could cause oxidative damage. Uniformly ^{15}N,^{13}C-labeled apo-CcmE-His$_6$ (residues 30–159) was expressed and purified as described in [17]. The NMR sample contained 350 μl of 1 mM protein solution in 20 mM sodium phosphate buffer at pH = 6.0 containing in addition 300 mM NaCl.

FkpA is a heat shock periplasmic peptidyl-prolyl cis/trans isomerase (PPIase) with chaperone activity. ^{15}N,^{13}C,^{2}H(90%)-labeled "shortened FkpA", sFkpA-His$_6$ (residues 10–224) was produced according to potocols described in [18]. The final ^{2}H,^{15}N,^{13}C-labeled NMR sample of the sFkpA protein is 1.2 mM in 20 mM MES, pH 6.0 buffer with 20 mM NaCl.

NMR experiments were performed at 303 K on a Bruker *AVANCE* 500 MHz and 600 MHz spectrometers equipped with a cryogenic Z-gradient DUL $^{13}C\{^{1}H\}$ probe, which is 6 times more sensitive than the conventional ^{1}H probe. NMR data were processed with the program NMRPipe [19]. Chemical shifts are reported relative to DSS (sodium 2,2-dimethyl-2-silapentane-5-sulfonate).

3. NMR Experiments

Figure 1 shows the experimental scheme of the 3D ^{13}C-detected HCC-TOCSY. The initial ^{1}H polarization is transferred to ^{13}C by an INEPT step generating ^{1}H–^{13}C multiple-quantum coherence during the chemical shift labeling period for the indirect $^{13}C(t_1)$ and $^{1}H(t_2)$ dimensions. In the IP experiment, τ_1 is set to 14 ms = $1/(2J_{CC})$ in order to maximize the pure ^{13}C in-phase operator at the beginning of the FLOPSY-16 ^{13}C-^{13}C mixing period [20]. The pulse field gradients (PFG) are

applied to select the in-phase operators before and after the TOCSY mixing. It should be noted that even with PFG based operator selection homonuclear ^{13}C zero-quantum terms insensitive to PFG can survive [21,22], resulting in distortion of the pure in-phase appearance of the crosspeaks [7,23]. Currently no efforts are made to prevent this. After a ^{13}C–read pulse, the NMR signal is detected during decoupling of ^{1}H or ^{2}H (for deuterated samples) spins. In the SE-HCC-TOCSY experiment, we aim to minimize relaxation and to preserve all magnetization transfer pathways. Therefore, τ_1 is set to much less than 14 ms = $1/(2J_{CC})$ (e.g. 5 ms was used in the spectrum of Figure 4). No gradients are applied before and after the mixing period to retain a superposition of in-phase and antiphase coherence significantly improving the signal-to-noise ratio. The product operator description of the coherence transfer pathway is given by Diagrams 1 and 2 for the IP and the SE experiment, respectively:

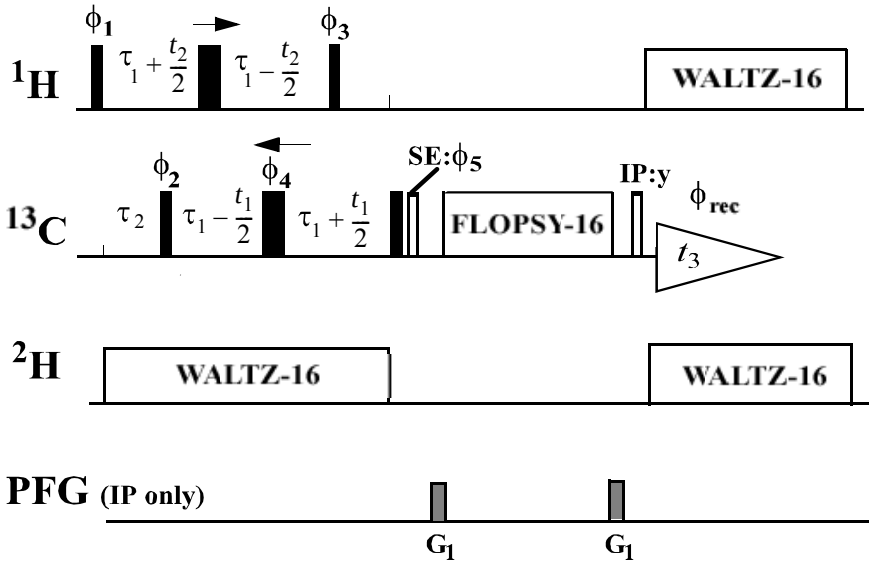


Figure 1. Pulse sequence of ^{13}C-detected 3D HCC TOCSY, the radio-frequency pulses on ^{1}H, ^{13}C and ^{2}H are applied at 2.5, 35 and 3.0 ppm, respectively. Narrow and wide bars indicate non-selective 90º and 180º pulses (black pulses are applied in both experiments; white pulses are only applied in the indicated experiment). ^{1}H and ^{2}H decoupling is achieved with WALTZ-16 [27] at a field strength around γB_1 = 2.5 kHz. Unless indicated otherwise, all radio-frequency pulses are applied with phase x. The phase cycle is: ϕ_1 = {x}; ϕ_2 = {x, -x, x, -x, x, -x, x, -x}; ϕ_3 = {y, y, y, y, -y, -y, -y, -y}; ϕ_4 = {x, x, -x, -x}; ϕ_{rec} = {x, -x, x, -x, -x, x, -x, x}. In the **IP experiment**: The delays are $\tau_1 = 1/(2J_{CC})$ = 14 ms, $\tau_2 = 1/(2J_{CH})$ = 4 ms. The FLOPSY-16 mixing time is 16.96 ms at γB_2 = 8.3 kHz. Pulsed field gradients indicated on the line marked PFG are applied along the z-axis with duration of 0.9 ms and strength of 40 G/cm. Quadrature detection in the indirect ^{13}C (t_1) dimension and ^{1}H (t_2) dimension is achieved by the States-TPPI method [28] applied to the phases ϕ_1 and ϕ_2, respectively. In the **SE experiment**: The delays are: τ_1 is set to smaller than $1/(2J_{CC})$ = 14 ms (depending on the relaxation properties of the multiple-quantum coherence), $\tau_2 = 1/(2J_{CH})$ = 4 ms. The FLOPSY-16 mixing time is shortened to 16.96/2 ms = 8.48 ms at γB_2 = 8.3 kHz. Pulsed field gradients are not applied. A phase-sensitive spectrum in the indirect $^{13}C(t_1)$ dimension is obtained using the ECHO/ANTIECHO method by recording two FIDs for each t_1 value with ϕ_5 = {x} and ϕ_5 = {-x}, respectively. Quadrature detection in the ^{1}H (t_2) dimension is achieved by the States-TPPI method applied to the phase ϕ_1.

$$H_z^i \rightarrow -H_y^i \rightarrow 2H_x^iC_z^i \rightarrow -2H_x^iC_y^i[t_1,t_2] \rightarrow C_z^i \xrightarrow{\text{FLOPSY-16}} C_z^j \rightarrow C_x^j[t_3] \quad (1)$$

$$H_z^i \rightarrow -H_y^i \rightarrow 2H_x^iC_z^i \rightarrow -2H_x^iC_\pm^i \prod_m (a^m E + b^m i * 2C_z^m)[t_1,t_2]$$
$$\rightarrow -C_\pm^i \prod_m (a^m[2\tau_1]E + b^m[2\tau_1]i * 2C_z^m) \xrightarrow{FLOPSY-16} C_\pm^j \prod_m (c^m E + d^m i * 2C_z^m) \quad (2)$$
$$\rightarrow C_-^j \prod_l (c^l E + d^l i * 2C_z^l)[t_3]$$

H^i and C^i stand for the spins of the hydrogen and carbon atom of the bond i, from which the magnetization pathway starts. C^j stands for the directly detected carbon atom. C^m are spins involved in the *J*-coupled network of the spins C^i and C^j, and C^l are spin(s) bonded to the directly detected carbon C^j. Due to the coherence order selective mixing achieved by isotropic FLOPSY sequence [24], the "minus" operators evolving during t_1 are transferred exclusively to "minus" operators, which is used for signal acquisition. Quadrature detection is achieved by flipping "plus" and "minus" operators with a ^{13}C 180^o pulse before TOCSY mixing. This results in the enhancement of sensitivity due to the coherence order selective transfer between adjacent ^{13}C spins. In addition, antiphase $2C_\pm^iC_z^j$ coherences developing during the t_1 period contribute to the C-C polarization transfer during isotropic mixing [25]. The mixture of in-phase or antiphase operators of C^i and C^j with respect to C^m and C^l is reflected by the time-dependent coefficients a^m, b^m and c^l, d^l with $\sqrt{(a^m)^2+(b^m)^2}=1$ and $\sqrt{(c^l)^2+(d^l)^2}=1$, respectively. In general, the splitting pattern becomes very complex due to the superposition of the in-phase and antiphase terms, as well as the different spin networks. Diagram 2 can be used to reconstruct the resonances of coupled ^{13}C spins by nonlinear fitting of the theoretical line shapes to the experimental ones by variation of the parameters of c^l, d^l.

4. Results and Discussion

As an example for the application to a protonated protein, Figure 2 shows data from a spectrum of the 3D ^{13}C-detected IP-HCC-TOCSY measured on 16 kDa uniformly ^{15}N,^{13}C-labeled apo-CcmE-His$_6$. Figures 2a and b show H-C strips from the 3D ^{13}C-detected IP-HCC-TOCSY spectrum and HN strips from the 3D HNCACB spectrum for Val80 and Ile84, respectively. After the sequence-specific backbone assignment, $^{13}C^\alpha$ and $^{13}C^\beta$ chemical shifts can be aligned for identification of the side-chain spin systems. The complete spin system of Val80 can be clearly recognized. In the case of Ile84, all matching 2D H-C strips are found with the sole exception of the strip corresponding to γ^1 ^{13}C resonance. As an application to a large highly deuterated protein, a 3D IP-HCC-TOCSY was

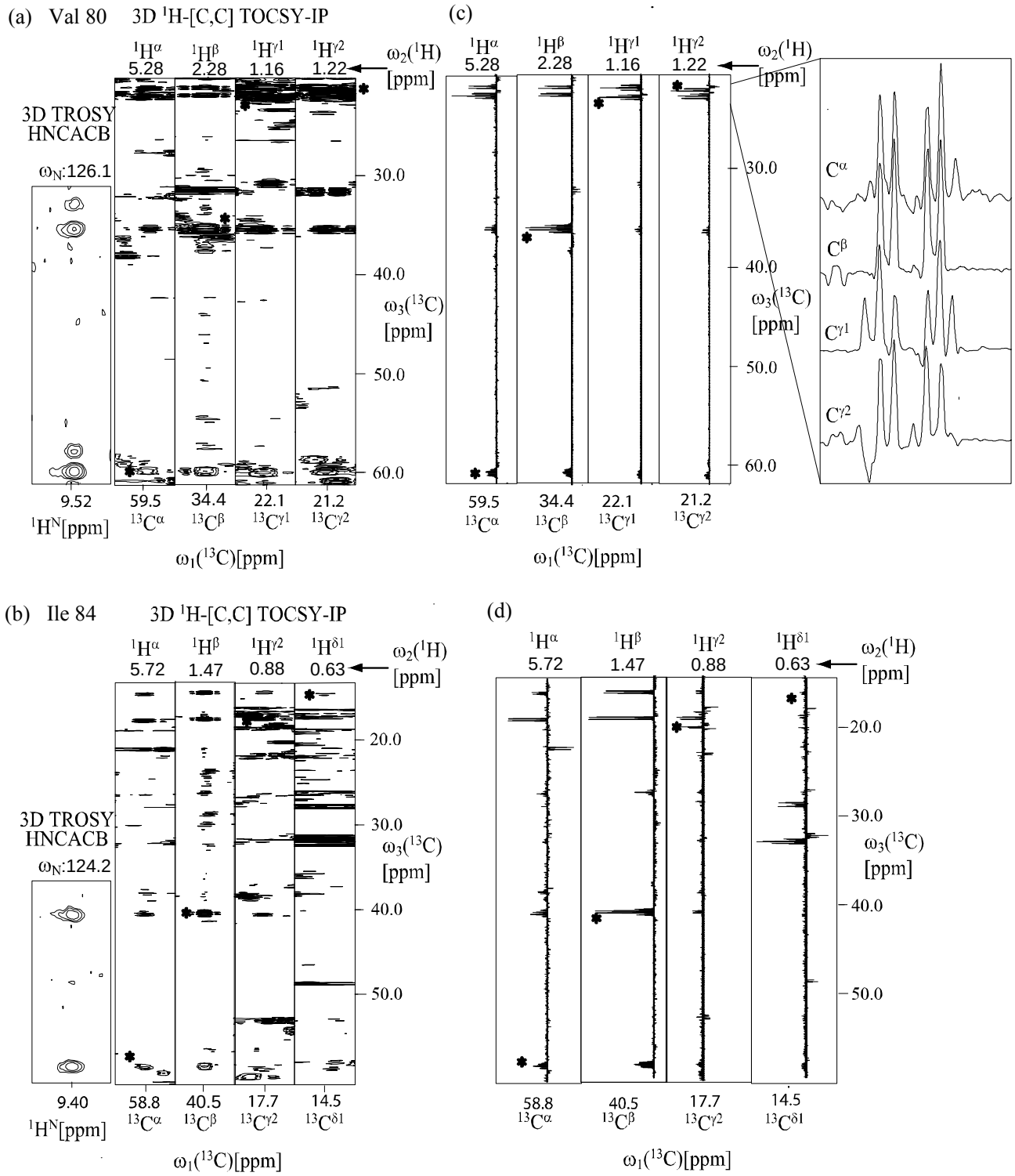
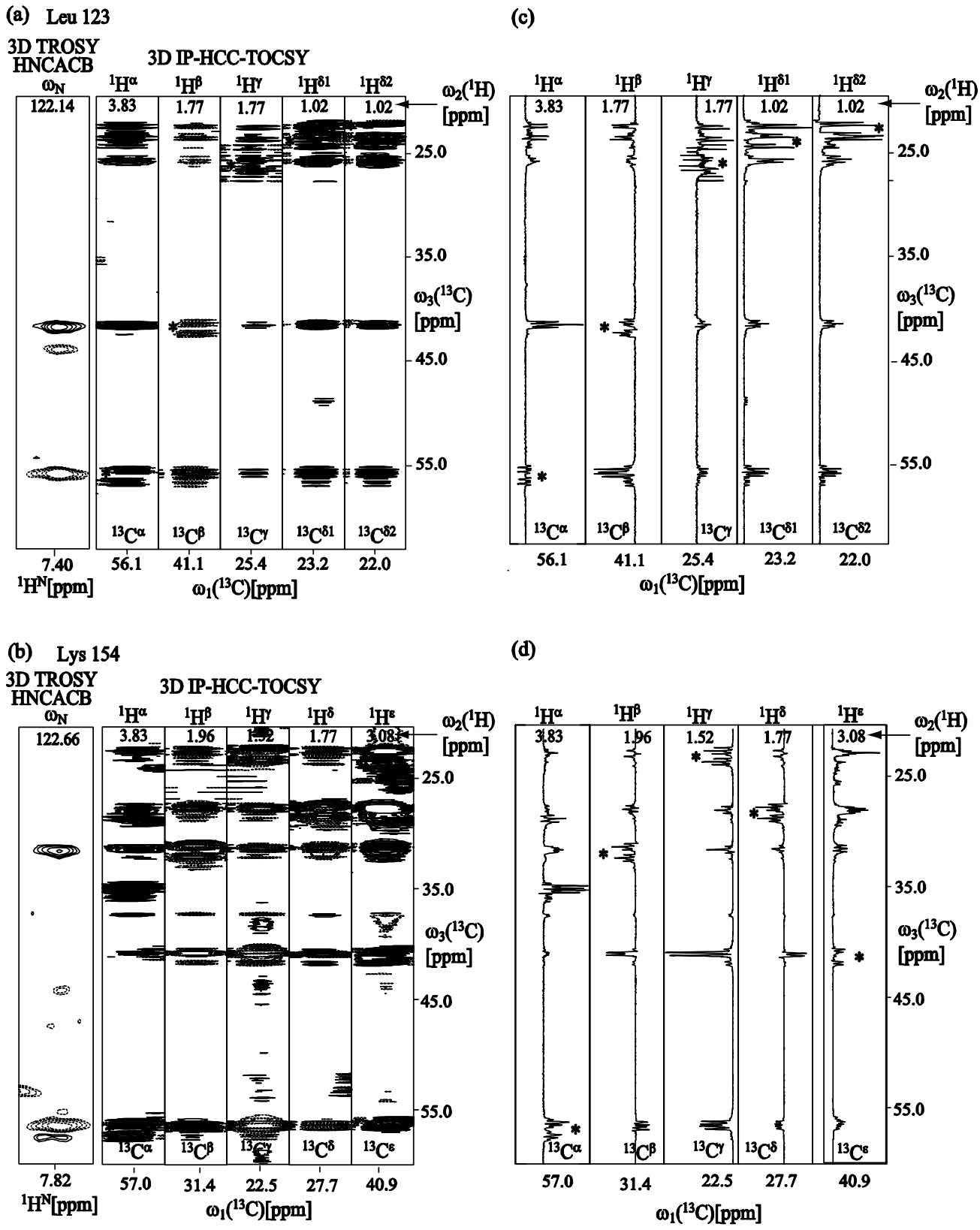


Figure 2. H-C strips taken from the 3D ^{13}C-detected IP-HCC-TOCSY spectrum and HN strips taken from the 3D HNCACB spectrum for (a) Val80 and (b) Ile84 of apo-CcmE-His$_6$. The corresponding indirect ^{13}C chemical shifts and their bonded ^{1}H chemical shifts are assigned and labeled for each strip. Diagonal peaks are marked with asterisks. $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts are aligned to corresponding assigned signals in the 3D HNCACB. Slices taken along the direct ^{13}C dimension are shown in c and d. The region of 20.5 – 22.5 ppm is magnified to present the resolved multiplets due to J_{CC} couplings. Alignment of the splitting pattern resolved at the high resolution obtained in the directly detected ^{13}C dimension, helps to confirm the assignment of the spin systems. The side chain ^{1}H and ^{13}C chemical shifts of Val80 are

completely assignable. The spin system of Ile84 can be identified without difficulty although the strip corresponding to $H^{\gamma 1}$-$C^{\gamma 1}$ is not visible. The experiment was performed at 600 MHz. 75(t_1) x 24(t_2) x 2048(t_3) complex points were accumulated, with t_{1max}(indirect ^{13}C) = 9.93 ms, t_{2max}(^{1}H) = 9.99 ms and t_{3max}(direct ^{13}C) = 168.7 ms, the interscan delay of 1 s and 56 scans per increment resulted in a total measurent time of 117 hours. The time domain data was multiplied with a cosine function in all dimensions and zero-filled to 256 x 128 x 2048 complex points.

recorded on 0.8 mM uniformly ^{15}N,^{13}C,^{2}H(~90%)-labeled dimeric 48 kDa sFkpA-His$_6$. Figures 3a and b show the 2D H-C strips from the 3D IP-HCC-TOCSY spectrum and HN strips from the 3D HNCACB spectrum of the complete spin systems of Leu123 and Lys154. As shown in Figure 2, the $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts can be aligned for identification of bonded H and C groups belonging to the same residue. The sign of the peaks in the H-C strips reports on the number of carbon neighbors helping the assignment. For example, the H-C strips for the β of Leu123 and the β, γ and δ of Lys154 are opposite in sign to those in the other strips [1].

Figure 3 (next page). H-C strips taken from the 3D ^{13}C-detected IP-HCC-TOCSY spectrum and HN strips taken from the 3D HNCACB spectrum for (a) Leu123 and (b) Lys154 of sFkpA-His$_6$. The corresponding indirect ^{13}C chemical shifts and their bonded ^{1}H chemical shifts are assigned and labeled for each strip. Diagonal peaks are marked with asterisks. $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shift are aligned to their assignments in the 3D HNCACB. Slices taken along the direct ^{13}C dimension are shown in c and d. The systems of both Leu123 and Lys154 are completely assigned. Broken lines indicate negative peaks. The peaks from the strips of $^{13}C^{\beta}$ of Leu123 and the $^{13}C^{\beta}$, $^{13}C^{\gamma}$ and $^{13}C^{\delta}$ of Lys 154 are opposite in sign to the other strips, which further confirms the assignment of the spin systems. The experiment was performed at 500 MHz. 75(t_1) x 24(t_2) x 2048(t_3) complex points were accumulated, with t_{1max}(indirect ^{13}C) = 11.93 ms, t_{2max}(^{1}H) = 11.99 ms and t_{3max}(direct ^{13}C) = 203 ms, the interscan delay of 1 s and 56 scans per increment resulted in a total measurement time of 117 hours. The time domain data was multiplied with a cosine function in all dimensions and zero-filled to 256 x 128 x 2048 complex points.


(a) Leu 123
3D TROSY HNCACB
3D IP-HCC-TOCSY
ωN
122.14
1Hα 1Hβ 1Hγ 1Hδ1 1Hδ2
3.83 1.77 1.77 1.02 1.02
ω2(1H) [ppm]
25.0
35.0
ω3(13C) [ppm]
45.0
55.0
13Cα 13Cβ 13Cγ 13Cδ1 13Cδ2
7.40
56.1 41.1 25.4 23.2 22.0
1HN[ppm]
ω1(13C)[ppm]
(c)
1Hα 1Hβ 1Hγ 1Hδ1 1Hδ2
3.83 1.77 1.77 1.02 1.02
ω2(1H) [ppm]
25.0
35.0
ω3(13C) [ppm]
45.0
55.0
13Cα 13Cβ 13Cγ 13Cδ1 13Cδ2
56.1 41.1 25.4 23.2 22.0
ω1(13C)[ppm]
(b) Lys 154
3D TROSY HNCACB
3D IP-HCC-TOCSY
ωN
122.66
1Hα 1Hβ 1Hγ 1Hδ 1Hε
3.83 1.96 1.52 1.77 3.08
ω2(1H) [ppm]
25.0
35.0
ω3(13C) [ppm]
45.0
55.0
13Cα 13Cβ 13Cγ 13Cδ 13Cε
7.82
57.0 31.4 22.5 27.7 40.9
1HN[ppm]
ω1(13C)[ppm]
(d)
1Hα 1Hβ 1Hγ 1Hδ 1Hε
3.83 1.96 1.52 1.77 3.08
ω2(1H) [ppm]
25.0
35.0
ω3(13C) [ppm]
45.0
55.0
13Cα 13Cβ 13Cγ 13Cδ 13Cε
57.0 31.4 22.5 27.7 40.9
ω1(13C)[ppm]

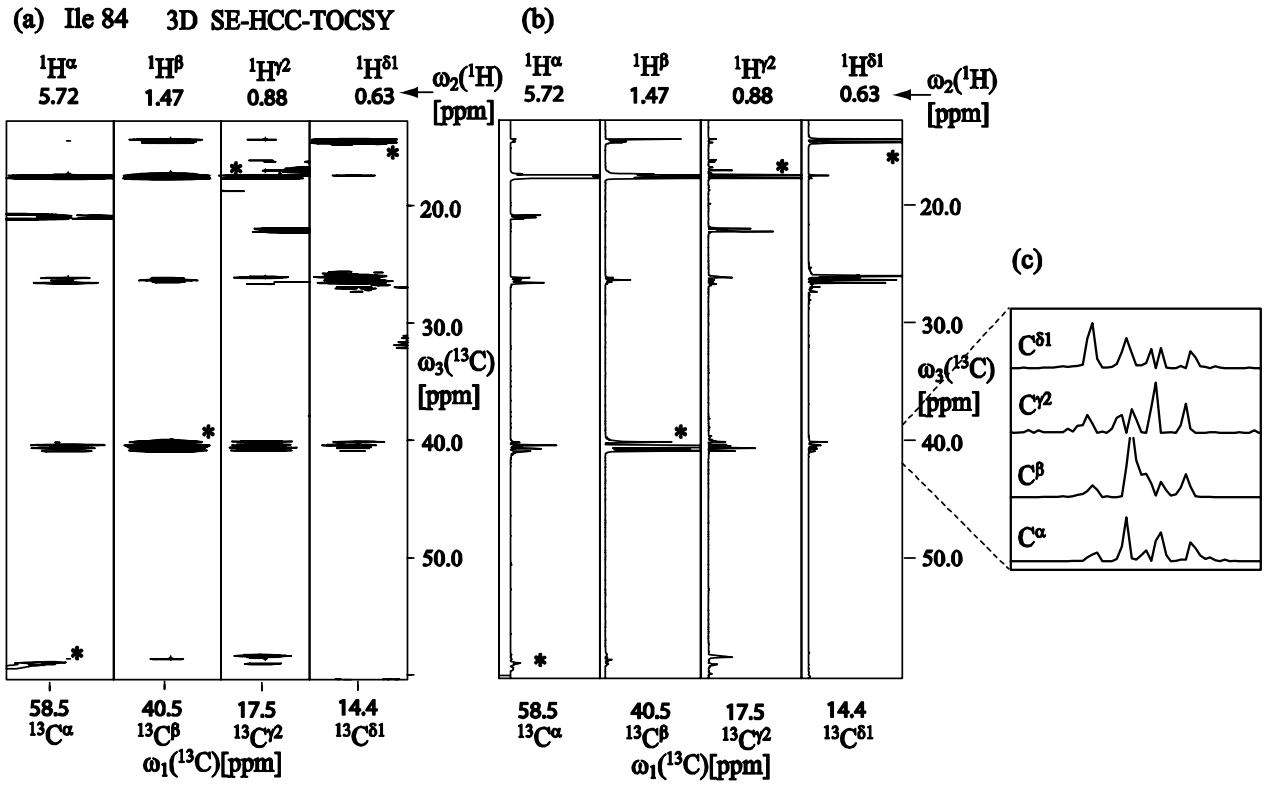


Figure 4. (a) H-C strips taken from the 3D ^{13}C-detected SE-HCC-TOCSY spectrum for Ile84 of apo-CcmE-His$_6$. The corresponding indirect ^{13}C chemical shifts and their bonded ^{1}H chemical shifts are assigned and labeled for each strip. Diagonal peaks are marked with asterisks. Compared with Figure 2 (b), the much cleaner spectrum demonstrates the significant gain in signal-to-noise ratio compared to the IP experiment, which facilitates the recognition of the spin systems. (b) Corresponding slices taken along the direct ^{13}C dimension. (c) Expanded regions showing the distorted lineshape of the peaks due to the mixture of in-phase and antiphase coherence. The experiment was performed at 600 MHz. 75(t_1) x 26(t_2) x 2048(t_3) complex points were accumulated, with t_{1max}(indirect ^{13}C) = 9.93 ms, t_{2max}(^{1}H) = 10.82 ms and t_{3max}(direct ^{13}C) = 168.8 ms, the interscan delay of 1 s and 56 scans per increment resulted in a total measurement time of 126 hours. The time domain data was multiplied with a cosine function in all dimensions and zero-filled to 256 x 128 x 2048 complex points and Fourier transform was applied in power mode.

The high sensitivity of the SE version of the 3D ^{13}C-detected HCC-TOCSY was demonstrated on the 16 kDa uniformly ^{15}N,^{13}C-labeled apo-CcmE-His$_6$ sample. Figure 4a shows H-C strips from the 3D ^{13}C-detected SE-HCC-TOCSY spectrum for Ile84. Although the maximum of the components may shift by up to several Hertz due to the mixture of the in-phase and antiphase coherences and concomitant lineshape distortion, identification of spin systems is obtained with little difficulty and the $^{13}C^\alpha$ and $^{13}C^\beta$ chemical shifts can still be aligned with the corresponding signals from the HNCACB. Slices taken along the direct dimension are shown in Figure 4b. An expansion shows the distorted multiplet patterns. As a comparison of Figures 2d and 4b reveals, using approximately same measurement time as for IP-HCC-TOCSY, the SE version affords approximate 8 times gain of signal-to-noise ratio. Shorter τ_1 and shorter TOSCY mixing time seems to play an important role in sensitivity improvement for the non-deuterated protein sample, as shown in the comparison of Figure

2 to Figure 4. However, there is not an obvious experimental sensitivity factor gain achieved for the larger deuterated 44 kDa protein using the SE version compared to its IP counterpart (data not shown for the SE version). In the protonated protein sample, there is still a large amount of protons in the proximity of the H-C group, which in fact is a strong relaxation factor through H^p-H and H^p-C dipole-dipole coupling mechanism (here H^p is the proton in the proximity of the active H-C group). Therefore, for the non-deuterated protein sample, a shorter τ_1 and a shorter TOSCY mixing time can generally greatly improve the sensitivity. However, for the larger deuterated protein, we can assume that there are not too many protons in the proximity of the H-C group and consequently the relaxation due to H^p-H and H^p-C dipole-dipole coupling could be negligible, thus further shortening the τ_1 and TOSCY mixing time can not bring an obvious sensitivity gain.

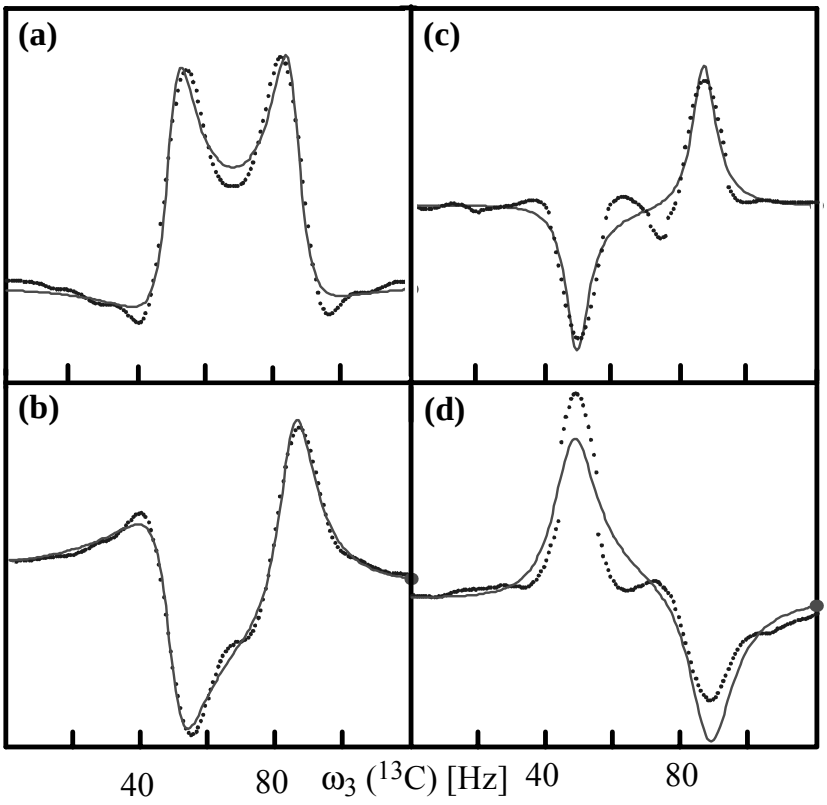


Figure 5. Slices along the direct dimension through $C^{\gamma 2}$ peaks of the spin system Ile84 of apo-CcmE-His_6 (dotted lines) are taken from the 3D ^{13}C-detected SE-HCC-TOCSY spectrum. The corresponding indirect ^{13}C chemical shifts and their bonded ^{1}H chemical shifts are assigned to (a) C^{α}, (b) C^{β}, (c) $C^{\gamma 2}$ and (d) $C^{\delta 1}$. The vertical scales are arbitrary. The solid lines represent the nonlinear fit of Equation 5 to the experimental line shapes using the program Matlab (The MathWorks, Inc.) in a 7-dimensional space comprising of: chemical shift ω_0, amplitude factor A, a uniform transverse relaxation rate R_2 (neglecting cross-correlated relaxation), the phases ϕ_1 and ϕ_2, the scalar coupling J, and an arbitrary baseline constant C. Source data are taken from 16.5 (set to 0 Hz in the plots above) to 17.3 ppm along the direct dimension. The values of the parameters obtained for ω_0, A, R_2, J, ϕ_1, ϕ_2 and C are: (a) 67.7 , 970.6 , 5.2 , 35.8 , 0.78 , -0.90 , -221.8; (b) 67.6 , 1490.8 , 6.8 , 35.6 , -0.40 , -4.30 , -221.7; (c) 68.2 , 405.7 , 7.8 , 40.6 , 2.93 , -0.27 , -211.4; (d) 68.2 , 988.6 , 4.2 , 37.9 , -0.06 , -3.32 , -220.0.

The identification of a spin system may be hampered by the nonuniform peak shapes recorded in the SE version. As can be inferred from Diagram 2, each cross-peak is a superposition of in-phase and antiphase components with particular phase values depending on the transfer pathway. The detected signals can be rewritten in the form of single transition operators:

$$C_-^j \prod_l (c^l E + d^l i * 2C_z^l)[t_3] = C_-^j \prod_l \{(c^l + id^l)(E/2 + C_z^l)[t_3] + (c^l - id^l)(E/2 - C_z^l)[t_3]\} \quad (3)$$

As can be seen from Equation 3, each single transition component has its specific phase, which depends on c^l and d^l. Thus, in general, the superimposed peak cannot be phased to an absorptive pattern for all components. After Fourier transform a doublet peak can be written as:

$$S(\omega) = \frac{AR_2\cos(\phi_1)}{R_2^2 + (\frac{J}{2} + \omega_0 - \omega)^2} + \frac{A(\frac{J}{2} + \omega_0 - \omega)\sin(\phi_1)}{R_2^2 + (\frac{J}{2} + \omega_0 - \omega)^2} + \frac{AR_2\cos(\phi_2)}{R_2^2 + (-\frac{J}{2} + \omega_0 - \omega)^2} + \frac{A(-\frac{J}{2} + \omega_0 - \omega)\sin(\phi_2)}{R_2^2 + (-\frac{J}{2} + \omega_0 - \omega)^2} + C \quad (4)$$

with assumption of a Lorentzian lineshape for each peak component. Here, ω_0 is the chemical shift, A is an amplitude factor, R_2 is a uniform transverse relaxation rate (neglecting cross-correlated relaxation), ϕ_1 and ϕ_2 describe the phases for the two doublet components, respectively, J is the J_{CC} scalar coupling constant, and C is an arbitrary constant counting for the baseline level. A nonlinear fit of Equation 4 to the multiplet peak pattern may reveal the chemical shift together with all other parameters.

Figure 5 shows slices through $C^{\gamma 2}$ peaks of the spin system Ile84 of apo-CcmE-His_6. Although the shapes vary significantly, a nonlinear fit yields chemical shift values differing only by up to 0.6 Hz. This deviation is smaller than the spectral resolution. Details on the fitting parameters are given in the figure caption. The power of this method strongly depends on the signal-to-noise ratio, the pattern complexity, the number of unknown parameters and the performance of the fitting procedure. Using this procedure, one could even decompose overlapping peaks, although we designate the complete analysis of this problem to further other work. Clearly, the presented approach represents an attempt to effectively extract the relevant spectral information from the experiment maximizing spectral sensitivity. Probably a combination of IP and SE experiments would be necessary to unequivocally establish resonance frequencies of ^{13}C spins.

Besides reducing the overlap by introducing the ^1H dimension, an additional advantage of this ^1H-start 3D experiment is a significant reduction of the interscan delay from 2.5 s to less than 1 s in comparison to its ^{13}C–start 2D counterpart [1] due to faster equilibrium magnetization recovery of ^1H compared to ^{13}C in the ^2H-^{13}C moieties [26] (see Chapter II). Neglecting transverse relaxation, theoretically, this new ^1H-start and ^{13}C-observe 3D HCC-TOCSY can be expected to have comparable sensitivity per unit time to its ^{13}C–start and ^{13}C-observe 2D version as follows:

$$\Gamma = (S/N)_{HCC}/(S/N)_{CC} = \sqrt{\frac{T_1^C}{T_1^H}} \bullet \frac{\gamma_H}{\gamma_C} \bullet P \quad (5)$$

where $(S/N)_{HCC}$ and $(S/N)_{CC}$ are signal-to-noise ratios in 3D HCC-TOCSY and 2D CC-TOCSY spectra, respectively. T_1^C and T_1^H are longitudinal relaxation times for ^{13}C and ^1H, with typical values for T_1^C of about 3-4 s and for T_1^H of 0.5-1 s. γ_C and γ_H are gyromagnetic ratios for ^{13}C and ^1H, P indicates the proton level in the partially deuterated protein sample. For 90% deuterated protein samples, P is equal to 0.1 and the resulting Γ ranges from 0.7 to 1.1. Optimal sensitivity for the given protein size could be obtained by varying the deuteration level [27].

For the evaluation of the optimal deuteration level, we consider an exemplary configuration of $^{13}CX_2$, in which the magnetization starting on a proton is first transferred to its bonded carbon ^{13}C and both ^1H and ^{13}C chemical shifts are labeled in a MQ mode during the constant time period, and subsequently transferred to another $^{13}C(X_2)$ (by FLOPSY mixing during which relaxation is not considered), and the signal of the $^{13}C(X_2)$ is finally detected. Here, X stands for H or D. Following the method proposed by Nietlispach et al. [28], we calculate the ratio of the signals of a partially deuterated sample to a fully protonated sample. The dipole-dipole coupling and chemical shift anisotropy (CSA) are assumed to be the only mechanisms for relaxation calculation and all cross-correlated relaxation terms are neglected. Relaxation due to dipole-dipole coupling with other protons in proximity is not taken into account in the analysis either. For proteins with large molecular size, only terms proportional to the spectral density function at the zero frequency, which make up the main contribution to the relaxation rate, are considered. As example, τ_1 = 4 ms (^1H single quantum period, SQ_H), τ_2 = 24 ms (^1H-^{13}C multiple quantum period, MQ_HC) and the following expressions for relaxation rates are used:

$$R_{SQ\_H} = R_{SQ\_H,CSA(H)} + R_{SQ\_H,D(HC)} + R_{SQ\_H,D(HX)} \quad (6)$$

$$R_{MQ\_HC} = R_{MQ\_HC,CSA(H)} + R_{MQ\_HC,CSA(C)} + R_{MQ\_HC,D(CX)} + R_{MQ\_HC,D(HX)} \quad (7)$$

Magnetization starting on D (for X=D) is not considered because it does not result in detectable signal in the HCC spectrum. For a partially deuterated protein sample, contributions to the signal are weighted with p^2 and $p(1\text{-}p)$ for CH_2 and CHD, respectively, where p is the protonation level. During detection, the average relaxation rate approximation is taken as [28]:

$$R_{SQ\_C} = R_{SQ\_C,CSA(C)} + 2(pR_{SQ\_C,D(CH)} + (1\text{-}p)R_{SQ\_C,D(CD)}) \quad (8)$$

Figure 6 a shows a 3D plot of the intensity gain of a partially deuterated sample over a fully protonated sample versus the correlation time τ_c and the protonation level p. Figure 6 b shows slices

at τ_c = 5, 20 and 40 ns. The optimal protonation level ranges from 0.15 to 0.20. This is close to the protonation level of the partially deuterated sFkpA protein ($\approx$ 0.1) used for demonstration. It should also be noted that the relative gain of sensitivity for a partially deuterated sample over a purely protonated sample increases dramatically as the protein size becomes larger. For τ_c = 5, the gain is around twofold. However, it can exceed 1000 times for larger size protein when τ_c is over 30 ns. During the FLOPSY mixing period, additional gain can also be expected due to the dipolar interactions between carbon and protons/deuterons. However, the detailed formalism of the relaxation during the FLOPSY mixing is very complicated and is out of the scope of the present work. The efficiency of the overall coherence transfer pathway depends on the H-C moieties with different configurations, such as CH, CH_2 and CH_3. The gain may be rather non-uniform for different H-C groups even at a certain deuteration level. Nonetheless, the model shows the strong advantage of deuteration, especially for large size protein samples.

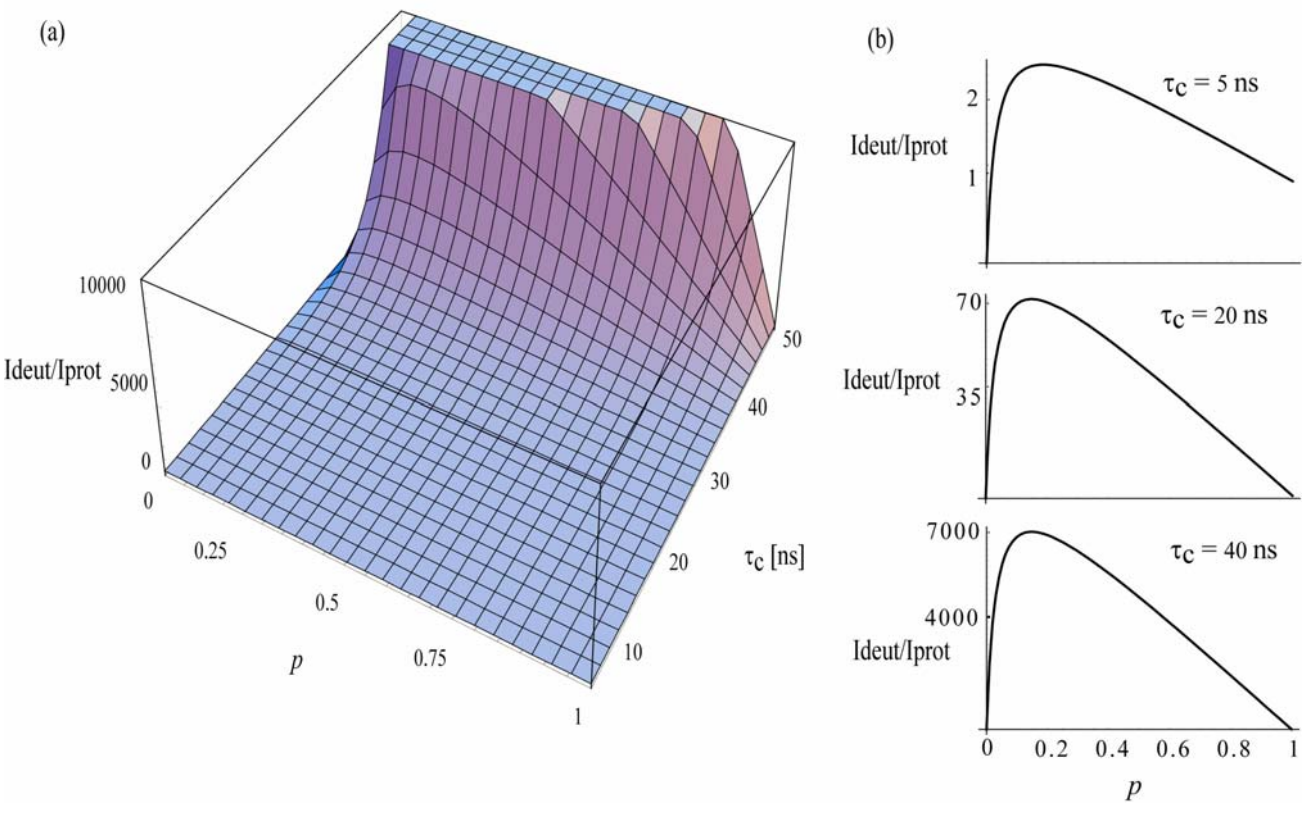


Figure 6. (a) 3D plot of the intensity gain of a partially deuterated sample over a fully protonated sample versus the correlation time τ_c and the protonation level *p*. (b) Slices taken at τ_c = 5, 20 and 40 ns from (a).

In conclusion, the new method presented here can serve as an attractive alternative to the standard ^{1}H-detected side-chain ^{1}H and ^{13}C assignment strategies. High resolution and good sensitivity can overcome problems associated with ^{13}C spectroscopy such as peak overlap. Especially for proteins

with high molecular weight that requires partial deuteration, this method proved to be very useful as exemplified with the 48 kDa sFkpA. Because the recovery delay can be considerably shortened and magnetization starts on ^{1}H instead of ^{13}C, the experiment can yield comparable or higher sensitivity per unit experimental time. In addition to chemical shifts, the clear splitting pattern of the peaks along the directly detected ^{13}C dimension in IP type 3D ^{13}C-detected HCC-TOCSY can also be used to identify and confirm peak alignment. If the splitting pattern of the peaks is not of interest, the SE experiment can greatly increase the signal-to-noise ratio. More sophisticated homonuclear decoupling can further increase the signal-to-noise ratio [8] (see the previous Chapter). The very high sensitivity indicates that the measurement time can be further shortened.

5. References

[1] Maudsley, A.A.; Müller, L.; Ernst, R. R.. J. *Magn. Reson.* **1977**, *28*, 463-469.
[2] Chan, T.M.; Westler, W.M.; Santini, R.E.; Markley, J.L. *J. Am. Chem. Soc.* **1982**, *104*, 4008-4010.
[3] Westler, W.M.; Kainosho, M.; Nagao, H.; Tomonaga, N.; Markley, J.L. *J. Am. Chem. Soc.* **1988**, *110*, 4093-4095.
[4] Oh, B.H.; Westler, W.M.; Darba, P.; Markley, J.L. *Science* **1988**, *240*, 908-911.
[5] Eletsky, A.; Moreira, O.; Kovacs, H.; Pervushin, K. *J. Biomol. NMR* **2003**, *26*, 167-169.
[6] Bermel, W.; Bertini, I.; Felli, I.C.; Kümmerle, R.; Pierattelli, R. *J. Am. Chem. Soc*. **2003**, *125*, 16423-16429.
[7] Vögeli, B.; Kovacs, H.; Pervushin, K. *J. Am. Chem. Soc.* **2004**, *126*, 2414-2420.
[8] Vögeli, B.; Kovacs, H.; Pervushin, K. *J. Biomol. NMR* **2005**, *31*, 1-9.
[9] Bax, A.; Clore, G.M.; Gronenborn, A.M. *J. Magn. Reson.* **1990**, *88*, 425-431.
[10] Olejniczak, E.T.; Xu, R.X.; Fesik, S.W. A. *J. Biomol. NMR* **1992**, *2*, 655-659.
[11] Tugarinov, V.; Kay, L.E. *J. Am. Chem. Soc.* **2003**, *125*, 13868-13878.
[12] Gardner, K.H.; Kay, L.E. *J. Am. Chem. Soc.* **1997**, *119*, 7599-7600.
[13] Goto, N.K.; Gardner, K.H.; Mueller, G.A.; Willis, R.C.; Kay, L.E. *J. Biomol. NMR* **1999**, *13*, 369-374.
[14] Serber, Z.; Richter, C.; Dötsch, V. *Chembiochem* **2001**, *2*, 247-251.
[15] Hu, K.F.; Eletsky, A.; Pervushin, K. *J. Biomol. NMR* **2003**, *26*, 69-77.
[16] Swapna, G.V.T.; Rios, C.B.; Shang, Z.G.; Montelione, G.T. *J. Biomol. NMR* **1997**, *9*, 105-111.
[17] Enggist, E.; Thöny-Meyer, L.; Güntert, P.; Pervushin, K. *Structure* **2002**, *10*, 1551-1557.
[18] Hu, K.F.; Plückthun, A.; Pervushin, K. *J. Biomol. NMR* **2004**, *28*, 405-406.
[19] Delaglio, F.; Grzesiek, S.; Vuister, G.W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277-293.
[20] Kadkhodaie, M.; Rivas, O.; Tan, M.; Mohebbi, A.; Shaka, A.J. *J. Magn. Reson.* **1991**, *91*, 437-443.
[21] Sørensen, O.W.; Rance, M.; Ernst, R.R. *J. Magn. Reson.* **1984**, *56*, 527-534.
[22] Braunschweiler, L.; Ernst, R.R. *J. Magn. Reson.* **1983**, *53*, 521-528.
[23] Kramer, F.; Luy, B.; Glaser, S.J. *Appl. Magn. Reson.* **1999**, *17*, 173-187.
[24] Parella, T. *J. Biomol. NMR* **2004**, *29*, 37-55.
[25] Zerbe, O.; Szyperski, T.; Ottiger, M.; Wüthrich, K. *J. Biomol. NMR* **1996**, *7*, 99-106.
[26] Pervushin, K.; Vögeli, B.; Eletsky, A. *J. Am. Chem. Soc.* **2002**, *124*, 12898-12902.
[27] Richter, G.; Kelly, M.; Krieger, C.; Yu, Y.H.; Bermel, W.; Karlsson, G.; Bacher, A.; Oschkinat,. *Eur. J. Biochem.* **1999**, *261*, 57-65.
[28] Nietlispach, D.; Clowes, R.T.; Broadhurst, R.W.; Ito, Y.; Keeler, J.; Kelly, M; Ashurst,J., Oschkinat, H.; Domaille, P.J.; Laue, E.D. *J. Am. Chem. Soc.* **1996**, *118*, 407-415.
[29] Shaka, A.J.; Keeler, J.; Freeman, R. *J. Magn. Reson.* **1983**, *53*, 313-340.
[30] Marion, D.; Bax, A. *J. Magn. Reson.* **1989**, *83*, 205-211.

Chapter VI

Measurements of Side-Chain ^{13}C-^{13}C Residual Dipolar Couplings in Uniformly Deuterated Proteins

Beat Vögeli[1], Helena Kovacs[2] and Konstantin Pervushin[1]

[1]Laboratorium für Physikalische Chemie, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

[2]Bruker BioSpin AG, Industriestrasse 26, CH-8117 Fällanden, Switzerland.

Journal of the American Chemical Society, 2004, 126, 2414-2420

1. Introduction

Uniform deuteration is frequently considered as a prerequisite for solution NMR structural studies of proteins with molecular weight larger than 30 kDa [1-6]. The use of deuteration in combination with transverse relaxation optimized spectroscopy (TROSY) enabled backbone resonance assignment in large soluble protein complexes [7-9] and integral membrane proteins solubilized in micelles [10-14]. Recently, uniform deuteration was proposed as a means to observe multiple long-range ^{1}H-^{1}H residual dipolar couplings (RDCs) even in small and medium-sized proteins [15,16]. Because protons are typically considered as the only structural probes located on the side chains, the depletion or complete replacement of side-chain protons with deuterons is frequently seen as an informational void in structural studies of large proteins [12]. Structural information for side chains not stemming from protons is available in the form of $^{13}C^{\alpha}$-$^{13}C^{\beta}$ RDCs [17,18]. Overall, the side-chain RDCs may prove indispensable for high-quality structure determinations of deuterated proteins. In addition, side-chain RDCs can potentially be employed to study side-chain dynamics [19,20].

The 2D ^{13}C-observe technology for isotopically enriched proteins was introduced by the Markley group [21,22]. On the basis of this, we proposed a new strategy for side-chain assignment in large uniformly deuterated proteins [23]. The key element in this approach is the nearly complete assignment of the side-chain ^{13}C resonances achieved using ^{13}C-start and ^{13}C-observe experiments in combination with the broadband homonuclear cross polarization. A notable feature of the ^{13}C-observe experiments is that multiple and redundant $^{1}J_{CC}$ scalar couplings are resolved as ^{13}C multiplets in the directly acquired ^{13}C dimension. In the current paper, we provide a theoretical and practical basis for measurements of homonulear ^{13}C-^{13}C RDCs stemming from deuterated side chains. We establish that, among different broadband polarization transfer schemes, the FLOPSY family [24] can be used to exchange magnetization between a *J*-coupled network of spins while largerly decoupling dipolar interactions between these spins. This property distinguishes FLOPSY from the WALTZ and MOCCA-SIAM rf-modulation schemes employed to exchange magnetization between dipolar coupled homonuclear spins [25,26]. Here, we demonstrate that (i) homonuclear dipolar decoupling during polarization transfer is important to eliminate dependence of line-shape distortions in TOCSY spectra of *J*-coupled spins [27,28] on the presence of dipolar interactions, which otherwise might result in biased ^{13}C-^{13}C RDC values, (ii) the homonuclear ^{13}C-^{13}C TOCSY provides an extensive set of side-chain ^{13}C-^{13}C RDCs which correlate well with the 3D structure and (iii) the ^{13}C-^{13}C RDCs can be effectively cross-validated using more conventional backbone ^{1}H-^{15}N RDCs measured at a lower degree of molecular alignment.

2. Experimental Section

RDCs were determined using a 1.4 mM u-^{2}H,^{13}C,^{15}N-labeled his_6-tagged human ubiquitin sample dissolved in 90%/10% H_2O/D_2O containing 10 mM potassium phosphate and 0.05% w/v sodium azide at pH 7.2. The sample with the aligned protein was obtained upon addition of Pf1 phage from a 50 mg/ml stock solution until the observed quadrupolar deuterium splitting of the $^{2}H_2O$ signal was 8 Hz. The [^{15}N,^{1}H]-correlation experiments were performed using a Bruker *AVANCE* 600 MHz spectrometer, equipped with a cryogenic Z-gradient TXI probe. The ^{13}C-start, ^{13}C-observe experiments were performed on a Bruker *AVANCE* 500 MHz spectrometer equipped with a cryogenic Z-gradient DUAL $^{13}C/^{1}H$ probe. All NMR experiments were performed at 20ºC. The 2D [^{13}C,^{13}C]-TOCSY experiment was run as described by Reference 23 except that FLOPSY-16 mixing was applied at γB_1 = 8.5 kHz and with τ_{mix} = 16.96 ms. The acquired 2D [^{13}C,^{13}C]-TOCSY spectra of u-^{2}H,^{13}C,^{15}N-labeled ubiquitin are shown in Figure 1. The spectra were analyzed using the program XEasy [29].

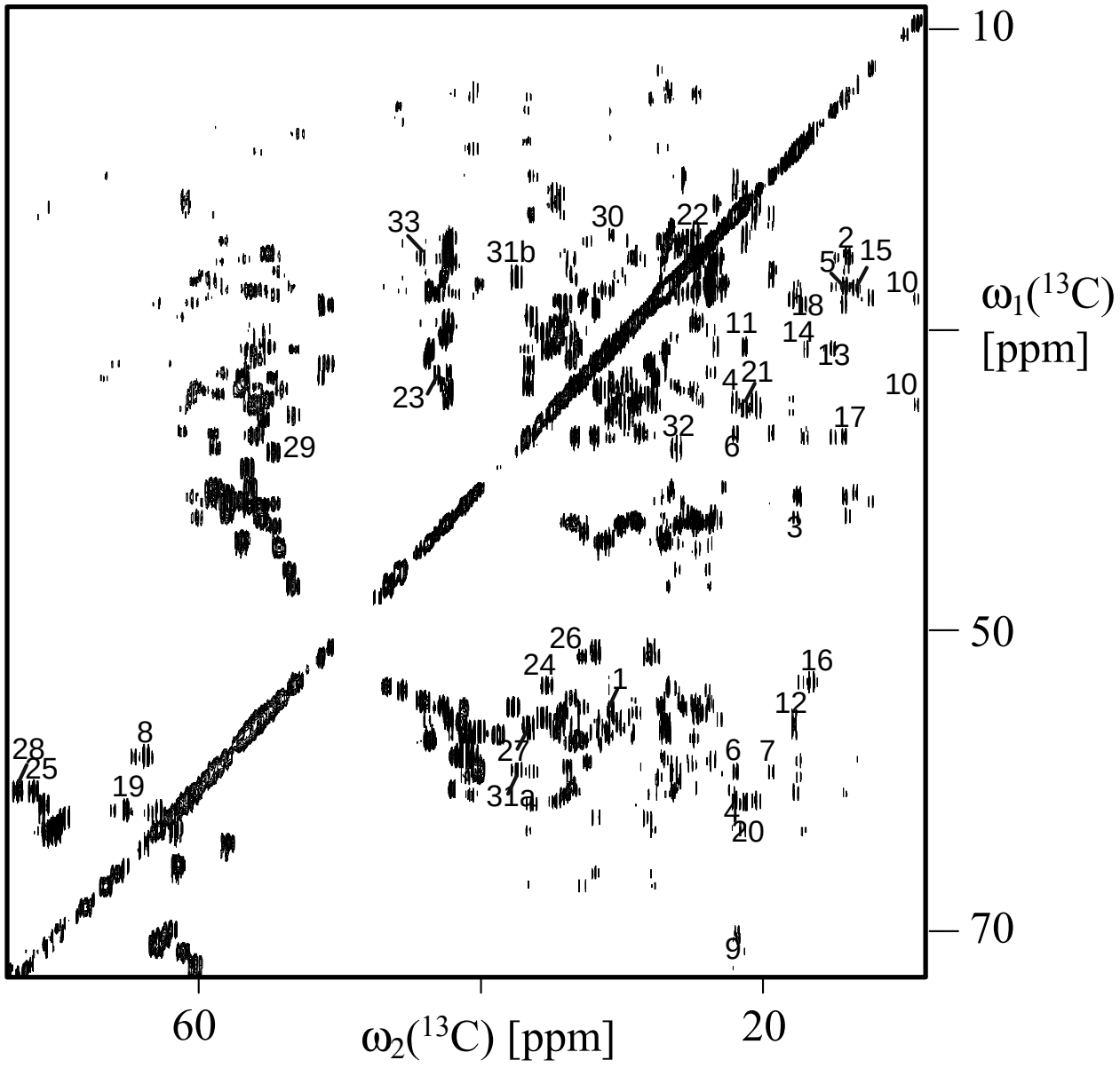


Figure 1. 2D [^{13}C,^{13}C]-TOCSY spectrum measured with uniformly ^{2}H,^{13}C,^{15}N-labeled ubiquitin (T = 293 K, 1.4 mM His-tagged human ubiquitin at pH 7.2 in 10 mM potassium phosphate buffer and 0.05% w/v sodium azide) recorded on a Bruker *Avance* spectrometer operating at a proton frequency of 500 MHz. The residual dipolar couplings were determined by comparing the splitting measured along the $^{13}C(\omega_2)$ dimension in a pair of the 2D [^{13}C,^{13}C]-TOCSY spectra recorded on samples without and with addition of Pf1 phage (quadrupolar deuterium splitting of $^{2}H_2O$ = 8 Hz). Selected cross-peaks whose multiplets were used in the analysis are marked with numbers, which correspond to the assignment listed in Tables 1 and 2. Letters count different cross-peaks which yield independently the same splitting values. In the experimental setup, the radio frequency carrier offsets were placed at 40 ppm (^{13}C), 173 ppm ($^{13}C'$), and 3.0 ppm (^{2}H). The 2D [^{13}C,^{13}C]-TOCSY was recorded with t_{1max} = 17.9 ms and t_{2max} = 203.2 ms, an interscan delay of 2 s, and 180 x 4096 complex points resulting in an acquisition time of 20 hours per spectrum. The mixing time was 16.96 ms. Before Fourier transform, the time domain data in the t_1 and t_2 dimensions were multiplied by a cosine function and zero-filled to 1024 and 16384 points, respectively.

3. Theoretical Basis

^{13}C-^{13}C and ^{1}H-^{15}N RDCs. For weak molecular alignment, the dipolar coupling between the spins *i* and *j*, D_{ij}, which depends on the axial component $A_a = (^3/_2)A_{zz}$ and the rhombicity $\eta = {}^2/_3(A_{xx} - A_{yy})/A_{zz}$ of the alignment tensor, where A_{xx}, A_{yy} and A_{zz} are the Cartesian components in the principle axis system, is defined as described by Reference 30:

$$D_{ij} = S(\mu_0/4\pi)\gamma_i\gamma_j /r_{ij}^3(h/2\pi^2)A_a \{(3\cos^2\theta - 1)/2 + {}^3/_4\ \eta(\sin^2\theta\cos 2\phi)\} \quad (1.1)$$

where θ and ϕ describe the polar angles of the mean orientation of the *i* - *j* interaction vector, S^2 is the generalized order parameter, μ_0 is the magnetic permeability of free space, γ_i is the gyromagnetic ratio of the spin *i*, r_{ij} is the *i*-*j* internuclear distance, and h is Planck's constant. Because on average the amplitude of D_{CC} is only $^1/_5$ of the D_{HN} values, it is instructive to measure ^{13}C-^{13}C and ^{1}H-^{15}N RDCs at higher and lower concentrations of the alignment media, respectively. In the irreducible tensor representation, the dipolar couplings are calculated as [31]:

$$D_{ij}^{clc} = \sum_{m=-2,2} A_m Y_{2m}(\theta_{ij}, \phi_{ij}) \quad (1.2)$$

where Y_{2m} is the spherical harmonics, and A_m corresponds to the irreducible components of the Saupe order matrix [32] and is obtained by a linear fit of the measured RDCs values against calculated *D* values. Because the phages are fully aligned at both concentrations [33,34], the Euler angles orienting the main axis system are expected to be the same, whereas the main components of the alignment tensor scale linearly with the Pf1 concentration. Under the assumption of a uniformly scaled S^2 for the C-C and H-N vectors, the independent measurements can be cross-validated: ^{13}C-^{13}C RDC values calculated from the scaled alignment tensor obtained from the ^{1}H-^{15}N measurements may be compared to the values calculated using the tensor obtained from the ^{13}C-^{13}C RDCs fit. In addition, the scalar product between the normalized five-dimensional alignment vectors (A_{-2}, A_{-1}, A_0, A_1, A_2) can be used as a measure of the degree of collinearity [31].

Line-Shape Distortions in TOCSY. In homonuclear 2D TOCSY spectra, line-shape distortions might arise due to the presence of the pulsed field gradient-insensitive zero-quantum coherence after the TOCSY mixing period along with the desired single spin polarization operators [27,28]. For a homonuclear dipolar and scalar coupled two-spin-½ system *IS*, the selected polarization transfer pathway under rf-irradiation resulting in the detectable signal is represented by Equation 2.

$$I_z \rightarrow T_{Iz\rightarrow Sz}(\tau) S_z + T_{Iz\rightarrow ZQ}(\tau) ZQ \quad (2)$$

where $ZQ \equiv -2(I_xS_y - I_yS_x)$, and the transfer functions $T_{Iz\rightarrow Sz}(\tau)$ and $T_{Iz\rightarrow ZQ}(\tau)$ are determined by spin-spin interaction Hamiltonians, chemical shifts of the involved spins, as well as the duration τ, strength, and modulation scheme of the rf irradiation [35].

A superposition of the residual dipolar coupling Hamiltonian, H_D, and the J-coupling Hamiltonian, H_J, renders $T_{Iz\rightarrow Sz}(\tau)$ and $T_{Iz\rightarrow ZQ}(\tau)$ as functions of J_{IS} and D_{IS} [36], thus resulting in different distortions of the line shapes in the presence or absence of spatial molecular alignment and, consequently, introducing a bias to the measured values of RDCs. To analyze these functional dependencies, the analytical effective chemical shift nonperturbed Hamiltonian H^{eff} is calculated using the exact effective Hamiltonian theory (EEHT) introduced by Untidt and Nielsen [37]. We demonstrate both analytically and computationally (see Appendix A) that the effective dipolar Hamiltonian is scaled down by a factor $s_D \approx -0.167$ when the FLOPSY pulse sequence is employed to exchange polarization between coupled ^{13}C spins. Thus, the problem of differential line-shape in the presence and absence of dipolar Hamiltonians is essentially avoided.

4. Results and Discussion

The 2D ^{13}C-^{13}C correlation spectroscopy enables detection of all types of ^{13}C spin systems found in proteins excluding methyl ^{13}C spins of methionine. Depending on the number of ^{13}C spins directly attached to the directly observed ^{13}C spin and neglecting effects of long range $^{>1}J_{CC}$-couplings, ^{13}C, ^{13}C cross-peaks in the 2D [^{13}C,^{13}C]-TOCSY spectrum are split in the ω_2 dimension into doublets, quadruplets (frequently with degenerate central components), and octets. ^{13}C-^{13}C RDCs can be effectively extracted from the doublet-split peaks stemming often from the structurally important methyl groups due to their good chemical shift dispersion and the splitting pattern. In addition, RDCs from quadruplet-split peaks stemming mostly from methylene groups can be measured. Figure 2 shows typical multiplet patterns extracted from 2D [^{13}C,^{13}C]-TOCSY spectra. In the first step of our analysis, 23 ^{13}C-^{13}C RDCs were determined by subtracting the splitting observed in doublets of the aligned sample from those of the unaligned sample (Table 1). Only the cross-peaks which exhibited a sufficiently high signal-to-noise ratio and absence of spectral overlap were analyzed. The alignment tensor was obtained by optimizing the fit of the measured ^{13}C-^{13}C RDCs values against D_{CC} values calculated on the basis of the X-ray crystal structure of human ubiquitin (PDB 1UBQ) [31,38], yielding a correlation coefficient of 0.988 and a root-mean-square deviation of 0.277 Hz (Figure 3a).

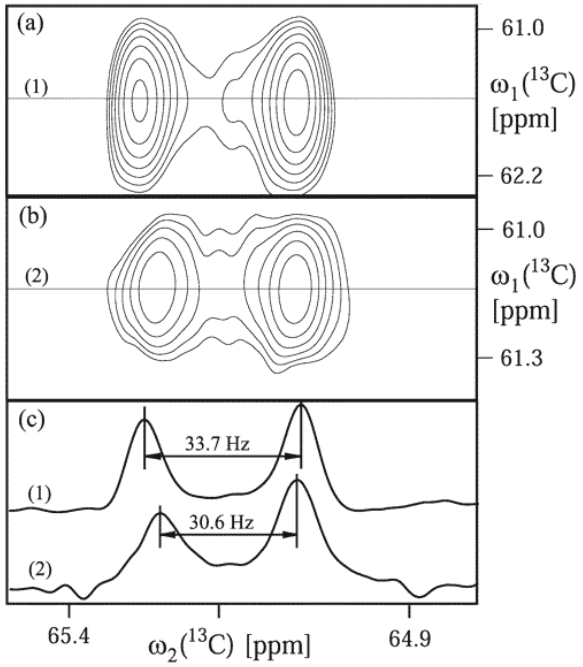


Figure 2. Expansions of the [$^{13}C^{\alpha}$,$^{13}C^{\beta}$] cross-peaks of Figure 1 (a) and upon addition of Pf1 (b). 1D cross-sections from the spectra (a) and (b) are shown in (c). The measured distances between the doublet components of Ser65 are $^{1}J_{C\alpha C\beta}$ = 33.7 Hz and ($^{1}J_{C\alpha C\beta} + D_{C\alpha C\beta}$) = 30.0 Hz, resulting in $D_{C\alpha C\beta}$ = -3.7 Hz.

In the second step, the alignment tensor calculated using the ^{13}C doublets was employed to predict the splitting of the outer components of the ^{13}C quadruplets. The measurements of the outer components yield the sum of the RDC contributions of each of the two attached ^{13}C spins 1 and 2, $D_{1CC} + D_{2CC}$ (Table 2). Theoretically, the contribution from each individual RDC can be calculated starting from the outermost ^{13}C atom, which has a coupling to only one neighbor carbon. From here, each successive carbon splitting can be decomposed into the already known contribution of the carbon further out on the side chain and the unknown contribution of the neighbor closer to the backbone. The relatively small values of ^{13}C-^{13}C RDCs and large uncertainties in the RDC measurements render this technique impractical. Instead, the sum $D_{1CC} + D_{2CC}$ is used as a structural constraint [39]. A correlation coefficient of 0.926 and a root-mean-square deviation of 0.772 Hz between predicted and experimental sums $D_{1CC} + D_{2CC}$ are obtained for the 10 cross-peaks, showing no overlap and a sufficient signal to noise ratio (see Figure 3b). Because these RDCs were not used in the fitting of the alignment tensor, the good fit between theoretical and experimental $D_{1CC} + D_{2CC}$ values indicates that the tensor is correct. In all cases where a given RDC could be measured independently from different cross-peaks, the deviation between the measurements was less than 0.3 Hz, and the average was used. The sole exception was the splitting of $^{13}C^{\gamma}$ of Glu64 which can be measured independently from two peaks which show equal signal-to-noise ratios and are both well resolved, where the determined values are 0 Hz and 1.9 Hz. In the case of 0 Hz, small peaks are

observed between the multiplet components, which might perturb the apparent couplings. This fact favors the choice of 1.9 Hz which also is closer to the theoretical value of about 1.5 Hz.

Table 1. Experimental and theoretical D_{CC} values derived from ^{13}C-doublets.

Cross-peak[1]	Residue	Atom	D_{CC}^{exp}, [Hz]	D_{CC}^{clc}, [Hz]
1	Met1	C^{γ}	-0.9	-0.804
2	Ile3	$C^{\delta1}$	-3.0	-2.885
3	Ile3	$C^{\gamma2}$	-1.0	-1.418
4	Val5	$C^{\gamma1}$	0.4	0.164
5	Ile13	$C^{\delta1}$	-0.6	-0.789
6	Val17	$C^{\gamma1}$	-1.1	-0.632
7	Val17	$C^{\gamma2}$	2.6	2.752
8	Ser20	C^{β}	3.2	3.453
9	Thr22	$C^{\gamma2}$	-0.1	-0.698
10	Ile23	$C^{\delta1}$	-3.0	-3.155
11	Val26	$C^{\gamma2}$	-1.1	-0.630
12	Ala28	C^{β}	1.1	1.001
13	Ile30	$C^{\delta1}$	2.0	1.891
14	Ile30	$C^{\gamma2}$	0.5	0.463
15	Ile36	$C^{\delta1}$	1.0	1.175
16	Ala46	C^{β}	-0.6	-0.308
17	Ile61	$C^{\delta1}$	0.9	0.433
18	Ile61	$C^{\gamma2}$	-3.2	-2.950
19	Ser65	C^{β}	-3.7	-3.638
20	Thr66	$C^{\gamma1}$	-0.7	-0.618
21	Val70	$C^{\gamma1}$	1.1	0.887
22	Leu73	$C^{\delta1}$	-0.5	-0.263
23	Arg74	C^{δ}	1.5	1.326

[1]Numbering of cross-peaks corresponds to the cross-peak annotation in Figure 1.

Table 2. Experimental and Theoretical $D_{1CC} + D_{2CC}$ Values Derived from the ^{13}C Quadruplets.

Cross-peak[1]	Residue	Atom	$(D_{1CC}+D_{2CC})^{exp}$, [Hz]	$(D_{1CC}+D_{2CC})^{clc}$, [Hz]
24	Glu18	C^{γ}	0.6	-0.568
25	Thr22	C^{β}	-1.7	-0.488
26	Pro38	C^{β}	2.9	2.667
27	Glu51	C^{γ}	-2.0	-2.104
28	Thr55	C^{β}	1.9	2.143
29	Asn60	C^{α}	0.5	-0.930
30	Lys63	C^{γ}	0.1	-0.491
31	Glu64	C^{γ}	1.9	1.504
32	Glu64	C^{β}	0.0	-0.626
33	Leu67	C^{β}	4.1	3.844

[1]Numbering of cross-peaks corresponds to the cross-peak annotation in Figure 1.

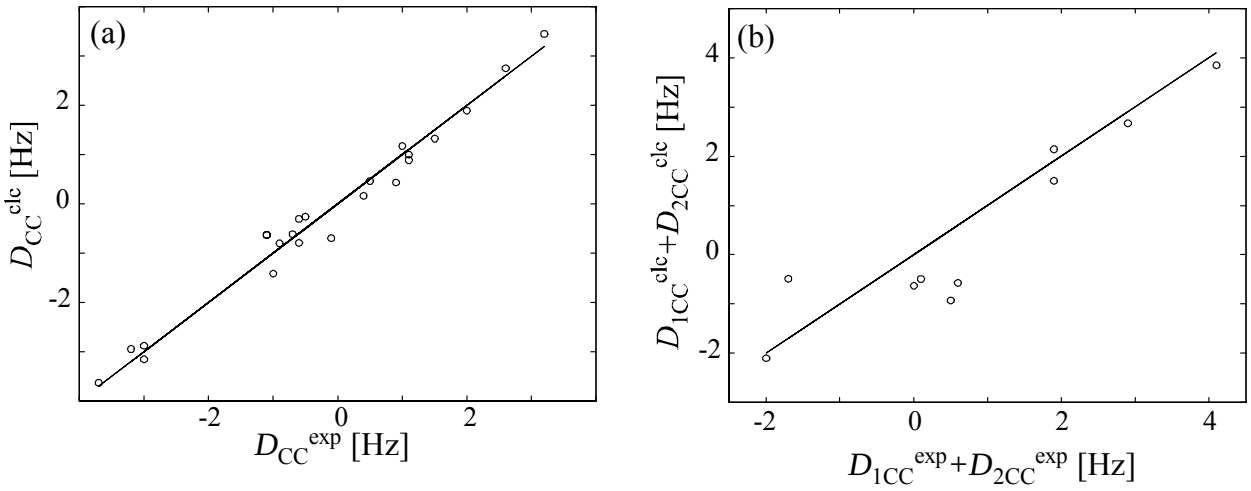


Figure 3. Calculated versus experimental values of D_{CC} of human ubiquitin measured using the 2D [^{13}C,^{13}C]-TOCSY spectra. In (a) and (b), the correlations between calculated D_{CC}^{clc} and D_{1CC}^{clc} + D_{2CC}^{clc} and experimental D_{CC}^{exp} and D_{1CC}^{exp} + D_{2CC}^{exp} values are derived from the ^{13}C-$^{13}\underline{\text{C}}$ doublets and ^{13}C$_1$-$^{13}\underline{\text{C}}$-^{13}C$_2$ quadruplets, respectively, where the $^{13}\underline{\text{C}}$ resonance is observed along the ω_2 dimension. The alignment tensor (see Table 3) was obtained by optimizing the fit between 23 D_{CC}^{exp} values obtained from ^{13}C-doublets against D_{CC} values calculated on the basis of the X-ray crystal structure of human ubiquitin [38]. In (a), the correlation coefficient is 0.988. In (b), the alignment tensor obtained with the use of ^{13}C-doublets was directly applied to calculate D_{1CC}^{clc} + D_{2CC}^{clc}, resulting in the correlation coefficient of 0.926.

The effects of the ^{13}C line-shape distortions due to the presence of the pulsed field gradient-insensitive zero-quantum coherencies after the TOCSY mixing period [27,28] on the ^{13}C-^{13}C RDCs measurements were analyzed by calculating the Taylor expansion of the exact effective Hamiltonian for a two spin system coupled by scalar and dipolar interactions. The FLOPSY pulse sequence largely decouples dipolar coupled spins, attenuating the effective dipolar coupling constant by a factor of –0.167. Although the ^{13}C shapes in the ω_2 dimension can be significantly perturbed by a zero-quantum coherence converted to the detectable magnetization by the last ^{13}C 90^o reading pulse, this perturbation is largely independent of the amplitude of RDCs, so that D_{CC} can still be reliably extracted as the difference between two measurements of the apparent maxima of the ^{13}C multiplets in the absence and presence of the alignment medium. For the parameters typically used for the ^{13}C-^{13}C TOCSY experiment, the residual, not suppressed part of the dipole/dipole coupling introduces a systematic bias to the measured values of D_{CC} at the level of 0.2%, which can be safely neglected.

The off-resonance propagation of the magnetization is analyzed by numerical calculations based on the H_D, H_J, and chemical shift Hamiltonians of the *IS* spin system in the rotating frame. Figure 4 shows the transfer functions $T_{Iz\rightarrow Sz}(\tau)$ and $T_{Iz\rightarrow ZQ}(\tau)$ calculated for τ = 16.96 ms with the variable offset frequency of the spin *S* while spin *I* was kept synchronized with the rotating frame using D_{IS} = 20 Hz, and J_{IS} = 35 Hz, and $\gamma_c B_1$ = 8.5 kHz. Although the off-resonance effects can increase the

relative efficiency of the polarization transfer under H_D and H_J Hamiltonians, the systematic bias to the values of D_{CC} does not exceed a level of 1% throughout the polarization transfer bandwidth.

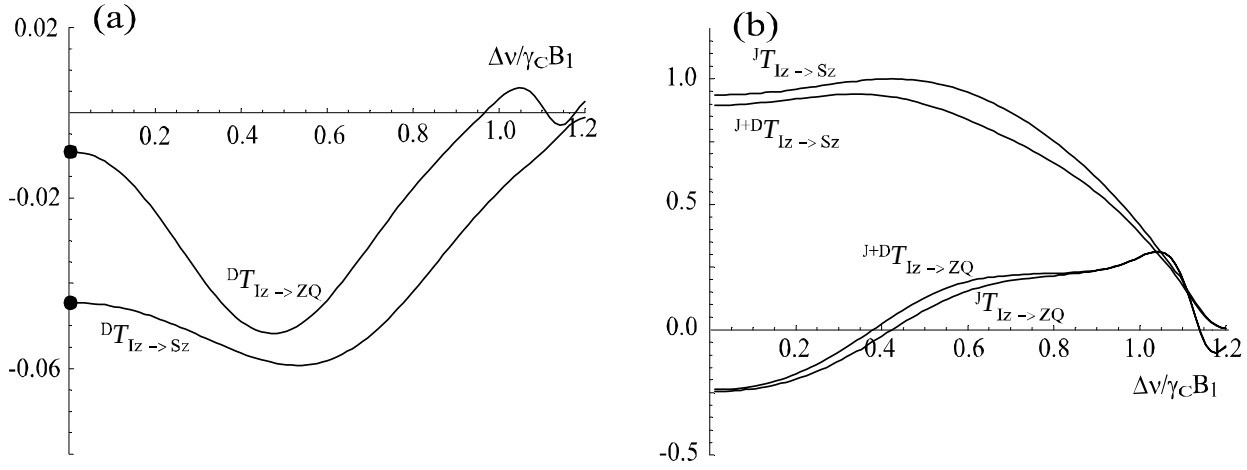


Figure 4. The polarization transfer functions $T_{Iz\rightarrow Sz}$ and $T_{Iz\rightarrow ZQ}$ calculated for the different offset frequencies of the spin *S* from the frequency of the spin *I* synchronized with the rotating frame. The evolution of the density operator I_{1z} via the [^{13}C,^{13}C]-TOCSY pulse sequence [23] with t_1= 0 ms, τ_{mix} = 16.96 ms, and γB_1 = 8.5 kHz was calculated in the presence of the H_D (a) and $H_D + H_J$ (b) Hamiltonians. The filled circles in (a) represent values of the transfer functions $T_{Iz\rightarrow Sz}$ and $T_{Iz\rightarrow ZQ}$ calculated using the effective Hamiltonian of Equations 5 and 6 for the FLOPSY-2 mixing sequence of the total duration $\tau_{\underline{RR}}$*24 = 16.96 ms at γB_1 = 8.5 kHz.

In general, for interacting spins *i* and *j* separated by the internuclear distance r_{ij}, D_{ij} is proportional to the product $\gamma_i \gamma_j / r_{ij}^3$. Therefore, one can predict that the maximal amplitude of D_{CC} in aliphatic side chains constitutes approximately $^1/_5$ of the amplitude of easily accessible backbone D_{HN} RDCs. Thus, to achieve optimal precision in both D_{CC} and D_{HN} measurements, it is reasonable to perform the latter at a somewhat lower concentration of alignment media. For example, backbone D_{HN} values in the range of ±20 Hz were achieved, which corresponds to an expected range of ±4 Hz for D_{CC} values. However, at these conditions, the signal-to-noise ratio in the [^{15}N,^{1}H]-correlation experiments was not sufficient for reliable alignment tensor calculations, so that the lower alignment degree was used to quantify D_{HN}. Therefore, a cross-validation of D_{HN} and D_{CC} required an appropriate scaling of the corresponding alignment tensors.

As an additional verification, we compared the alignment tensor to the corresponding alignment tensor derived from the independently measured backbone ^1H-^{15}N RDCs. Because of the severe line broadening observed in the [^{15}N,^{1}H]-correlation experiments, the lower Pf1 concentration was used. A comparison of the parameters of the tensors listed in Table 3 indicates a good correspondence between the two independently determined tensors. The Euler angles differ by less than 10º, and the scaling factor is 0.56±0.01 for all of the main components under the assumption of uniformly scaled

order parameters S^2. The theoretical values of the ^{13}C-^{13}C RDCs were calculated using the scaled up alignment tensor obtained from the ^{1}H-^{15}N measurements. A comparison with theoretical values obtained from the original alignment tensor yields a correlation coefficient of 0.967 and a root-mean-square deviation of 0.484 Hz (Figure 5). The high degree of collinearity of these two tensors is also reflected in the 0.976 value for the scalar product between the normalized five-dimensional alignment vectors corresponding to the two alignment tensors.

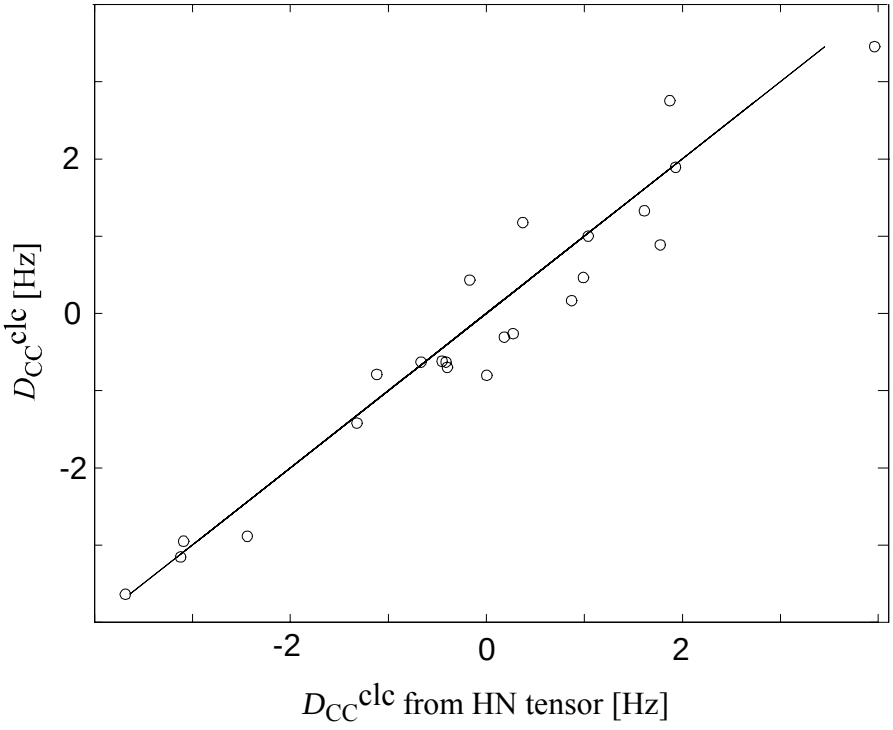


Figure 5. Calculated values of ^{13}C-^{13}C RDCs from the corresponding alignment tensor versus calculated values of ^{13}C-^{13}C RDCs using a scaled up alignment tensor obtained from the ^{1}H-^{15}N measurements. The correlation coefficient is 0.967, and the root-mean-square deviation 0.484 Hz.

Table 3. Alignment Tensor Parameters.

Measured RDCs	$(^3/_2)A_{zz}S$= A_aS [10^{-4}]	$(^3/_2)A_{xx}S$ [10^{-4}]	$(^3/_2)A_{yy}S$ [10^{-4}]	Rhombicity	α [deg]	β [deg]	γ [deg]
^{13}C-^{13}C	10.72	-1.19	-9.53	0.519	147.7	56.0	45.9
^{1}H-^{15}N	5.96	-0.68	-5.29	0.515	139.0	48.0	49.0

Because we introduce a scaling factor to correlate ^{13}C-^{13}C RDCs obtained from a ^{13}C-^{13}C tensor with ^{13}C-^{13}C RDCs derived from a ^{1}H-^{15}N tensor, possible systematic differences between averaged order parameters S_{CC}^2 and S_{HN}^2 are absorbed into the scaling factor. Although we expect certain

dynamical averaging of measured RDCs [40,41], the high degree of correlation between two alignment tensors independently derived from ^{13}C-^{13}C RDCs and ^{1}H-^{15}N RDCs might indicate applicability of our simplified approach.

The effect of mobility of individual side chains can be evaluated using the generalized order parameters of C^{γ} spins independently derived from experimental $^{3}J_{NC\gamma}$ and $^{3}J_{C'C\gamma}$ (S_J^2), dipolar couplings (S_D^2) [41], and ^{2}H relaxation rates (S_{rel}^2) [42]. For the long side chains of Ile3 ($C^{\gamma 2}$), Val5 ($C^{\gamma 1}$), Val17 ($C^{\gamma 1}$ and $C^{\gamma 2}$), Thr22 ($C^{\gamma 2}$), Val26 ($C^{\gamma 2}$), Ile30 ($C^{\gamma 2}$), Ile61 ($C^{\gamma 2}$), and Val70 ($C^{\gamma 1}$), for which all of these measurements are available, the averaged S^2 values usually do not deviate by more than 15% from an overall average value of 0.81 except for Val26 (+20%) and Val70 (-58%). Although a significant discrepancy between S_J^2, S_D^2 and S_{rel}^2 is observed for some amino acids (e.g., for Ile3, these values cover a range from 0.38 (S_D^2) to 0.98 (S_{rel}^2)), the overall mobility of the core side chains is rather restricted. Note that in eq 1.1 S instead of S^2 is used. The averaged S values all stay within ±8% except for Val26 (+10%) and Val70 (-35%).

To simulate the effect of nonuniform S^2 values, all of the measured values in Table 1 were randomly changed within a range of ±20% (this corresponds to changes of S^2 values within ±40%). The main components of the alignment tensor change by less than 3%, and the Euler angles change by less than 1.5º. The calculated RDCs do not change more than 0.25 Hz, resulting in the correlation coefficient of 0.982. Thus, with the precision of our ^{13}C-^{13}C RDCs measurements of around 15%, the use of uniform S^2 values for theoretical RDC calculations might be justified. It should be noted that, indeed, a weak correlation between $|D_{CC}^{exp}| - |D_{CC}^{clc}|$ and the average between S_J^2, S_D^2 and S_{rel}^2 was observed for all side chains with a sole exception of Val70, where the low order parameter does not correlate with a good match between D_{CC}^{exp} and D_{CC}^{clc}. A possible source of this discrepancy might be a slightly shorter ubiquitin construct lacking residue 76 used in the current studies as compared to the construct used for the S^2 measurements and X-ray analysis, calling for a more detailed investigation. In a second simulation, we assumed that two arbitrarily selected side chains are highly mobile. The measured RDCs of the cross-peaks 3 and 13 in Table 1 were artificially reduced in input to calculations by 50%, corresponding to the reduction of S^2 by 75%. Even in this case, the orientation of the alignment tensor is only weakly affected (the Euler angles deviate by less than 3º), but the main components are downscaled by 22%, 7%, and 9%. Thus, calculated RDCs showed a variation within 0.5 Hz as compared to the unperturbed input, yielding a slightly lower correlation coefficient of 0.972. It can be concluded that, although the alignment tensor is determined rather correctly, the angular information pertinent to the specific bonds may be biased, so that care has to be taken when highly mobile side chains are analyzed.

The presently described approach to measure ^{13}C-^{13}C RDCs is limited, on one hand, by spectral overlap and, on the other hand, by the complex multiplet structure of many ^{13}C-correlation cross-

peaks. This is usually observed for the ^{13}C spins covalently bound to ^{15}N or ^{13}C' spins. The availability of ^{13}C-detection cryo-probes and improvements in electronics enabling decoupling ^{15}N and ^{13}C' spins during signal acquisition shall significantly increase the number of cross-peaks amenable to extraction of useful RDCs. To alleviate spectral overlap, the use of spectrometers operating at higher magnetic field strength would be advantageous. A 2D [^{13}C,^{13}C]-TOCSY spectrum recorded at the proton frequency of 900 MHz using a ^{13}C-detection, ^{2}H-,^{15}N-decoupling probe-head is available in Appendix B (see also Chapter IV). Decoupling of ^{2}H, ^{15}N and ^{13}C' during signal acquisition significantly simplifies the multiplet structure of cross-peaks. New experiments can be developed which utilize several magnetization transfer steps. We expect that the application of spin-state editing techniques such as homonuclear E.COSY will further simplify spectra and increase the number of side-chain RDCs obtained in highly deuterated proteins. The applicability of E.COSY to large proteins can be estimated on the basis of the line-width at half-peak height ($\Delta\nu_{1/2}$) of the $^{13}C^{aliphatic}$ spins of about 15 Hz observed in spectra of 44 kDa uniformly ^{2}H,^{13}C,^{15}N-labeled BsCM [23]. Assuming that useful spectral information can be extracted as long as the condition $\Delta\nu_{1/2} < {}^{1}J_{CC}$ is fulfilled, one would estimate that E.COSY experiments will be feasible up to 100 kDa proteins. Thus, direct detection of ^{13}C resonances in perdeuterated proteins for extracting a large number of side-chain ^{13}C-^{13}C RDC constraints shows promise as a new method of facilitating high-quality structure determination even for proteins larger than 40 kDa.

5. Appendix A

Effective Hamiltonian for FLOPSY Mixing. The analytical expression is obtained for the effective chemical shift nonperturbed Hamiltonian H^{eff} using the exact effective Hamiltonian theory (EEHT) [37]. Because H_D and H_J commute and H_J is invariant to coordinate frame transformations for on-resonance spins, only the dipolar effective Hamiltonian is considered. A minimal supercycle averaging out the chemical shift Hamiltonian is the FLOPSY-2 sequence, of the form R$\underline{\text{R}}$, where the element R is a symmetric composite pulse, and the underline indicates a phase shift of all constituent pulses by 180º [24]. The action of the sequence can be described by the product of 18 propagators, $U_{R\underline{R}} = U_{18}\ldots U_2U_1$, where

$$U_i = exp[-i(\beta_i\cos\varphi_i(I_x + S_x) + \beta_i\sin\varphi_i\,(I_y + S_y) + 2\beta_i a_D\sqrt{6}T_{2,0})] \quad (3)$$

using the irreducible spherical tensor operator $T_{2,0}= (3I_zS_z-\boldsymbol{IS})/\sqrt{6}$, and where $a_D = \pi(2\pi D_{IS})/(2\omega_{rf})$ expresses the dependency on the angular frequencies for the dipolar coupling and the rf field amplitude $\omega_{rf} = 2\pi\gamma B_1$, and β_i is the rotational angle of the ith pulse applied with the phase φ_i. The

homonuclear dipolar coupling Hamiltonian in the rotating frame using the high field approximation is defined as

$$H_D = 2\pi D_{IS}\sqrt{6}T_{2,0} \quad (4)$$

The individual propagators are transformed into the matrix representation followed by transformation of the concatenated propagator $U_{R\underline{R}}$ into the coupled basis. The coupled basis is commonly used to reduce the dimensionality of the problem and facilitate interpretations of the obtained operators in terms of spectral transitions. To avoid excessively complicated expressions, $U_{R\underline{R}}$ was expanded into a Taylor series around $a_D = 0$, and the expansion was truncated at the 10th order of a_D. Using the closed solution to the usual infinite series expansion of the logarithm of the $U_{R\underline{R}}$ propagator [37], the effective Hamiltonian for the FLOPSY-2 sequence is obtained in the matrix form. The projection of the obtained effective Hamiltonian to the standard irreducible spherical tensor operators is given by Equation 5, where for simplicity the Taylor expansion around $a_D = 0$ is truncated at the 3rd order:

$$-iH^{eff}\tau_{R\underline{R}} = b_{2,1}T_{2,1} + (-b_{2,1})^*T_{2,-1} + b_{2,2}T_{2,2} + (-b_{2,2})^*T_{2,-2} + b_{2,0}\sqrt{6}T_{2,0} \quad (5)$$

and the symbol "*" denotes complex conjugation and $\tau_{R\underline{R}} = 23.55\,\tau_{90} = 23.55\pi/(2\omega_{rf})$, $-T_{2,\pm1} = \pm(I^{\pm}S_z + I_zS^{\pm})$ and $T_{2,\pm2} = \frac{1}{2}I^{\pm}S^{\pm}$.

$$b_{2,1} = (-0.05 + 2.50i)\,a_D + (3.99 + 0.37i)\,a_D^3 + O(a_D^5) \quad (6.1)$$

$$b_{2,2} = (-0.59 + 4.83i)\,a_D + (8.82 + 7.48i)\,a_D^3 + O(a_D^5) \quad (6.2)$$

$$b_{2,0} = -1.59i\,a_D - 1.26i\,a_D^3 + O(a_D^5) \quad (6.3)$$

Due to the reflection symmetry of the pulse sequence, all even-order terms of a_D vanish in $H^{eff}\tau_{R\underline{R}}$ [43], which is confirmed by an inspection of the Taylor expansion of Equation 6 to higher orders.

Expressions similar to Equations 5 and 6 can also be derived using the zero-order average Hamiltonian theory formulated for the windowless cyclic pulse sequences [44]. Although not essential for the current application, the EEH theory has a clear advantage of providing a Taylor series of the exact effective Hamiltonian even in the case of a complicated pulse sequence, whereas the higher-order approximations of the effective Hamiltonian would be increasesingly difficult to produce using the average Hamiltonian approach.

The form of the dipolar coupling Hamiltonian of Equation 4 is not strictly conserved in the effective Hamiltonian of Equations 5 and 6, as is the case for the DIPSI-2, WALTZ, MLEV, and MOCCA pulse sequences. Nonetheless, it is instructive to find a new, tilted coordinate frame, where the term proportional to $\sqrt{6}T'_{2,0}$ has the largest amplitude. This frame is established by

diagonalization of the normalized projection matrix $\langle H^{eff} | I_i S_j \rangle$ with i, j = x, y, and z [45] resulting in the effective Hamiltonian in the tilted frame with the z' axis tilted by 83.97º, 83.45º, and 8.92º from z, x, and y, respectively, and x' tilted by 148.48º from x:

$$H^{eff} = (-0.167)\, 2\pi D_{IS} \sqrt{6} T'_{2,0} - (0.017)\, 2\pi D_{IS} (2I_{x'}S_{x'} - 2I_{y'}S_{y'}) \quad (7)$$

Neglecting the second term in the Hamiltonian of Equation 7, the dipolar scaling factor [45] $s_D \approx -0.167$ is obtained, which can be compared to the s_D values of –0.5, 1, and 0.25 calculated for the DIPSI-2, MOCCA-XY16, and MLEV-16 schemes.

6. Appendix B

Figures 6-8 show a ^{2}H-, ^{15}N-, and $^{13}C'$-decoupled 2D [^{13}C,^{13}C]-TOCSY spectrum recorded on a Bruker *Avance* spectrometer operating at the proton frequency of 900 MHz. Details on the decoupling technique are presented in Chapter IV.

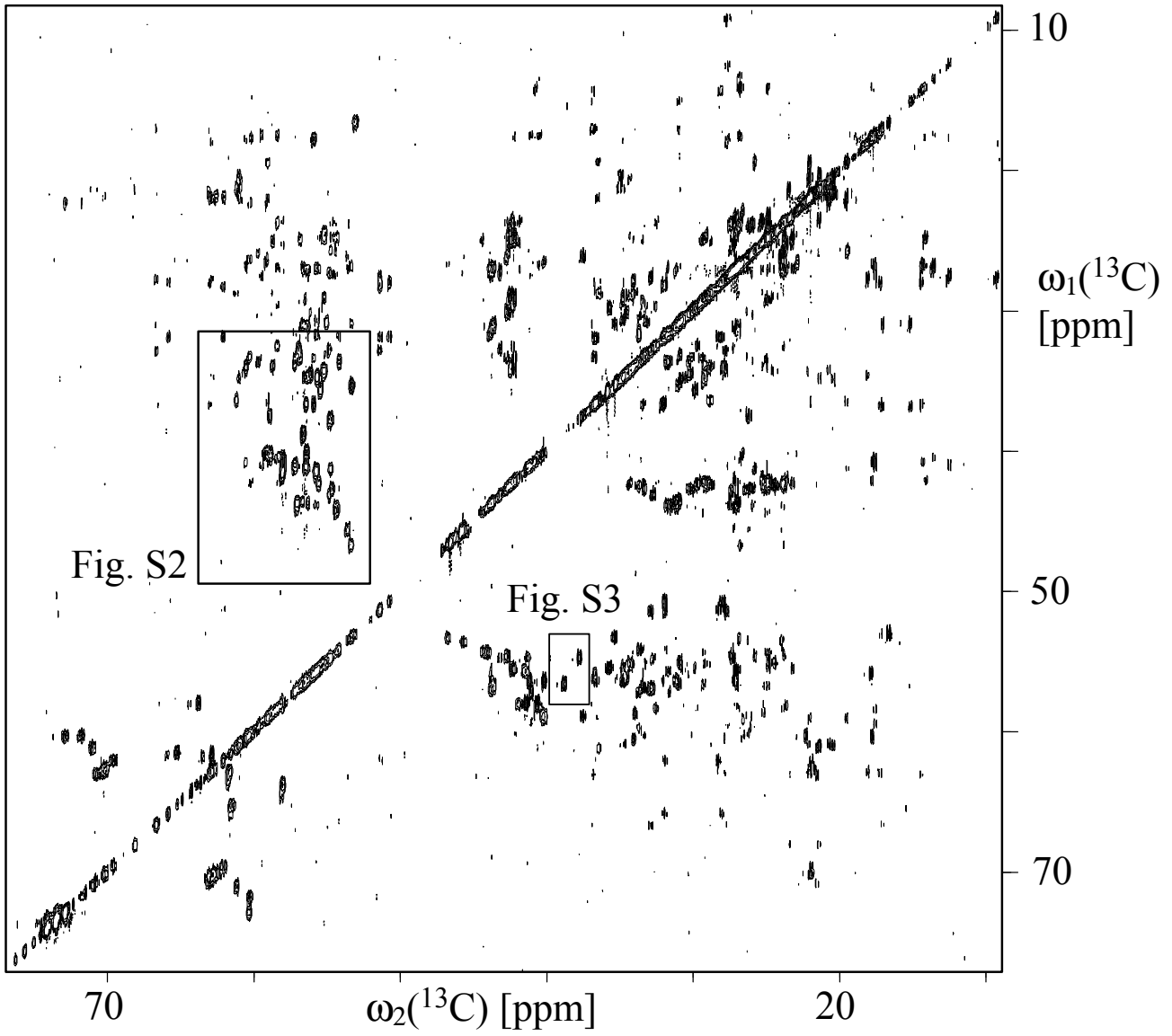


Figure 6 (next page). 2D [^{13}C,^{13}C]-TOCSY spectrum measured with uniformely ^{2}H,^{13}C,^{15}N-labeled ubiquitin (T=293 K, 1.4 mM human ubiquitin at pH 7.3 in 10 mM potassium phosphate buffer and 0.05% w/v sodium azide) recorded on a Bruker *Avance* spectrometer, equipped with a 5mm DUX 13C-1 probehead and operating at a proton frequency of 900 MHz. In the experimental setup the radio-frequency carrier offsets were placed at 40 ppm (^{13}C) and 174 ppm ($^{13}C'$). ^{2}H and ^{15}N was decoupled during both the indirect and direct evolution period. Additionally, bandselective homonuclear $^{13}C'$ decoupling was applied during the acquisition period. A detailed protocol of the decoupling technique is presented in Chapter IV. The FLOPSY-16 mixing was applied at γB_1 = 8.5 kHz and with τ_{mix} = 16.96 ms. The 2D [^{13}C,^{13}C]-TOCSY was recorded with t_{1max} = 14.1 ms and t_{2max} = 147.5 ms, an interscan delay of 1.5 s and 256 x 4096 complex points, resulting in an acquisition time of 18 hours. Before Fourier transform the time domain data in the t_1 and t_2 dimensions were multiplied with a cosine function and zero-filled to 1024 and 16384 points, respectively. The boxes show the regions which are expanded in Figures 7 and 8.

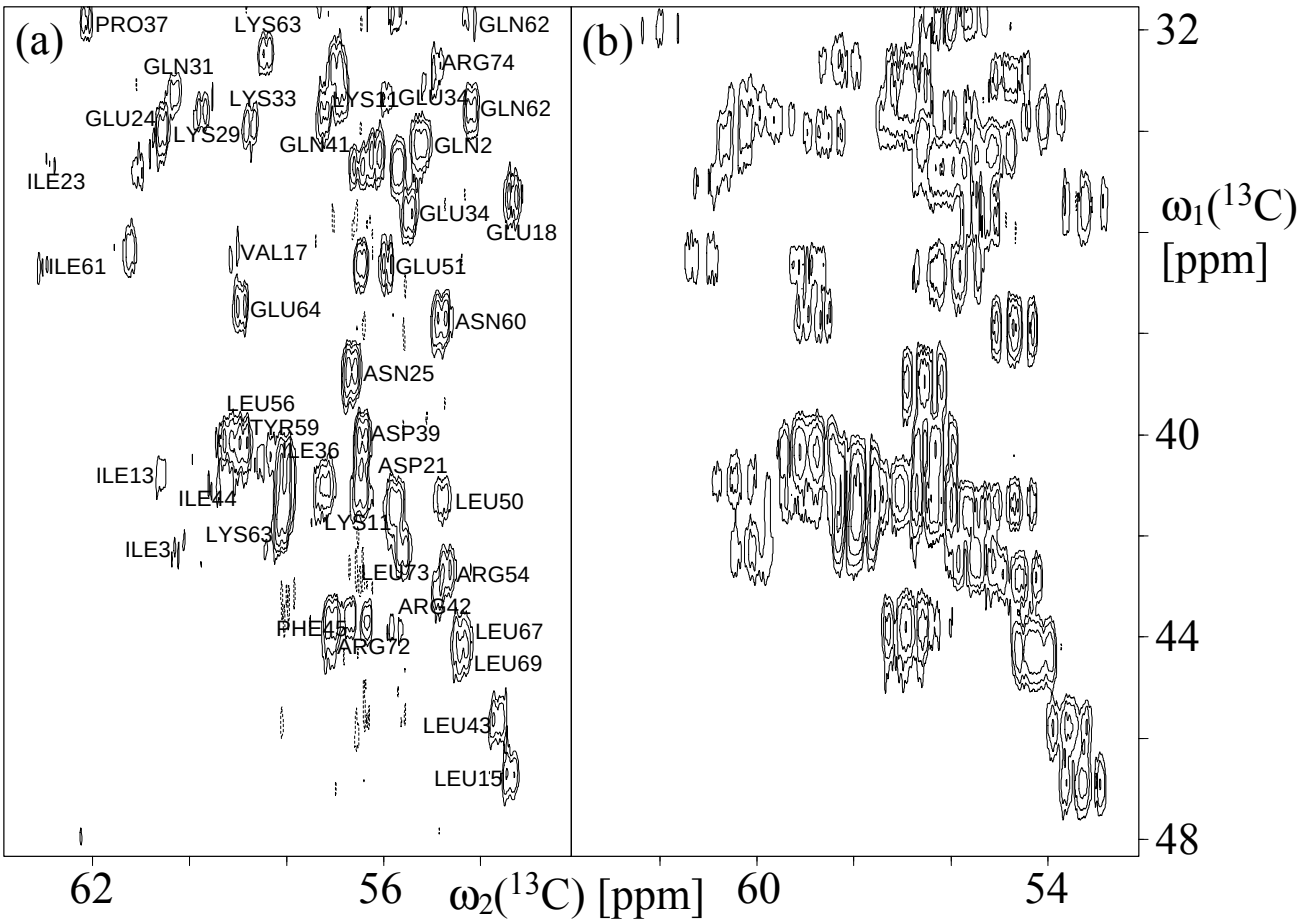


Figure 7. Expansions of the 2D [^{13}C,^{13}C]-TOCSY spectra of Figure 6 (a) and Figure 1 (b) of a region with cross-peaks at C^{α} frequencies along the direct dimension, measured at 900 MHz and 500 MHz, respectively. Decoupling of ^{2}H, ^{15}N and $^{13}C'$ applied in (a) results in simplified patterns in the direct dimension.

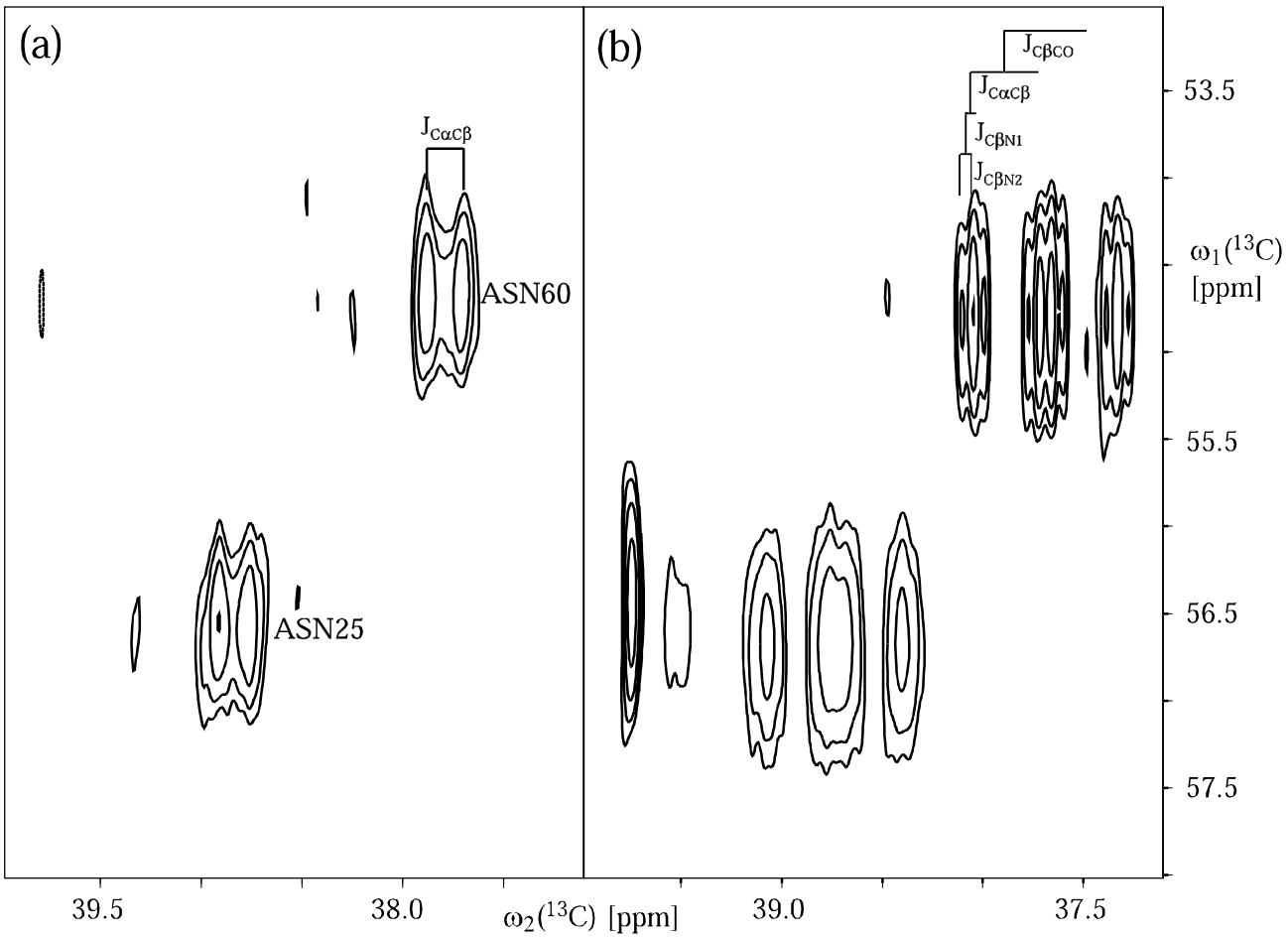


Figure 8. Expansions of the 2D [^{13}C,^{13}C]-TOCSY spectra of Figure 6 (a) and Figure 1 (b) showing the C^{β}/C^{α}-cross-peaks of Asn25 and Asn60 measured at 900 MHz and 500 MHz, respectively. Decoupling of ^{2}H, ^{15}N and $^{13}C'$ simplifies the splitting pattern along the direct dimension due to $^{1}J_{C\beta CO}$, $^{1}J_{C\alpha C\beta}$, $^{2}J_{C\beta N1}$ and $^{2}J_{C\beta N2}$ (b) to a doublet split by $^{1}J_{C\alpha C\beta}$ only (a), where N1 and N2 stand for ^{15}N in the side-chain and backbone, respectively.

7. References

[1] Grzesiek, S.; Anglister, J.; Ren, H.; Bax, A. *J. Am. Chem. Soc.* **1993**, *115*, 4369-4370.
[2] Browne, D.T.; Kenyon, G.L.; Packer, E.L.; Sternlicht, H.; Wilson, D.M. *J. Am. Chem. Soc.* **1973**, *95*, 1316-1323.
[3] LeMaster, D.M. *Method Enzymol.* **1989**, *177*, 23-43.
[4] LeMaster, D.M. *Ann. Rev. of Biophysics and Biophysical Chemistry* **1990**, *19*, 243-266.
[5] LeMaster, D.M. *Q. Rev. Biophys.* **1990**, *23*, 133-174.
[6] Venters, R.A.; Farmer, B.T.; Fierke, C.A.; Spicer, L.D. *J. Mol. Biol.* **1996**, *264*, 1101-1116.
[7] Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *J. Am. Chem. Soc.* **2000**, *122*, 7543-7548.
[8] Takahashi, H.; Nakanishi, T.; Kami, K.; Arata, Y.; Shimada, I. *Nat. Struct. Biol.* **2000**, *7*, 220-223.
[9] Tugarinov, V.; Muhandiram, R.; Ayed, A.; Kay, L.E. *J. Am. Chem. Soc.* **2002**, *124*, 10025-10035.
[10] Fernandez, C.; Adeishvili, K.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 2358-2363.
[11] Fernandez, C.; Hilty, C.; Bonjour, S.; Adeishvili, K.; Pervushin, K.; Wüthrich, K. *FEBS Lett.* **2001**, *504*, 173-178.
[12] Arora, A.; Tamm, L.K. *Curr. Opin. Struct. Biol.* **2001**, *11*, 540-547.
[13] Hwang, P.M.; Choy, W.Y.; Lo, E.L.; Chen, L.; Forman-Kay, J.D.; Raetz, C.R.; Prive, G.G.; Bishop, R.E.; Kay, L.E. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13560-13565.
[14] Schubert, M.; Kolbe, M.; Kessler, B.; Oesterhelt, D.; Schmieder, P. *Chembiochem* **2002**, *3*, 1019-1023.
[15] Meier, S.; Haussinger, D.; Jensen, P.; Rogowski, M.; Grzesiek, S. *J. Am. Chem. Soc.* **2003**, *125*, 44-45.
[16] Wu, Z.R.; Bax, A. *J. Am. Chem. Soc.* **2002**, *124*, 9672-9673.
[17] Permi, P.; Rosevear, P.R.; Annila, A. *J. Biomol. NMR* **2000**, *17*, 43-54.
[18] Evenas, J.; Mittermaier, A.; Yang, D.W.; Kay, L.E. *J. Am. Chem. Soc.* **2001**, *123*, 2858-2864.

[19] Chou, J.J.; Bax, A. *J. Am. Chem. Soc.* **2001**, *123*, 3844-3845.
[20] Mittermaier, A.; Kay, L.E. *J. Am. Chem. Soc.* **2001**, *123*, 6892-6903.
[21] Oh, B.H.; Westler, W.M.; Darba, P.; Markley, J.L. *Science* **1988**, *240*, 908-911.
[22] Westler, W.M.; Kainosho, M.; Nagao, H.; Tomonaga, N.; Markley, J.L. *J. Am. Chem. Soc.* **1988**, *110*, 4093-4095.
[23] Eletsky, A.; Moreira, O.; Kovacs, H.; Pervushin, K. *J. Biomol. NMR* **2003**, *26*, 167-179.
[24] Kadkhodaie, M.; Rivas, O.; Tan, M.; Mohebbi, A.; Shaka, A.J. *J. Magn. Reson.* **1991**, *91*, 437-443.
[25] Möglich, A.; Wenzler, M.; Kramer, F.; Glaser, S.J.; Brunner, E. *J. Biomol. NMR* **2002**, *23*, 211-219.
[26] Hansen, M.R.; Rance, M.; Pardi, A. *J. Am. Chem. Soc.* **1998**, *120*, 11210-11211.
[27] Braunschweiler, L.; Ernst, R.R. *J. Magn. Reson.* **1983**, *53*, 521-528.
[28] Sørensen, O.W.; Rance, M.; Ernst, R.R. *J. Magn. Reson.* **1984**, *56*, 527-534.
[29] Bartels, C.; Xia, T.H.; Billeter, M.; Güntert, P.; Wüthrich, K. *J. Biomol. NMR* **1995**, *6*, 1-10.
[30] Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111-1114.
[31] Sass, J.; Cordier, F.; Hoffmann, A.; Cousin, A.; Omichinski, J.G.; Lowen, H.; Grzesiek, S. *J. Am. Chem. Soc.* **1999**, *121*, 2047-2055.
[32] Saupe, A. *Z. Naturforsch.* **1964**, *19A*, 161-171.
[33] Hansen, M.R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065-1074.
[34] Clore, G.M.; Gronenborn, A.M. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5891-8.
[35] Kramer, F.; Luy, B.; Glaser, S.J. *Appl. Magn. Reson.* **1999**, *17*, 173-187.
[36] Luy, B.; Glaser, S.J. *J. Magn. Reson.* **2001**, *148*, 169-181.
[37] Untidt, T.S.; Nielsen, N.C. *Phys. Rev. E* **2002**, *65*, art. no.-021108.
[38] Vijaykumar, S.; Bugg, C.E.; Cook, W.J. *J. Mol. Biol.* **1987**, *194*, 531-544.
[39] Ottiger, M.; Delaglio, F.; Marquardt, J.L.; Tjandra, N.; Bax, A. *J. Magn. Reson.* **1998**, *134*, 365-369.
[40] Tolman, J.R.; Al-Hashimi, H.M.; Kay, L.E.; Prestegard, J.H. *J. Am. Chem. Soc.* **2001**, *123*, 1416-1424.
[41] Chou, J.J.; Case, D.A.; Bax, A. *J. Am. Chem. Soc.* **2003**, *125*, 8959-8966.
[42] Lee, A.L.; Flynn, P.F.; Wand, A.J. *J. Am. Chem. Soc.* **1999**, *121*, 2891-2902.
[43] Mansfiel, P.; Orchard, M.J.; Stalker, D.C.; Richards, K.H. *Phys. Rev. B* **1973**, *7*, 90-105.
[44] Burum, D.P.; Linder, M.; Ernst, R.R. *J. Magn. Reson.* **1981**, *44*, 173-188.
[45] Kramer, F.; Glaser, S.J. *J. Magn. Reson.* **2002**, *155*, 83-91.

Chapter VII

Observation of Individual Transitions in Magnetically Equivalent Spin Systems

*Konstantin Pervushin and Beat Vögeli**

Laboratorium für Physikalische Chemie, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

* Spectroscopy

Journal of the American Chemical Society, 2003, 125, 9566-9567

1. Introduction

Individual transitions of magnetically equivalent spin systems such as methyl groups residing in tumbling molecules in isotropic solution cannot be observed as multiplet-split NMR lines. Construction of an NMR experiment overcoming this limitation and thus enabling direct and selective observation of individual ^{1}H transitions in methyl spin systems would bring a number of advantages. This would enable (i) TROSY-type resolution enhancement [1] in the directly acquired $^1H^{Methyl}$ dimension of multi-dimensional NMR experiments by constructive use of the interference between $^1H^{Methyl}$ chemical shift anisotropy (CSA) and $^1H^{Methyl}$-$^1H^{Methyl}$ dipole-dipole (DD) interactions [2], (ii) direct measurement of cross-correlation between Curie spin relaxation (CSR) and $^1H^{Methyl}$-$^1H^{Methyl}$ DD relaxation in paramagnetic proteins [3-6], from which the distance and orientation of the methyl group relative to the paramagnetic center can be calculated [7,8]; (iii) measurement of the values and sign of residual $^1H^{Methyl}$-$^1H^{Methyl}$ dipolar couplings [9,10], D_{HH}, in proteins weakly aligned by means of either an external medium [11-13], the natural anisotropy of diamagnetic susceptibility [14,15] or an internal orienting device such as a paramagnetic center [16,17]. The absence of scalar splitting in methyl groups eliminates the need to acquire an otherwise indispensable reference spectrum in isotropic phase or in the diamagnetic state of the protein as it the case for D_{CH} measurements in methyl groups [18]. In this Chapter, we describe such an experiment and demonstrate its use with two proteins weakly aligned by means of either Pf1 phages or a naturally present paramagnetic heme group.

The presented technique allows spin-state selection in CH_3 groups, which is not the case in previous methods. In the work of Sibille et al. [9] spin-state selection is achieved in CH_2D methyl groups, and in the work of Kaikkonen et al. [10], which applies to CH_3 groups, the ^{1}H-^{1}H coupling is measured from an antiphase splitting. The acquisition of the resolved ^{1}H-^{1}H antiphase spectrum is crucial for the successful measurements of D_{HH}. If D_{HH} values are smaller than the $^1H^{Methyl}$ line-width, the negative and positive ^{1}H-^{1}H multiplet components in the antiphase spectrum largely cancel each other. This might result in strongly biased values of D_{HH} or the total absence of spectrum as it is the case in magnetically equivalent ^{1}H methyl protons. The pulse sequence of Figure 1 was designed to alleviate the need to acquire the resolved ^{1}H-^{1}H antiphase spectrum by directly polarizing chosen individual components of the generally degenerate (under conditions of isotropic tumbling) $^1H^{Methyl}$ triplet. The individual components are not postacquisitionally reconstructed from in-phase and antiphase-type of subspectra [18] but rather are selectively polarized and directly observed as positive spectral intensity using a single application of the pulse sequence. Thus, the very small values of D_{HH} can be measured.

2. Methods

Considering evolution only under the large $^1J_{HC}$ = 125 Hz couplings in CH_3 groups, at the time point *a* one obtains the density operator $\sigma_a = (2S_xI_{1z} + 2S_xI_{2z} + 2S_xI_{3z}) \cos(\omega_S t_1)$, where *S* and *I* stand for the ^{13}C and ^1H spins, respectively. The effect of the subsequent unitary transformations of the density operator during the pulse sequence element enclosed in the box (Figure 1) can be analyzed by calculating the projections of the density operators at time points *a* and *b*, $<\sigma_a|\sigma_b>$ [19]. For this purpose, the program POMA [20] modified to account for the effects of pulsed magnetic field gradients was used. As target operators we choose the single-transition operators $I_{1i}^{\alpha\alpha} = I_{1i}(E/2 + I_{2z})(E/2 + I_{3z})$, $I_{1i}^{\beta\beta} = I_{1i}(E/2 - I_{2z})(E/2 - I_{3z})$ defining two "outer" transitions and $I_{1i}^{\alpha\beta} = I_{1i}(E/2 + I_{2z})(E/2 - I_{3z})$ and $I_{1i}^{\beta\alpha} = I_{1i}(E/2 - I_{2z})(E/2 + I_{3z})$ defining two "central" transitions, where i = x,y. Because all three methyl protons are equivalent, a complete description of the ^1H triplet includes summation over the single-transition operators with indices 1, 2 and 3, where the relevant operators can be derived by circular permutation of their indices. For all values of the phase γ and factor κ (Figure 1) the "x"-projections $<2S_xI_{1z}|I_{1x}^{ij}>$, $i,j = \alpha,\beta$ vanish and, thus, will not be considered further. Numerical analysis of the "*y*"-projection $<2S_xI_{1z}|I_{1y}^{\alpha\beta}+I_{1y}^{\beta\alpha}>$ as a function of γ and κ shows that at κ = 6.577 the central transitions are not polarized with any choice of γ. Calculating $<2S_xI_{1z}|I_{1y}^{\alpha\alpha}>(\gamma)|_{\kappa=6.577}$ and $<2S_xI_{1z}|I_{1y}^{\beta\beta}>(\gamma)|_{\kappa=6.577}$ as a function of γ (Figure 2) shows that at γ= 150º and 30º, the ^{13}C antiphase magnetization evolving during the t_1 period is transferred exclusively to the $I_{1y}^{\alpha\alpha}$ and $I_{1y}^{\beta\beta}$ operators, respectively. Thus, in the absence of relaxation and the effects of passive couplings the experimental scheme of Figure 1 exclusively polarizes a chosen individual transition in a degenerate multiplet of magnetically equivalent ^1H. The element in the box in Figure 1 represents a new type of polarization transfer scheme, where the manifold of transitions of an insensitive spin is funneled exclusively to a selected transition of a system of sensitive spins, enabling its direct spectral observation even in the case of the multiplet degeneracy. The experimental scheme of Figure 1 can be run using the constant-time ^{13}C chemical shift evolution when the higher spectral resolution along the ^{13}C dimension is needed.

For molecules weakly aligned in the magnetic field a new RDC term in the Hamiltonian, namely $H = -D_{HH}\Sigma I_z^i I_z^j$, where $i < j$ and $i,j = \{1,2,3\}$, should be considered, where D_{HH} is the amplitude of the residual dipolar couplings [9,10]. The RDC Hamiltonian term does not interchange individual transitions of the ^1H multiplet. Therefore, the evolution of σ_{acq} during the acquisition time can be interpreted as a spectral separation of the $I^{\alpha\alpha}$ and $I^{\beta\beta}$ lines by $2D_{HH}$. Since these spectral transitions are observed in different experiments, the sign of D_{HH} can be inferred from the relative shifts of the $I^{\alpha\alpha}$ and $I^{\beta\beta}$ components.

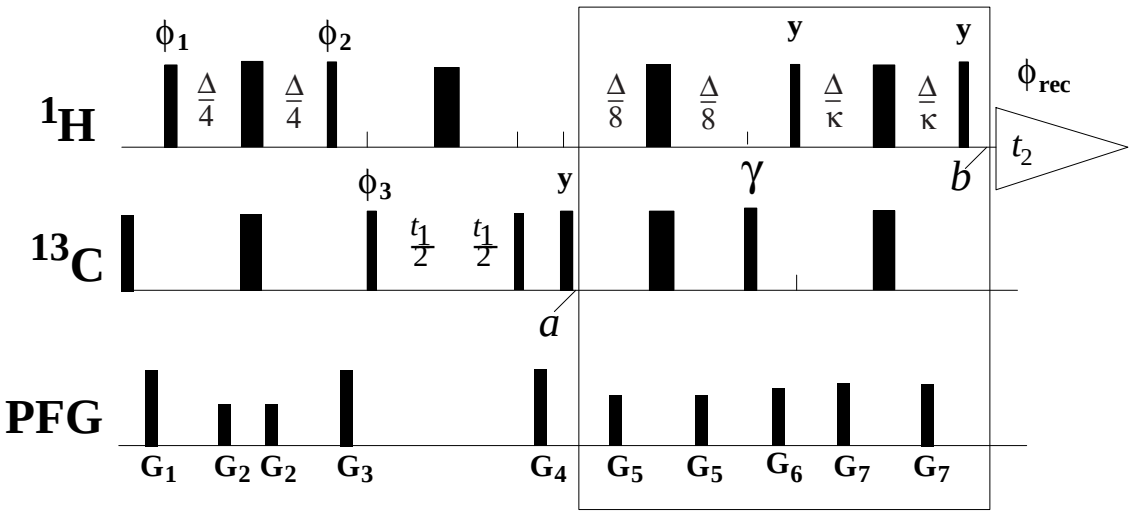


Figure 1. Experimental scheme of the 2D $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$ and $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$ HSQC experiments. The radio frequency pulses on ^{13}C and ^{1}H are applied at 20 and 4.8 ppm, respectively. Narrow and wide black bars indicate nonselective 90° and 180° pulses, respectively. The line marked PFG indicates the duration and strength of pulsed magnetic field gradients applied along the z-axis: G_1: 800 μs, 80 G/cm, G_2: 800 μs, 20 G/cm, G_3: 1 ms, -90 G/ cm, G_4: 1.2 ms, 80 G/cm, G_5: 400 μs, 55 G/cm, G_6: 1 ms, 65 G/cm and G_7: 400 μs, 70 G/cm. The delay $\Delta = 1/(^{1}J_{CH})$ is set to 2 ms. The factor $\kappa = 6.577$. The phases are defined as follows: $\phi_1 = \{x, -x\}$, $\phi_2 = \{2y,2(-y)\}$, $\phi_3 = \{4x, 4(-x)\}$ and $\phi_{rec} = \{y, -y,-y, y,-y, y, y,-y\}$. The phase γ is set to 150° and 30° in order to obtain the 2D $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$ and $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$ HSQC spectrum, respectively. Quadrature detection in the $^{13}C(t_1)$ dimension is achieved by the States-TPPI method applied to the phase ϕ_3.

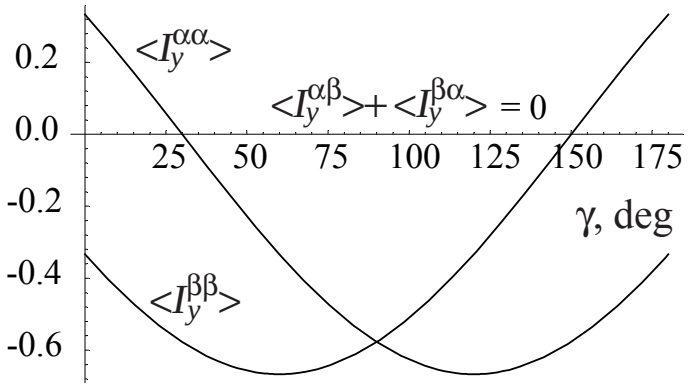


Figure 2. Amplitudes of projections of the $2S_xI_{1z}$ operator to the single-transition operators $\langle I_y^{ij}\rangle = \langle 2S_xI_{1z}|I_y^{ij}\rangle$ where $i,j = \alpha,\beta$ as a function of the phase γ of the ^{13}C 90° pulse.

The effect of the RDC Hamiltonian on the quality of selection of individual components is estimated by numerical evaluation of the density operator throughout the pulse sequence of Figure 1 using the program NMRSIM (Bruker BioSpin). In simulations of the 2D $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$ HSQC with D_{HH} in the range of 0-20 Hz and $D_{HC} = {}^{1}/_{2.3}D_{HH}$, the $I^{\alpha\alpha}$ component was selected, and spectral intensities at the central and $I^{\beta\beta}$ components were monitored. At D_{HH} = 20 Hz, the intensity of the

central and ββ transitions are 15% of the intensity of the selected component. Thus, we conclude that the proposed experiment should exhibit a good tolerance for the presence of RDC and passive scalar coupling interactions described by similar Hamiltonians.

3. Results and Discussion

The performance of the experiments of Figure 1 is demonstrated with two proteins, the diamagnetic heme chaperone apo-CcmE [21] weakly aligned in the presence of Pf1 phages [22], and its paramagnetic, heme group-containing counterpart, holo-CcmE, complexed with KCN. The sensitivity (*S/N*) of 2D $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$- and $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$-HSQC spectra measured with both proteins is on average a factor of 0.5 smaller to that observed in the regular $[^{13}C,^{1}H]$-HSQC (a theoretical attenuation factor is 0.58 without considering relaxation effects). At the achieved maximal *S/N* = 20 no cross-talk between $I^{\alpha\alpha}$ and $I^{\beta\beta}$ components of the ^{1}H triplet was detected in the spectra. The values and signs of D_{HH} can be easily extracted from the 1D ^{1}H slices taken through the corresponding peaks in the two spectra. These values are reported in the form of histograms in Figure 4 in the Appendix. By repeating the measurements the statistical variation of the obtained values of D_{HH} was estimated to be in the range of 0.4 Hz. It should be noted that for holo-CcmE the proposed experiment provided values of D_{HH} which are significantly smaller than the line-width of ^{1}H resonances, so that alternative methods designed to detect the individual transitions in the form of the ^{1}H antiphase magnetization would fail [5]. This fact bears special importance for paramagnetic proteins where ^{1}H lines are typically broadened by dipolar interactions with electronic spin. Very attractive is the absence of the need for a diamagnetic form as a reference state, which might be difficult or impossible to prepare as is the case for CcmE [12].

For many cross-peaks stemming from the rigid core of the CcmE proteins, significantly different line-width and concomitant spectral intensities were observed (for example compare γ2 Leu34 and Leu43, Figure 3, a and b). Due to the relatively large $^{1}H^{Methyl}$-CSA values of $\Delta\sigma$ = 10 ppm [4], smaller line-width is expected for the $I^{\beta\beta}$ component of the ^{1}H triplet at high polarizing magnetic fields, thus making the 2D $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$-HSQC an attractive building block for TROSY-type spectroscopy with methyl groups.

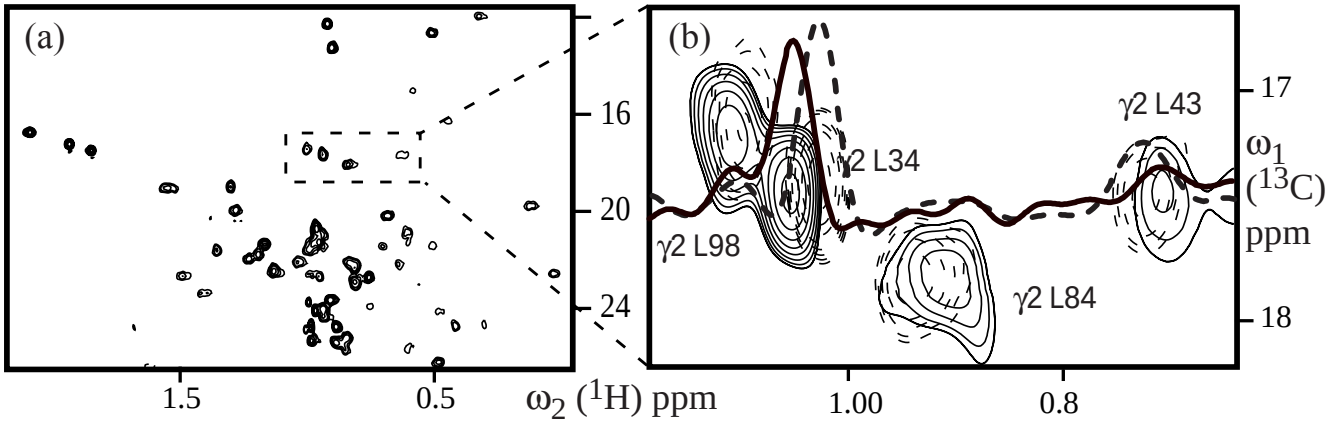


Figure 3. (a) 2D $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$-HSQC spectrum and (b) an expansion showing $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$-(solid lines) and $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$-(dashed lines) HSQC spectra of the 121 amino-acid fragment of uniformly $^{13}C,^{15}N$-labeled apo-CcmE weakly oriented in the magnetic field by addition of 20 mg/ml Pf1 phages measured with the experimental scheme of Figure 1. Sample conditions were 0.1 mM apo-CcmE in 20 mM Tris buffer, pH = 7, T = 293 K. For each spectrum 1024(t_2), 45(t_1) complex points were recorded with 64 scans per increment and 1 s recycle delay resulting in 1.2 hours of measuring time.

4. Appendix

Figure 4 contains histograms of D_{HH} values of apo-CcmE weakly oriented by Pf1 phage and paramagnetic holo-CcmeE and Figure 5 of a comparison of $[^{13}C,^{1}H^{\alpha\alpha/\beta\beta}]^{Methyl}$-HSQC with regular $[^{13}C,^{1}H]$-HSQC.

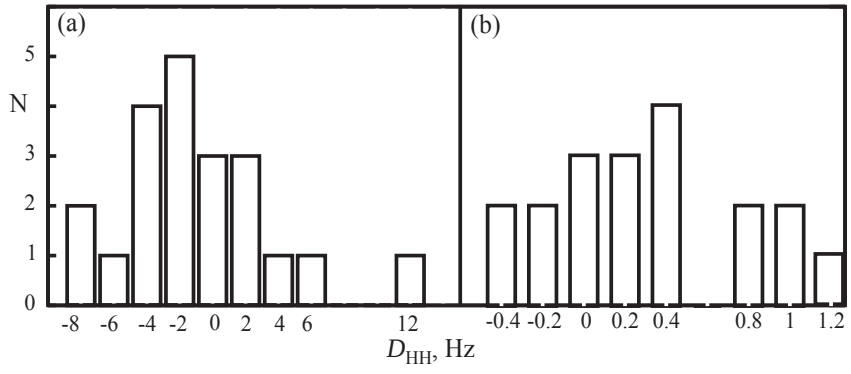


Figure 4. (a): Histogram of D_{HH} values of apo-CcmE weakly oriented by Pf1 phage were obtained as a statistical average of a half the distance in Hz between ^{1}H resonances in 3 pairs of 2D $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$- and $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$-HSQC spectra. The experimental conditions are the same as in Figure 3. (b): Histogram of D_{HH} values of holo-CcmE/KCN complex containing a paramagnetic heme group covalently bound to His130 [21] were obtained with the identical setup. Sample conditions were 0.5 mM holo-CcmE in 20 mM Tris buffer, 10 mM KCN, pH = 7, T = 293 K.

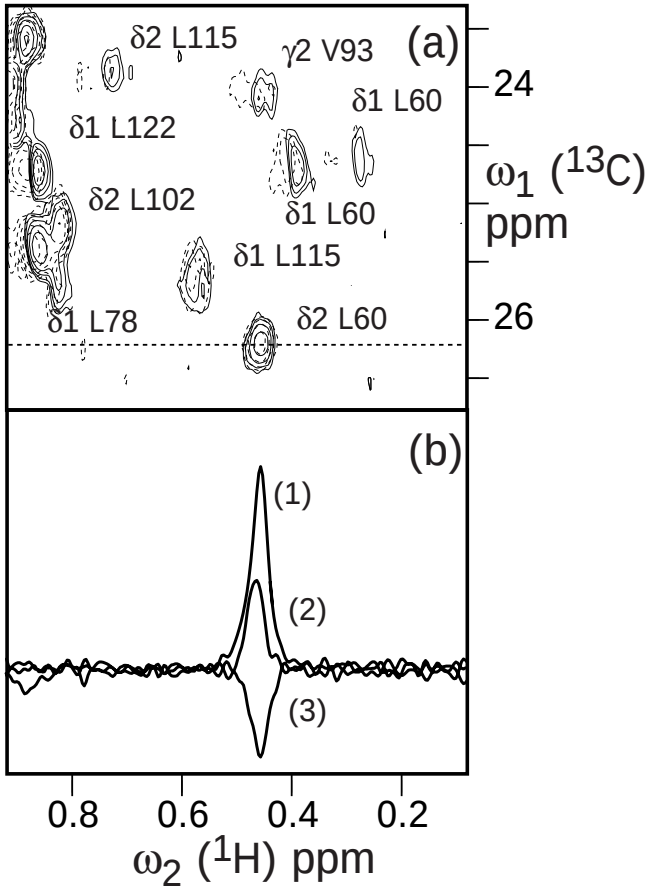


Figure 5. (a) Superimposed expansion of 2D $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$- (solid lines) and $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$- (dashed lines) HSQC spectra of the 121 amino-acid fragment of uniformly $^{13}C,^{15}N$-labeled apo-CcmE weakly oriented in the magnetic field by addition of 20 mg/ml Pf1 phage measured with the experimental scheme of Figure 1.(b) 1D slices along $\omega_2(^{1}H)$ dimensions at the position indicated by the dashed line in (a): (1) regular $[^{13}C,^{1}H]$-HSQC, (2) $[^{13}C,^{1}H^{\alpha\alpha}]^{Methyl}$- and (3) $[^{13}C,^{1}H^{\beta\beta}]^{Methyl}$-HSQC. To simplify the comparison the slice (3) was multiplied with (–1). Sample conditions were 0.1 mM apo-CcmE in 20 mM Tris buffer, pH = 7, T = 293 K. For each spectrum 1024(t_2), 45(t_1) complex points were recorded with 64 scans per increment and 1 s recycle delay resulting in 1.2 hours of measuring time.

5. References

[1] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366-12371.
[2] Müller, N.; Bodenhausen, G. *J. Chem. Phys.* **1993**, *98*, 6062-6069.
[3] Bertini, I.; Luchinat, C.; Tarchi, D. *Chem. Phys. Lett.* **1993**, *203*, 445-449.
[4] Qin, J.; Delaglio, F.; Lamar, G.N.; Bax, A. *J. Magn. Reson. Ser. B* **1993**, *102*, 332-336.
[5] Ghose, R.; Prestegard, J.H. *J. Magn. Reson.* **1997**, *128*, 138-143.
[6] Boisbouvier, J.; Gans, P.; Blackledge, M.; Brutscher, B.; Marion, D. *J. Am. Chem. Soc.* **1999**, *121*, 7700-7701.
[7] Mandal, P.K.; Madhu, P.K.; Müller, N. *Chem. Phys. Lett.* **2000**, *320*, 269-276.
[8] Madhu, P.K.; Mandal, P.K.; Müller, N. *J. Magn. Reson.* **2002**, *155*, 29-38.
[9] Sibille, N.; Bersch, B.; Coves, J.; Blackledge, M.; Brutscher, B. *J. Am. Chem. Soc.* **2002**, *124*, 14616-14625.
[10] Kaikkonen, A.; Otting, G. *J. Am. Chem. Soc.* **2001**, *123*, 1770-1771.
[11] Tolman, J.R.; Flanagan, J.M.; Kennedy, M.A.; Prestegard, J.H. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 9279-9283.
[12] Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111-1114.
[13] Choy, W.Y.; Tollinger, M.; Mueller, G.A.; Kay, L.E. *J. Biomol. NMR* **2001**, *21*, 31-40.

[14] Bothnerby, A.A.; Gayathri, C.; Vanzijl, P.C.M.; Maclean, C. *J. Magn. Reson.* **1984**, *56*, 456-462.
[15] Bothnerby, A.A.; Dadok, J.; Mishra, P.K.; Vanzijl, P.C.M. *J. Am. Chem. Soc.* **1987**, *109*, 4180-4184.
[16] Tolman, J.R.; Flanagan, J.M.; Kennedy, M.A.; Prestegard, J.H. *Nat. Struct. Biol.* **1997**, *4*, 292-297.
[17] Barbieri, R.; Bertini, I.; Cavallaro, G.; Lee, Y.M.; Luchinat, C.; Rosato, A. *J. Am. Chem. Soc.* **2002**, *124*, 5581-5587.
[18] Kontaxis, G.; Bax, A. *J. Biomol. NMR* **2001**, *20*, 77-82.
[19] Glaser, S.J.; Schulte-Herbrüggen, T.; Sieveking, M.; Schedletzky, O.; Nielsen, N.C.; Sørensen, O.W.; Griesinger, C. *Science* **1998**, *280*, 421-424.
[20] Güntert, P.; Schaefer, N.; Otting, G.; Wüthrich, K. *J. Magn. Reson. Ser. A* **1993**, *101*, 103-105.
[21] Enggist, E.; Thöny-Meyer, L.; Güntert, P.; Pervushin, K. *Structure* **2002**, *10*, 1551-1557.
[22] Hansen, M.R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065-1074.

Chapter VIII

Measuring ^{1}H-^{1}H and ^{1}H-^{13}C RDCs in Methyl Groups: Example of Pulse Sequences with Numerically Optimized Coherence Transfer Schemes

Konstantin Pervushin, Beat Vögeli, Tim Heinz and Philippe Hünenberger*

Laboratorium für Physikalische Chemie, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

* Spectroscopy

Journal of Magnetic Resonance, 2005, 172, 36-47

1. Introduction

The design of a particular NMR experiment is dictated by the type of spectral information desired. For complex spin systems in large biomolecules, specific experiments may be exceptionally difficult to construct when high sensitivity, suppression of spurious signals, and tolerance to relaxation effects are required. For example, ^{13}C-labeled methyl groups represent one of the most complicated systems of *J*-coupled homo- and heteronuclear spins found in solvated proteins and nucleotides. Previously an experiment recovering individual ^{1}H transitions was proposed in order to measure ^{1}H-^{1}H residual dipolar couplings in methyl groups [1]. Although proven to be practically useful this experiment does not deliver theoretically possible sensitivity and the time duration of its critical coherence transfer element (CTE) was not evaluated against theoretical benchmarks. A general tool enabling constructing of coherence transfer- and pulse-sequence duration-optimized complex NMR experiments especially in applications to methyl groups would enable e.g.: (i) the enhancement of the resolution in TROSY-based experiments [2-5]; (ii) the investigation of the spin dynamics associated with the fast rotation of methyl groups in proteins with high molecular weights [6]; (iii) the direct measurement of cross-correlations between Curie spin relaxation (CSR) and ^{1}H-^{1}H dipole-dipole relaxation within methyl groups of paramagnetic proteins [7-14]; (iv) the measurement of the absolute values and signs of residual ^{1}H-^{1}H (D_{HH}) and ^{1}H-^{13}C (D_{HC}) dipolar couplings [15-17] within methyl groups of proteins that are weakly aligned in solution [18-20].

Heteronuclear chemical shift correlation experiments typically rely on several types of basic building blocks [21,22]. These include a polarization transfer element (PTE), followed by an element indirectly encoding the chemical shifts of the heteronuclear spins (CSE), and a coherence transfer element (CTE) providing a density operator suitable for signal detection. The PTE and CSE elements can be combined into a so-called coherence excitation element (CEE). Several attempts have been made to systematically optimize these elements based on analytical or numerical approaches. For example, a spin-state selective PTE was optimized for the presence of residual dipolar couplings using a target-function minimization algorithm [23]. The theoretical upper bound for the PTE efficiency in a two spin ½ system subjected to relaxation was established [24,25], and experimentally feasible PTEs were proposed to drive the density operator along the theoretically optimal pathway. An example of optimized CSE is the TROSY experiment (transverse relaxation optimization) [2]. Efforts have also been dedicated to the optimization of CTEs [26-31], resulting in a significant enhancement of the spectral sensitivity. The upper bound for an arbitrary coherence transfer [32-34] and the minimal time required to accomplish this transfer (the latter only for special cases) [35-37] were established analytically, providing very useful benchmarks for further optimizations of the CTEs involved in specific experiments.

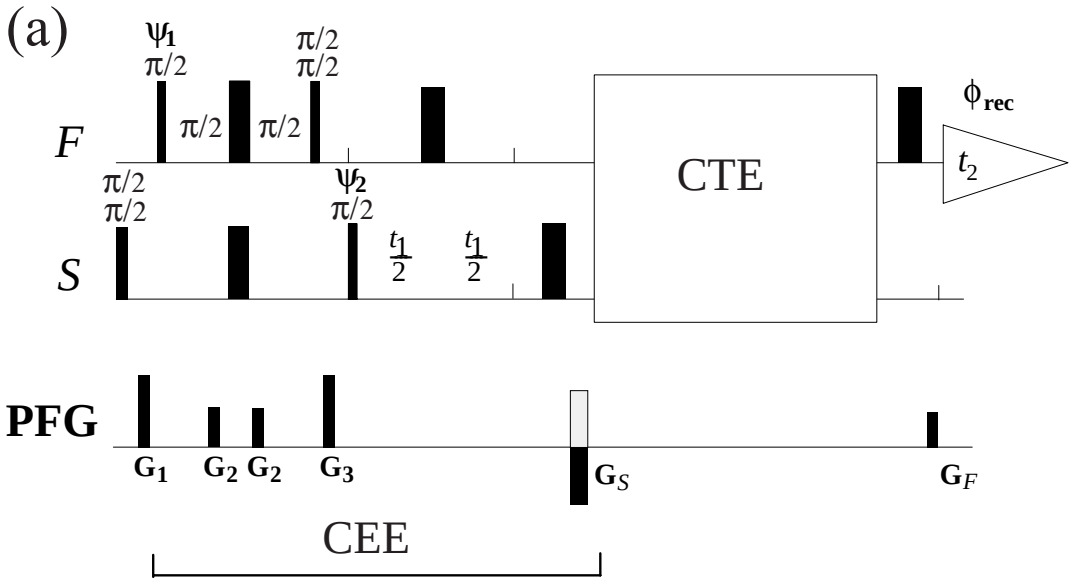
In this Chapter, we present an approach suitable for systematic construction of CTEs capable of transferring an arbitrary (generally non-Hermitian) operator *A* encoding the chemical shift of heteronuclear spins to another arbitrary operator *C* suitable for signal detection in methyl groups. The maximization of the efficiency for this transfer is undertaken simultaneously with the minimization of the pulse-sequence duration (to avoid problems related to relaxation) and of the transfer to an operator *D* (or to a set of such operators) representing undesired or spurious signals in the spectrum. This formulation translates into a complex optimization problem in a space of high dimensionality, the degrees of freedom being the pulse-sequence variables (pulse flip angles and phases, as well as delays between pulses) defining the CTE. As an example, we selected a problem of measuring ^{1}H-^{1}H and ^{1}H-^{13}C RDCs in methyl groups. However, although complex, this problem is tractable numerically because all components of the above optimization statement (efficiency of the desired transfer, suppression of undesired transfers, pulse-sequence duration) can be translated into computationally tractable mathematical functions. Due to the availability of analytical gradients for these functions, methods like gradient-based function minimization (e.g. steepest descent) or molecular dynamics [38] (e.g. simulated annealing [39]) can be applied to find a near-optimal solution to this problem.

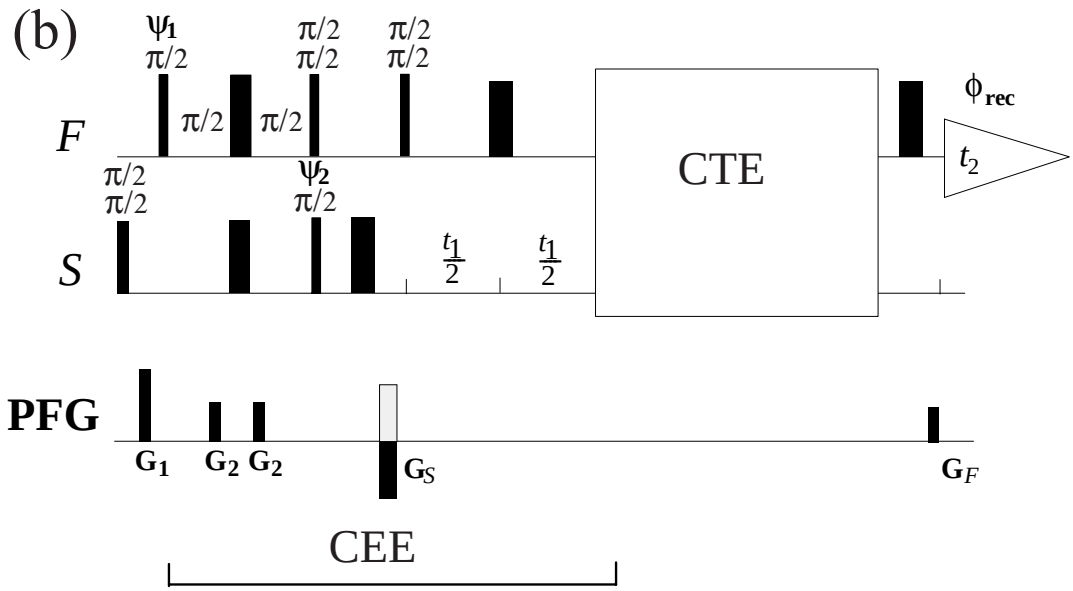
The performance of CTEs constructed using this approach is evaluated against the benchmarks provided by the theoretical unitary bound for coherence transfer [33] and the corresponding minimal transfer time (when available) [35].

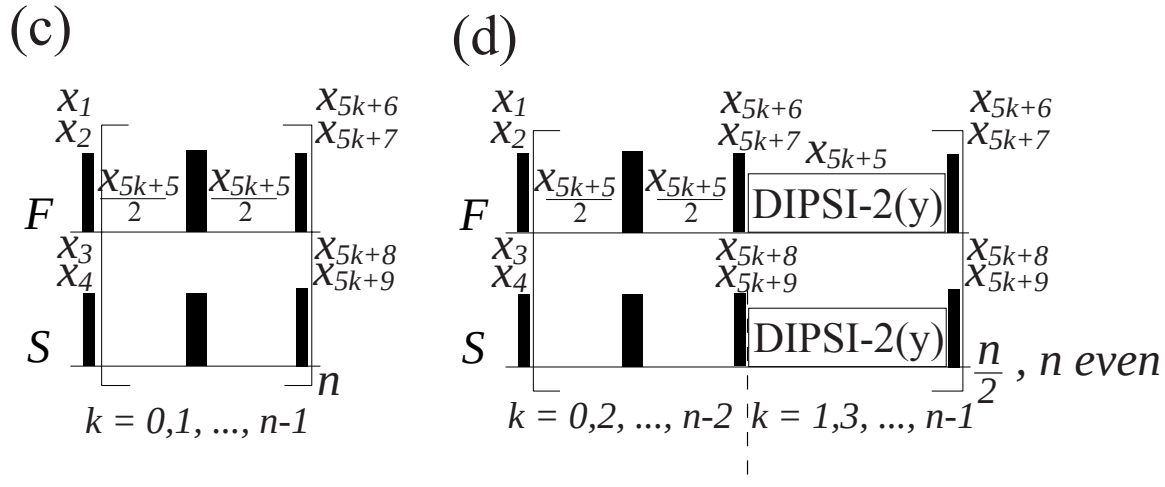
2. Theoretical Part

Heteronuclear Coherence Transfer in I_3S Spin Systems. The present discussion is restricted to the I_3S spin system, where I_3 denotes three magnetically equivalent spin ½ entities *I* (e.g., ^{1}H) and *S* a heteronuclear spin (e.g., ^{13}C), coupled via a scalar coupling constant J_{IS}. This situation represents one of the most difficult spin systems encountered, e.g. for the methyl groups in biomolecules. However, the present formalism is easily extended to other spin systems (e.g., I_2S in methylene groups or systems of coupled ^{13}C and ^{1}H homonuclear spins). The spins of type *I* are combined for convenience into a pseudospin $F = I^{(1)} + I^{(2)} + I^{(3)}$. It is also assumed that the coherence transfer is achieved by a sequence of unitary transformations neglecting the effect of relaxation and involving non-selective radiofrequency (rf) pulses.

Figure 1 (next page). General pulse sequences of two-dimensional [I_N,S]–correlation experiments for spin systems consisting of N magnetically equivalent spin ½ entities $I^{(i)}$, i = 1,...,N, coupled to a heteronuclear spin S by the coupling constant J_{IS}. For convenience, the spins I are combined into a pseudospin $F_k = \sum_i^N I^{(i)}{}_k$, where $k \in$ {x,y,z}. In (a) and (b), an excitation of the coherences $2F_zS_x$ and $2F_xS_x$, respectively, is performed and used for chemical-shift encoding of the spin S. The sequences consist of a coherence excitation (CEE) element followed by a coherence transfer element (CTE). Two types of CTEs, based either on a sequence of n pulse-interrupted free precession (INEPT) elements, or an alternating sequence of n INEPT and planar-mixing elements, are shown in (c) and (d), respectively. Wide black bars indicate non-selective π rf-pulses applied with the zero phase corresponding to the x-axis, according to the usual notation for the direction of the oscillating magnetic field in the rotating frame [45]. In this notation, the phases $\pi/2$, π and $3\pi/2$ rad correspond to the y-, x- and -y-axes, respectively. Narrow black bars indicate arbitrary non-selective rf-pulses characterized by the phase (first row) and the flip angle (second row) indicated above the bar. The planar-mixing sequence consists of a DIPSI-2 pulse train applied with the phase $\pi/2$ simultaneously to both the F and S channels. The length ϕ of the free-precession delays and planar-mixing sequences is given in radian, and is related to the metric time τ according to the equation $\phi = 2\pi J_{IS}\tau$. The pulses and delays in square brackets are repeated n times (c) or $n/2$ times with n even (d). Thus, a CTE of either type is entirely defined by a vector $\boldsymbol{x}$ of variables x_i (in radian), with i = 1, 2, …, $5n$+4. In the present work, when applied to [^{13}C,^{1}H]-correlation experiments on methyl groups, the rf-pulses on the ^{13}C (spin S) and ^{1}H (spin I) nuclei are centered at 20 and 3 ppm, respectively. In this case the durations and strengths of the pulsed magnetic field gradients (PFG) applied along the z-axis are selected as G_1: 800 μs, 80 G/cm; G_2: 800 μs, 20 G/cm; G_3: 1 ms, -60 G/ cm; G_S: 1 ms, 90 G/cm; G_F: 251 μs, 90 G/cm. A uniform value of J_{CH} = 125 Hz is assumed for all methyl groups to determine the actual duration of the free-precession delays and planar-mixing sequences. The phases in (a) and (b) are defined as: ψ_1 = {x, -x}; ψ_2 = {2y, 2(-y)}; ϕ_{rec} = {x, -x, -x, x}. Quadrature detection in the ^{13}C(t_1)-dimension is achieved by the echo-anti-echo method [26,51]. The anti-echo signal is obtained by inversion of the phase x_3 in the CTE, with simultaneous inversion of the sign of the G_S pulse.


(a)
ψ1
π/2
π/2
π/2
F
π/2
π/2
π/2
ψ2
π/2
π/2
S
t1/2
t1/2
CTE
ϕrec
t2
PFG
G1
G2
G2
G3
GS
GF
CEE


(b)
ψ1
π/2
π/2
π/2
π/2
F
π/2
π/2
π/2
ψ2
π/2
S
t1/2
t1/2
CTE
ϕrec
t2
PFG
G1
G2
G2
GS
GF
CEE


(c)
x1
x2
x5k+6
x5k+7
F
x5k+5/2
x5k+5/2
x3
x4
x5k+8
x5k+9
S
n
k = 0,1, ..., n-1
(d)
x1
x2
x5k+6
x5k+7
x5k+5
F
x5k+5/2
x5k+5/2
DIPSI-2(y)
x3
x4
x5k+8
x5k+9
S
DIPSI-2(y)
n/2 , n even
k = 0,2, ..., n-2
k = 1,3, ..., n-1

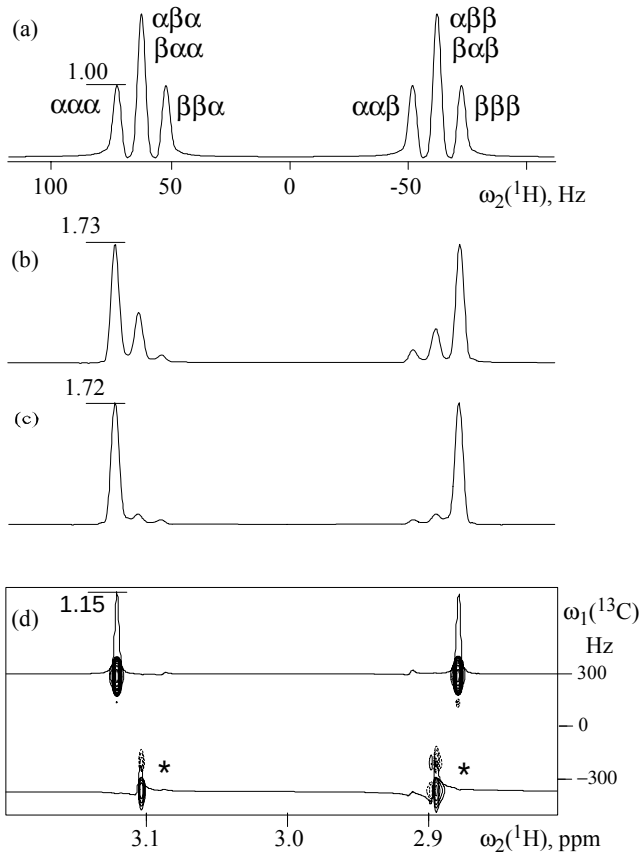


Figure 2. Simulated two-dimensional [^{13}C,^{1}H]-correlation spectra (and corresponding one-dimensional ^{1}H-slices) for a methyl group with $\omega(^{13}C)$ = 300 Hz, $\omega(^{1}H)$ = 3 ppm, J_{CH} = 125 Hz, and D_{HH} = 10 Hz. The spectra were calculated using the program NMRSIM (Bruker Biospin AG). (a) Slice at $\omega_I(^{13}C)$ = 300 Hz of a [^{13}C,^{1}H]-sensitivity-enhanced HSQC spectrum [51] calculated without ^{13}C decoupling during signal acquisition. The individual components of the ^{1}H multiplet detected via $I^{(1)-}$ are designated according to the state of the spins $I^{(2)}$, $I^{(3)}$ and S. (b) and (c), slices at $\omega_I(^{13}C)$ = 300 Hz of the [^{13}C,^{1}H]-correlation spectra calculated using the experimental scheme of Figure 1a based on $OCTE_1$ and $OCTE_2$ (Table 1), respectively. The inner components (ββα and ααβ) of the ^{1}H multiplet can be obtained by repeating the calculations with inversion of the phases of the first pulse on the channel S and the last pulse on the channel F of the CTE, together with the inversion of the sign of G_S. (d) Contour plot of the [^{13}C,^{1}H]-correlation spectrum superimposed with slices at $\omega_I(^{13}C)$ = 300 Hz and –310 Hz, calculated using the experimental scheme of Figure 1b, based on $OCTE_3$ (Table 1) and the strength of both G_S and G_F set to zero. The signals designated by asterisks can be suppressed using non zero values of G_S and G_F. All one-dimensional slices are drawn to the scale and the amplitudes relative to the ααα component in (a) are explicitly indicated.

The aim of the method is to maximize a coherence transfer $A \rightarrow C$, while simultaneously minimizing an undesired transfer $A \rightarrow D$ and the pulse-sequence duration Θ, based on experimentally accessible propagators U resulting from the generic CTEs shown in Figures 1c and d. The source operators A considered here will be restricted to

$$A_1 = 2F_zS^- \text{ and } A_2 = (F_xS^-)_{\alpha\beta} + (F_xS^-)_{\beta\alpha}\,. \qquad (1)$$

In terms of the individual spins, the latter operator may be rewritten as $A_2 = I^{(1)}{}_x S^- (E - 4I^{(2)}{}_z I^{(3)}{}_z) + I^{(2)}{}_x S^- (E - 4I^{(1)}{}_z I^{(3)}{}_z) + I^{(3)}{}_x S^- (E - 4I^{(1)}{}_z I^{(2)}{}_z)$, where E is the identity operator. These two operators, used to record the chemical shift of the spin S, can be generated by the CEEs displayed in Figures 1a and b. These are transferred to the target operators

$$C_1 = F^-_{\alpha\alpha\alpha} + F^-_{\beta\beta\beta} \text{ and } C_2 = F^-_{\alpha\alpha\beta} + F^-_{\beta\beta\alpha}, \tag{2}$$

while the corresponding suppressed-target operators are

$$D_1 = F^-_{\alpha\beta\alpha} + F^-_{\beta\alpha\alpha} + F^-_{\alpha\beta\beta} + F^-_{\beta\alpha\beta} + F^-_{\alpha\alpha\beta} + F^-_{\beta\beta\alpha} \tag{3.1}$$

and

$$D_2 = F^-_{\alpha\beta\alpha} + F^-_{\beta\alpha\alpha} + F^-_{\alpha\beta\beta} + F^-_{\beta\alpha\beta} + F^-_{\alpha\alpha\alpha} + F^-_{\beta\beta\beta}. \tag{3.2}$$

In the above Equations, the subscript triplets $i,j,k \in \{\alpha, \beta\}$ designate the spin states of the ^{1}H, ^{1}H, and ^{13}C spins (in this order). The pseudospin operators can be rewritten in terms of the individual spin operators as, $F^-_{\alpha\alpha\alpha/\beta\beta\beta} = I^{(1)-}(E \pm 2I^{(2)}{}_z)(E \pm 2I^{(3)}{}_z)(E \pm 2S_z)$ + two isomorphous terms with circularly permuted indeces 1, 2 and 3, where the state α in a given position of the subscript list maps to the "+" sign in the corresponding position of the right-hand side. A spectral representation of individual spin transitions is provided in Figure 2a. The choice of the specific operators in Equations 1-3 will be discussed in the Results and Discussion.

The operators defined above, as well as the unitary-transformation propagator U resulting from an experimentally accessible CTE, can be represented as matrices in a basis conforming with the symmetry properties of the spin system [40,41]. In the case of the I_3S system, the corresponding matrices A, C, D and U assume a block structure along the diagonal, consisting of one square block of dimensions 8 x 8 and 2 identical square blocks of dimensions 4 x 4, as shown in Figure 3. Since, in the absence of relaxation, there is no "cross-talk" between individual blocks, it is sufficient to consider the unitary transformations separately for the 8 x 8 and one of the 4 x 4 blocks [41], thereby reducing dimensionality of the problem.

In the absence of relaxation, the transfer $A \rightarrow C$ can be described by [31]

$$UAU^\dagger = b(U)C + Q \text{ with } \mathrm{Tr}\{C^\dagger Q\} = 0, \tag{4}$$

where the dagger denotes Hermitian conjugation, Q is a residual operator, and $b(U)$ is the transfer efficiency given by

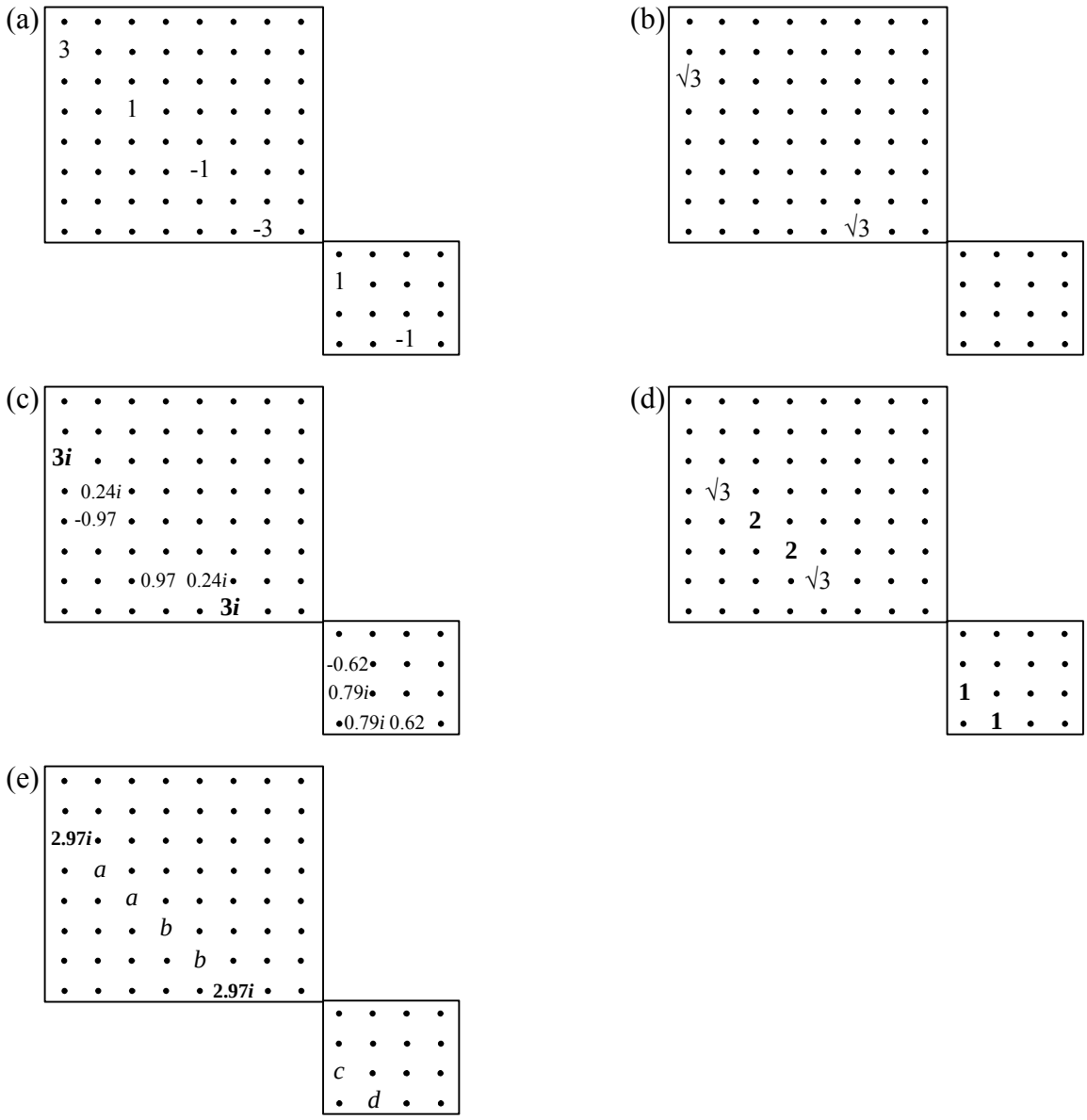


Figure 3. Schematic representation of the matrices associated with specific operators. Examples of matrices are given for (a) the source operator A_1 (Equation 1), (b) the target operator $C_1^\dagger$ (c) the operator $UA_1U^\dagger$ corresponding to the optimized propagator $OCTE_1$ (Table 1), (d) the operator $D^\dagger$ of the suppressed target (Equation 3.1). (e) The operator $UA_1U^\dagger$ corresponding to $OCTE_2$ (Table 1). The actual 16 x 16 matrices consist of 8 x 8 + 4 x 4 + 4 x 4 blocks along the diagonal. Only the 8 x 8 block and one of the two identical 4 x 4 are displayed. The elements outside the squares are zero. Points in (a), (b) and (d) represent zero and in (c) and (e) irrelevant numbers. In (c) the components of the matrix $UA_1U^\dagger$ corresponding to the matrix $C_1^\dagger$ are highlighted. In (d) the components of the suppression matrix D_1 representing the central transitions of the ^{1}H-multiplet given by the spin operators $F^-_{\alpha\beta\alpha} + F^-_{\beta\alpha\alpha} + F^-_{\alpha\beta\beta} + F^-_{\beta\alpha\beta}$ are shown in bold. In (e) a = −0.04 −0.01i, b = 0.04 + 0.03i, c = −0.21 + 0.10i, and d = 0.23 + 0.11i.

$$b(U) = \mathrm{Tr}\{U\,A\,U^\dagger\,C\,\}/\mathrm{Tr}\{C^\dagger C\}. \qquad (5)$$

The maximum value b_{max} of the complex norm of $b(U)$, the so-called unitary bound value, can be determined numerically for arbitrary source and target operators [33]. This quantity represents an important benchmark for the construction and evaluation of optimized CTEs. It has been demonstrated [34] that in the case of I_3S spin systems with a quantum evolution governed by non-selective rf irradiation and heteronuclear J-coupling Hamiltonians, it is always possible to construct a CTE which achieves the maximum polarization-transfer efficiency b_{max}.

A CTE can be constructed experimentally using either a series of n pulse-interrupted free-precession periods [42] (INEPTs, [43,44]), as shown in Figure 1c, or an alternating series of n INEPT and planar-mixing sequences (the latter usually implemented via the DIPSI-2 supercycle [31]), as shown in Figure 1d. Since in the I_3S spin system the operators F_kS_k, with $k \in \{x,y,z\}$, do not commute, it is generally not possible to convert a pure INEPT-based CTE to an alternating INEPT/planar-based CTE and vice versa. Thus, both types of CTEs should be tried out for the optimization process. Note, however, that the INEPT-based CTE is easier to implement in practice since it imposes no restriction on the length of the free-precession delays, while short INEPT/planar-based sequences are restricted by the minimal duration of the DIPSI (or any other) supercycle.

The propagator U corresponding to the CTEs of Figures 1c and d can be constructed using the skew-Hermitian Hamiltonians

$$\tau H_{rf} = i\beta_I (\cos(\alpha_I)F_x + \sin(\alpha_I)F_y) + i\beta_S (\cos(\alpha_S)S_x + \sin(\alpha_S)S_y) , \quad (6.1)$$

$$\tau H_J = i\gamma F_zS_z , \quad (6.2)$$

$$\tau H_{planar} = i\gamma/2(F_zS_z + F_xS_x) , \quad (6.3)$$

as building blocks. Here, α_I and α_S are the phase shifts of I and S pulses relative to the carrier rf-radiation, $\beta_I = \tau\omega_I^{rf}$ and $\beta_S = \tau\omega_S^{rf}$ are the rotation angles (in radian) of I and S pulses with the strength ω^{rf} and duration τ (defined as in [45]), and $\gamma = 2\pi\tau J_{IS}$. The operator H_{planar} represents the effective approximate planar-mixing Hamiltonian resulting from the DIPSI-2 sequence applied along the y-axis. For both INEPT-based and INEPT/planar-based CTEs, the concatenated propagator $U(\boldsymbol{x})$ is a function of a vector $\boldsymbol{x}$ of variables x_i, i = 1, 2,..., $5n$+4 containing all the information about the pulse flip angles and phases, the delays between the pulses, and the length of the DIPSI-2 supercycle (see Figures 1c and d). This vector contains the parameters determining the M individual events (rf-pulse, free precession or planar mixing) in the CTE. The propagator can thus be written

$$U(\boldsymbol{x}) = \prod_{k=1}^{M} \exp(\tau H_k) . \quad (7)$$

The quantum evolution of A to C proceeds via infinitely fast rotations under the H_{rf} Hamiltonian (assuming that the rf-pulses are infinitely strong) and the finite-time transformations under the H_{planar}

and H_J Hamiltonians. Note that although the events in the pulse sequence of Figure 1 are represented along time according to a "from-left-to-right" rule, the individual propagators in Equation 7 are multiplied in the reversed order.

For the numerical optimization of CTEs, it is convenient to define the real and positive target function

$$f_{AC}(\boldsymbol{x}) = b(\boldsymbol{x})b^*(\boldsymbol{x}), \tag{8.1}$$

with the short notation $b(\boldsymbol{x}) = b(U(\boldsymbol{x}))$ for the transfer efficiency defined in Equation 5. The gradient of this function given by

$$\nabla f_{AC}(\boldsymbol{x}) = (\mathrm{Tr}\{(\nabla U)AU^\dagger C + UA(\nabla U)^\dagger C\}b^*(\boldsymbol{x}) + b(\boldsymbol{x})(\mathrm{Tr}\{(\nabla U)AU^\dagger C + UA(\nabla U)^\dagger C\})^*)/\mathrm{Tr}\{C^\dagger C\}, \tag{8.2}$$

where the star denotes complex conjugation, and ∇U is obtained by applying the chain rule for derivatives to the product in Equation 7.

As illustrated in Figure 3c, even in the case of a maximally efficient transformation U of the source operator A (e.g. A_1 in Figure 3a) to the target operator C (e.g. C_1 in Figure 3b), the operator Q in Equation 4 may contain observable magnetization leading to spurious signals in the spectrum. These signals can be combined into the operator D (e.g. D_1 in Figure 3d), transfer to which can be suppressed by minimizing a function $f_{AD}(\boldsymbol{x})$ analogous to the function $f_{AC}(\boldsymbol{x})$, simultaneously with the maximization of $f_{AC}(\boldsymbol{x})$ itself. However, it turns out that a more efficient reduction of the spurious signals is achieved by suppressing individually a set of matrices $D_i^{(k)}$, $k = 1,\ldots,K$, each containing only one of the K non-zero elements of D_i (e.g. elements $D^{(4,2)}$, $D^{(5,3)}$, $D^{(6,4)}$, $D^{(7,5)}$, $D^{(11,9)}$, $D^{(12,10)}$ in Figure 3d), because it avoids artificial cancellation of positive and negative contributions to the penalty function stemming from the different transitions involved in D. In this case, the suppression is achieved by minimizing the sum of functions $f_{ADk}(\boldsymbol{x})$, rather than a simple function $f_{AD}(\boldsymbol{x})$, simultaneously with the maximization of $f_{AC}(\boldsymbol{x})$ itself. The matrix representation of the source operator transformed by a unitary propagator optimized through this procedure is given in Figure 3e.

To complete the specification of the NMR experiment, the matrix $UAU^\dagger$ representing the density operator after the unitary transformation by the CTE is transformed into a detectable spectrum using the free precession Hamiltonian

$$H_{obs} = \mathrm{i}2\pi\,(({}^1J_{IS} + {}^1D_{IS})F_zS_z + ({}^1J_{II} + {}^1D_{II})(I^{(1)}{}_zI^{(2)}{}_z + I^{(1)}{}_zI^{(3)}{}_z + I^{(2)}{}_zI^{(3)}{}_z)). \tag{9}$$

The spectrum is then obtained by Fourier transform of the FID function

$$FID(t) = \mathrm{Tr}\{\exp(H_{obs}t)\,C\exp(-H_{obs}t)\,F^-\}\exp(-R\,t)\,, \tag{10}$$

where R is an arbitrary line-broadening parameter. This spectral representation of the target matrix can be used for the direct analysis of the polarization transfer efficiency and of the level of undesired signal suppression achieved by CTE.

Optimization of the Transfer Function using Molecular Dynamics. A CTE composed of a series of n INEPT or n alternating INEPT/planar units (Figures 1c and d) is entirely defined by the N-dimensional vector $\boldsymbol{x}$ introduced above. Given a source matrix A and a number of units n, the problem of simultaneously (i) maximizing the efficiency of transfer to a target matrix C, (ii) minimizing the transfers to a set of suppressed target matrices D_k, $k = 1 \ldots K$ and (iii) minimizing the pulse-sequence duration may be rewritten in the form of an N-dimensional minimization problem for the target function

$$V(\boldsymbol{x}) = -f_{\mathrm{AC}}(\boldsymbol{x}) + \mu \sum_{k=1}^{K} f_{\mathrm{ADk}}(\boldsymbol{x}) + \nu g(\boldsymbol{x}), \tag{11}$$

where μ (unitless) and ν (units of radian) are weighting coefficients, f_{AC} and f_{AD} are defined by Equation 8.1, and

$$g(\boldsymbol{x}) = \begin{cases} \Theta^2(\boldsymbol{x}) & \text{if } \Theta(\boldsymbol{x}) \le \Theta_0 \\ 2\Theta_0\Theta(\boldsymbol{x}) - \Theta_0^2 & \text{otherwise} \end{cases} \quad \text{with } \Theta(\boldsymbol{x}) = \sum_{k=0}^{n-1} x_{5k+5} \tag{12}$$

The two last terms in Equation 11 should be viewed as penalty functions imposed during the minimization of the first term. The penalty function associated with the duration Θ of the pulse sequence (sum of delays and DIPSI-2 length variables) is harmonic below a given threshold Θ_0 and linearized beyond, in such a way that the function is continuous and continuously differentiable over the whole range of durations. The negative gradient $\boldsymbol{F}(\boldsymbol{x})$ of $V(\boldsymbol{x})$ is given on a component-by-component basis

$$F_i(\boldsymbol{x}) = -\frac{\partial V(\boldsymbol{x})}{\partial x_i} = \frac{\partial f_{\mathrm{AC}}(\boldsymbol{x})}{\partial x_i} - \mu \sum_{k=1}^{K} \frac{\partial f_{\mathrm{ADk}}(\boldsymbol{x})}{\partial x_i} - \nu \frac{\partial g(\boldsymbol{x})}{\partial x_i} \tag{13}$$

Because this gradient can be easily calculated (see Equation 8.2) together with V, one may try to minimize this function by means of a gradient-based function minimization (e.g. steepest descent) algorithm (see Supplementary Material of the online publication). However, the dimensionality of the problem increases rapidly with the number of INEPTs or alternating INEPT/planar units, while the hypersurface described by V is likely to involve an increasingly larger number of local minima and barriers. For this reason the steepest-descent algorithm, which only leads to the local minimum closest to the a given starting point $\boldsymbol{x}(0)$, is not likely to be efficient (i.e., numerous attempts with different random starting points are needed to possibly reach the global minimum). A more efficient strategy is to use a molecular-dynamics (simulated-annealing) algorithm to solve the minimization problem. Speaking of molecular-dynamics is slightly misleading in the present context, because the

degrees of freedom involved are not Cartesian coordinates of atoms in molecules, but rather angular variables characterizing a pulse sequence. The term is nevertheless retained here.

In the molecular-dynamics approach, a common "mass" m is assigned to the N degrees of freedom of the problem. These variables are then propagated along a "time" coordinate t according to Newton's second law

$$d^2\boldsymbol{x}(t)/dt^2 = d\boldsymbol{v}(t)/dt = m^{-1}F(\boldsymbol{x}(t)), \quad (14)$$

where $\boldsymbol{v}$ stands for the "velocity" vector associated with $\boldsymbol{x}$. In this case, V can be interpreted as a "potential energy" to be minimized. In practice, Equation 14 is integrated iteratively based on a finite time-step size Δt. To enhance the search power of molecular dynamics, the technique of simulated annealing is applied here. In this case, the system is coupled to a thermostat, so that the "kinetic energy" associated to the velocity $\boldsymbol{v}$ (related to a system "temperature"), rather than the total energy, is kept (on average) constant at a specified value. In the simulated-annealing protocol, the imposed value of the kinetic energy is linearly decreased from a large value (efficient crossing of potential energy barriers, search biased towards high-entropy regions) to a small value (poorer search, biased towards low potential energy regions) over the given number of time steps of a run. This search is then followed by a steepest-descent minimization.

All the searches (steepest descent as well as simulated annealing) are carried out under periodic boundary conditions. The pulse flip angle and phase-shift variables in $\boldsymbol{x}$ are restricted to the range $[0, 2\pi]$ by application of 2π-periodicity, while the corresponding delay variables are restricted to the range $[0, 4\pi]$ by application of "mirroring" boundary conditions. The latter choice ensures that the function g in Equation 11 is continuous throughout the search (periodic boundary conditions would lead to discontinuities in this case).

3. Results and Discussion

Selection of the Source, Target and Suppression Operators. Figure 2a shows a simulated one-dimensional ^{1}H-spectrum of a methyl group of a hypothetical protein slightly aligned in solution (e.g., with residual dipolar coupling D_{HH} = 10 Hz). Individual transitions of the multiplet pattern are labeled according to the spin states of the coupled ^{1}H, ^{1}H, and ^{13}C spins involved (in this order). Note that individual transitions involving the magnetically equivalent ^{1}H-spins could not be observed as multiplet-split NMR lines in the absence of alignment (D_{HH} = 0 Hz). To measure the value of the D_{HH}-coupling constant, the ^{1}H-transitions $\alpha\alpha$ and $\beta\beta$ can be observed individually in separate spectra under conditions of ^{13}C decoupling [1]. In this case, observation of the $\alpha\alpha\alpha$ and $\beta\beta\beta$ transitions

representing the outer components of the ^{1}H multiplet in one spectrum, and of the $\alpha\alpha\beta$ and $\beta\beta\alpha$ transitions representing the corresponding inner components in the other spectrum, permits the direct estimation of the D_{HH}-coupling constant as well as the sum of the D_{HC}- and J_{HC}-coupling constants.

These considerations result in the choice of the target (detection) operators C_1 and C_2 (Equation 2). Note that the insertion of a ^{13}C-π-pulse before signal acquisition in an experiment applying the C_1 operator is equivalent to the application of the C_2, eliminating the need for the separate construction of two complementary experiments. The matrix representation of the target operator C_1 in the symmetry-adapted basis set is shown in Figure 3b.

To produce a two-dimensional [^{1}H,^{13}C]-correlation spectrum, a coherence suitable for the detection of the ^{13}C chemical shifts has to be excited, which will be subsequently transferred (using a suitably constructed CTE) to the target operator for detection. The simplest choice for this source operator is the antiphase operator A_1 (Equation 1), which can be easily generated using INEPT [46]. The pulse sequence exciting the coherence associated with the operator A_1 is depicted in Figure 1a, while the corresponding matrix representation in the symmetry-adapted basis set is shown in Figure 3a.

A major disadvantage of using the complete antiphase operator A_1 is its decomposition during the chemical-shift evolution periods, caused by the differential relaxation of the individual transitions comprising the operator [3]. In a coupled heteronuclear single-quantum correlation (HSQC) spectrum, the evolution of A_1 under conditions of negligible relaxation results in four spectral lines with relative intensities 3:1:1:3, comprising eight transitions within the energy-level diagram [4,31]. The transverse magnetization corresponding to the (most intense) outer lines of the quartet relaxes significantly faster than the magnetization associated with the inner ones. As a consequence, the contributions from the outer lines can become quite small in systems of high molecular weight [4]. This relaxation-induced decomposition results in numerous spurious transfers and, therefore, in significantly biased measurements.

In contrast, the coupled heteronuclear multiple-quantum correlation (HMQC) spectrum measured using the $2F_xS^-$ source operator for detection of ^{13}C chemical shifts exhibits (in the absence of relaxation) only three lines with relative intensities 1:2:1 comprising four transitions. For the two central transitions, dipolar auto- and cross-correlation contributions completely cancel out. This represents a completely optimized TROSY effect, resulting from the combination of fast methyl rotation and very slow overall molecular tumbling [3-6,47]. These TROSY transitions are defined by the operator A_2 (Equation 1), which is therefore an attractive source operator for the construction of quantitative NMR experiments. The pulse sequence exciting the coherence associated with the operator A_2 is depicted in Figure 1b, while the corresponding matrix representation in the symmetry-adapted basis set is shown in the Supplementary Material (see the online publication). Note that the

excitation of this TROSY operator requires active filtering of the two outer transitions, which otherwise can result in spurious signals. Fortunately, the signals stemming from the undesired outer transitions can be suppressed without introduction of extra delays in the experiment. To achieve this filtering, the experimental scheme of Figure 1b takes an advantage of the facts that: (i) the operators representing the inner and outer transitions of the triplet evolve with opposite frequencies (see simulated spectrum of Figure 2d); (ii) the constructed CTE preserves the coherence order (e.g., in one experiment the "minus" source operators are transferred exclusively to "minus" target operators). Therefore, adjusting the signs of the G_S and G_F gradients (see Figure, either the inner or the outer components of the triplet can be refocused (see caption of Figure2).

Table 1. Theoretical maximal bound, b_{max}, relative experimental transfer efficiency b/b_{max}, ratio of the absolute values of amplitudes of the suppressed central $\|a^{\alpha\beta\alpha} + a^{\beta\alpha\alpha}\|$ and inner $\|a^{\beta\beta\alpha}\|$ components to the amplitude of the selected $\|a^{\alpha\alpha\alpha}\|$ component of the ^{1}H-multiplet, and the duration Θ of designed coherence transfer elements (CTE).

Type of transfer	b_{max}	b/b_{max} (%)	$\|a^{\alpha\beta\alpha} + a^{\beta\alpha\alpha}\|/\|a^{\alpha\alpha\alpha}\|$, $\|a^{\beta\beta\alpha}\|/\|a^{\alpha\alpha\alpha}\|$ (%)	Θ (rad)
$A_1 \rightarrow C_1$ (OCTE$_1$) [a]	$\sqrt{3}$	100	30, 8	3.62
$A_1 \rightarrow C_1$ (OCTE$_2$) [b]	$\sqrt{3}$	98	9, 3	5.09
$A_2 \rightarrow C_1$ (OCTE$_3$) [c]	$2/\sqrt{3}$	90	0, 3	6.73
$A_1 \rightarrow F_{\alpha\alpha}^{-}[1]$ [d]	$\sqrt{3}$	62	0, 0	3.45

[a] The pulse sequence OCTE$_1$ relies on an alternating INEPT/planar sequence with n = 2, and was optimized by direct maximization of f_{AC} in Equation 8.1 for the transfer $A_1 \rightarrow C_1$ without additional restraints. The corresponding pulse-sequence vector (Figure 1e) is $\boldsymbol{x}$ = {π, π/2, 0, π/2, 3.62, π, π/2, π, π/2,0,0,0,0,0}. See also the matrix of Figure 3c and spectrum of Figure 2b.

[b] The pulse sequence OCTE$_2$ relies on an alternating INEPT/planar sequence with n = 2 (Figure 1e) and was optimized for the transfer $A_1 \rightarrow C_1$ by minimization of V in Equation 11 with Θ_0 = 7 rad, μ = 1 and v = 25 rad^{-1} through simulated annealing. The corresponding pulse-sequence vector is $\boldsymbol{x}$ = {π, 3.03, 0, 3π/4, 1.93, 0, 1.44, 6.12, 0, π, 0, 4.84, π, π/2}. See also the matrix in Figure 3e and spectrum in Figure 2c.

[c] The pulse sequence OCTE$_3$ relies on an INEPT sequence with n = 4 (Figure 1d) and was optimized for the transfer $A_2 \rightarrow C_1$ by minimization of V in Equation 11 with Θ_0 = 7 rad, μ = 1 and v = 25 rad^{-1} through simulated annealing. The corresponding pulse-sequence vector is $\boldsymbol{x}$ = {4.66, 4.37, 0.97, 3π/4, 1.74, 4.88, 0.20, 0.60, 5.50, 1.81, 0.14, 3.96, 2.54, 4.26, 2.21, 1.17, 2.03, 3.74, 1.97, 0.96, 5.32, 4.59, 0.97, 1.57}. See also the matrix in Supplementary Material (see online publication) and spectrum in Figure 2d.

[d] The detection operator $F_{\alpha\alpha}^{-} = I^{(1)-}(E + 2I^{(2)}{}_z)(E + 2I^{(3)}{}_z)$ + two isomorphous terms with circularly permuted indeces 1, 2 and 3 [1].

A first attempt to construct a CTE achieving the transfer $A_1 \rightarrow C_1$ with high efficiency by direct maximization of the function f_{AC} (Equation 8.1) without additional restraints resulted in an alternating INEPT/planar-based CTE with $n = 2$ projecting the source operator on the desired target operator with maximal theoretical efficiency b_{max}. This optimized CTE will be referred as $OCTE_1$. The corresponding parameter vector $\boldsymbol{x}$ is given in Table 1, the matrix representation of the operator $UAU^{\dagger}$ in Figure 3c, and the simulated spectrum in Figure 2b. As can be seen both from the matrix and the spectrum, the undesired multiplet components can reach a magnitude as high as one third of the signal of interest. This can be particularly inconvenient for practical applications involving high molecular weights, where the unwanted central component of the multiplet is TROSY-enhanced. Therefore it is essential to be able to actively minimize transfers to the undesired transitions.

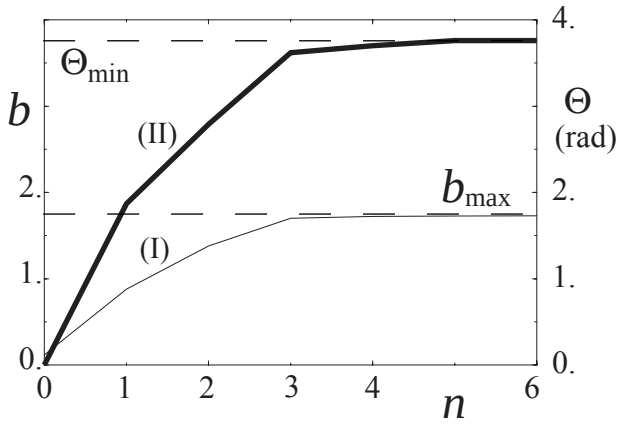


Figure 4. Numerical estimates of the maximal achievable coherence transfer b (curve I) and corresponding minimal pulse sequence length Θ required to achieve this transfer (curve II), displayed as functions of the complexity n of the INEPT-based CTE (Figure 1c) for the transfer $A_1 \rightarrow C_1$. For each n 10,000 unconstrained maximizations of f_{AC} (Equation 8.1) starting with random $\boldsymbol{x}$ were performed resulting in a pool of NMR experiments each characterized by specific b and Θ. The pool was doubly sorted according to b and Θ to obtain the maximal and minimal values (curves I and II, respectively) of the corresponding parameters attained within one experiment. The theoretical unitary bound for this transfer, $b_{max} = \sqrt{3}$, is shown as a horizontal line. The estimated minimal pulse sequence length, $\Theta_{min} = 3.62$, to achieve this transfer is also shown as a horizontal line.

Boundaries and Criteria of Optimization. The theoretical bounds b_{max} characterizing the maximal achievable efficiency for the unitary transformation of the source operator to the target operator were evaluated numerically [33,34], and are listed in Table 1 for the transfers $A_1 \rightarrow C_1$ and $A_2 \rightarrow C_1$. The corresponding bounds for the $A_1 \rightarrow C_2$ and $A_2 \rightarrow C_2$ transfers, respectively, evaluate to the same two values. Analytical estimates Θ_{min} for the minimal pulse sequence length required for the desired coherence transfer are currently only available for a limited set of special cases [35-37]. In the present case, this value is only available for the $A_1 \rightarrow C_1$ transfer ($\Theta_{min} = 3.62$, the proof to be

published elsewhere) and it is indeed achieved in the experiment based on $OCTE_1$ (see Table 1). For the transfer of $A_2 \rightarrow C_1$, the analytical minimal transfer time remains to be established.

An important criterion in the construction of practically useful CTEs is their complexity, defined by the number n of repeating units (Figure 1c) or alternate repeating units (Figure 1d) they consist of. Direct maximization of the function f_{AC} in Equation 8.1 without additional restraints were performed for INEPT-based CTEs with different levels of complexity n. These results, shown in Figure 4, suggest that for any transfer $A \rightarrow C$, an optimized CTE capable of reaching the theoretical unitary bounds b_{max} should possess a minimal complexity level n_{min}. For example, for the transfer $A_1 \rightarrow C_1$ this number is four for the INEPT-based and two for INEPT/planar-based CTEs, respectively. In the former case, the minimal pulse-sequence length required to achieve near-optimal transfer for $n \geq n_{min}$ is indeed the estimated value $\Theta_{min} = 3.62$.

Optimization of the Transfer Function using Molecular Dynamics. The search for local minima of the function $V(\boldsymbol{x})$ of Equation 11 in the space of pulse-sequence variables $\boldsymbol{x}$, using either a gradient-based (steepest-descent) or a molecular-dynamics (simulated-annealing) approach, was implemented in a C++ program. Because the steepest-descent algorithm only leads to the local minimum closest to a given starting point, the global minimum of V can only be located by performing a large number of minimizations initiated from random starting points. In contrast, the simulated-annealing approach with a suitably-optimized protocol often permits the determination of the global minimum in a single run. A simulated-annealing protocol is defined by four parameters, namely the (common) mass assigned to the pulse-sequence variables, the initial (high) and final (low) kinetic energies, and the duration of the run. An appropriate minimal mass was determined by ensuring energy conservation (no drift, negligible fluctuations) in simulations without temperature coupling. The use of a larger mass leads to slower sampling without further improving energy conservation, and is thus not recommended. Values for the various other parameters involved in the two algorithms were optimized by trial and error (Supplementary Materials, see online publication). The evolution of the kinetic (linearly decreased) and potential (function V in Equation 11) energies along a typical simulated-annealing run are shown in Figure 5.

The threshold for linearization of the penalty term associated with the pulse-sequence duration (Equation 12) was set to $\Theta_0 = 7$ rad. The biasing parameters μ and ν in Equation 11 were varied in the ranges $\mu = 0.01 - 5$ and $\nu = 10 - 100$ rad^{-1}. The resulting optimized CTEs were evaluated by calculation of the corresponding simulated spectra (see Figure 2). Indicative values of $\mu = 1$ and $\nu = 25$ rad^{-1} were found to provide a good compromise between transfer efficiency and CTE duration.

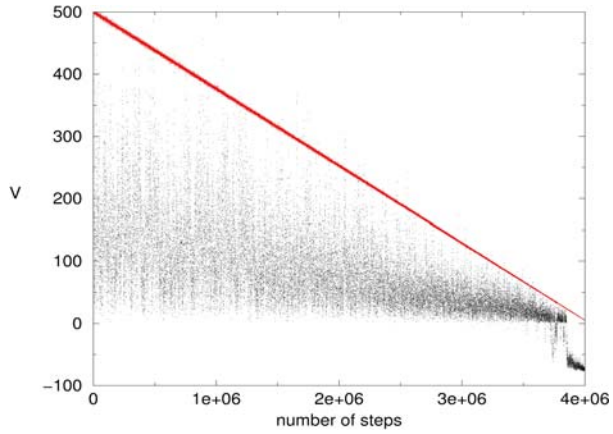


Figure 5. Example of evolution of "kinetic" energy (straight line, gradually decreased from 500 to 5) and potential energy (Equation 11, with Θ_0 = 7 rad, μ = 1 and ν = 25 rad^{-1}) in reduced units (Supplementary Materials, see online publication) during a typical simulated-annealing optimization run. The curve corresponds to the transfer $A_2 \rightarrow C_1$, optimized with simultaneous suppression of the spurious target D_1 for an INEPT-based CTE (Figure 1c) with n = 4.

In general, pulse sequences of similar quality can also be produced using the steepest-descent algorithm. However, although simulated annealing nearly systematically reaches the global minimum in a single run, steepest descent typically requires the minimization of hundreds of random starting points before reaching this global minimum. For pulse sequences of limited complexity ($n \leq 3$), the efficiency of the two methods in terms of computer time is comparable. However, simulated annealing tends to outperform steepest descent upon increasing the complexity of the pulse sequence.

The optimization of complex pulse sequences often results in the appearance of delay lengths close to zero. For example, most of the optimized INEPT-based CTEs, with n = 4 or 5 could be reduced to a simpler CTE with n = 3 possessing similar properties. Similarly, INEPT/planar-based CTEs with n = 4 could generally be reduced to CTEs with n = 2. Note, finally, that optimized pulse sequences are often degenerate, i.e., identical values of V could often be achieved by multiple pulse sequences. In general, the degenerate sequences were characterized by similar delay lengths and pulse phase shifts differing through integer fractions of π.

As discussed previously, a CTE optimized for efficiency by direct maximization of the function f_{AC} in Equation 8.1 without additional restraints is generally characterized by spurious transitions in the spectrum, corresponding to unwanted elements in the transformation matrix. These effects are illustrated in Figures 2b and 3c for the $A_1 \rightarrow C_1$ transfer through the optimized pulse sequence OCTE$_1$ (Table 1). Performing the optimization of an INEPT/planar-based CTE with n = 2 by minimizing the function V in Equation 11 (with the suppressed matrix D_1, Θ_0 = 7 rad^{-1}, $\mu = 1$ and $\nu = 25$ rad^{-1}), i.e., including active suppression of unwanted transitions and minimization of the pulse-sequence duration, results in a "cleaner" transfer operator $UAU^{\dagger}$ and spectrum, without significantly compromising the high transfer efficiency for the desired transition and the short

duration of the pulse sequence. This optimized CTE will be referred to as $OCTE_2$. The corresponding vector $\boldsymbol{x}$ is given in Table 1, the matrix representation of $UAU^{\dagger}$ in Figure 3c, and the simulated spectrum in Figure 2c. A similar optimization (with the suppressed matrix D_2, and identical parameters) was performed of an INEPT-based CTE with $n = 4$ for the transfer $A_2 \rightarrow C_1$, that will be referred to as $OCTE_3$. The corresponding vector $\boldsymbol{x}$ is given in Table 1, the matrix representation of $UAU^{\dagger}$ in Figure 1S in Supplementary Material, and the simulated spectrum in Figure 2d. For the two optimized CTEs, 98% and 90% of the maximal unitary bounds are achieved (Table 1). Note that optimized INEPT/planar-based CTEs were generated for both the transfers of $A_1 \rightarrow C_1$ and $A_2 \rightarrow C_1$. However, the experimental implementation of the planar Hamiltonian of Equation 6.3 using a DIPSI-2 pulse train imposes a technical limitation on the minimal length of the individual planar-mixing elements. So far, no experimentally attractive solution based on an alternating INEPT/planar-based CTE was found for the transfer of $A_2 \rightarrow C_1$. On the other hand, INEPT-based CTEs are free from these constraints, although, they tend to produce on average somewhat longer experiments achieving the same transfer efficiency. Clearly, more ample numerical statistics as well as analytical work are needed to comparatively evaluate these two types of CTEs.

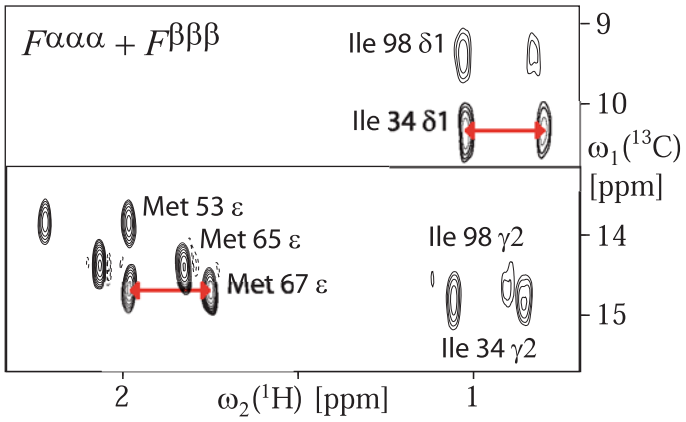


Figure 6. Sample of the two-dimensional [^{13}C, ^{1}H]-correlation spectrum selecting the C_1 components of methyl groups. The spectrum was measured using the source operator A_1 (Figure 2a) with an INEPT-based $OCTE_2$ (Table 1) for 0.1 mM apo-CcmE in 20 mM Tris buffer (pH 7.5, T = 298 K) in a 10 mg/ml solution of Pf1 phages. The experiment was run on a Bruker *Avance* 600 spectrometer with t_{1max} = 16.55 ms and t_{2max} = 106.5 ms, an interscan delay of 1 s, 100 x 512 complex points resulting in an acquisition time of 6 hours. The time domain data was multiplied with a cosine function in the t_1 and t_2 dimensions and zero-filled to 512 and 8192 points, respectively. Radio-frequency carrier offsets were placed at 0.5 ppm (^{1}H) and 20 ppm (^{13}C).

Measurements of ^{1}H-^{1}H and ^{1}H-^{13}C Residual Dipolar Couplings in the Methyl Groups. As an illustrative application of the method, optimized CTEs are used to quantitatively measure residual ^{1}H-^{1}H (D_{HH}) and ^{1}H-^{13}C (D_{HC}) dipolar couplings (RDC) in a protein weakly aligned by means of Pf1 phages [48] at the optimal sensitivity using a minimal number of spectra. Two optimized CTEs were selected for this experimental implementation, $OCTE_2$ and $OCTE_3$, which correlate ^{13}C and ^{1}H chemical shifts in methyl groups using the pathways $A_1 \rightarrow C_1$ (complemented with $A_1 \rightarrow C_2$) and $A_2 \rightarrow C_1$ (complemented with $A_2 \rightarrow C_2$), respectively.

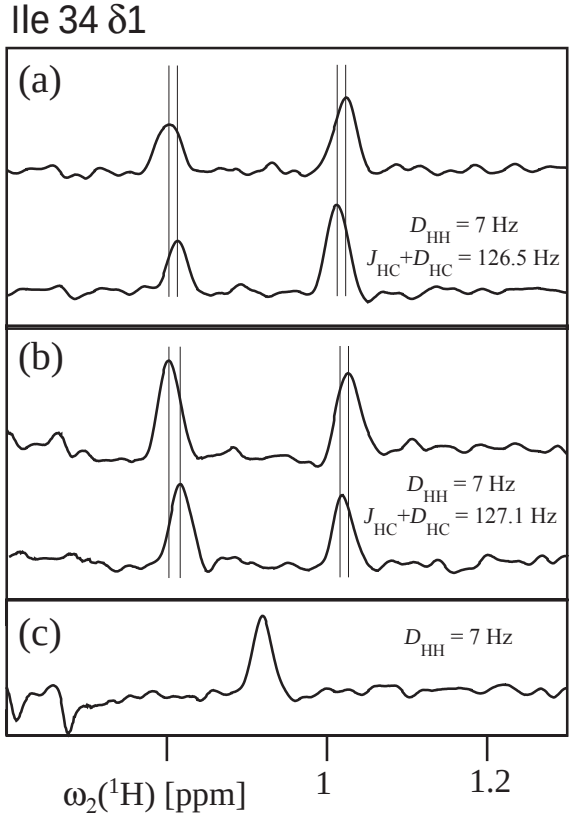


Figure 7. One-dimensional slices along the ω_2 dimension of the 2D [^{1}H,^{13}C]-correlation spectra, taken at the corresponding cross-peaks of the methyl groups δ1 of ILE34. In (a) slices are taken from the two-dimensional [^{13}C,^{1}H]-correlation spectra measured using the transfer $A_1 \rightarrow C_1$ (upper slice) and $A_1 \rightarrow C_2$ (lower slices) with $OCTE_2$ (see Table 1) and in (b) the transfer $A_2 \rightarrow C_1$ (upper slice) and $A_2 \rightarrow C_2$ (lower slices) with $OCTE_3$ (see Table 1). For comparison, the corresponding slice taken from the ($H^{\alpha\alpha}$)-selected two-dimensional [^{13}C,^{1}H]-spectrum [1] acquired in the same total experimental time is shown in c. The extracted RDC values are shown.

The performance of the newly designed experiments is demonstrated using diamagnetic heme chaperone apo-CcmE (17 kDa) [49] as a model system. Figure 6 shows a two-dimensional [^{13}C,^{1}H]-correlation spectrum obtained with $OCTE_2$. At the achieved average signal-to-noise ratio of about 20, the suppressed components are below the noise level. The values and signs of the residual dipolar coupling constant D_{HH} can be easily extracted from one-dimensional ^{1}H-slices taken through the corresponding peaks in the spectra, as shown in Figure 7. By repeating the measurements, the

statistical variation of the obtained values of D_{HH} was estimated to be in the range of 0.4 Hz. The experiment using the pathway of $A_2 \rightarrow C_1$ appears to be preferable, considering its tolerance to differential-relaxation effects during the chemical shift evolution period. Note that for holo-CcmE, the proposed experiment provided values of D_{HH} which are significantly smaller than the linewidth of ^{1}H-resonances, so that alternative methods designed to detect the individual transitions in the form of the ^{1}H antiphase magnetization would fail [15].

In summary, the use of the presently developed CTE optimization algorithm permitted the construction of two new complete NMR experiments achieving maximal possible coherence transfer efficiency and nearly complete suppression of undesired signals. Both features are critical for the quantitative and unbiased determination of residual dipolar couplings in methyl groups, and the sensitivity of the present experiments is doubled as compared to that of the previous experiment [1]. A variety of other experiments can be designed using the present optimization approach. For example, the inphase-type spectra can be augmented with the corresponding antiphase spectra to reduce spectral overlap. In this case, an optimized experiment should be particularly useful in combination with isotope-labeling schemes, where the protein is perdeuterated except for the methyl groups [47,50], thereby reducing losses due to relaxation and passive scalar and dipolar couplings.

Balance between Transfer Efficiency and Pulse Sequence Duration. The results of numerical CTE optimizations shown in Figure 4 suggest that it is feasible to trade some of the transfer efficiency for a shorter pulse-sequence duration. The corresponding closed-form analytical relationships between the efficiency of the polarization (coherence) transfer and the minimal time required to achieve it are established only for a few particular types of transfers [37]. Using the present optimization method, shorter pulse sequences can be produced by imposing stronger penalties on the duration of CTE (weighting factor ν in Equation 11). Unfortunately, however, the shorter pulse sequences are usually affected by a poorer suppression of undesired multiplet components, so that a balance between the CTE duration and artifact suppression should be maintained. If shorter CTEs with suboptimal transfer efficiencies are to be generated, an alternative to the penalty-function approach for artifact suppression (employed in the presented study) should be considered. For example, the most prominent (TROSY-enhanced) undesired central transitions of the ^{13}C-decoupled ^{1}H-triplet (Figure 2a) can be eliminated by the appropriate choice of the heteronuclear gradient-based echo/anti-echo pathway (Figure 2d). The previously suggested two-dimensional [^{13}C,^{1}H$^{\alpha\alpha/\beta\beta}$]^{Methyl}-HSQC experiment [1] is yet another practical example of a shorter and suboptimal CTE, nonetheless delivering a comparable spectral sensitivity even for the small protein considered above (Figure 7c). In this experiment, magnetic field gradients inserted between the constituent INEPT elements are employed to rebalance parallel coherence transfer pathways. This approach

provides an additional capability to control quantum evolution and can be used for systematic CTE design in future work

4. Supporting Information

In the published online version of this Chapter, the following data can be found at doi:10.016/j.jmr.2004.09.016: The matrix representation associated with the coherence transfer $A_2 \rightarrow C_1$ is shown. The document "Optimization of CTEs using steepest descent or simulated annealing" providing the details of the optimization algorithms applied in the present study is presented.

5. References

[1] Pervushin, K; Vögeli, B. *J. Am. Chem. Soc.* **2003**, *125*, 9566-9567.
[2] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366-12371.
[3] Ollerenshaw, J.E.; Tugarinov, V.; Kay, L.E. *Magn. Reson. Chem.* **2003**, *41*, 843-852.
[4] Tugarinov, V.; Hwang, P.M.; Ollerenshaw, J.E.; Kay, L.E. *J. Am. Chem. Soc.* **2003**, *125*, 10420-10428.
[5] Tugarinov, V.; Sprangers, R.; Kay, L.E. *J. Am. Chem. Soc.* **2004**, *126*, 4921-4925.
[6] Korzhnev, D.M.; Kloiber, K.; Kanelis, V.; Tugarinov, V.; Kay, L.E. *J. Am. Chem. Soc.* **2004**, *126*, 3964-3973.
[7] Bertini, I.; Luchinat, C.; Tarchi, D. *Chem. Phys. Lett.* **1993**, *203*, 445-449.
[8] Qin, J.; Delaglio, F.; Lamar, G.N.; Bax, A. *J. Magn. Reson. Ser. B* **1993**, *102*, 332-336.
[9] Ghose, R.; Prestegard, J.H. *J. Magn. Reson.* **1997**, *128*, 138-143.
[10] Boisbouvier, J.; Gans, P.; Blackledge, M.; Brutscher, B.; Marion, D. *J. Am. Chem. Soc.* **1999**, *121*, 7700-7701.
[11] Mandal, P.K.; Madhu, P.K.; Müller, N. *Chem. Phys. Lett.* **2000**, *320*, 269-276.
[12] Bertini, I.; Kowalewski, J.; Luchinat, C.; Parigi, G. *J. Magn. Reson.* **2001**, *152*, 103-108.
[13] Bertini, I.; Cavallaro, G.; Cosenza, M.; Kümmerle, R.; Luchinat, C.; Piccioli, M.; Poggi, L. *J. Biomol. NMR* **2002**, *23*, 115-125.
[14] Madhu. P.K.; Mandal, P.K.; N. Müller, N. *J. Magn. Reson.* **2002**, *155*, 29-38.
[15] Kaikkonen, A.; Otting, G. *J. Am. Chem. Soc.* **2001**, *123*, 1770-1771.
[16] Kontaxis. G.; Bax, A. *J. Biomol. NMR* **2001**, *20*, 77-82.
[17] Sibille, N.; Bersch, B.; Coves, J.; Blackledge, M.; Brutscher, B. *J. Am. Chem. Soc.* **2002**, *124*, 14616-14625.
[18] Tolman, J.R.; Flanagan, J.M.; Kennedy, M.A.; Prestegard, J.H. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 9279-9283.
[19] Tjandra, N.; Bax, A. *Science* **1997**, *278*, 1111-1114.
[20] Tjandra, N.; Omichinski, J.G.; Gronenborn, A.M.; Clore, G.M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732-738.
[21] Ernst, R.R.; Bodenhausen, G.; Wokaun, A. *The Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. **1987**, Oxford, Clarendon.
[22] Cavanagh, J.; Fairbrother, W.J.; Palmer, A.G.; Skleton, N.J. . *Protein NMR Spectroscopy. Principles and Pratice*. **1996**, Academic Press, San Diego.
[23] Schulte-Herbrüggen, T.; Untidt, T.S.; Nielsen, N.C.; O.W. Sørensen, O.W. *J. Chem. Phys.* **2001**, *115*, 8506-8517.
[24] Khaneja, N.; Luy, B.; Glaser, S.J. Boundary of quantum evolution under decoherence, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13162-13166.
[25] Khaneja, N.; Reiss, T.; Luy, B.; Glaser, S.J. *J. Magn. Reson.* **2003**, *162*, 311-319.
[26] Cavanagh, J.; Palmer, A.G.; Wright, P.E.; Rance, M. *J. Magn. Reson.* **1991**, *91*, 429-436.
[27] Barbato, G.; Ikura, M.; Kay, L.E.; Pastor, R.W.; Bax, A. *Biochemistry* **1992**, *31*, 5269-5278.
[28] Schleucher, J.; Schwendinger, M.; Sattler, M.; Schmidt, P.; Schedletzky, O.; Glaser, S.J.; Sørensen, O.W.; Griesinger, C. *J. Biomol. NMR* **1994**, *4*, 301-306.
[29] Sattler, M.; Schmidt, P.; Schleucher, J.; Schedletzky, O.; Glaser, S.J.; Griesinger, C. *J. Magn. Reson. Ser. B* **1995**, *108*, 235-242.
[30] Pervushin, K.; Wider, G.; Wüthrich, K. *J. Biomol. NMR* **1998**, *12*, 345-348.
[31] Untidt, T.; Schulte-Herbrüggen, T.; Luy, B.; Glaser, S.J.; Griesinger, C.; Sørensen, O.W.; Nielsen, N.C. *Mol. Phys.* **1998**, *95*, 787-796.

[32] Stoustrup, J.; Schedletzky, O.; Glaser, S.J.; Griesinger, C.; Nielsen, N.C.; Sørensen, O.W. *Phys. Rev. Lett.* **1995**, *74*, 2921-2924.
[33] Glaser, S.J.; Schulte-Herbrüggen, T.; Sieveking, M.; Schedletzky, O.; Nielsen, N.C.; Sørensen, O.W.; Griesinger, C. *Science* **280**, 421-424.
[34] Untidt, T.S.; Glaser, S.J.; Griesinger, C.; Nielsen, N.C. *Mol. Phys.* **1999**, *96*, 1739-1744.
[35] Khaneja, N.; Brockett, R.; Glaser, S.J. *Phys. Rev. A* **2001**, *63*, art. no.-032308.
[36] Khaneja, N.; Glaser, S.J.; Brockett, R. *Phys. Rev. A* **2002**, *65*, art. no.-032301.
[37] Reiss, T.O.; Khaneja, N.; Glaser, S.J. *J. Magn. Reson.* **2002**, *154*, 192-195.
[38] Vangunsteren, W.F.; Berendsen, H.J.C. *Angew. Chem.-Int. Edit. Engl.* **1990**, *29*, 992-1023.
[39] Kirkpatrick, S.; Gelatt, C.D.; Vecchi, M.P. *Science* **1983**, *220*, 671-680.
[40] Nielsen, N.C.; Schulte-Herbrüggen, T.; Sørensen, O.W. *Mol. Phys.* **1995**, *85*, 1205-1216.
[41] Nielsen, N.C.; Sørensen, O.W. *J. Magn. Reson. Ser. A* **1995**, *114*, 24-31.
[42] Untidt, T.S.; Schulte-Herbrüggen, T.; Sørensen, O.W.; Nielsen, N.C. *J. Phys. Chem. A* **1999**, *103*, 8921-8926.
[43] Müller, L. *J. Am. Chem. Soc.* **101**, 4481-4484 (1979).
[44] Bodenhausen, G.; Ruben, D.J. *Chem. Phys. Lett.* **1980**, *69*, 185-189.
[45] Levitt, M.H. *J. Magn. Reson.* **1997**, *126*, 164-182.
[46] Morris, G.A.; Freeman, R. *J. Am. Chem. Soc.* **1979**, *101*, 760-762.
[47] Tugarinov, V.; Kay, L.E. *J. Biomol. NMR* **2004**, *28*, 165-172.
[48] Hansen, M.R.; Müller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065-1074.
[49] Enggist, E.; Thöny-Meyer, L.; Güntert, P.; Pervushin, K. *Structure* **2002**, *10*, 1551-1557.
[50] Rosen, M.K.; Gardner, K.H.; Willis, R.C.; Parris, W.E.; Pawson, T.; Kay, L.E. *J. Mol. Biol.* **1996**, *263*, 627-636.
[51] Kay, L.E.; Keifer, P.; Saarinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663-10665.

Chapter IX

Discussion of NMR Techniques

In the following paragraphs we discuss the developed experiments in a context of NMR applications and alternative strategies to the tackled problems.

1. TROSY

Longitudinal ^{1}H Relaxation Optimization. Optimization of longitudinal ^{1}H relaxation is particularly useful for experiments in which the total acquisition time is critical. The predicted gain, up to 2.5-fold, is especially important for experiments with more than one indirect evolution periods. The simplest case is the LTROSY-HNCA experiment presented in Chapter II. An example for a non-[^{15}N,^{1}H]-TROSY based experiment is the ^{13}C-detected MQ-HACACO experiment [1].

Recently, longitudinal ^{1}H relaxation optimization has been proven to be worthwhile in reduced dimensionality NMR designed for applications in structural genomics. L-(4,3)D HN*N*C$^{\alpha,\beta}$C$^{\alpha}$ and L-(4,3)D HN*N*(CO)C$^{\alpha,\beta}$C$^{\alpha}$ are 4-dimensional "out-and-back" experiments with only two indirect evolution periods (t_1: ^{13}C$^{\alpha,\beta}$ and ^{13}C$^{\alpha}$; t_2: ^{15}N) and a sampling of two chemical shifts in a G-matrix Fourier transform (GFT) dimension (^{13}C$^{\alpha,\beta}$ and ^{13}C$^{\alpha}$ yielding ω(^{13}C$^{\alpha}$) and ω(^{13}C$^{\beta}$), respectively) [2]. For the purpose of assignment, these sequences were run 1.5 hours each on a 2 mM ubiquitin sample. Another approach to reduced dimensionality NMR is chemical shift coding in line shapes through apparent residual scalar coupling [3]. A signal-to-noise gain of about 10% was achieved in an HNCACB^{coded}HAHB experiment with a 77 residues protein by longitudinal ^{1}H relaxation optimization. There, two Gaussian pulses at water frequency and 1.7 ppm, respectively, were employed instead of ^{1}H^N-selective E-BURP pulses in order to keep water and part of the aliphatic protons along the +z-axis [4].

Transverse Relaxation Optimization. The TROSY technique has found wide application in pulse sequence design. Its use has largely been restricted to two spin systems where the CSA and dipole tensors nearly cancel. Recently, however, TROSY experiments for CH_2- and CH_3-groups were developed. The relaxation optimization for methylene groups makes use of favorable interferences between ^{1}H-^{13}C and ^{1}H-^{1}H dipole and ^{1}H and ^{13}C CSA relaxation mechanisms [5]. A gain of up to 3 times was observed in [^{13}C, ^{1}H]-correlation spectra. Kay and co-workers showed that the [^{13}C,^{1}H]-HMQC experiment is a TROSY experiment for methyl groups due to cancellation of intra-methyl dipolar relaxation rates in large proteins [6,7]. The rapidly relaxing coherences do not mix with those that are relaxing slowly throughout the duration of the pulse sequence. Using a [^{13}C,^{1}H]-HZQC (^{13}C/^{1}H zero-quantum) improves resolution even further [8]. The pulse sequences presented in the Chapters VII and VIII can also be used as TROSY-type experiments by proper selection of transition. For example, the sequence shown in Figure 1b of Chapter VIII offers a tool

for selection of exclusively slowly relaxing transitions in a similar way as described by References 6 and 7. All relevant transverse relaxation rates R_2 of C_+ for methyl CH_2D groups are listed in Reference 9. Therefore, one can theoretically design a TROSY-type pulse sequence for partially deuterated proteins.

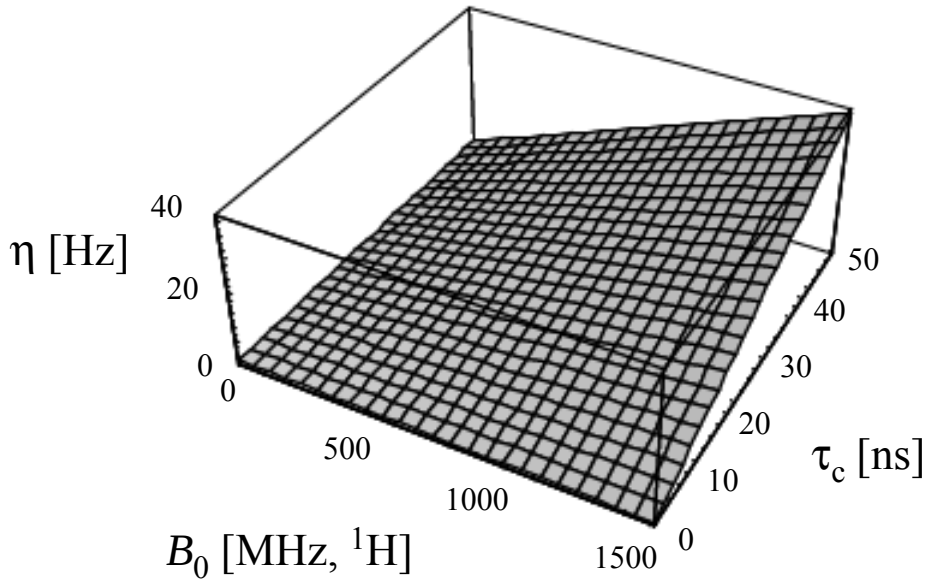


Figure 1. Cross-correlated relaxation rate η of CO, scalar coupled to spin C^{α} being in the α or β state as a function of the polarizing field B_0 and correlation time τ_c. The angle between the main axis of the CSA tensor and the C^{α}-CO vector is assumed to be 153°, $\Delta\delta(^{13}\text{CO CSA})$ = 130 ppm and the CSA asymmetry parameter is 0.7. For more elaborate values see References 15-17.

Another mechanism affecting cross-correlated relaxation is the Curie spin relaxation (CSR) involving unpaired electron spins in paramagnetic proteins [10,11] and an additional TROSY effect is caused by cross-correlated relaxation between *i*-*j* dipole and *i*-e^- CSR interactions for a nuclear spin *i* *J*-coupled to a nuclear spin *j* [12]:

$$\eta^{\text{param}ij} = \frac{1}{30\pi}\left(\frac{\mu_0}{4\pi}\right)^2 \frac{B_0\gamma_i{}^2\gamma_j\mu_B{}^2 g_e{}^2 hS(S+1)}{r_{e\text{-}i}{}^3 r_{ij}{}^3 k_B T} Q(\Theta_{eij})\left[4\tau_D + \frac{3\tau_D}{1+\omega_i{}^2\tau_D{}^2}\right], \quad (1)$$

$$1/\tau_D = 1/\tau_c + 1/\tau_M\ . \quad (2)$$

Here, g_e is the electronic g factor (assumed to be isotropic), μ_B is the electronic Bohr magneton, *S* is the electronic spin multiplicity, *T* is the temperature, k_B is Boltzmann constant, τ_M is the residence time of the ion in cases of chemical exchange and a third term on the right side of Equation 2 reflecting the electron spin-lattice relaxation time is safely neglected. Unfortunately, the position of the narrowest component within a multiplet depends on the bond orientation of the involved spins due to the dependence on $Q(\Theta_{eij})$ and the effect drops with $1/r_{e\text{-}i}{}^3$ as the distance $r_{e\text{-}i}$ between the

paramagnetic center and *i* increases [13]. This leads to the fact that only peaks from spins close to the paramagnetic center exhibit a measurable effect. However, in parallel, the paramagnetic autorelaxation scales with $1/r_{e\text{-}i}^6$ making these peaks weak or indistinguishable from the background.

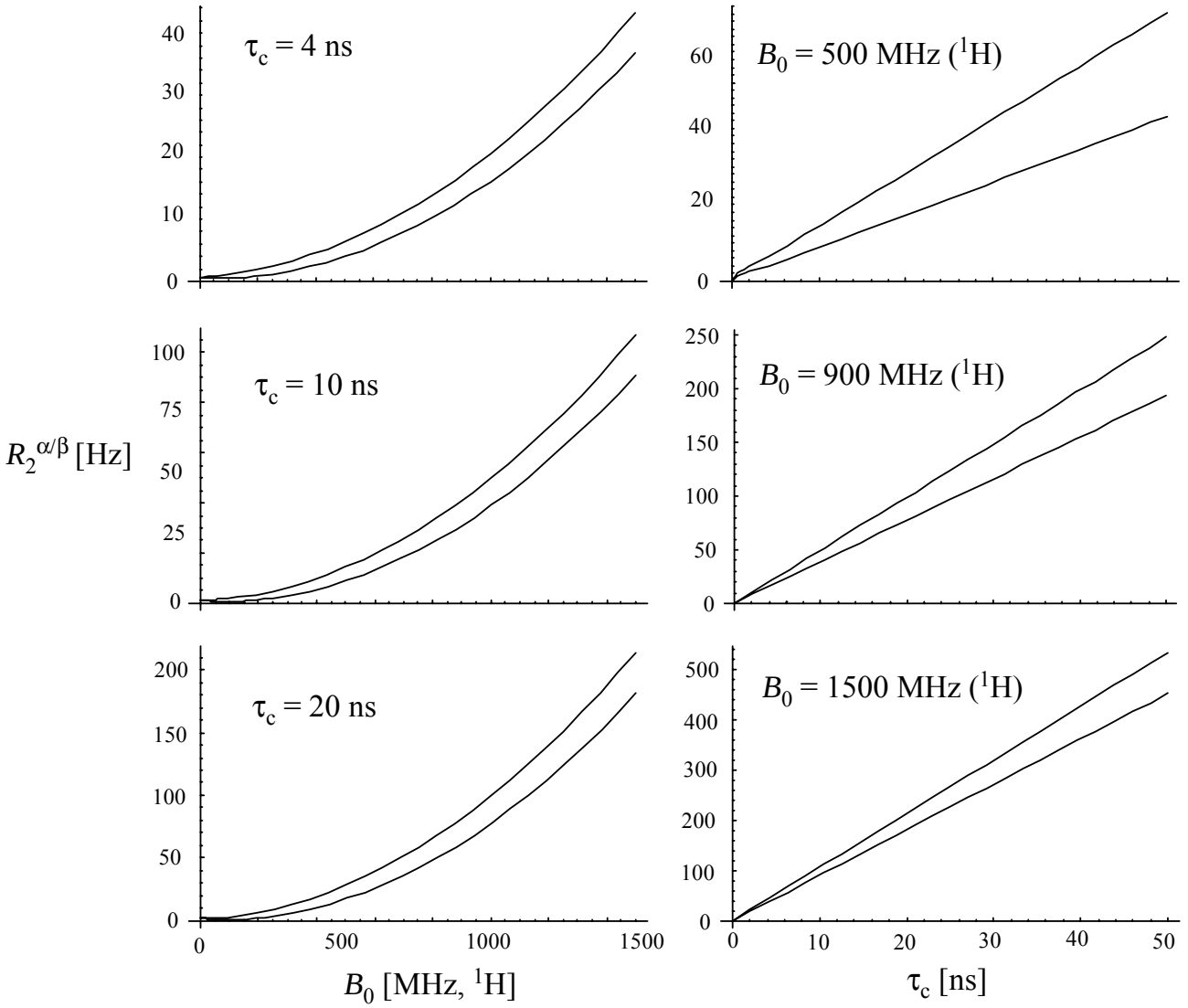


Figure 2. Single quantum relaxation rates $R_2^{\alpha/\beta}$ of CO, scalar coupled to spin C^α being in the α or β state. Left panel: $R_2^{\alpha/\beta}$ as a function of B_0 at τ_c = 5, 10 and 20 ns. Right panel: $R_2^{\alpha/\beta}$ as a function of τ_c at B_0 = 500, 900 and 1500 MHz (^{1}H). R_2^α is the higher rate in all cases. All other parameters are the same as given in the caption of Figure 1.

Because ^{13}C spectroscopy is one of the main topics of this thesis, it is tempting to analyze its potential for TROSY effects. No ^{13}C-detection-relevant spin pair (including spins like $^{13}C^\alpha$, ^{13}CO or ^{15}N) features an adequate cancellation of CSA and dipole/dipole relaxation rates as is observed in the [^{15}N,^{1}H]-TROSY [14]. In Figures 1 and 2, respectively, the cross-correlated relaxation rate η and the transverse relaxation rates $R_2^{\alpha/\beta}$ of CO (*I*), scalar coupled to spin C^α (*S*) being in the α or β state are

calculated using formulae analoguous to the expressions in Chapter III. The non-uniform values for the relaxation rates depend on the specific surrounding and only a rough guess can be given. η scales approximately linearly with B_0 and the correlation time τ_c within the range of experimental relevance (Figure 1).

However, since the autorelaxation rates are increased by comparable factors, the relaxation rates of the single quantum states do not exhibit a strong TROSY effect in the experimental range (as shown in Figure 2), though a reasonably strong effect may be observed with B_0 fields of approximately 200 MHz (left panel of Figure 2). A [$^{13}C^{\alpha}$,^{13}CO]-HSQC spectrum with $J_{C\alpha CO}$-coupling during the acquisition period is presented in the ^{13}C-spectroscopy part of this Chapter (Figure 4). This spectrum was recorded at 900 MHz using human ubiquitin as a test sample (correlation time $\tau_c \approx 4$ ns). As predicted (Figure 2), the doublet components feature only slight intensitiy differences.

In conclusion, the TROSY technique does not promise to become a pathbreaking tool in ^{13}C spectroscopy. However, a ^{13}C-^{13}H spin pair detected using ^{13}C spins does offer a more favorable trade-off between dipole-CSR cross-correlated relaxation and paramagnetic autorelaxation than the commonly used ^{1}H-^{15}N spin pair detected on ^{1}H. In a ^{1}H-detected spectrum peaks of spins as close as 10 Å to the paramagnetic center are observed. ^{13}C-detected spectra, in turn, show even peaks of spins separated by only 5 Å [18,19]. The cross-correlated relaxation rates may be compared as:

$$C(r_{eC}, r_{eH}) = \eta^{paramCH} / \eta^{paramHN} = [\gamma_C{}^2 \gamma_H / (r_{eC} r_{CH})^3] / [\gamma_H{}^2 \gamma_N / (r_{eH} r_{HN})^3] \quad (3)$$

At equal distances from the paramagnetic center, $\eta^{paramHN}$ is twice as strong as $\eta^{paramCH}$, but for the closest observable peaks in each case Equation 3 yields

$$C(5\text{ Å}, 10\text{ Å}) \approx 4.3\ . \quad (4)$$

Exactly in the sphere within which only ^{13}C-detected spins are observable, the TROSY effect may boost the advantage of ^{13}C-spectroscopy further.

Cross-Correlated Relaxation. The [^{15}N,^{13}C]-DQ/ZQ-[^{15}N,^{1}H]-TROSY-E.COSY experiment is one of many triple-resonance experiments correlating $^{1}H^{N}$, ^{15}N and $^{13}C^{\alpha}$ [20]. These experiments provide spectra for backbone assignment, backbone short-range angular information by measurement of *J*-couplings or cross-correlated relaxation rates, and global angular information by measurement of residual dipolar couplings. The conversion of these triple resonance experiments to their [^{15}N,^{1}H]-TROSY based counterparts doubles the upper molecular weight limit of accessible proteins. However, the linear combination of the cross-correlated relaxation rates of multiple quantum of ^{15}N and $^{13}C^{\alpha}$ *J*-coupled to $^{1}H^{N}$ and $^{1}H^{\alpha}$, respectively, no longer provides a single-term expression as derived in Reference 21. The selection of the TROSY state with respect to $^{1}H^{N}$ results in the

Expressions 10 and 13 of Chapter III. For the non-proline case (Expression 10), two terms are left with the following angular dependencies:

$$Const_1 Q(\Theta_{D(HC)D(HN)}) + Const_2 Q(\Theta_{D(HC)CSA(N)}) = R_\beta^{DQ} - R_\alpha^{DQ} - R_\beta^{ZQ} + R_\alpha^{ZQ} \quad (5)$$

In Chapter III it is proven that it is still possible to obtain angular restraints for the dihedral angle formed by the $^1H^\alpha$-$^{13}C^\alpha$ and $^1H^N$-^{15}N vectors because a uniform value can be assigned to the main axis of the ^{15}N CSA tensor and its orientation is only slightly dependent on the chemical environment. In addition, measurement of the residual dipolar couplings $D_{H\alpha HN}$ and $D_{H\alpha C\alpha}$ from the E.COSY pattern could be yet another source of structural information.

2. ^{13}C-Detection Spectroscopy

Higher dimensionality. In search for the inclusion of a third dimension in a [^{13}C,^{13}C]-correlation experiment with partially or fully deuterated proteins, different pulse sequences were proposed. Independently of the channel configuration, a J_{CC}-resolved dimension can be implemented in a [^{13}C,^{13}C]-TOCSY experiment as depicted in Figure 3a. After the FLOPSY-mixing period, a 90º pulse flips the magnetization along z into the transverse plane. The evolution under the J_{CC}-coupling Hamiltonian is registered by incrementation of t_2, while the chemical shift evolution is refocused by a 180º pulse and ^{13}CO is decoupled by selective pulses. Just before the last two pulses, the density operator σ is:

$$\sigma(t_1, t_2) \sim \cos(\pi J_{CiCk} t_2) I_y^i - \sin(\pi J_{CiCk} t_2) 2 I_x^i I_z^k , \quad (6)$$

where J_{CiCk} is the coupling between atom *i* and its neighbor *k*. Similar expressions consisting of four and eight terms are obtained for two and three coupled spins, respectively. If gradient 4 is omitted, each component of Expression 6 is detected. However, the spectrum cannot be phased in the *J*-resolved dimension. Otherwise, gradient 4 dephases all terms except for the one proportional to I_y^i which describes the in-phase magnetization.

The dispersion along the additional dimension of the strongly overlapped peaks in the 2D [^{13}C,^{13}C]-TOCSY is poor. Figure 3b shows a [^{13}C,^{13}C]-TOCSY plane at 36 Hz in the *J*-resolved dimension of a BsCM sample, in which mostly resonances of C^β with two coupling partners are visible (unpublished results). All peaks with the same number of coupled nuclei lie in the same planes because of the high uniformity of J_{CC}-couplings. In particular, the doublet region does not overlap with the quadruplet regions. Therefore the separation in the *J*-resolved dimension does not offer significant advantage over the 2D TOCSY [22]. The resolution of the splitting pattern is higher

in the direct dimension because only a few points can be run in the *J*-resolved dimension due to time limitations. Thus, *J*-measurement in the indirect dimension is of limited practicality.

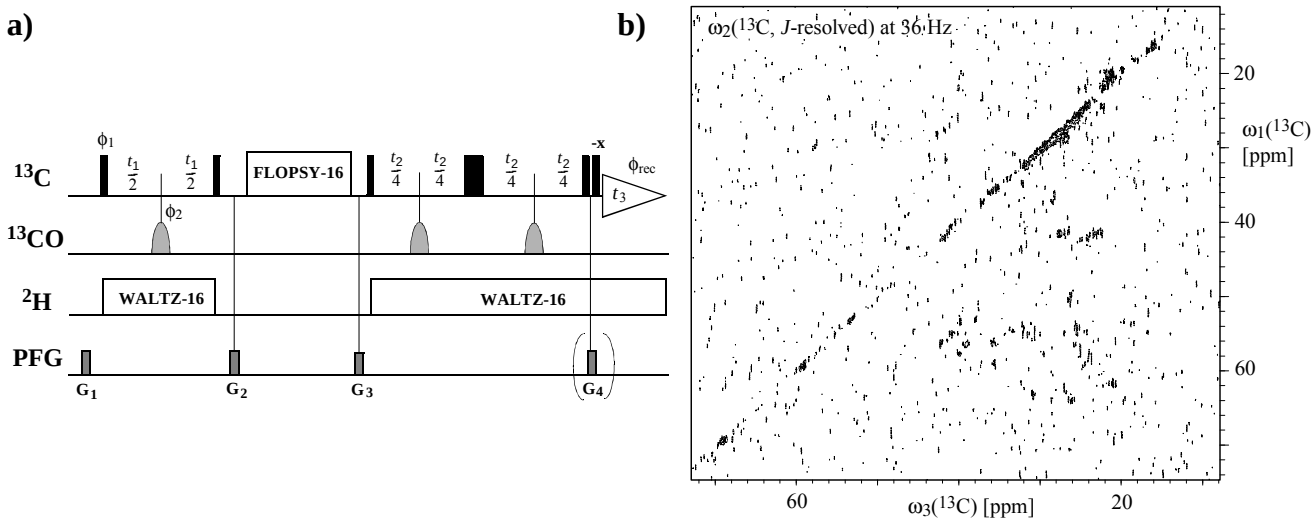


Figure 3. a) The experimental scheme of the J_{CC}-resolved 3D [^{13}C,^{13}C]-TOCSY experiment. Narrow (wide) black bars represent non-selective 90° (180°) pulses on ^{13}C. The grey curved shapes indicate selective 180° Gaussian pulses on the ^{13}CO resonance with the duration of 100 ms. The ^{13}C, ^{13}CO and ^{2}H carriers are positioned at 40.0 ppm, 174.0 ppm and 3.0 ppm, respectively. The boxes on the line labeled PFG indicate sine-shaped pulsed magnetic field gradients along the z-axis with 1.0 ms duration except for G_2, which has a duration of 0.8 ms: G_1, amplitude 30 G/cm; G_2, 33 G/cm; G_3, 19 G/cm; G_4, 21 G/cm. The phase cycle is ϕ_1 = {x, -x}, ϕ_2 = {x, x, -x, -x}, ϕ_{rec} = {x, -x} and x for all pulses without phase specification. The WALTZ-16 and FLOPSY-16 pulse trains are applied as described in Chapter VI. Gradient 4 is optionally applied for selection of the in-phase term in the J_{CC}-resolved dimension only. **b)** [^{13}C,^{13}C]-TOCSY plane at ω_2 (^{13}C, *J*-resolved) = 36 Hz of ^{2}H,^{13}C,^{15}N-labeled BsCM (See Chapter II). The observed resonances stem mainly from $^{13}C^{\beta}$ atoms with two coupled spins. The data was recorded on a Bruker *AVANCE* 500 MHz spectrometer equipped with a cryogenic Z-gradient DUAL ^{13}C/^{1}H probe. t_{2max} = 31.8 ms, the relaxation delay d_1 = 4.5 s resulting in 2 days measuring time, 180 x 8 x 1024 complex points were multiplied by a cosine function and zero-filled to 1024 x 32 x 2048 complex points. All other parameters are set as given in Chapter VI.

A second way of using exclusively ^{13}C dimensions is the use of ^{13}CO as a polarizing nucleus followed by an INEPT magnetization transfer to $^{13}C^{\alpha}$ and a [^{13}C,^{13}C]-TOCSY element. ^{13}CO promises stronger dispersion than J_{CC}, but since all magnetization starts from the backbone an intensity drop is expected for long side chains.

If a ^{15}N channel is available, a ^{15}N dimension can be implemented as demonstrated by the Bertini group with the CANCO experiments [23,24]. Clearly, it is of disadvantage to use ^{15}N as polarizing nucleus or as detection nucleus because of its low gyromagnetic ratio. Therefore, in proteins ^{15}N-correlated experiments are best suited for detection of backbone resonances.

Conclusively, the introduction of a ^{1}H dimension, as presented in Chapter V, offers the most powerful approach to high-dimensional ^{13}C spectroscopy. A ^{1}H channel is required and a ^{2}H channel

would allow the most advantageous experimental setup, because decoupling could be performed on both ^{1}H and ^{2}H simultaneously on a partially deuterated sample. The most severe obstacle of 3D ^{13}C-detection NMR, the practical limit on experiment time, is met best with the HCC-TOCSY. A typical 2D HCC-TOCSY requires ~ 20 hours. 10 points in a third dimension extends the experiment time to about one week. However, the high signal-to-noise ratio of the HCC-TOCSY makes it possible to reduce the experiment time to two days. The same arguments also hold for the introduction of a third dimension to other types of [^{13}C,^{13}C]-correlation experiments such as NOESY and COSY.

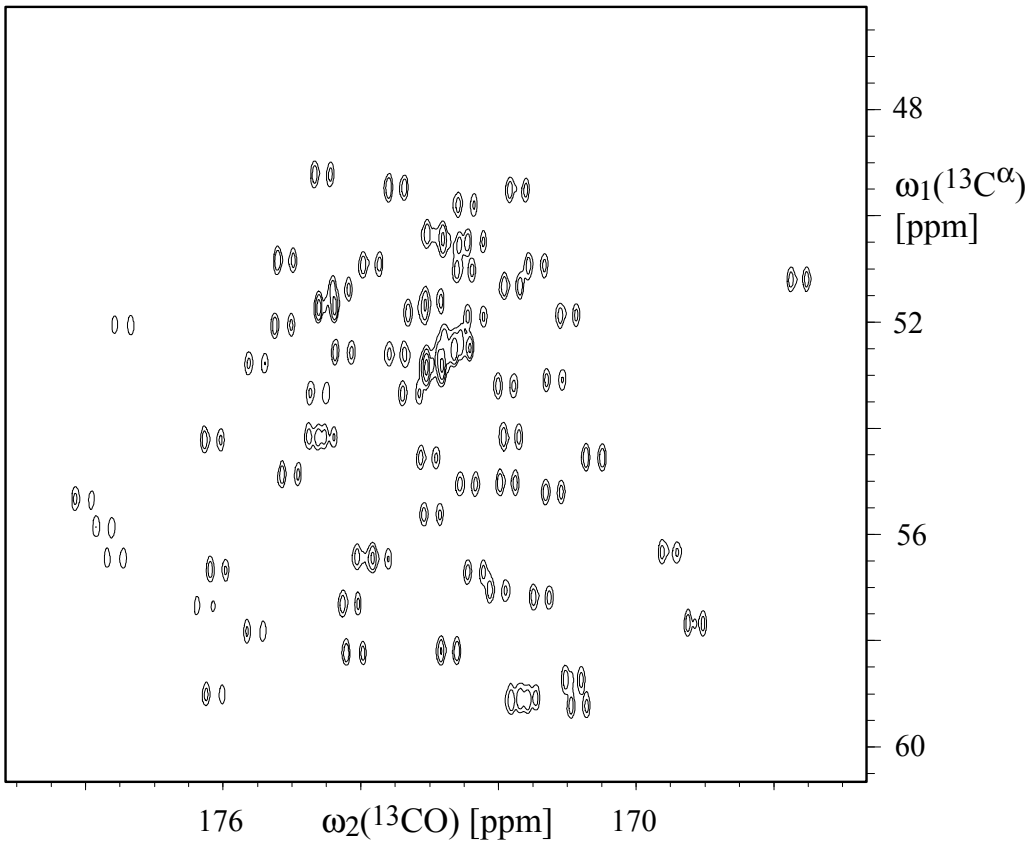


Figure 4. [$^{13}C^{\alpha}$,^{13}CO]-HSQC spectrum of a ^{2}H,^{13}C,^{15}N-labeled human ubiquitin sample [20]. The data was recorded on a Bruker *AVANCE* 900 MHz spectrometer equipped with a conventional Z-gradient DUX $^{13}C\{^{15}N,^{2}H\}$ probe, which is optimized for ^{13}C detection. The carrier offsets placed at 174 ppm (^{13}CO), 56 ppm ($^{13}C^{\alpha}$), 118 ppm (^{15}N) and 4 ppm (^{2}H). ^{2}H was decoupled using WALTZ-16 during the whole sequence, ^{15}N during the indirect evolution period with a Gaussian shaped 180º pulse and during acquisition with GARP-4 employing standard parameters. Other parameters were t_{1max} = 13.8 ms, t_{2max} = 113.1 ms and the relaxation delay d_1 = 1.5 s. 200 x 512 complex points resulting in 5.5 hours measuring time were multiplied by a cosine function and zero-filled to 1024 x 1024 complex points.

Peak Overlap. Two strategies for peak overlap reduction are proposed within the scope of this thesis. First, decoupling of involved ^{13}C and ^{15}N nuclei, outlined in Chapter IV, is a powerful tool for both protein and RNA spectroscopy. Multiple-band-selective homonuclear decoupling during acquisition especially offers the possibility for using decoupling combinations designed for the

specific experimental demand. Modern hardware allows high selectivity in multiple-band-selective pulses. Second, an additional dimension as discussed above can disperse overlapping peaks efficiently, especially if the third dimension encodes ^{1}H frequencies. Currently, in this laboratory, a third possibility is being developed. Spin-state-selective spectroscopy eliminates selected peaks in multiplets. Similar to TROSY, the COCAINE experiment chooses one component of a doublet in the direct dimension by exclusively polarizing of an individual spin transition [25]. Although designed for [$^{13}C^{\alpha}$,^{13}CO]-spin pairs, it can be applied to any spin pair such as [^{15}N,^{1}H]. Further reduction of overlap is achieved, if the selected component is significantly less broad than the de-selected component. Figure 4 shows a [$^{13}C^{\alpha}$,^{13}CO]-HSQC spectrum with $J_{C\alpha CO}$-coupling during the acquisition period (unpublished results). The experiment was run at 900 MHz using a ^{2}H,^{13}C,^{15}N-labeled human ubiquitin sample (correlation time $\tau_c \approx 4$ ns). The doublet components exhibit only slight intensitiy difference as predicted by Figure 2. It can be concluded that the TROSY effect is not expected to have high impact on ^{13}C-spectroscopy.

Paramagnetic Molecules. As calculated in Equation 2 of Chapter I, ^{13}C relaxation due to paramagnetic mechanisms is attenuated 16 times compared to ^{1}H relaxation. The radius of the "dark sphere" around the paramagnetic center is reduced to 5 Å compared to a radius of 10 Å for ^{1}H [18,19]. There, the paramagnetic TROSY effect becomes stronger than that observed at the limiting distance of 10 Å using in ^{1}H-detection NMR. In turn, the evaluation of the cross-correlated relaxation rate $\eta^{paramCH}$ may be used to restrain bond vectors in structure calculations.

Covariance Spectroscopy. Indirect covariance spectroscopy can overcome the intrinsically low sensitivity associated with ^{13}C detection [26-28]. A typical pulse sequence features an indirect evolution period on ^{13}C (t_1), a magnetization transfer step to ^{1}H, a [^{1}H,^{1}H]-TOCSY mixing period, and detection on ^{1}H (t_2) [28]. The standard 2D frequency matrix $S(\omega_1,\omega_2)$ obtained by Fourier transform of the signal is used to calculate the covariance matrix C:

$$C(\omega_1,\omega_1') = \int S(\omega_1,\omega_2)S(\omega_1',\omega_2)\mathrm{d}\omega_2 = S \bullet S^{T} \qquad (7)$$

The final spectrum $F(\omega_1,\omega_1') = C^{1/2}$ displays the same spin-connectivities as a ^{13}C-detected [^{13}C,^{13}C]-TOCSY spectrum because cross-peaks (ω_1,ω_1') appear if both ω_1 and ω_1' form a cross-peak with ω_2 in S. However, the spectral resolution in both dimensions of F is determined by the unfavorable sampling in t_1.

^{13}C-^{13}C RDCs. So far, only a few ^{13}C-^{13}C RDC values are extractable from side chains. The use of homonuclear decoupling for RDC measurements from simplified splitting patterns is still challenging. Attempts to extract reliable RDCs from [^{13}C,^{13}C]-TOCSY spectra of RNA as shown in Chapter VI have been unsuccessful so far due to the high sensitivity of small RDCs to distortional

effects. Further investigation is expected to make RDC measurements tractable. Inclusion of a third dimension, e.g. ^{1}H, together with appropriate homonuclear decoupling in experiments such as ^{1}H-[^{13}C,^{13}C]-TOCSY or ^{1}H-[^{13}C,^{13}C]-COSY, is expected to make possible the extraction of a majority of the side-chain ^{13}C-^{13}C RDC values. They are a powerful tool for precise bond constraints of side chains and dynamical analysis.

Overall, especially with regard to the increasing availability of cryogenic ^{13}C-optimized probeheads, the advantages of ^{13}C spectroscopy show much promise for biomolecular NMR applications.

3. Methyl Groups

Transfer Optimization. So far, relaxation effects are not considered in the analysis. A more sophisticated optimization may be derived from the Liouville-Von Neumann equation describing the evolution of the density operator $\sigma(t)$ under the Hamiltonians H (given by Equations 6.1-6.3 in Chapter VIII) and a relaxation superoperator $\hat{R}$ [29]:

$$\frac{\mathrm{d}\sigma(t)}{\mathrm{d}t} = -\mathrm{i}[H,\sigma(t)] - \hat{R}(\sigma(t)-\sigma(0)). \tag{8}$$

While the addtitional terms can safely be neglected in Equation 6.1, the contributions to Equations 6.2 and 6.3 can modify the propagator $U(\boldsymbol{x})$ given by Equation 7 significantly.

Second, pulses on different nuclei and evolution periods with J-coupling is considered. However, a more complete analysis would include phase cycling, which can get rid of undesired components of the density matrix, and use of gradients, which can delete or select certain terms. This optimization would be implemented in additional discrete dimensions similar to the dimension for the element count.

A substantial loss in the signal-to-noise ratio may come from the non-uniformity of the coupling constant. If the experiments are used for RDC measurement, the $\boldsymbol{x}$ vector is optimized for J, whereas in general it is mismatched for $J + D$. It would be instructive to find CTE which are as insensitive as possible (under conservation of the selection and efficiency performance) [30]. An optimization for a effective coupling,

$$J^{\mathrm{eff}}(\theta, \phi) = J + D(\theta, \phi), \tag{9}$$

does not solve the problem, because $D(\theta, \phi)$ may ranges from negative to positive values with roughly the same absolute values depending on the polar angles θ and ϕ. For experiments on a

protein with an expectedly usual distribution of $J^{eff}(\theta, \phi)$, optimization for J still meets the demand best.

The optimization of magnetization transfers on the matrix level comprises a fundamental treatment and therefore is a general method which may find widespread application in complex spin network calculations.

Khaneja et al. have applied an algorithm called GRAPE (gradient ascent pulse engineering) used in laser spectroscopy to NMR pulse-sequence optimization [31]. This approach is orders of magnitude faster than the method used here. GRAPE can optimize a much higher number of pulse parameters. However, it does not necessarily converge to the global minimum.

***D*HH RDCs.** In the Appendix of Chapter VII, the histogram in Figure 4 presents 21 D_{HH} couplings obtained at 600 MHz from an apo-CcmE sample which is partially aligned with Pf1 bacteriophage. To increase the signal-to-noise ratio and decrease the peak linewidth, the experiment was repeated at 900 MHz. 16 D_{HH} values, for which the resonance assignment is unambiguous, were extracted by measuring half the relative shift between the $H^{\alpha\alpha}$ and $H^{\beta\beta}$ peaks. In the dipolar Hamiltonian for alike spins numbered with *i* and *j* [32],

$$H_D = -\left(\frac{\mu_0}{4\pi}\right)\frac{h\gamma_H^2}{\pi r_{HH}^3}\sum_{i<j}\left\{I_z^i I_z^j - \tfrac{1}{4}\left(I_+^i I_-^j + I_-^i I_+^j\right)\right\}\left(3\cos^2\psi_{ij} - 1\right)/2 \qquad (10)$$

The angular dependence on the angle ψ_{ij} between the polarizing field B_0 and the *i*-*j* vector can be replaced by the angle ψ_{C3} between B_0 and the C_3 symmetry axis in the effective Hamiltonian part [33]. D_{HH} then takes the same form as the splitting between two weekly coupled spins ½ scaled by a factor of ¾ [34]:

$$D_{HH} = S(\mu_0/4\pi)\gamma_H^2/r_{HH}^3(h/2\pi^2)\,{}^3/_4\,A_a\,\{(3\cos^2\theta_{C3} - 1)/2 + {}^3/_4\,\eta\,(\sin^2\theta_{C3}\cos 2\phi_{C3})\}. \qquad (11)$$

Here θ_{C3} and ϕ_{C3} are the angles defining the mean orientation of the C_3 symmetry axis in the principle axis system of the aligment tensor $\tilde{A}$. Therefore, in the same way as described in Chapter VI, the measured values, designated D_{HH}^{exp}, can be fit against the values D_{HH}^{calc}, which are calculated from the mean NMR solution structure obtained from 20 conformers [35] and the simultaneously determined aligment tensor. Error may stem from the assumed uniformity of the order parameter S^2. In fact, a formula has been derived recently which describes the decrease of the order paramater as the number of consecutive mobile dihedral angles between the methyl group and the protein backbone increases [36]. As shown in Chapter VI, a fit is still possible for side-chain RDCs. Indeed, one obtains a correlation coefficient of 0.970 and a root-mean-square deviation of 0.47 Hz (Figure 5). It should be mentioned that the RDC of γ2 of Ile34 is discarded in the fit because the measured value of ~ −5 Hz contradicts strongly to a theoretical positive value. This discrepancy

can only be explained by an incorrect resonance assignment or a poorly defined orientation of the methyl group in the structure. So far, structural refinement by RDCs was carried out mainly on protein backbones. Here, it is demonstrated that measurements of D_{HH} in methyl groups can also serve as a powerful refinement tool in structure determination.

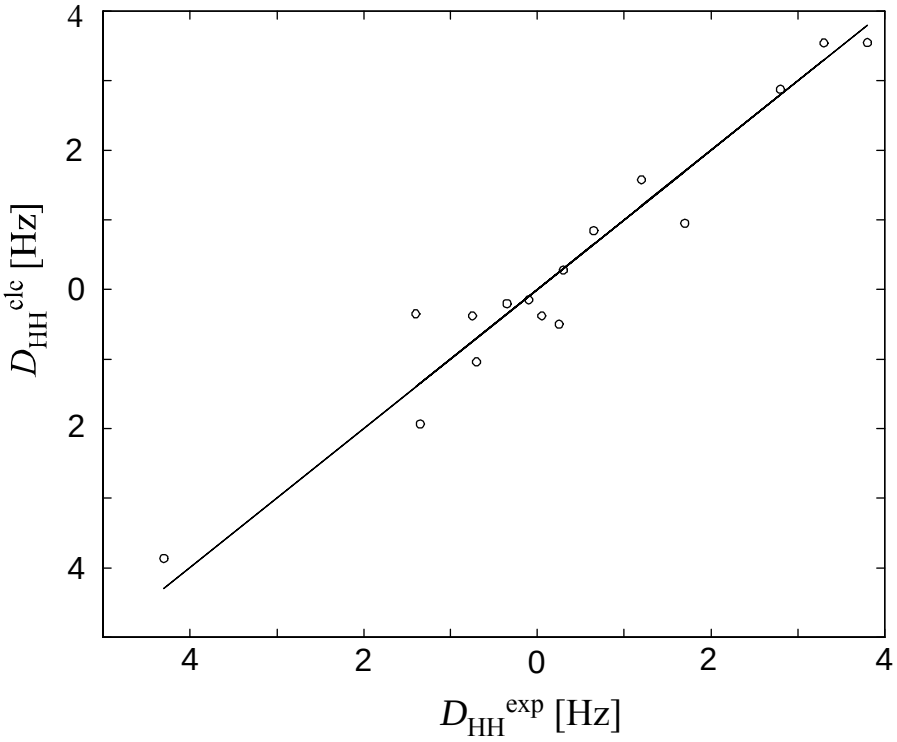


Figure 5. Calculated versus experimental values of methyl D_{HH} of apo-CcmE. 16 values were used in the fit yielding a correlation coefficient of 0.970 and a root-mean-square deviation of 0.47 Hz. The Cartesian components in the principle axis system are (A_{xx}, A_{yy}, A_{zz}) x ($^3/_2$)S = (-0.53, -3.18, 3.72) x 10^{-4} and the Euler angles orienting the principle axis system are (α, β, γ) = (73.8°, 28.3°, 26.7°).

A considerable number of RDC values are not extractable from the 2D spectra because of peak overlap. Van Melckebeke at al. designed a 3D methyl [^{13}C,^{1}H]-correlation experiment, which is Hadamard amino-acid-type encoded in the third dimension [37,38]. This concept may be included in spin-state-selective experiments in order to measure RDCs from selectively specific amino-acid types.

4. References

[1] Pervushin, K.; Eletsky, A. *J. Biomol. NMR* **2003**, *25*, 147-152.
[2] Atreya, H.S.; Szyperski, T. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9642-9647.
[3] Kwiatkowski, W.; Riek, R. *J. Biomol. NMR* **2003**, *25*, 281-290.

[4] Pegan, S.; Kwiatkowski, W.; Choe, S.; Riek, R. *J. Magn. Reson.* **2003**, *165*, 315-319.
[5] Miclet, E.; Williams, D.C.; Clore, G.M.; Bryce, D.L.; Boisbouvier, J.; Bax, A. *J. Am. Chem. Soc.* **2004**, *126*, 10560-10570.
[6] Tugarinov, V.; Hwang, P.M.; Ollerenshaw, J.E.; Kay, L.E. *J. Am. Chem. Soc.* **2003**, *125*, 10420-10428.
[7] Ollerenshaw, J.E.; Tugarinov, V.; Kay, L.E. *Magn. Reson. Chem.* **2003**, *41*, 843-852.
[8] Tugarinov, V.; Sprangers, R.; Kay, L.E. *J. Am. Chem. Soc.* **2004**, *126*, 4921-4925.
[9] Marion, D. *J. Magn. Reson.* **2003**, *166*, 202-214.
[10] Guéron, M. *J. Magn. Reson.* **1975**, *19*, 58-66.
[11] Vega, A.J.; Fiat, D. *Mol. Phys.* **1976**, *31*, 347-355.
[12] Bertini, I.; Luchinat, C.; Tarchi, D. *Chem. Phys.Lett.* **1993**, *203*, 445-449.
[13] Madhu, P.K.; Grandori, R.; Hohenthanner, K.; Mandal, P.K.; Müller,N. *J. Biomol. NMR* **2001**, *20*, 31-37.
[14] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366-12371.
[15] Oas, T.G.; Hartzell, C.J.; McMahon, T.J.; Drobny, G.P.; Dahlquist, F.W. *J. Am. Chem. Soc.* **1987**, *109*, 5956-5962.
[16] Dayie, K.T.; Wagner, G. *J. Am. Chem. Soc.* **1997**, *119*, 7797-7806.
[17] Cornilescu, G.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 10143-10154.
[18] Bermel, W.; Bertini, I; Felli, I.C.; Kümmerle, R.; Pierattelli, R. *J. Am. Chem. Soc.* **2003**, *125*, 16423-16429.
[19] Babini, E.; Bertini, I.; Capozzi, F.; Felli, I.C.; Lelli, M.; Luchinat, C. *J. Am. Chem. Soc.* **2004**, *126*, 10496-10497.
[20] Cavanagh, J.; Fairbrother, W.J.; Palmer, A.G.; Skelton, N.J., *Protein NMR Spectroscopy: Principles and Practice*. **1996**, Academic Press, New York.
[21] Reif, B.; Diener, A.; Hennig, M.; Maurer, M.; Griesinger, C. *J. Magn. Reson.* **2000**, *143*, 45-68.
[22] Eletsky, A.; Moreira, O.; Kovacs, H.; Pervushin, K. *J. Biomol. NMR* **2003**, *26*, 167-179.
[23] Bertini, I.; Duma, L.; Felli, I.C.; Fey, M.; Luchinat, C.; Pierattelli, R.; Vasos, P.R. *Angew. Chem. Int. Ed. Engl.* **2004**, *43*, 2257-2259.
[24] Bermel, W.; Bertini, I.; Felli, I.C.; Pierattelli, R.; Vasos, P.R. *J. Magn. Reson.* **2005**, *172*, 325-329.
[25] Lee, D.; Vögeli, B.; Pervushin, K. *J. Biomol. NMR* **2005**, *accepted*.
[26] Brüschweiler, R.; Zhang, F. *J. Chem. Phys.* **2004**, *120*, 5253-5260.
[27] Brüschweiler, R. *J. Chem. Phys.* **2004**, *121*, 409-414.
[28] Zhang, F.; Brüschweiler, R. *J. Am. Chem. Soc.* **2004**, *126*, 13180-13181.
[29] Ernst, R.R.; Bodenhausen, G.; Wokaun, A. *The Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. **1987**, Oxford, Clarendon.
[30] Schulte-Herbrüggen, T.; Untidt, T.S.; Nielsen, N.C.; Sørensen, O.W. *J. Chem. Phys.* **2001**, *115*, 8506-8517.
[31] Khaneja, N.; Reiss, T.; Kehlet, C.; Schulte-Herbrüggen, T.; Glaser, S.J. *J. Magn. Reson.* **2004**, *172*, 296-305.
[32] Abragam, A. *The Principles of Nuclear Magnetism.* **1961**, *17*, Oxford, Clarendon.
[33] Allen, P.S. *J. Chem. Phys.* **1968**, *48*, 3031-3036.
[34] Kaikkonen, A.; Otting, G. *J. Am. Chem. Soc.* **2001**, *123*, 1770-1771.
[35] Enggist, E.; Thöny-Meyer, L.; Güntert, P.; Pervushin, K. *Structure* **2002**, *10*, 1551-1557.
[36] Ming, D.; Brüschweiler, R. *J. Biomol. NMR* **2004**, *29*, 363-368.
[37] Van Melckebeke, H.; Simorre, J.P.; Brutscher, B. *J. Am. Chem. Soc.* **2004**, *126*, 9584-9591.
[38] Van Melckebeke, H.; Simorre, J.P.; Brutscher, B. *J. Magn. Reson.* **2004**, *170*, 199-205.

Chapter X

An Enzymatic Molten Globule: Efficient Coupling of Folding and Catalysis

Katherina Vamvaca[1], Beat Vögeli[2], Peter Kast[1], Konstantin Pervushin[2] and Donald Hilvert[1]*

[1]Laboratorium für Organische Chemie and

[2]Laboratorium für Physikalische Chemie, Swiss Federal Institute of Technology, ETH-Hönggerberg, CH-8093 Zürich, Switzerland.

* NMR Spectroscopy

Proceedings of the National Academy of Sciences of the United States of America,

2004, 101, 12860-12864

1. Introduction

Conformational diversity is essential to the function of many proteins. For example, numerous regulatory proteins in the cell are intrinsically unfolded, adopting a well-defined structure only in the presence of a suitable target molecule [1]. In the immune system conformational plasticity appears to provide antibodies with enhanced functional diversity, enable their binding to multiple antigens through different preexisting conformers [2]. Moreover, enzyme activity is often fine-tuned through local substrate-induced conformational changes ("induced fit" mechanisms) and allosteric transitions [3]. Nevertheless, excessive conformational diversity can lead to protein dysfunction. The etiology of certain diseases, most notably those involving amyloid formation such as prion disorders, is intimately linked to pronounced structural changes in protein molecules [4], and the unfolded states of proteins, especially those of enzymes, are completely inactive [5-7].

The sensitivity of enzyme activity to even seemingly minor perturbations in active-site structure [6] makes the (re)design of protein catalysts a major challenge. This challenge is increasingly being met through the application of evolutionary strategies that do not require detailed prior knowledge of structure-function relationships [8]. For instance, we have successfully exploited genetic selection in combination with structure-based design to convert an entwined dimeric enzyme [the AroQ chorismate mutase from *Methanococcus jannaschii* (MjCM)] into a monomeric four-helix-bundle protein with near-native enzymatic activity (Figure 1a) [9]. The monomer mMjCM was isolated from a library of 20^6 variants containing random hinge-loop sequences inserted into the dimer-spanning H1 helix of the parent enzyme. It contains the interhelical turn Ala-Arg-Trp-Pro-Trp-Ala-Glu-Lys and efficiently catalyzes the rearrangement of chorismate into prephenate with a k_{cat} value (3.2 s^{-1}) that is indistinguishable from that of the parent dimeric enzyme and a K_m value (170 μM) that is elevated by only 3-fold [9].

Here, we present evidence that the resting state of mMjCM, unlike that of its dimeric precursor, has the properties of a molten globule, a nonnative fluxional state that is normally associated with protein-folding intermediates and functional inactivity [5,6,10,11]. The surprising finding that this highly active catalyst is structurally disordered has potentially general implications for understanding how catalysis occurs in natural biological systems and for future progress toward the design of novel enzymes.

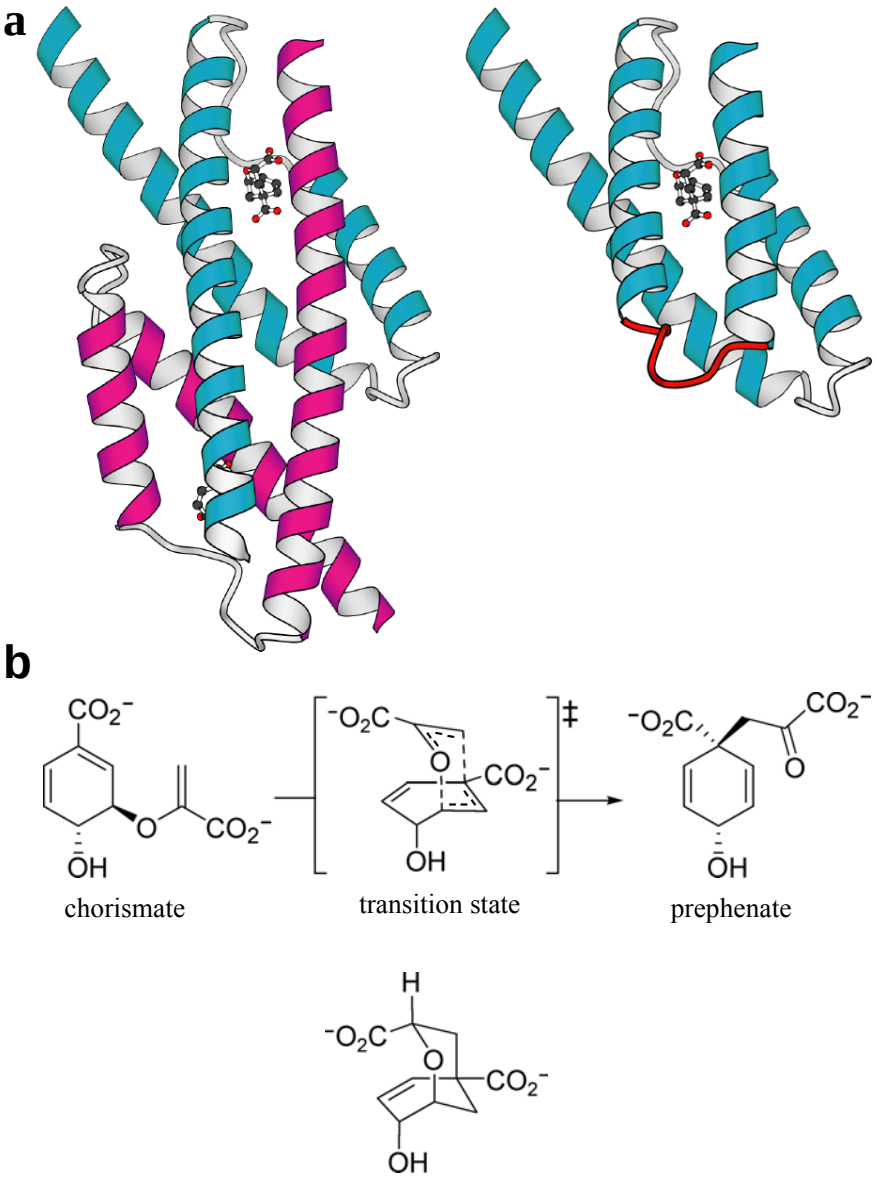


Figure 1. Topological redesign of the enzyme MjCM. (**a**) The thermostable MjCM homodimer (left) was converted into a monomer (mMjCM, right) by inserting a flexible hinge loop (red) into the long H1 helix [9]. The models are based on the X-ray structure [20] of a related *E. coli* chorismate mutase domain complexed with a transition state analog, **1** [14], which is shown in the models in ball-and-stick representation. (**b**) Both enzymes efficiently catalyze the rearrangement of chorismate to prephenate.

2. Materials and Methods

Chorismate Mutase Production and Purification. Truncated versions of mMjCM and MjCM, in which the six C-terminal amino acids were replaced by LEHHHHHH, were used for all experiments [9,12]. We produced ^{15}N-labeled proteins in *Escherichia coli* strain KA13 [12] by using 20%

Bioexpress growth medium (Cambridge Isotope Laboratories, Cambridge, MA) supplemented with M9 salts (containing $^{15}NH_4Cl$) and glucose. The enzymes were purified by using standard methods [9,12].

NMR Spectroscopy. All spectra were obtained at 20°C with a 600 MHz Bruker NMR spectrometer (Bruker, Billerica, MA); the proteins were in PBS buffer (20 mM sodium phosphate/50 mM NaCl, pH 6.5). Backbone resonances for the complex between mMjCM and a transition-state analog (**1**, Figure 1b) were assigned by conventional 3D NMR techniques, including transverse relaxation-optimized spectroscopy (TROSY), using ^{13}C,^{15}N-labeled mMjCM.

CD Spectroscopy. Near-UV spectra of protein samples (0.15 mM, plus 0.28 mM compound **1**) were collected at 20˚C using an Aviv 202 spectropolarimeter (Aviv Associates, Lakewood, NJ). Thermal denaturation of proteins (16 μM) in the absence and presence of **1** (80 μM) was monitored at 222 nm.

Binding of 1-Anilinonaphthalene 8-Sulfate (ANS). Proteins (2 μM) in the absence and presence of **1** (60 μM) were analyzed in PBS containing 2 μM ANS at 20˚C with an Aminco-Bowman series 2 luminescence spectrometer (excitation wavelength, 370 nm).

H/D Exchange and MS. Isotopic exchange was initiated at room temperature by 20-fold dilution of an aqueous protein solution (25 μM protein in 2.5 mM sodium phosphate/40 mM NaCl, pH 6.7, in the absence and presence of 0.8 mM **1**) into D_2O (pD 7.4). At appropriate times after dilution, 150 μL aliquots were removed, quenched by decreasing the pD to 2.6 (with the addition of 150 μL acetonitrile, containing 0.6% formic acid) and by lowering the temperature to 0°C, and analyzed on a TSQ 7000 electrospray-ionization MS spectrometer (Finnigan-MAT, San Jose, CA) [13]. The monomer contains 109 amino acid residues and a total number of 254 labile protons, whereas each dimer subunit, which lacks the 8-aa engineered loop, has 101 residues and 231 labile H atoms. The calculation of the percentage of protected protons [percentage of protected H atoms = $(M_D\text{-}M)/(M_D\text{-}M_H) \times 100\%$, where M_H, M, and M_D are the average molecular masses of the nondeuterated, partially deuterated, and fully deuterated protein, respectively] allows direct comparison of the H/D exchange behavior of the monomer and the wild-type dimer.

3. Results

Structural studies were initiated to gain insight into the properties of mMjCM. The topologically redesigned enzyme is highly helical, as judged by CD spectroscopy, and its monomeric topology has been confirmed by gel filtration and analytical ultracentrifugation [9]. Because attempts to crystallize the monomer have been unsuccessful to date, we produced it in labeled form for NMR spectroscopic

analysis. Surprisingly, the [^{15}N,^{1}H]-TROSY spectrum of the monomer exhibits far fewer signals than expected for a native-like protein, and the dispersion of the observed peaks is poor (Figure 2a). These features are characteristic of a poorly packed protein with multiple energetically accessible conformations. However, addition of the bicyclic dicarboxylic acid **1** (Figure 1b) [14], a transition state analog inhibitor for the enzyme (K_i = 2.5 μM), induces a dramatic transition from a disordered to an ordered state. At saturating concentrations of ligand, a well resolved spectrum is obtained with signal dispersion, which is typical for a conventionally folded protein (Figure 2b). This behavior is unique to the designed monomer because the NMR data show that the parent MjCM dimer has native-like structure both in the absence and presence of ligand (Figures 2c and 2d).

Spectra of the monomer obtained at intermediate concentrations of ligand feature two sets of resonances corresponding to a superposition of the spectra measured with and without saturating amounts of ligand (see Figure 5 in the Appendix). The amplitudes of the resonances attributed to the ordered state are directly proportional to the concentration of protein present as a complex with transition-state analog. This observation supports the existence of a dynamic equilibrium between two distinct states of the protein, disordered and ordered, which is a function of ligand concentration and is slow on the NMR time scale. The poor dispersion of resonances in the disordered state can be attributed to rapid interconversion of an ensemble of (many) energetically accessible conformations, possibly including a fraction of native-like species. Assuming that native structures would exhibit a similar chemical-shift dispersion as the complex with transition-state analog, we attempted to stabilize them in the absence of ligand by changing the temperature (see Figure S5), adding osmolytes such as sodium sulfate and tetramethylammonium chloride, and reducing protein concentration (to minimize dimerization of monomers). Because only minor changes were evident in the resulting NMR spectra, we conclude that the fraction of native-like conformers in the population of unliganded enzyme is small.

Various additional biophysical methods confirm that unliganded mMjCM is disordered but becomes more structured upon ligand binding. First, thermal denaturation of the protein is noncooperative in the absence of **1** but shows modest cooperativity in its presence (T_m = 55°C) (Figure 3a). For comparison, the unliganded MjCM dimer unfolds with a high degree of cooperativity (T_m = 88°C) [12]. Second, large ligand-induced changes in the near-UV CD spectrum of mMjCM are observed [15]; analogous conformational changes are not evident in MjCM by this technique (Figure 3b). Third, as is typical for molten globules, the engineered monomer binds the environmentally sensitive hydrophobic dye ANS, giving rise to significantly enhanced fluorescence with a characteristic blue shift of the emission maximum (Figure 3c) [15,16]. ANS, which does not inhibit the enzyme, also binds to mMjCM in the presence of **1** to some extent, suggesting that the mMjCM complex itself retains considerable dynamic character. By comparison, the fluorescence

increase seen for ANS in the presence of the wild-type dimer is more typical for a conventionally folded protein [17].

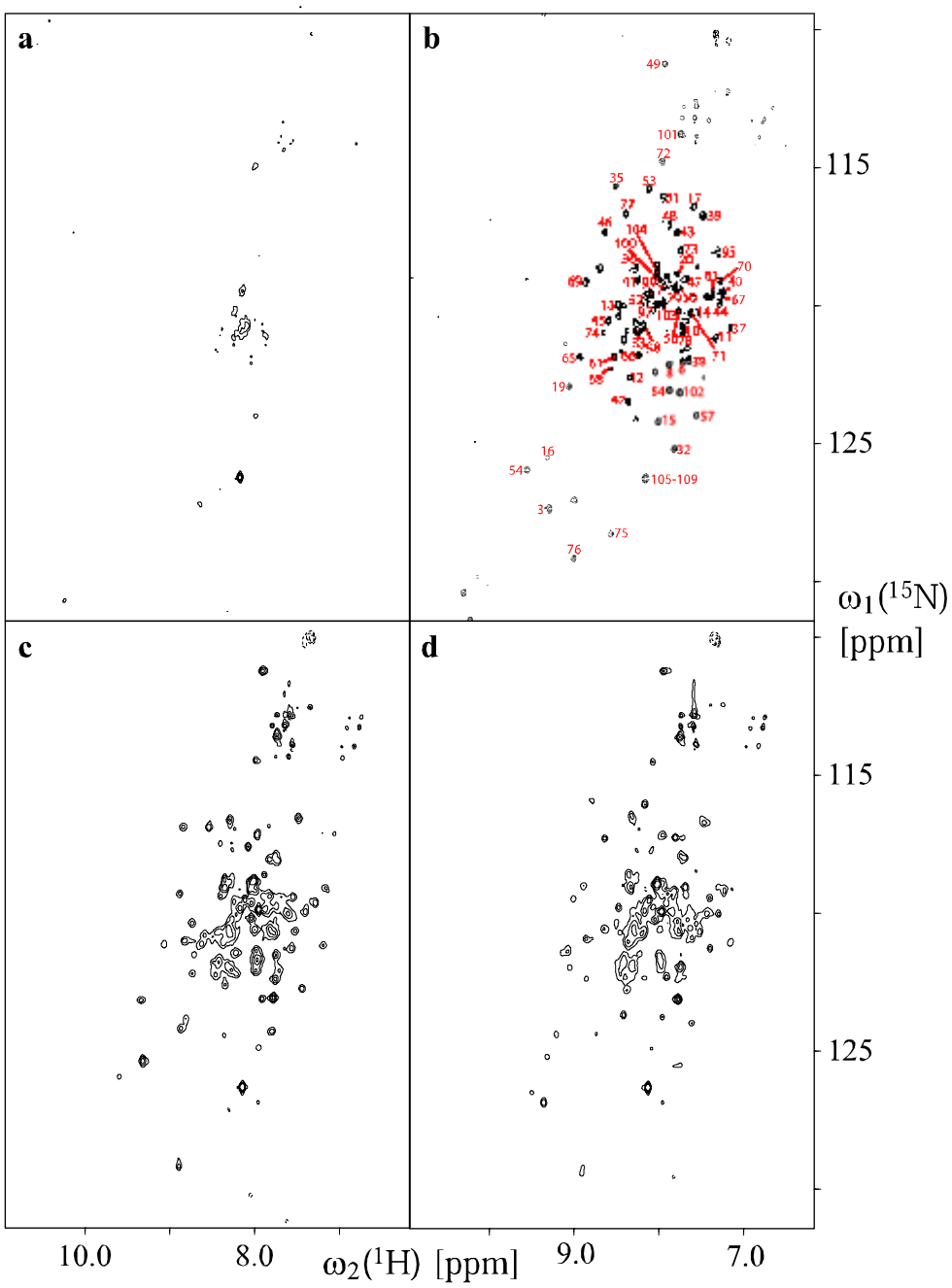


Figure 2. NMR spectra of ^{15}N-mMjCM and ^{15}N-MjCM in the absence and presence of **1**. (**a**) The [^{15}N,^{1}H]-TROSY spectrum of 40 µM mMjCM. Peak dispersion did not change appreciably when the protein concentration was varied in the range 10 µM to 0.6 mM, ruling out oligomerization-induced line-broadening effects. (**b**) The [^{15}N,^{1}H]-TROSY spectrum of the ligand-bound mMjCM (sample from **a**, supplemented with 1.2 mM **1**). The red numbers next to the peaks indicate the corresponding residues in the protein. The secondary structure of the monomer, assessed based on the chemical shifts of N, C^{α}, and C^{β} spins (P. Anikeeva & K.P., unpublished results), is in agreement with the proposed structural model (Figure 1a). Residues 105-109 comprise part of the C-terminal His tag, which is unstructured and gives rise to degenerate resonances. (**c**) The [^{15}N,^{1}H]-TROSY spectrum of 0.6 mM MjCM. (**d**) The [^{15}N,^{1}H]-TROSY spectrum of the ligand-bound MjCM (sample from **c**, supplemented with 2 mM **1**).

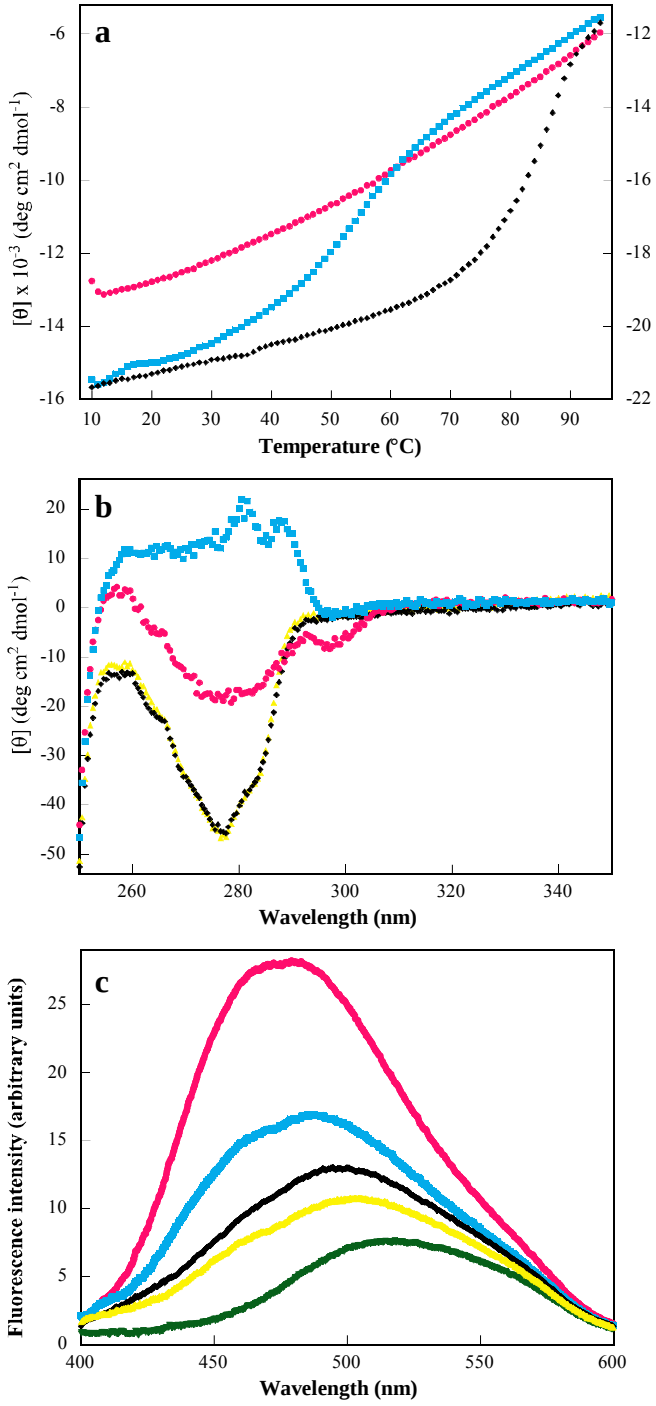


Figure 3. Characterization of mMjCM in the absence (magenta circles) and in the presence (cyan squares) of **1**, and of MjCM in the absence (black diamonds) and in the presence (yellow triangles) of ligand. (**a)** Thermal denaturation curves for mMjCM with and without **1**. The published [12] denaturation curve for MjCM is provided for comparison (scale on the right). (**b**) Near-UV CD spectra of the free and ligand-bound forms of mMjCM and MjCM. The well resolved peaks (280.5 and 287.5 nm) for the mMjCM•**1** complex indicate that the two tryptophan residues located in the engineered loop are in a more asymmetric environment compared to the free protein [31]. In the near-UV CD spectrum of MjCM (which contains no tryptophans) these peaks are not observed. (**c**) Fluorescence emission spectra of the free and ligand-bound proteins in the presence of ANS and of ANS alone (green triangles).

Further compelling evidence for the high conformational flexibility of the monomer comes from H/D exchange experiments [18]. The rate of exchange was monitored by electrospray ionization MS in the presence and absence of transition state analog (Figure 5 and Table 1 in the Appendix). Shortly after initiation of exchange (10 s), all but 17% of the labile protons in mMjCM become deuterated. Addition of **1** initially protects a significant fraction of the protons in the monomer from exchange (≈35%), but within 5 minutes the degree of deuteration in the complex is comparable with that for the unliganded enzyme. Thus, these results provide additional evidence for a highly mobile structure of MjCM, even in the presence of ligand. In contrast, the wild-type dimer is considerably more resistant to H/D exchange. Of the exchangeable protons in the MjCM dimer, ≈35% are shielded from exchange at the initial time point; ≈25% of the labile protons do not exchange for several hours. As in the case of the monomer, ligand binding to the dimer protects additional NH groups from exchange, but the magnitude of the effect (≈3-5%) is much smaller.

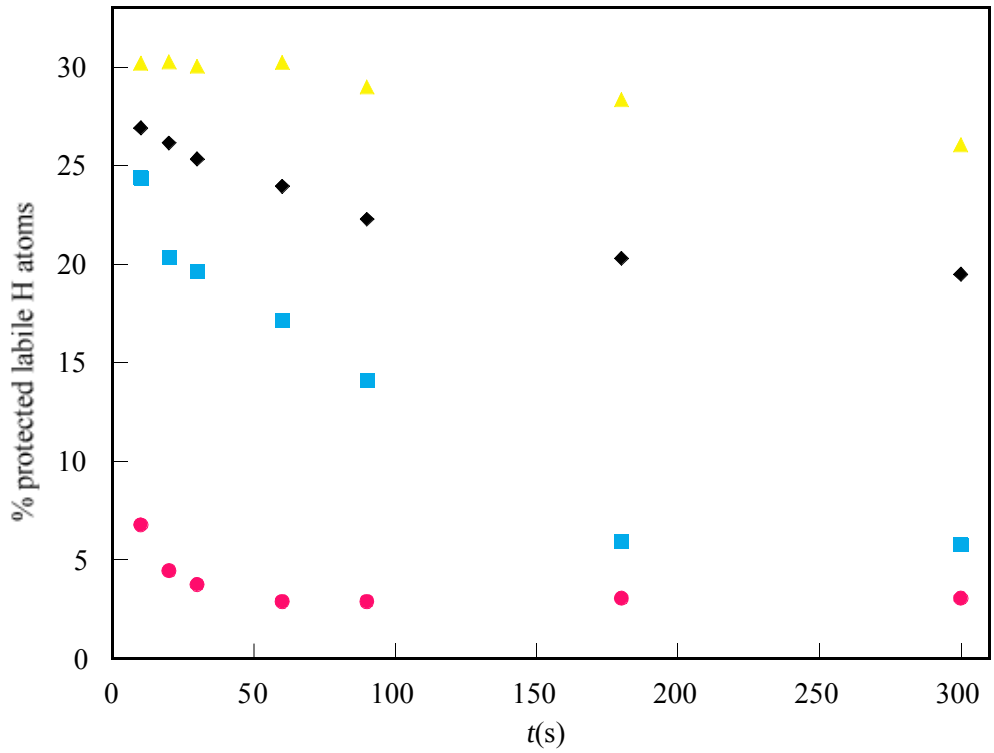


Figure 4. H/D exchange of mMjCM (magenta circles), mMjCM•**1** (cyan squares), MjCM (black diamonds), and MjCM•**1** (yellow triangles). The substrate chorismate protects mMjCM from exchange to a similar extent as **1** (data not shown).

4. Discussion

Poor amide-signal dispersion, noncooperative thermal denaturation, ANS binding and rapid H/D exchange are the hallmarks of molten globules, relatively compact states of proteins having near-native secondary structure but fluctuating tertiary structure [5,10,11,17,19]. The monomeric mutase apparently exists as an ensemble of conformers in dynamic equilibrium and adopts an ordered structure only in the presence of ligand. Although native-like species may be present to some extent in the absence of ligand, they are not detected bv NMR or near-UV CD, suggesting that they constitute only a small fraction of the total population.

The molten globular character of unliganded mMjCM may be attributable to the absence of a well-packed hydrophobic core. The interior of mMjCM must accommodate the highly polar active site [20]. The large number of charged residues at the core of the protein likely destabilizes a native-like folded state in the absence of ligand. However, the transition-state analog or substrate can fill the pocket, providing a template around which the protein can "crystallize". The extensive H-bonding and electrostatic interactions that are formed between **1** and the protein can then propagate globally to improve packing throughout the structure. Because the parent MjCM dimer contains a conventional hydrophobic core between its two active sites [20], it is considerably more stable than mMjCM, even in the absence of **1**, and substantially less structural ordering is required for catalysis.

The enhanced flexibility of mMjCM compared with MjCM has an amazingly small energetic cost. The monomer and dimer differ in catalytic efficiency (k_{cat}/K_m) only by a factor of three ($\approx$0.6 kcal/mol) entirely because of differences in K_m [9]. Preliminary calorimetric measurements show that the unfavorable entropy associated with preorganizing the enzyme is offset by enthalpic gains from the noncovalent interactions between **1** and mMjCM [K.V. & I. Jelesarov, unpublished results], in accord with the proposal of Williams et al. [21] that catalytic efficiency can originate from improved packing interactions in the transition state. In a sense, the structural ordering seen here can be considered an extension of the classic induced-fit model of enzyme catalysis [22] in which the substrate causes structural rearrangement at the active site that lead to a high affinity complex. In the case of mMjCM, however, the substrate appears to induce global ordering, rather than local restructuring, taking the classical model to new extremes. Although precise positioning of catalytic residues is generally considered to be a prerequisite for efficient catalysis, our findings suggest there is apparently little penalty associated with locking the catalytically active conformation of mMjCM in place as the reaction proceeds.

The absence of a significant population of natively structured protein in the absence of ligand argues against an alternative "preequilibrium" mechanism [23] in which addition of ligand shifts the equilibrium between the molten globule and a scarcely populated native-like species competent for ligand binding and catalysis in favor of the native state. The high activity of the catalyst would

require that ≈30% of the free protein be ordered, which is not observed. However, presteady-state kinetic measurements will be necessary to exclude this mechanistic possibility definitifely.

Many de novo designed proteins have turned out to be molten globules [17]. In fact, protein engineers have expended considerable effort over the years to convert these molecules into native-like proteins [17,24-27]. The tacit assumption has been that a preorganized native-like structure is required for biological function. However, our results with mMjCM show that the molten globular state does not necessarily preclude efficient catalysis. Although the molten globules formed during the (un)folding of many naturally occurring enzymes are completely inactive [5,6], there have been hints in the literature that other engineered molten globules might retain function under some conditions. For example, two circularly permuted dihydrofolate reductase variants, although less active than the parent enzyme, were shown to possess some properties of the molten globule state and recover native-like tertiary structure in the presence of NADPH and methotrexate [28]. A staphylococcal nuclease mutant also displays certain features of the molten globule state and adopts a functional structure upon binding Ca^{2+} and thymidine 3',5'-bisphosphate [29]. Additional characterization of these systems is clearly needed, but in light of our findings, the available data suggest that the coupling of folding and catalysis may be a more general phenomenon than has been previously appreciated.

Conceivably, many if not most modern-day enzymes could have evolved from molten globules [30]. Insofar as structural plasticity confers relaxed substrate specificity and/or catalytic promiscuity [23], a molten globule could have had a significant evolutionary advantage in the primordial world, performing several functions in the primitive cell. The similar activities of mMjCM and MjCM, and the ability of the monomer to assume the metabolic role of the dimer in the cell [9], suggest that catalytic efficiency may not have been the (sole) driving force for the evolution of native protein structures. Instead, the need to control individual functions separately, minimize intrinsic instability, or prevent aggregation, proteolytic degradation, or improper cellular trafficking would have favored the well defined structures of modern enzymes. Given their possible evolutionary heritage, it is tempting to speculate that molten globules might be superior to highly evolved enzymes as starting points for the directed evolution of tailored catalysts in the laboratory. By applying random mutagenesis and powerful selection or screening technologies to molecules like mMjCM, it may be possible to channel their inherent conformational diversity [23] down multiple evolutionary trajectories and generate a variety of interesting new structures and functions.

5. Appendix

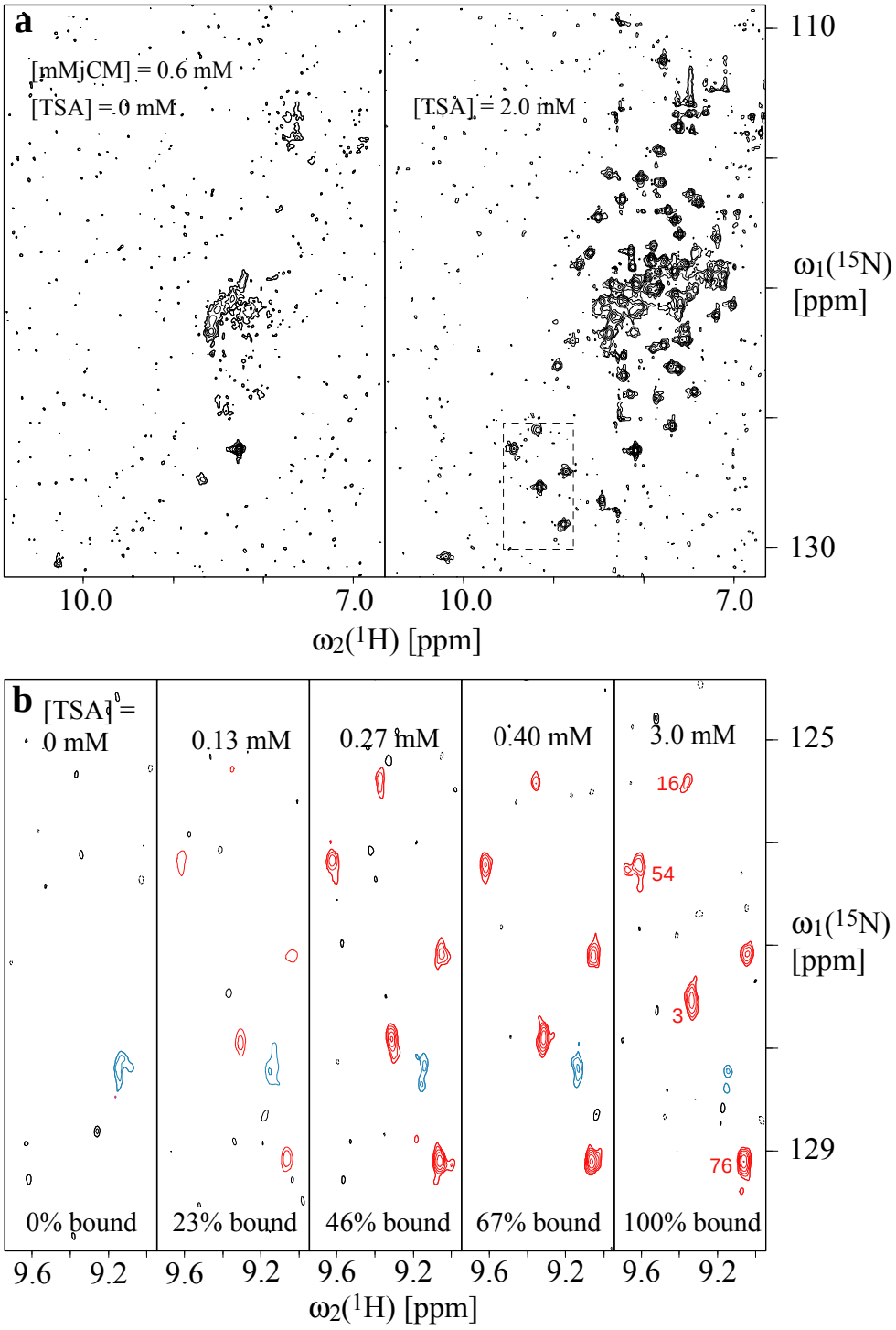


Figure 5. Ligand-induced structural changes in the engineered monomeric enzyme mMjCM. (**a**) The [^{15}N,^{1}H]-TROSY NMR spectra of mMjCM at 10°C in the absence and presence of compound **1**. (**b**) Titration of mMjCM with **1**, monitored by [^{15}N,^{1}H]-transverse relaxation-optimized spectroscopy (TROSY) NMR. Peaks corresponding to the molten globular and native states of mMjCM are blue and red, respectively.

Table 1. Raw H/D exchange data

t, s	*M* (mMiCM)	*M* (mMiCM+**1**)	*M* (MiCM)	*M* (MiCM+**1**)
0	13,170	—	12,099	—
10	13,385	13,345	12,253	12,246
20	13,391	13,354	12,255	12,246
30	13,392	13,356	12,257	12,247
60	13,394	13,361	12,259	12,246
90	13,394	13,368	12,263	12,249
180	13,394	13,387	12,267	12,250
300	13,394	13,387	12,269	12,255
∞	13,401	—	12,310	—

Both proteins were fully deuterated ($t = \infty$) by heat denaturation (10 min at 95° C) in the presence of 95% D_2O. The D_2O/H_2O ratio was 95:5 in all H/D exchange experiments.

6. References

[1] Dyson, H.J.; Wright, P.E. *Cur. Opin. Struct. Biol.* **2002**, *12*, 54-60.
[2] James, L.C.; Roversi, P.; Tawfik, D.S. *Science* **2003**, 299, 1362-1367.
[3] Hammes, G.G. *Biochemistry* **2002**, *41*, 8221-8228.
[4] Dobson, C.M. *Trends Biochem. Sci.* **1999**, *24*, 329-332.
[5] Ptitsyn, O.B.; Pain, R.H.; Semisotnov, G.V.; Zerovnik, E.; Razgulyaev, O.I. *FEBS Lett.* **1990**, *262*, 20-24.
[6] Tsou, C. *Biochim. Biophys. Acta* **1995**, *12534*, 151-162.
[7] Dunker, A.K.; Brown, C.J.; Lawson, D.J.; Iakoucheva, L.M.; Obradovic, Z. *Biochemistry* **2002**, *41*, 6573-6582.
[8] Taylor, S.V.; Kast, P.; Hilvert, D. *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 3310-3335.
[9] MacBeath, G.; Kast, P.; Hilvert, D. *Science* **1998**, *279*, 1958-1961.
[10] Dolgikh, D.A.; Gilmanshin, R.I.; Brazhnikov, E.V.; Bychkova, V.E.; Semisotnov, G.V.; Venyaminov, S.Y.; Ptitsyn, O.B. *FEBS Lett.* **1981**, *136*, 311-315.
[11] Kuwajima, K. *Proteins Struct. Funct. Genet.* **1989**, *6*, 87-103.
[12] MacBeath, G.; Kast, P.; Hilvert, D. *Biochemistry* **1998**, *37*, 10062-10073.
[13] Smith, D.L.; Deng, Y.; Zhang, Z. *J. Mass Spectrom.* **1997**, *32*, 135-146.
[14] Bartlett, P.A.; Johnson, C.R. *J. Am. Chem. Soc.* **1985**, *107*, 7792-7793.
[15] Demarest, S.J.; Boice, J.A.; Fairman, R.; Raleigh, D.P. *J. Mol. Chem. Biol.* **1999**, *294*, 213-221.
[16] Semisotnov, G.V.; Rodionova, N.A.; Razgulyaev, O.I.; Uversky, V.N.; Gripas, A.F.; Gilmanshin, R.I., A. *Biopolymers* **1991**, *31*, 119-128.
[17] Betz, S.F.; Raleigh, D.P.; DeGrado, W.F. *Curr. Opin. Struct. Biol.* **1993**, *3*, 601-610.
[18] Englander, S.W.; Mayne, L.; Bai, Y.; Sosnick, T.R. *Protein Sci.* **1997**, *6*, 1101-1109.
[19] Ohgushi, M.; Wada, A. *FEBS Lett.* **1983**, *164*, 21-24.
[20] Lee, A.Y.; Karplus, P.A.; Ganem, B; Clardy, J. *J. Am. Chem. Soc.* **1995**, *117*, 3627-3628.
[21] Williams, D.H.; Stephens, E.; Zhou, M. *J. Mol. Biol.* **2003**, *329*, 389-399.
[22] Koshland, D.E. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2375-2378.
[23] James, L.C.; Tawfik, D.S. *Trends Biochem. Sci.* **2003**, *28*, 361-368.
[24] Bryson, J.W.; Betz, S.F.; Lu, H.S.; Suich, D.J.; Zhou, H.X.; O'Neil, K.T.; DeGrado, W.F. *Science* **1995**, *270*, 935-941.
[25] Cordes, M.H.J.; Davidson, A.R.; Sauer, R.T. *Curr. Opin. Struct. Biol.* **1996**, *6*, 3-10.
[26] Gibney, B.R.; Rabanal, F.; Skalicky, J.J.; Wand, A.J.; Dutton, P.L. *J. Am. Chem. Soc.* **1999**, *121*, 4952-4960.
[27] Wei, Y.; Kim, S.; Fela, D.; Baum, J.; Hecht, M.H. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 13270-13273.
[28] Uversky, V.N.; Kutyshenko, V.P.; Protasova, N.Y.; Rogov, V.V.; Vassilenko, K.S.; Gudkov, A.T. *Protein Sci.* **1996**, *5*, 1844-1851.

[29] Li, Y.; Jing, G. *J. Biochem.* **2000**, 128, 739-744.
[30] DeGrado, W.F. *Nature* **1993**, *365*, 488-489.
[31] Woody, W.R.; Dunker, A.K. *Circular Dichroism and the Conformational Analysis of Biomolecules.* **1996**, Plenum, New York.

Chapter XI

NMR Studies of Structure and Dynamics of a Partially Folded Enzyme

Unpublished

1. Introduction

The experimental design of monomeric MjCM has been based on a homology model [1]. Its structure is unknown to date. Because of the molten globular nature of monomeric MjCM, attemps to crystallize it have been unsuccessful [2] (see Chapter X). NMR is one of the few means amenable to structural studies of disordered proteins. Therefore, samples were produced in $^{15}N,^{13}C$-labeled form for NMR spectroscopy. In addition, analysis of intramolecular dynamics is critical for understanding the folding process and ligation of TSA. Here, we present an analysis of the secondary structure of TSA-bound mMjCM and model-free fits of parameters describing internal motions.

2. NMR Spectroscopy

For assignment and structure determination 2D [$^{15}N,^{1}H$]-TROSY, 2D [$^{13}C^{aliphatic},^{1}H$]-HSQC, 2D [$^{13}C^{aromatic},^{1}H$]-HSQC, 2D NOESY (τ_{mix} = 150 ms), 2D TOCSY, 3D TROSY-HNCA, -HNCACO and –HNCACB, ^{15}N-edited NOESY (τ_{mix} = 70 ms), $^{13}C^{aliphatic}$-edited NOESY, $^{13}C^{aromatic}$-edited NOESY (τ_{mix} = 70 ms), ^{15}N-edited NOESY (τ_{mix} = 70 ms) and ^{15}N-edited TOCSY [3,4] were performed at 293 K on a Bruker *Avance* 600 spectrometer equipped with a cryogenic Z-gradient TXI probe and on a Bruker *Avance* 900 spectrometer equipped with a Z-gradient TXI probe. The ^{15}N- and $^{15}N,^{13}C$-labeled samples were prepared as described previously [1,2] and TSA was added up to complete saturation. For dynamical analysis the following data was acquired with a ^{15}N-labeled sample at 293 K and 600 MHz fields: R_1, R_2^{eff} and *NOE* with modified [$^{15}N,^{1}H$]-TROSY sequences with longitudinal and transversal relaxation periods and NOE-elements [5,6], respectively. η_{xy} was obtained from the difference in peak intensities between TROSY and anti-TROSY experiments [5].

3. Assignment and Secondary Structure of TSA-bound mMjCM

The spectra were Fourier transformed with the program XWIN-NMR 3.5 (Bruker BioSpin) and analysed with the program XEasy [7]. The spin systems of some residues could not be assigned due to missing peaks. These are mainly peaks from residues in the loop regions and a part of the α-helix designated H3 in the original dimer structure [3]. The secondary structure of mMjCM is identified by a comparison of the experimental $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts with the corresponding chemical shifts found in random structures using the program MAPPER [8]. Figure 1 shows α-helical regions on top of the amino sequence. The deviations of the experimental $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts from

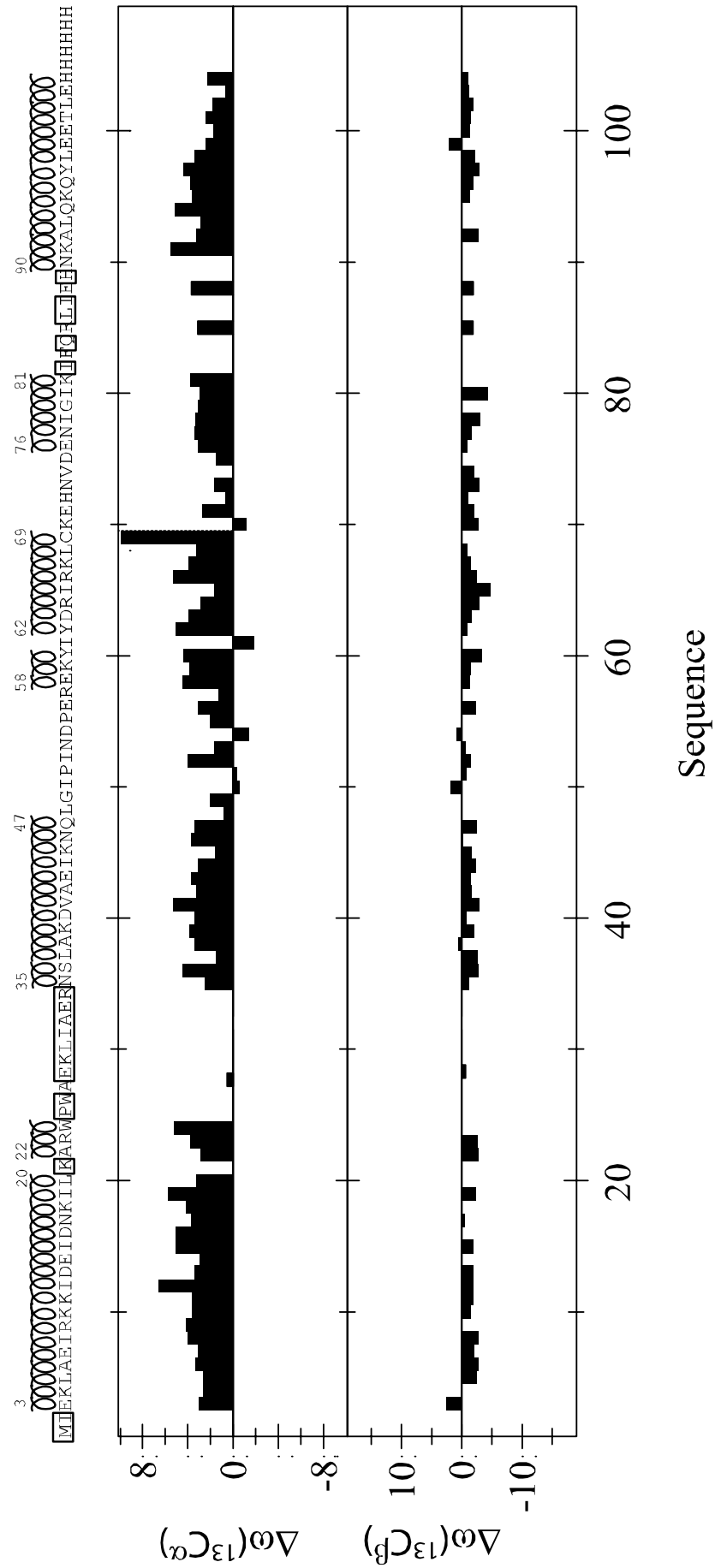
the random coil shifts are represented with black boxes. The H1 helix of the dimer is divided into two shorter α-helices separated by the engineered six-residue turn. The N-terminal α-helix extends to residue 24, which covers 3 residues of the engineered six-residue turn (residues 22-27). The second α-helix ranges from residues 35 to 47. The former H2 α-helix is conserved from residue 58 to residue 69. The helix is rather disordered from residue 82 to 89 since most of the peaks stemming from this region are broadened beyond detection by conformational exchange processes (see Dynamics). α-helical structure is observed from residue 90 to the His-tag.

4. Oligomeric States of mMjCM

An averaged $S^2\tau_c$ value, $\langle S^2\tau_c \rangle$, was determined at concentrations of 0.6 mM and 50 μM at 293 K from exchange-free cross-correlated relaxation rates. The value at 0.6 mM of mMjCM was measured using 2D experiments as described above. The values at lower concentration were extracted from 1D [^{15}N-storage,^{1}H]-experiments measuring the decay of the amide proton resonance bulk upon increasing the magnetization storage time on transverse ^{15}N [9] (data not shown).

At high concentration of 0.6 mM TSA-bound mMjCM, $\langle S^2\tau_c \rangle$ was about 12.7 ns whereas it was only around 5 ns at 50 μM. We conclude that the equilibrium is shifted from a monomeric state at low concentration to a dimeric state at high concentration. It's likely that the dimerization proceeds not via domain swapping, but rather is mediated by the exposed hydrophobic face. This hypothesis corresponds well to data obtained from analytical ultracentrifugation [10]. Unbound mMjCM at low concentration yields yet even lower $\langle S^2\tau_c \rangle$ value of about 2 ns. Due to the assumption that τ_c of an unfolded or molten globule protein is not expected to be lower than that of the folded protein, we conclude that the unliganded mMjCM has on average much lower order parameters S^2 than TSA-bound mMjCM reflecting much higher bond-specific mobility in the molten globule form. In the following, mMjCM dynamics at high concentration, which is amenable to NMR studies, is analyzed.

Figure 1 (following page). The sequence and secondary structure of mMjCM. α-helices are identified by comparison of experimental $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shifts with the chemical shift index (CSI). Residues with ambiguously assigned or missing backbone resonances are enclosed in boxes. The deviations of the experimental $^{13}C^{\alpha}$ and $^{13}C^{\beta}$ chemical shift values from the random coil shift values are represented with black boxes.


MIEKLAEIRKKIDEIDNKILKARWPWAEKLIAERNSLAKDVAEIKNQLGIPINDPEREKYIYDRIRKLCKEHNVDENIGIKLHQHLLHHNKALQKQYLEETLEHHHHHH
3
20
22
35
47
58
62
69
76
81
90
Δω(13Cβ)
Δω(13Cα)
8
0
-8
10
0
-10
20
40
60
80
100
Sequence

5. Studies of Intramolecular Dynamics of TSA-bound mMjCM

5.1. Theory

In a two spin ½ system, the longitudinal relaxation rate R_1 of the density operator I_z, the transverse relaxation rate R_2 and the transverse cross-correlated relaxation rate η_{xy} for the density operator $I_{xy}(E/2 \pm S_z) = I_{xy}^{\alpha/\beta}$, and the nuclear Overhauser enhancement *NOE* are given by [11,12]:

$$R_1 = A_{D(IS)}^2(\tfrac{1}{3} J(\omega_I - \omega_S) + J(\omega_I) + 2J(\omega_I + \omega_S)) + 4A_{CSA(I)}^2 B_0^2 J(\omega_I), \quad (1)$$

$$R_2 = A_{D(IS)}^2(\tfrac{2}{3} J(0) + \tfrac{1}{6} J(\omega_I - \omega_S) + \tfrac{1}{2} J(\omega_I) + J(\omega_S) + J(\omega_I + \omega_S)) + A_{CSA(I)}^2 B_0^2(\tfrac{8}{3} J(0) + 2J(\omega_I)), \quad (2)$$

$$\eta_{xy} = A_{D(IS)} A_{CSA(I)} Q(\Theta_{D(IS)\,CSA(I)}) B_0(\tfrac{8}{3} J(0) + 2J(\omega_I)), \quad (3)$$

$$NOE = 1 + (\gamma_S/\gamma_I) A_{D(IS)}^2(2J(\omega_I + \omega_S) - \tfrac{1}{3} J(\omega_I - \omega_S))/R_1, \quad (4)$$

where the definitions are the same as for Equations 7-9 in Chapter III. The standard approach to the interpretation of relaxation data is the model-free method [11]. It assumes that the overall and internal molecular motions are independent. For isotropic overall rotation, the spectral density function is then given by

$$J(\omega_k) = \frac{S^2\tau_C}{1+(\omega_k\tau_C)^2} + \frac{(1-S^2)\tau}{1+(\omega_k\tau)^2}, \quad (5)$$

where S^2 is the generalized order parameter reflecting the degree of spatial restriction and τ is defined as:

$$1/\tau = 1/\tau_c + 1/\tau_e, \quad (6)$$

where τ_e is an effective correlation time reflecting fast internal motion.

S^2 can be decomposed into contributions from internal motions S_f^2 and S_s^2 on the fast and slow time scales, respectively [13]:

$$S^2 = S_f^2 S_s^2. \quad (7)$$

Thus the spectral density function can be expressed as

$$J(\omega_k) = S_f^2\left[\frac{S_s^2\tau_C}{1+(\omega_k\tau_C)^2} + \frac{(1-S_s^2)\tau}{1+(\omega_k\tau)^2}\right]. \quad (8)$$

An extended spectral density function is obtained for anisotropic overall rotation [11,12]. Assuming that the diffusion tensor ***D*** is symmetric, the spectral density function is dependent on the angle θ formed by the I-S bond and the tensor main axis. The spectral density function is:

$$J(\omega_k) = S_f{}^2 \sum_{j=1}^{3} A_j \left[\frac{S_s{}^2 \tau_j}{1+(\omega_k \tau_j)^2} + \frac{(1-S_s{}^2)\tau'_j}{1+(\omega_k \tau'_j)^2} \right], \quad (9)$$

where $1/\tau'_j = 1/\tau_j + 1/\tau_e$, $1/\tau_1 = 6D\perp$, $1/\tau_2 = 5D\perp + D\parallel$, $1/\tau_3 = 2D\perp + 4D\parallel$, $A_1 = (3\cos^2\theta - 1)^2/4$, $A_2 = 3\sin^2\theta\cos^2\theta$ and $A_3 = {}^3/_4\sin^4\theta$.

Fast chemical or conformational exchange between sites *l* and *m*, characterized by

$$k_{ex,lm} >> |\omega_l - \omega_m| , \quad (10)$$

yields a single spectral peak and contributes to the effective transverse relaxation rate [12]:

$$R_2{}^{eff} = R_2 + R_{2ex} , \quad (11)$$

$$R_{2ex} = A_{ex(I)}{}^2 B_0{}^2 J_{ex(I)}(0). \quad (12)$$

For the simplest case of two site exchange

$$J_{ex(I)}(\omega_k) = \frac{k_{ex,12}}{k_{ex,12}{}^2 + \omega_k{}^2} , \quad (13)$$

$$A_{ex(I)} = \sqrt{p_1 p_2}\gamma_I(\omega_1 - \omega_2) , \quad (14)$$

where p_1 and p_2 are the populations of state 1 and 2, respectively.

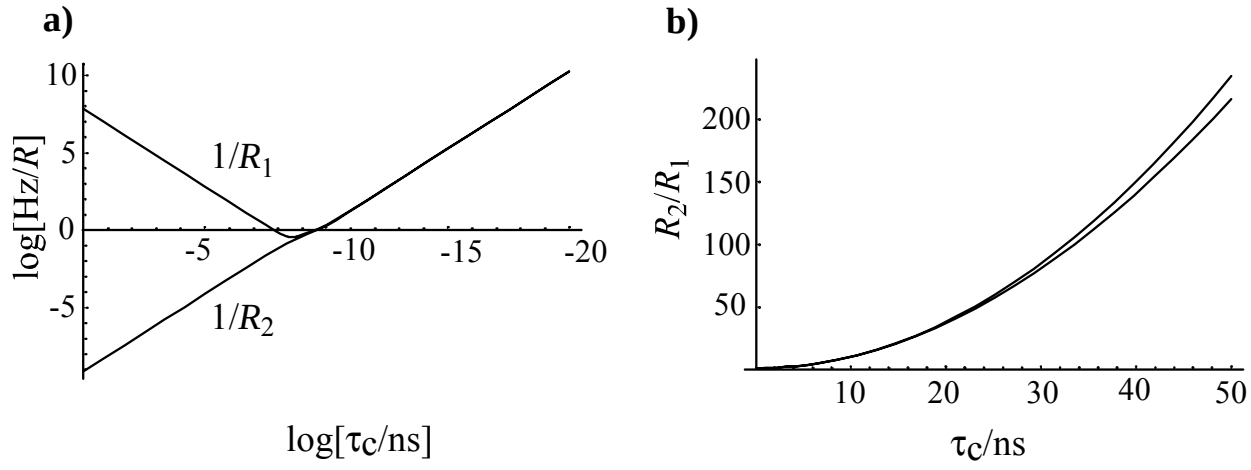


Figure 1. a) R_1 and R_2 versus τ_c in logarithmic scale (base 10). b) R_2/R_1 versus τ_c. In all calculations the following values were used: B_o = 600 MHz, $\Delta\delta(^{15}N$ CSA) = -160 ppm and r_{HN} = 1.02 x 10^{-10} m. In a) the second term of $J(\omega_k)$ was neglected. In b) $S^2 = 1$ (bottom curve) and $S^2 = 0.9$ and $\tau_c = 40$ ps (top curve).

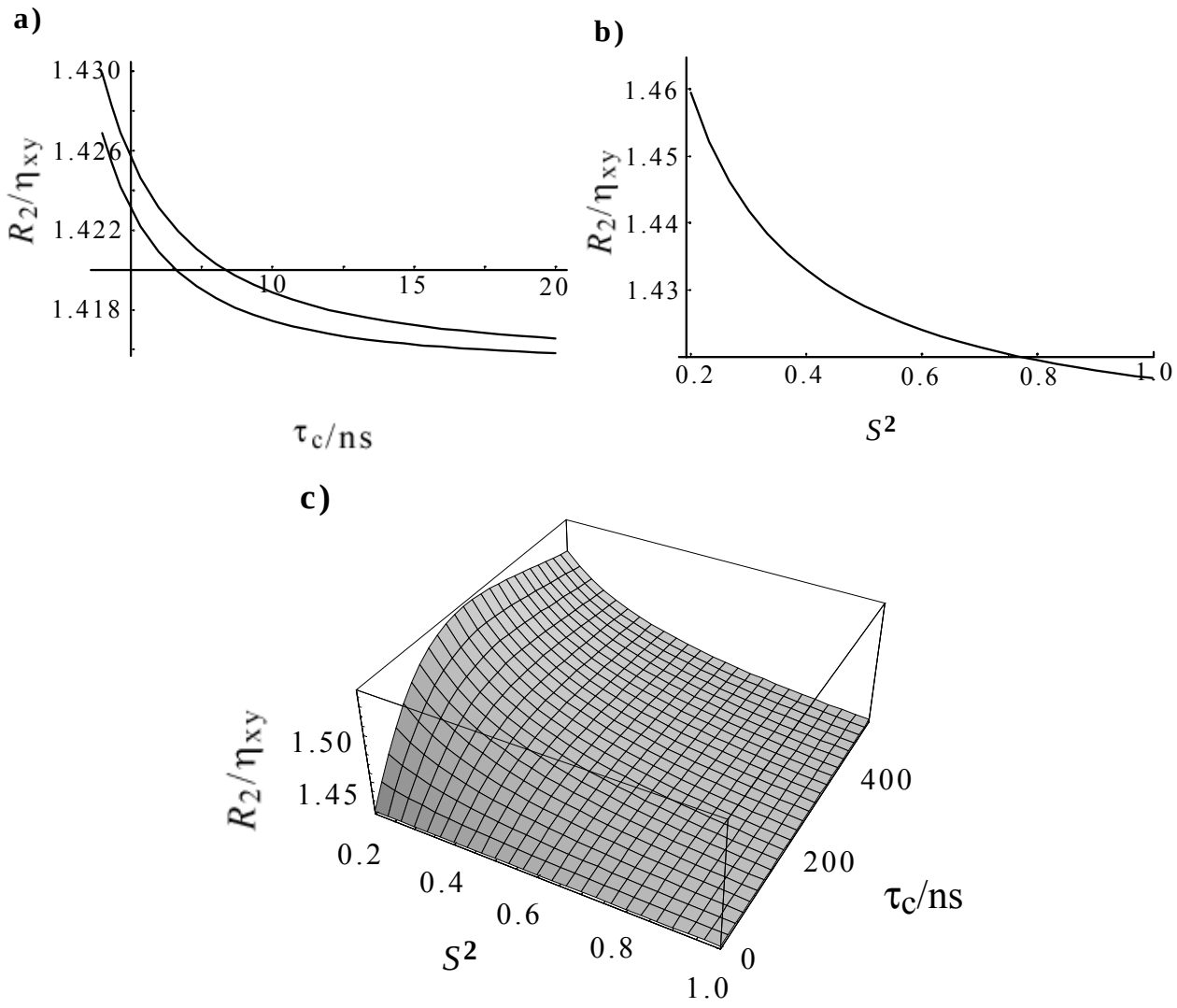


Figure 2. a) $R_2/\eta_{xy} = A(B_0)$ versus τ_c. The same values as in Figure 1 were used. $\Theta_{D(IS)\,CSA(I)} = 20°$. $S^2 = 1$ (bottom curve) and $S^2 = 0.9$ and $\tau_c = 40$ ps (top curve). b) $R_2/\eta_{xy} = A(B_0)$ versus S^2 with $\tau_c = 40$ ps and $\tau_c = 12$ ns. c) $R_2/\eta_{xy} = A(B_0)$ versus S^2 and τ_c with $\tau_c = 12$ ns.

The relaxation data is analyzed in two steps. First a proper model with the minimal number of parameters is chosen and τ_c is determined neglecting the second term of $J(\omega_k)$. It is instructive to use ratios of the Equations 1-3 because they are independent of S^2. τ_c can be estimated from the ratio R_2/R_1 [14]. Figure 1 shows $1/R_1$ and $1/R_2$ versus $\log(\tau_c/\text{ns})$ and R_2/R_1 versus τ_c. If no exchange is present this ratio is equal to R_2^{eff}/R_1. In the presence of exchange R_2/R_1 can be obtained by making use of the redundancy of R_2 and η_{xy}:

$$R_2 = A(B_0)\,\eta_{xy} \quad . \tag{15}$$

The scaling factor $A(B_0)$ is nearly τ_c-independent as shown in Figure 2.

Once the model is chosen and τ_c determined, the parameters describing the spectral density function are then fit to the model. These parameters provide information on the spatial restriction and the time scale of internal motions.

5.2. Choice of Model and Determination of Correlation Time

R_1, R_2^{eff} and *NOE* for 74 residues and η_{xy} for 70 residues were obtained. Values versus residue number are shown in Figure 4a-d. The following averages (denoted by brackets) were determined:

$$<R_2^{eff}/R_1> = 23.9 , \quad (16)$$

$$<R_2^{eff}/\eta_{xy}> = 2.0 , \quad (17)$$

$$<\eta_{xy}/R_1> = 12.2 , \quad (18)$$

$$<\eta_{xy}> = 12.7\ s^{-1} . \quad (19)$$

Isotropic Overall Motion Without Exchange. In a simple approximation, $R_2^{eff} = R_2$ and Equation 16 corresponds to a rotational correlation time τ_c = 16 ns under the assumption of isotropical rotation in solution and exclusion of possible exchange [14] as shown in Figure 1.

However, Figure 2 shows that $A(B_0)$ is almost independent of τ_c and equals to 1.416. Even taking into account an error of up to 10% for $\Delta\delta(^{15}N$ CSA) and another 10% for the nonuniformity of $\Theta_{D(HN)CSA(N)}$ can not make up for the difference between the theoretical and experimental factor given by Equation 17. Local mobility introduces no significant changes as demonstrated by Figures 2b and c.

Isotropic Overall Motion With Exchange. Longitudinal relaxation and cross-correlated relaxation terms are not affected by conformational exchange [15]. Therefore, Equation 18 is independent of conformational exchange and may be used to rewrite R_2/R_1 as

$$<R_2/R_1> = 1.416<\eta_{xy}/R_1> = 17.3. \quad (20)$$

This corresponds to a rotational correlation time τ_c = 13 ns as shown in Figure 1. $<S^2\tau_c>$ can be estimated as:

$$<S^2\tau_c> = <\eta_{xy}>/(\tfrac{8}{3} A_{D(HN)} A_{CSA(N)} Q(\Theta_{D(HN)CSA(N)}) B_0) = <\eta_{xy}>/0.97s^{-2} = 13.1\ ns , \quad (21)$$

where all parameters are chosen as described in the captions of Figures 1 and 2. It is assumed that $\omega^2\tau_c^2$ is much smaller than 1. Equation 21 implies that:

$$<S^2> \approx <S^2\tau_c>/\tau_c \approx 1.01 , \quad (22)$$

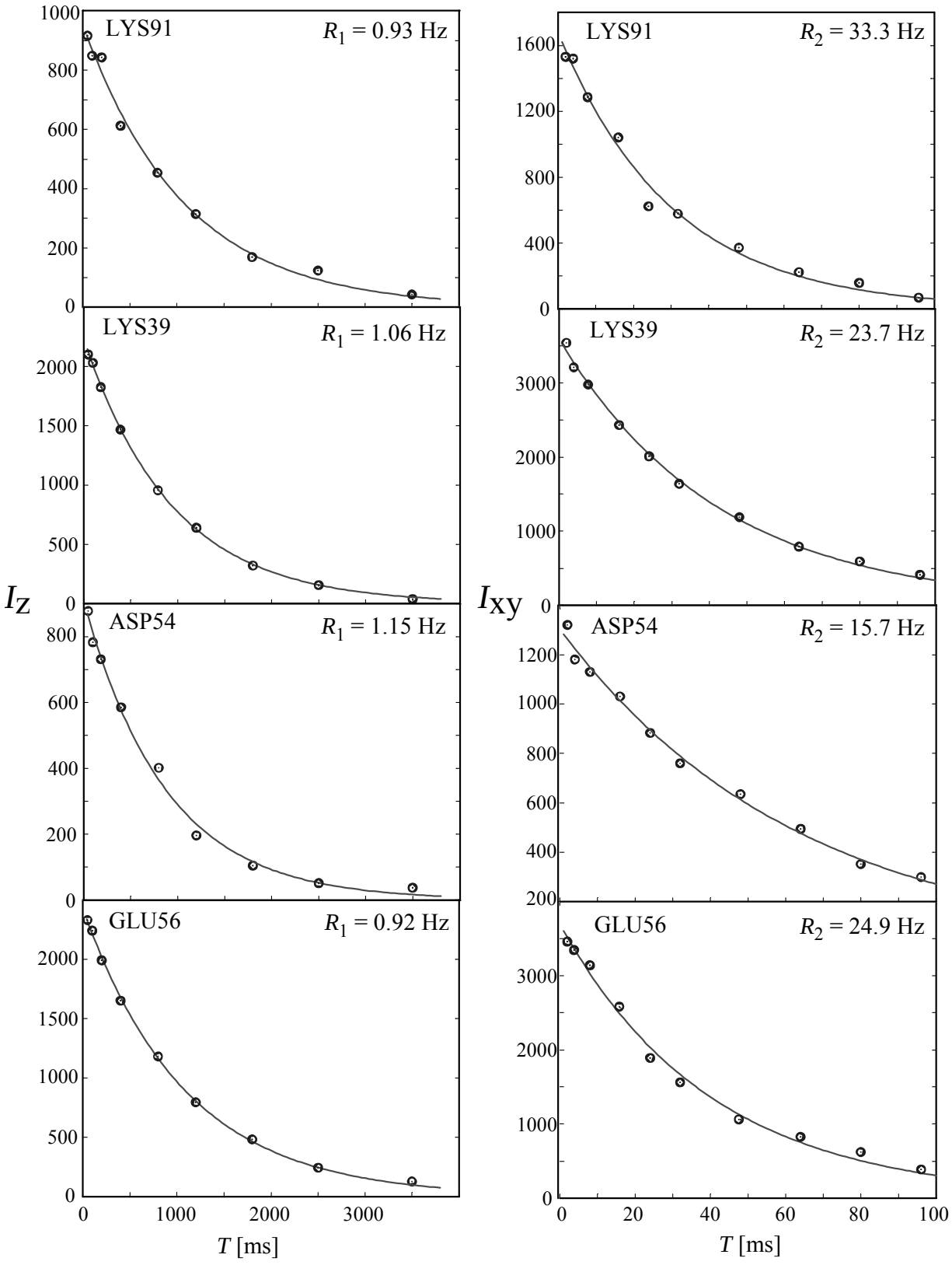


Figure 3. Representative R_1 and R_2 fits for Lys91, Lys39, Asp54 and Glu56. Left panel: Peak intensities I_z in a [^{15}N,^{1}H]-TROSY spectrum versus the longitudinal relaxation delay T. Right panel: Peak intensities I_{xy} in a [^{15}N,^{1}H]-TROSY spectrum versus the transversal relaxation delay T. Black dots represent measured data, the grey curves are the exponential fits carried out as $I = A\exp(-RT)$ in the program Matlab (The MathWorks Inc.).

which would correspond to a perfectly rigid protein. In Equation 21, $\Theta_{D(IS)\ CSA(I)}$ was 20º. If it is instead given a value of 15º, $<S^2>$ becomes 0.92, which is only slightly higher than the values obtained from fits (see below).

Rotational Diffusion Anisotropy. The ratio in Equation 1 may be further increased by rotational diffusion anisotropy [11].

Choice of Model. The relaxation data is analyzed with the program Modelfree Version 4.15 [16,17]. 8 Models comprising different combinations of fitting parameters are shown in Table 1. τ_c and S^2 (or S_s^2 if S_f^2 was used) were employed in any model. The fitted data is evaluated against the experimental data using the variance

$$\chi^2 = \sum_i \sum_j \left(X_{ij\,\mathrm{exp}} - X_{ij\mathrm{clc}} \right)^2 / \sigma_{\mathrm{X}ij}{}^2 \quad , \tag{23}$$

where i sums over all types of relaxation rates and j sums over all evaluated H-N bonds, X_{ij} is a relaxation rate of type i and bond j, and σ_{Xij} is its experimental uncertainty. X is R_1, R_2 and NOE except for model 8, where R_2 is replaced by $1.416\eta_{xy}$. The following parameters are used in all fits: σ_{R1ij} = 0.03 s^{-1}, σ_{R2ij} = 0.3 s^{-1} and σ_{NOEij} = 0.04. The experimental uncertainties were obtained from error propagation of the spectral noise for *NOE*s and from Monte Carlo simulations with random addition of values in the spectral noise range to each measured point in the T_1 and T_2 fits. Though, the uncertainties can vary strongly from residue to residue, here average values are used for simplicity. The upper limits of τ_e and R_{2ex} were set to 1 ns and 30 s^{-1}, respectively. For testing different models, relaxation rates of 74 residues from rigid parts of the protein were used (70 for model 8).

Table 1. Models comprising different combinations of fitting parameters and χ^2.

Model	τ_c [ns]	S^2 or S_S^2	τ_e [ns]	S_f^2	R_{2ex} [s^{-1}]	$D\perp/D\parallel$	χ^2
1	18.1	√					16’000
2	20.7	√	√				9’080
3	6.8	√			√		466
4	10.6	√	√		√		8.27
5	10.2	√	√	√	√		3.34
6	14.6	√	√			0.04	6’850
7	6.0	√			√	16.9	276
8[1]	14.4	√	√				2’040

[1] R_2 was replaced by $1.416\eta_{xy}$.

Models 1-5 assume isotropic overall rotation. χ^2 of the simplest model (model 1) is reduced by less than an order of magnitude by introduction of τ_e (model 2). However, it is reduced by more than an order of magnitude by introduction of the conformational exchange term (model 3). As has already been elaborated above, inclusion of the exchange term is necessary to proceed with the choice of the model. χ^2 of model 3 is reduced by almost 2 orders of magnitude upon inclusion of τ_e (model 4). Only slight improvement is achieved by additional use of the extended model (model 5). To check the impact of anisotropic models, models 2 and 3 were extended with the rhombicity of the diffusion tensor $\boldsymbol{D}$, $D\perp/D\parallel$. Both models yield very improbable rhombicities which hardly improve χ^2. The data, however, were fit against a model structure (Di Jorio et al., unpublished results), and this may partially explain the lack of improvement. Model 8 strongly improves the fit compared over the exchange-free model 2, further supporting the inclusion of the exchange term. However, the fit to model 8 is not as good as the fit to model 4, which only differs by the inclusion of the exchange term. Clearly, the dependence of η_{xy} on the nonuniform values of $\Delta\delta(^{15}N$ CSA) and $\Theta_{D(HN)CSA(N)}$ is too strong as outlined above. It is therefore concluded that model 4 is most adequate to describe the intramolecular dynamics. It should be noted that model 4 cannot be justified by the F-statistics [16], because in this model the number of experimental relaxation parameters equals the number of free internal motional parameters. Model 5 is underdetermined because 4 parameters are fit against 3 measured values and cannot be applied. If R_{2ex} is to be included, the only alternative to model 4 is model 3, which is obviously too simple to adequately describe intramolecular dynamics.

5.3. S^2 and R_{2ex} Fits

The correlation time τ_c obtained from model 4 is 10.6 ns. This value is 2 ns lower than the value estimated from the Equations 19 and 21 using $<\eta_{xy}>$. Nevertheless, a value around 11 ns is in reasonable agreement with both calculations. The values for S^2, R_{2ex}, and τ_e are given in Table 2. Figures 4 and 5 show plots of S^2 and R_{2ex} versus residue numbers. The error bars reflecting the simulation errors are rather uniformly clustered around 0.03 and 0.65 s^{-1} for S^2 and R_{2ex}, respectively. Almost all S^2 values range between 0.75 and 0.90. Values lower than 0.75 are seen for residues 11, 56, 57, 64, 77, 91, 92, 94 and 96, but they are all still higher than 0.71. Values higher than 0.90 are observed for residues 49, 63, 73 and 75. Overall, no significant deviations in S^2 are observed.

Most τ_e values are smaller than 40 ps. However, there are significant deviations, such as for residues 3, 5, 49, 66, 73, 75, 101, 102 and 103. R_{2ex} ranges from 2 s^{-1} to 25 s^{-1} with an average value of 12.7 s^{-1}. The distribution of R_{2ex} values throughout the sequence is rather non-uniform. However some tendencies can be observed: Residues 5-38 show values mostly well above average, 39-54

mostly below average, 56-61 above average, 62-90 mostly below average (of which 72-76 strongly), 91-99 slightly above average and 101-103 significantly below average.

Table 2. S^2 ,R_{2ex} and τ_e values obtained from the fits using model 4.

Residue	S^2	ΔS^2	R_{2ex} [s^{-1}]	ΔR_{2ex} [s^{-1}]	τ_e [ns]	$\Delta\tau_e$ [ns]
3	0.889	0.028	5.162	0.443	426.741	160.410
5	0.871	0.027	13.151	0.677	54.321	73.692
6	0.771	0.030	15.702	0.688	20.678	8.889
7	0.784	0.028	15.575	0.663	24.393	11.506
8	0.832	0.031	13.347	0.727	19.654	15.796
9	0.878	0.031	11.353	0.708	28.756	56.715
10	0.801	0.030	15.177	0.694	15.969	12.012
11	0.749	0.028	14.156	0.703	10.204	7.148
12	0.832	0.032	9.158	0.760	0.000	7.720
13	0.828	0.033	9.262	0.716	3.735	9.262
14	0.772	0.027	20.482	0.643	11.594	7.326
15	0.838	0.032	12.805	0.743	12.172	13.373
16	0.867	0.032	10.544	0.755	0.000	13.225
17	0.785	0.030	16.038	0.714	6.677	7.892
18	0.824	0.032	22.895	0.727	8.560	11.831
19	0.778	0.029	10.060	0.703	13.883	9.162
20	0.843	0.029	24.438	0.697	27.326	19.628
22	0.802	0.028	15.331	0.710	0.000	6.581
23	0.764	0.031	16.449	0.690	4.200	5.934
24	0.871	0.032	11.240	0.720	0.000	13.517
35	0.826	0.031	7.728	0.710	16.289	14.562
36	0.859	0.027	13.924	0.684	30.913	35.276
37	0.793	0.029	15.133	0.700	23.690	11.695
38	0.874	0.029	13.333	0.704	20.194	27.881
39	0.807	0.029	11.377	0.694	28.135	15.139
41	0.873	0.030	11.092	0.709	0.000	13.905
42	0.865	0.031	8.422	0.711	0.000	9.418
43	0.836	0.029	12.843	0.696	3.950	9.944
44	0.776	0.028	12.426	0.623	15.618	8.419
45	0.867	0.031	7.741	0.742	18.841	27.918
46	0.84	0.028	7.158	0.708	9.570	12.409
47	0.811	0.031	13.273	0.735	14.671	11.054
48	0.764	0.027	14.520	0.670	11.100	7.819
49	0.910	0.025	9.807	0.655	58.591	215.219
50	0.859	0.030	12.546	0.700	4.737	12.484
52	0.805	0.031	7.817	0.690	39.687	17.000
53	0.759	0.027	12.985	0.675	15.895	8.149
54	0.894	0.026	2.069	0.592	15.458	56.535
56	0.714	0.028	14.058	0.677	5.866	5.168
57	0.737	0.028	15.573	0.632	6.594	5.970
58	0.803	0.030	16.161	0.700	13.958	9.931

Residue	S^2	ΔS^2	R_{2ex} [s^{-1}]	ΔR_{2ex} [s^{-1}]	τ_e [ns]	$\Delta\tau_e$ [ns]
59	0.776	0.027	16.552	0.664	23.320	10.769
61	0.900	0.030	14.211	0.741	2.307	60.762
62	0.772	0.026	9.502	0.652	11.594	8.024
63	0.902	0.030	11.700	0.712	16.837	132.644
64	0.733	0.029	18.841	0.665	9.413	6.503
65	0.783	0.028	11.783	0.682	0.000	5.330
66	0.782	0.029	11.179	0.697	51.885	14.129
68	0.787	0.029	8.940	0.685	0.000	5.429
69	0.857	0.033	6.773	0.760	7.807	14.051
70	0.809	0.027	15.200	0.640	26.109	16.396
71	0.788	0.028	15.673	0.666	20.807	11.874
72	0.862	0.030	10.875	0.730	38.979	29.482
73	0.904	0.026	9.283	0.636	49.018	155.373
75	0.944	0.019	8.911	0.453	50.089	297.311
76	0.776	0.030	7.536	0.698	15.618	8.687
77	0.721	0.028	14.109	0.623	18.783	7.065
80	0.779	0.029	9.830	0.641	0.000	5.451
84	0.761	0.028	14.640	0.661	14.318	6.782
85	0.797	0.028	18.918	0.702	3.042	6.952
87	0.809	0.03	9.322	0.693	7.715	10.790
90	0.774	0.031	9.297	0.666	9.881	7.646
91	0.738	0.031	22.093	0.690	0.000	3.822
92	0.748	0.029	17.586	0.701	3.859	5.868
93	0.809	0.028	14.040	0.669	26.109	13.357
94	0.729	0.029	13.013	0.643	19.520	7.476
95	0.767	0.027	13.808	0.669	16.588	8.697
96	0.747	0.029	16.150	0.685	11.700	7.364
97	0.807	0.032	12.029	0.762	9.833	9.766
98	0.811	0.031	14.293	0.715	24.050	12.309
99	0.759	0.03	16.195	0.696	15.895	8.214
101	0.841	0.024	9.633	0.597	91.535	63.867
102	0.888	0.02	7.656	0.491	163.949	179.918
103	0.849	0.026	6.040	0.451	572.570	121.661

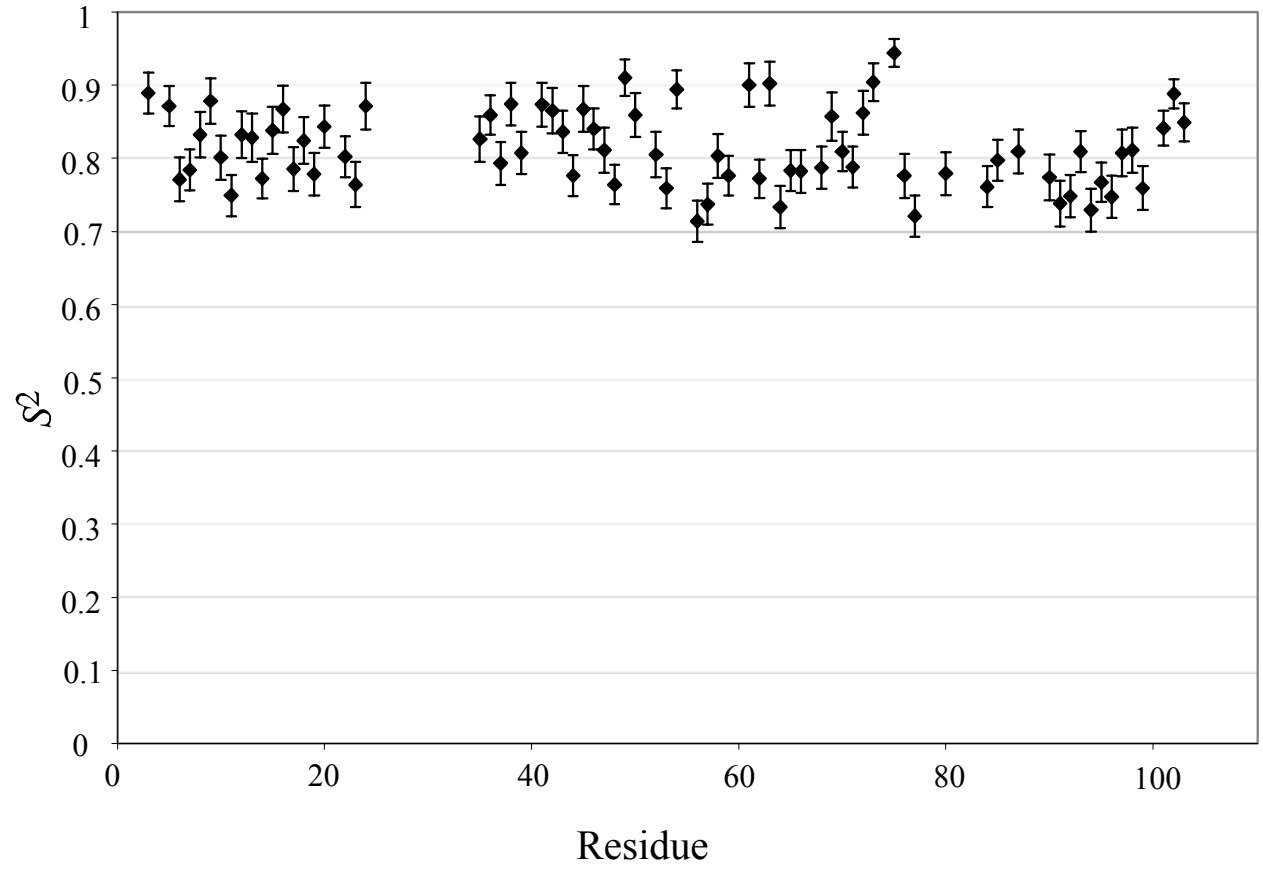


Figure 4. S^2 versus residue numbers.

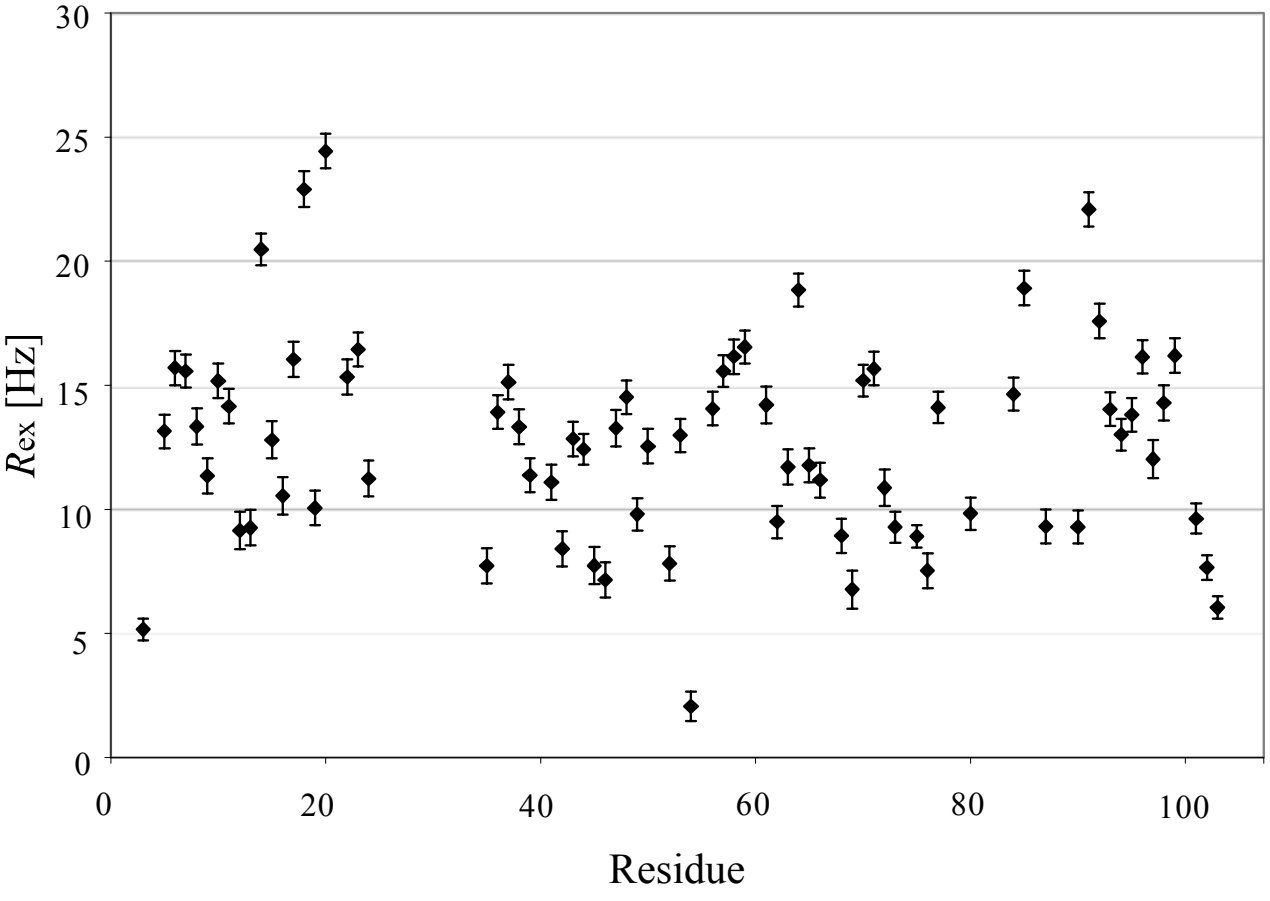


Figure 5. R_{2ex} versus residue numbers.

6. Discussion

The predominance of α-helical structure in monomeric MjCM is confirmed. As suggested in Reference 3, mMjCM has a four α-helix-bundle motif. By introduction of the six-residue turn into H1 of the dimer, the upper half becoming the N-terminus of the monomer folds down and forms the active site together with the second half.

Fast internal motion is observed for the loop 72-75, the C'-terminal parts 3-5, the N'-terminal part 101-103. These parts undergo rather small conformational exchange line-broadening compared to the average rate. All other residues show rather restricted internal motion. Residues 5-38 show strong exchange broadening. This region includes the first α-helix and the engineered six-residue turn and its neighboring residues. Active site residues are 9, 34, 45, 54, 57, 58, 90 and 94. None of these residues show exchange or fast internal mobility, which is significantly above the average. Residue 54 is additionaly nearly exchange-free. 34 and 57 belong to the group with the lowest order. The order of this motion is, however, insignificant. Overall, the protein undergoes large global exchange. Unbound mMjCM is in the molten globular state (see previous Chapter) and does not show any evidence of tertiary structure. Furthermore, relaxation analysis of unbound mMjCM shows that the spatial bond restriction is low. Although complexation with TSA introduces structural order and restricts spatial bond mobility to a level of regular protein, the TSA binding equilibrium allows for global exchange on a timescale faster than is detectable by NMR spectroscopy. Therefore, only one resonance per atom is observed.

We are currently working on an NMR solution structure based on the spectra from the experiments listed above. This will complete the picture of MjCM together with the secondary structure and intramolecular analysis presented in this Chapter.

7. References

[1] MacBeath, G.; Kast, P.; Hilvert, D. *Science* **1998**, *279*, 1958-1961.
[2] Vamvaca, K.; Vögeli, B.; Kast, P.; Pervushin, K.; Hilvert, D. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 12860-12864.
[3] Cavanagh, J.; Fairbrother, W.J.; Palmer, A.G.; Skleton, N.J. . *Protein NMR Spectroscopy. Principles and Pratice*. **1996**, Academic Press, San Diego.
[4] Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 13585-13590.
[5] Pervushin, K.; Riek, R.; Wider, G.; Wüthrich, K. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 12366-12371.
[6] Renner, C.; Schleicher, M.; Moroder, L.; Holak, T.A. *J. Biomol. NMR.* **2002**, *23*, 23-33.
[7] Güntert, P.; Dötsch, V.; Wider, G.; Wüthrich, K. *J. Biomol. NMR.* **1992**, *2*, 619-629.
[8] Güntert, P.; Salzmann, M.; Braun, D.; Wüthrich, K *J. Biomol. NMR* **2000**, *18*, 129-137.
[9] Lee, D.; Hilty, C.; Wider, G.; Wüthrich *J. Am. Chem. Soc.*, submitted.
[10] MacBeath, G.; Kast, P.; Hilvert, D. *Science* **1998**, *279*, 1958-1961.
[11] Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546-4559.

[12] Korzhnev, D. M.; Billeter, M.; Arseniev, A. S.; Orekhov, V.Y. *Prog. Nucl. Magn. Reson. Spectrosc.* **2001**, *38*, 197-266.
[13] Clore, G.M.; Szabo, A.; Bax, A.; Kay, L.E.; Driscoll, P.C.; Gronenborn, A.M. *J. Am. Chem. Soc.* **1990**, *112*, 4989-4991.
[14] Kay, L.E.; Torchia, D.A.; Bax, A. *Biochemistry* **1989**, *28*, 8972-8979.
[15] Tjandra, N.; Bax, A. *J. Am. Chem. Soc.* **1996**, *118*, 6986-6991.
[16] Mandel, A.M.; Akke, M.; Palmer, A.G. *J. Mol. Biol.* **1995**, *246*, 144-163.
[17] Palmer, A.G.; Rance, M.; Wright, P.E. *J. Am. Chem. Soc.* **1991**, *113*, 4371-4380.

Curriculum Vitae

Personal Data

Name:	Beat Vögeli
Date of birth:	May 13, 1976
Place of birth:	Zürich, Switzerland
Citizenship:	Swiss

Education

1996	Swiss Graduation Diploma for High School (Eidgenössische Matura Typus B).
1998	Vordiploma in Physics at ETH Zürich, Switzerland.
2000-2001	Diploma Thesis under the Supervision of Prof. Kurt Wüthrich, Institute of Molecular Biology and Biophysics, ETH Zürich.
2001	Diploma in Physics at ETH Zürich (Dipl. Phys. ETH).
2001-2005	Ph.D. Thesis under the Supervision of Prof. Konstantin Pervushin, Laboratory of Physical Chemistry, ETH Zürich (to be awarded).

Publications

Beat Vögeli, Roland Riek (Supervisor) and Kurt Wüthrich (Supervisor), *NMR Solution Structure of the Fragments of the Human Prion Protein hPrP(121-226) and hPrP(130-230)*, 2001, Diploma Thesis ETH Zürich.

Konstantin Pervushin, Beat Vögeli and Alexander Eletsky**,** 2002, J. Am. Chem. Soc., 124, 12898-12902.

Beat Vögeli and Konstantin Pervushin, 2002, J. Biomol. NMR, 24, 291-300.

Konstantin Pervushin and Beat Vögeli, 2003, J. Am. Chem. Soc., 125, 9566-9567.

Beat Vögeli, Helena Kovacs and Konstantin Pervushin, 2004, J. Am. Chem. Soc., 126, 2414-2420.

Katherina Vamvaca, Beat Vögeli, Peter Kast, Konstantin Pervushin and Donald Hilvert, 2004, Proc. Natl. Acad. Sci. USA, 101, 12860-12864.

Konstantin Pervushin, Beat Vögeli, Tim Heinz and Philippe Hünenberger, 2005, J. Magn. Reson., 172, 36-47.

Beat Vögeli, Helena Kovacs and Konstantin Pervushin, 2005, J. Biomol. NMR, 31, 1-9.

Kaifeng Hu, Beat Vögeli and Konstantin Pervushin, J. Magn. Reson., in press.

Donghan Lee, Beat Vögeli and Konstantin Pervushin, J. Biomol. NMR, in press.

Talks

EENC European Experimental Nuclear Magnetic Resonance Conference, Prague, Czech Republic, 2002.

Lecture Summer School, Otocec, Slowenia, 2003.

Group Representative for Evaluation Committee, ETH Zürich, 2004.

Colloquium Laboratory of Physical Chemistry, ETH Zürich, 2004.

Poster Presentations

5. Igler NMR Tage, Igls, Austria, 2002.

EENC European Experimental Nuclear Magnetic Resonance Conference, Prague, Czech Republic, 2002.

ENC 44th Experimental Nuclear Magnetic Resonance Conference, Savannah, USA, 2003.

D-BIOL Symposium ETH Zürich Congress-Center, Davos, Switzerland, 2004.

EMBO Young Investigator Programme, EMBL Heidelberg, Germany, 2004.

International Greek Biotechnology Forum, Athens, Greece, 2004.

Protein Society Meeting, San Diego, USA, 2004.

Joining Forces Symposium, ETH Zürich, 2004.