

Department Chemie, Lehrstuhl II für Organische Chemie  
der Technischen Universität München

# **NMR Investigations on Structure, Dynamics and Function of VAT-N and DOTATOC**

Mandar Vinayakrao Deshmukh

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität  
München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Johannes Buchner

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Horst Kessler
2. Univ.-Prof. Dr. Frank H. Köhler

Die Dissertation wurde am 29.06.2004 bei der Technischen Universität München eingereicht  
und durch die Fakultät für Chemie am 29.07.2004 angenommen.







तद् गर्वो मयि मास्तु किंतु तदहं तच्छास्त्रतो वृद्धधीः॥

– ऋग्वेद पृ. ३६

*Let me never feel superior to my preceding scientists  
because my interpretation of facts depends on the scientific path they have developed.*

*Rig Veda (Page. 36), 3000 B.C.*



प्रिय बाबा आणि आई,  
तुमच्या माझ्या वरच्या विश्वासामुळे व तुमच्या आशिर्वादामुळेच मला इथवर पोहचता आले. त्यामुळेच हा प्रबंध तुम्हा दोघांना अर्पण करित आहे.

*Dear Aai and Baba (parents),  
Your blessings have helped me to reach up to here  
and that is why this work is dedicated to you both.*





## Acknowledgements

Research described in this thesis was carried out in the NMR Laboratory of Prof. Dr. Horst Kessler of Technische Universität München and supervised by Prof. Dr. Horst Kessler during October 2000 till June 2004.

I am grateful to Prof. Dr. Horst Kessler for his support, earnest guidance, extremely cordial nature and parental care. His innovative ideas and guidance made me appreciate his inborn capabilities as a leading international scientist and a teacher. Because of his jovial and friendly nature, my association with him is a great endeavor.

I thank Prof. Dr. Steffen Glaser who taught me mathematical background of NMR and of RDCs. The memories of our collaborated work on the calculation of the Dipolar Coupling constant will be cherished by me.

My sincere thanks to PD Dr. Gerd Gemmecker, who introduced me to the fascinating world of biomolecular NMR spectroscopy. He inspired, guided, supported and encouraged me during all these years. It has been a pleasure to work with him.

Dr. Rainer Haessner had not only made spectrometers and computers running for us but made my stay in Munich comfortable. I thank him for his always cheerful face and extremely helping nature.

I acknowledge Prof. Dr. Bernd Reif for regular discussions and very clever suggestions.

My colleague, Michael John, stood firmly with me as a friend for all these years. He was my constant supporting factor at TUM and a major collaborator of VAT-N project. I am grateful to Michael for the immense help extended to me right from the first day (picking me up at the airport) till date and for correcting this thesis meticulously.

I appreciate Saravanakumar Narayanan for “walking with me on most of the weekends” and of course for the discussion about the structural biology on the way.

I thank Georg Voll for his immense help during my stay at TUM and also for a splendid collaboration on DOTATOC.

I remember the people who have helped me, those are,

- Dr. Frank Kramer for collaboration on the derivation of RDC equation.

- Andreas Enthart for correcting some chapters and overall help.
- Dr. Murray Coles for the collaboration in the RDC-refined structure calculation of VAT-N and overall discussion on VAT-N projects and for correcting three chapters of this thesis.
- Prof. Dr. Baumeister and Dr. Jürgen Peters for providing three  $^{15}\text{N}$  VAT-N samples.
- Dr. Angelika Kühlewein for collaboration on Ga-DOTATOC.
- Prof. Mäcke for providing DOTATOC samples.
- All the former and current members of AK Kessler with whom I interacted and the secretaries of Prof. Kessler.

I am obliged to my parents and my in-laws for their encouragement and blessings that made me to reach until here. I express my gratitude to my sister-in-law Anagha for her well-wishes. My wife Mrunal also needs to be acknowledged for her love, affection and support extended to me during the time of this Ph.D. work. The acknowledgement wouldn't complete, if I forget to show my affection towards my younger brother, Manjeet, for his support, faith and encouragement to me throughout the period of my scientific life.

München, 29 June 2004

Mandar V. Deshmukh





Though the Sun is small in the size, it lights whole universe. Similarly, one should use words and give maximum meaning out of minimum words.

*Dyaneshwari*

*(Translation of Bhagwad Gita into Marathi by Saint Dyaneshwar at the age of 16)*

जैसे बिंब जरी बचके एवढे ।  
परी प्रकाशा त्रैलोक्य थोकडे ।  
शब्दाची व्याप्ति तेणे पाडे ।  
अनुभवावी ॥

ज्ञानेश्वरी, अध्याय



## **Index**

ABBREVIATIONS .....	XVII
<b>1. PREAMBLE.....</b>	<b>1</b>
<b>2. APPLICATIONS OF MODERN NMR SPECTROSCOPY TO BIOLOGICAL SYSTEMS.....</b>	<b>4</b>
<b>2.1. Relaxation Mechanisms in NMR.....</b>	<b>4</b>
<b>2.2. Spin Interactions in NMR Spectroscopy .....</b>	<b>5</b>
2.2.1. Chemical Shift and Chemical Shift Anisotropy (CSA) .....	5
2.2.2. Scalar Coupling Interactions.....	7
2.2.3. Dipolar Interactions .....	8
<b>2.3. Some Important Concepts in NMR Spectroscopy .....</b>	<b>9</b>
2.3.1. Density Matrix and Product Operator Formalism .....	9
2.3.2. Pulse Fourier-Transform NMR.....	9
2.3.3. Two Dimensional NMR Spectroscopy .....	10
2.3.4. Coherence and Coherence Transfer .....	11
<b>2.4. Experiments Necessary for Biomolecular NMR.....</b>	<b>12</b>
2.4.1. The INEPT Experiment .....	12
2.4.2. The HSQC Experiment.....	13
2.4.3. Sequence Specific Assignments .....	14
2.4.4. Structural Constraints in Protein Structural Calculation.....	15
2.4.5. The TROSY Experiment.....	17
<b>3. RESIDUAL DIPOLAR COUPLINGS: INTRODUCTION AND THEORY .....</b>	<b>20</b>
<b>3.1. Historical Background and Development of RDCs .....</b>	<b>20</b>
3.1.1. First Observation of Dipolar Couplings in Solution .....	20
3.1.2. Alignment of Molecules by External Magnetic Field.....	22
3.1.3. Alignment by External Alignment Media.....	22
<b>3.2. The Concept of the Alignment Tensor.....</b>	<b>23</b>
3.2.1. Static Dipolar Coupling Hamiltonian .....	24

3.2.2.	Time Dependent and Average Dipolar Coupling Hamiltonian.....	25
3.2.3.	Outline of the Key Results .....	26
<b>3.3.</b>	<b>Derivation of the Probability and Alignment Tensors.....</b>	<b>30</b>
<b>3.4.</b>	<b>Expression of Dipolar Coupling Constant in Various Formats.....</b>	<b>35</b>
3.4.1.	Representation of a Highly Rhombic Tensor .....	37
<b>3.5.</b>	<b>Alignment Tensor in the Presence of Internal Motion .....</b>	<b>37</b>
<b>3.6.</b>	<b>Generalised Degree of Order .....</b>	<b>38</b>
<b>3.7.</b>	<b>Conclusion.....</b>	<b>39</b>
<b>4.</b>	<b>PRACTICAL ASPECTS OF RESIDUAL DIPOLAR COUPLINGS: SUBDOMAIN ORIENTATION IN VAT-N.....</b>	<b>41</b>
<b>4.1.</b>	<b>Alignment Media .....</b>	<b>41</b>
4.1.1.	DMPC-DHPC Bicelles.....	45
4.1.2.	Experimental Procedure for the Preparation of Bicelles .....	47
4.1.3.	Filamentous Phage Pfl .....	48
4.1.4.	Polyacrylamide Gel .....	48
<b>4.2.</b>	<b>NMR Experiments for Measuring RDCs.....</b>	<b>49</b>
4.2.1.	Coupled HSQC.....	49
4.2.2.	$^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC .....	49
4.2.3.	TROSY-SemiTROSY (Tr-SmTr) .....	51
4.2.4.	Comparison of the $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC and TROSY-SemiTROSY Approach.....	52
<b>4.3.</b>	<b>Calculation of the Alignment Tensor .....</b>	<b>53</b>
<b>4.4.</b>	<b>Structure Calculation.....</b>	<b>54</b>
<b>4.5.</b>	<b>Application of RDCs: Determination of Subdomain Orientation of VAT-N .....</b>	<b>55</b>
4.5.1.	Introduction to the VAT Complex .....	55
4.5.2.	Solution Structure and the Subdomain Orientation in VAT-N .....	56
4.5.3.	Experimental Section .....	57
4.5.4.	Alignment of VAT-N with Phages.....	58
4.5.5.	Effect of Temperature Change (from 320 K to 313 K) on VAT-N .....	60



4.5.6.	Alignment of VAT-N in Bicelles.....	60
4.5.7.	Subdomain Orientation from the Residual Dipolar Coupling .....	63
4.5.8.	The RDC Refined Structure.....	64
<b>4.6.</b>	<b>Conclusion .....</b>	<b>65</b>
<b>5.</b>	<b>BACKBONE RELAXATION AND INTERNAL DYNAMICS OF VAT-N .....</b>	<b>67</b>
<b>5.1.</b>	<b>Spin Relaxation and Motions in a Protein.....</b>	<b>67</b>
<b>5.2.</b>	<b>Spectral Density Function .....</b>	<b>68</b>
<b>5.3.</b>	<b>Model-free Formalism.....</b>	<b>69</b>
5.3.1.	Dynamical Model Selection and Parameter Estimation .....	71
<b>5.4.</b>	<b>Estimation of Correlation Time and Diffusion Tensor .....</b>	<b>72</b>
5.4.1.	Hydrodynamic Calculations.....	72
5.4.2.	$R_2/R_1$ Ratio.....	73
<b>5.5.</b>	<b>Analysis of the Relaxation Data.....</b>	<b>74</b>
5.5.1.	Error Estimation on the Model Selection.....	74
<b>5.6.</b>	<b>Experimental Measurement of Relaxation Data.....</b>	<b>75</b>
5.6.1.	Error Estimation on the Relaxation Data .....	76
<b>5.7.</b>	<b>Relaxation Measurement for VAT-N.....</b>	<b>77</b>
5.7.1.	Experimental Section .....	78
5.7.2.	Unspecific Oligomerization of VAT-N .....	79
<b>5.8.</b>	<b>Residue Specific Relaxation Rate Analysis.....</b>	<b>81</b>
5.8.1.	Estimation of the Diffusion Tensor and the Model-free Calculation .....	83
5.8.2.	Residue Specific Model and Squared Order Parameter Selection.....	85
5.8.3.	Small Amplitude Motions between sub-domains of VAT-N .....	87
<b>5.9.</b>	<b>Conclusions.....</b>	<b>87</b>
<b>6.</b>	<b>SUBSTRATE BINDING STUDIES OF VAT-N .....</b>	<b>89</b>
<b>6.1.</b>	<b>Ligand screening, the Nature of the Binding and Location of the Binding Site .</b>	<b>89</b>
6.1.1.	Primary Approaches.....	89

6.1.2. Chemical Shift Mapping .....	90
6.1.3. Distance Measurements between the Ligand and the Substrate.....	90
<b>6.2. Substrate Binding Studies of VAT-N .....</b>	<b>91</b>
6.2.1. Assignment of VAT-N Amide Resonances at pH 6.7.....	91
6.2.2. VAT-N:SsrA .....	93
6.2.3. VAT-N: Ubiquitin .....	94
6.2.4. VAT-N:casein .....	95
6.2.5. VAT-N:Barstar .....	97
<b>6.3. Conclusions .....</b>	<b>102</b>
<b>7. INVESTIGATION OF THE STRUCTURAL DIFFERENCES IN Ga<sup>III</sup>- AND Y<sup>III</sup>- DOTATOC .....</b>	<b>104</b>
<b>7.1. Somatotropin Release Inhibiting Factor (SRIF) .....</b>	<b>104</b>
7.1.1. Receptor Subtypes .....	105
7.1.2. Structural Investigations of Somatostatin Analogues .....	106
7.1.3. Radio Labeled Analogues .....	107
7.1.4. Scope of the Present Work .....	108
<b>7.2. Experimental Conditions .....</b>	<b>108</b>
<b>7.3. Results and Discussion .....</b>	<b>110</b>
7.3.1. <sup>1</sup> H NMR and Spectral Assignments .....	110
7.3.2. Characterization of Eu <sup>III</sup> -DOTATOC.....	110
7.3.3. Characterization of Ga <sup>III</sup> - and Y <sup>III</sup> -DOTATOC.....	112
7.3.4. Identification of the Two Conformations of Y <sup>III</sup> -DOTATOC.....	116
7.3.5. The Coalescence between the Conformations.....	118
7.3.6. Cis-trans Isomerization in Y <sup>III</sup> -DOTATOC .....	120
7.3.7. Structure Calculations and MD Simulations .....	122
<b>7.4. Conclusions .....</b>	<b>125</b>
<b>8. REFERENCES.....</b>	<b>126</b>
<b>9. APPENDIX.....</b>	<b>137</b>

<b>9.1.</b>	<b>RDC Studies on VAT-N .....</b>	<b>137</b>
9.1.1.	Pulse Program Implemented: $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC.....	137
9.1.2.	Comparison of RDCs Obtained by $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC and Tr-SmTr Approach	140
9.1.3.	Example of RDC Input File for CNS Calculation .....	143
<b>9.2.</b>	<b><math>^{15}\text{N}</math> Backbone Relaxation Rates (<math>R_1</math> and <math>R_2</math>) and <math>^{15}\text{N}</math>-<math>^1\text{H}</math> Heteronuclear-NOE Used for the Model-free Analysis of VAT-N .....</b>	<b>144</b>
<b>9.3.</b>	<b>Chemical Shift and NOE Tables of <math>\text{Ga}^{\text{III}}</math> and <math>\text{Y}^{\text{III}}</math>-DOTATOC .....</b>	<b>148</b>
9.3.1.	Proton Chemical Shifts for $\text{Ga}^{\text{III}}$ -DOTATOC (290 K) .....	148
9.3.2.	Carbon Chemical Shifts for $\text{Ga}^{\text{III}}$ -DOTATOC (290 K) .....	149
9.3.3.	Proton Chemical Shifts for $\text{Y}^{\text{III}}$ -DOTATOC (Major-290 K).....	149
9.3.4.	Proton Chemical Shifts for $\text{Y}^{\text{III}}$ -DOTATOC (Minor-290 K).....	150
9.3.5.	Carbon Chemical Shifts for $\text{Y}^{\text{III}}$ -DOTATOC (Major-290 K).....	150
9.3.6.	Carbon Chemical Shifts for $\text{Y}^{\text{III}}$ -DOTATOC (Minor-290 K).....	151
9.3.7.	Proton Chemical Shifts for $\text{Y}^{\text{III}}$ -DOTATOC (Major-275 K).....	151
9.3.8.	Proton Chemical Shifts for $\text{Y}^{\text{III}}$ -DOTATOC (Minor-275 K).....	152
9.3.9.	Carbon Chemical Shifts for $\text{Y}^{\text{III}}$ -DOTATOC (Major-275 K).....	152
9.3.10.	Carbon Chemical Shifts for $\text{Y}^{\text{III}}$ -DOTATOC (Minor-275 K).....	153
9.3.11.	NOEs Used for the Peptide Structure Calculation of $\text{Ga}^{\text{III}}$ -DOTATOC.....	154
9.3.12.	NOEs Used for the Peptide Structure Calculation of $\text{Y}^{\text{III}}$ -DOTATOC (Major) .....	157
9.3.13.	NOEs Used for the Peptide Structure Calculation of $\text{Y}^{\text{III}}$ -DOTATOC (Minor).....	158
<b>9.4.</b>	<b>List of Publications .....</b>	<b>159</b>
<b>9.5.</b>	<b>Zusammenfassung für das Jahrbuch der TUM.....</b>	<b>160</b>
<b>9.6.</b>	<b>Lebenslauf.....</b>	<b>161</b>



**ABBREVIATIONS**

AAA, ATPases associated with various cellular activities;  
ADC, analog to digital converter;  
AER, atomic element radius;  
AP, anti-phase;  
ATP, adenosine tri-phosphate;  
CHAPSO, 3-[(3-cholamidopropyl)-dimethylammonio]-2-hydroxy-1-propane sulfonate;  
CNS, crystallography and NMR systems;  
COSY, correlated spectroscopy;  
CO(TTPB)<sub>2</sub>, bis[toluyltris(pyrazolyl)borato]cobalt(II);  
CPMG, Carr Purcell Meiboom Gill;  
CRINEPT, cross-correlated relaxation-enhanced polarization transfer;  
CRIPT, cross relaxation-induced polarization transfer;  
CROP, cross-correlated relaxation optimized polarization transfer;  
CSA, chemical shift anisotropy;  
CT, constant time;  
CTAB, cetyltrimethyl-ammonium bromide;  
CW, continuous wave;  
DBPC, 2,6-di-tertiary-butyl-para-cresol;  
DFS, depth first search;  
DG, distance geometry;  
DHPC, dihexanoylphosphatidylcholine;  
DIODPC, 1,2-di-O-dodecyl-*sn*-glycero-3-phosphocholine;  
DLPC, dilauroyl phosphatidylcholine;  
DMPC, dimyristoylphosphatidylcholine;  
DMPX, 3,7-dimethyl-1-propargylxanthine;  
DMSO, dimethylsulfoxide;  
DNA, deoxyribonucleic acid;  
DOTA, 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid;  
DOTATOC, DOTA-Tyr<sup>3</sup>-octreotide;  
DQF-COSY, double quantum filtered COSY;  
*E. coli*, *Escherichia coli*;  
E.COSY, exclusive correlation Spectroscopy;  
EM, electron microscopy;

## XVIII

Eq., equation;

ESFF, extensible systematic force field;

ESI-MS, electron spray ionization mass spectroscopy;

EXSY, exchange spectroscopy;

FFT, fast Fourier transform;

FT, Fourier transform;

Fmoc, 9-fluorenylmethyloxycarbonyl;

GDO, generalized degree of order;

GFP, green fluorescence protein;

HMBC, heteronuclear multiple-bond correlation;

HMQC, heteronuclear multiple quantum correlation;

HPLC, high performance liquid chromatography;

HSQC, heteronuclear single quantum correlation;

IP, in-phase;

IC<sub>50</sub>, inhibitory concentration 50%;

INEPT, insensitive nucleus enhancement by polarization transfer;

kDa, kilo Dalton;

MAS, magic angle sample spinning;

MD, molecular dynamics;

MEXICO, measurement of exchange rates in isotropically labeled compounds;

MHz, megahertz;

MW, molecular weight;

*n*D, (*number*) dimensional;

NMR, nuclear magnetic resonance;

NOE, nuclear Overhauser enhancement;

NOESY, nuclear Overhauser and exchange spectroscopy;

PAS, principle Axis System;

PET, positron emission tomography;

ppm, parts per million;

RDCs, residual dipolar couplings;

r.f, radio frequency;

ROESY, rotating frame Overhauser effect spectroscopy;

RP-HPLC, reverse phase HPLC;

S<sup>3</sup>E, spin state selective excitation;

SAR, structure activity relationship;

SDS, sodium dodecyl sulfate;

SRIF, somatotropin release inhibiting factors;

SSTR, somatostatin seven transmembrane receptor;

STD, saturation transfer difference

TOCSY, total correlation spectroscopy;

TPPI, time proportionate phase incrementation;

TROSY, transverse relaxation optimized spectroscopy;

Tr-SmTr, TROSY-SemiTROSY;

TSPA, 3-(trimethylsilyl) propionic acid;

VAT, valosin containing protein like ATPases of *Thermoplasma Acidophilium*;

VCP, valosin containing protein;

WATERGATE, water suppression by gradient- tailored excitation;





## 1. Preamble

The independent and simultaneous discovery of the NMR phenomenon was accomplished in 1946, at Harvard by Purcell *et al.* <sup>[1]</sup>, and at Stanford by Bloch *et al.* <sup>[2]</sup>. Since then, NMR spectroscopy has been established as a prominent tool in chemistry and biochemistry. Particularly successful has been the application of NMR to biomolecules. One of the main reasons behind this great success has been the clear understanding of the fundamental principles of various spin interactions from the very early stages of its inception, *e. g.* the discovery of the property of nuclear spins by Wolfgang Pauli in 1924 and the concept of the dipolar interactions between spins in 1927 <sup>[3, 4]</sup>. In modern NMR spectroscopy and methodological developments, the theoretical understanding of spin interactions has been of great importance.

An introduction to the NMR phenomenon, various spin interactions and several relaxation mechanisms, some basic concepts of the formalism, Fourier transform NMR, multidimensional NMR and the concept of coherence are presented in chapter 2. Much of biomolecular NMR spectroscopy practiced today involves two or more nuclei, because information derived from two or more nuclei and their interactions under a tailored Hamiltonian yield a direct correlation with the structure of the bio-macromolecule. However, the low sensitivity of most of the NMR active nuclei is an intrinsic problem. This can be reduced by achieving magnetization transfer from sensitive nuclei to less sensitive nuclei *via* coherent or non-coherent pathways. Magnetization transfer therefore has a great importance in NMR spectroscopy. Such a magnetization transfer is achieved by techniques like INEPT <sup>[5]</sup>, CRIPT <sup>[6]</sup>, CRINEPT <sup>[7]</sup> and CROP sequence <sup>[8]</sup>. Introduction to these methods as well as introduction of the important experiments like HSQC <sup>[9, 10]</sup>, TROSY <sup>[11]</sup> and triple resonance experiments for sequence specific assignment are also part of chapter 2.

Residual dipolar couplings (RDCs) have recently found a wide range of applications in high resolution NMR of biomolecules in the liquid state <sup>[12-14]</sup>. A non-isotropic orientational distribution of a molecule of interest results in non-zero average dipolar coupling constants. Such residual dipolar couplings (RDCs) can be produced by addition of external alignment medium in the sample solution and constitute information on the structure of the molecule in the form of bond orientations. The alignment tensor is the crucial concept for the interpretation of RDCs. In chapter 3, an intuitive introduction to the alignment tensor and an elementary derivation of the key equations is presented, in the course of which, the importance of the probability tensor is also discussed. Practical aspects of RDCs, such as,

choice of appropriate alignment media, pulse program and the fitting algorithm are discussed in chapter 4.

Chapter 4 also describes the application of RDCs for evaluating the inter-domain orientation of VAT-N. VAT-N (20.5 kDa) is the N-terminal domain of the VAT, an AAA ATPases. VAT (Valosine-containing protein-like ATPases of *Thermoplasma acidophilum*) displays a tripartite domain structure, N-D1-D2, commonly found for proteins belonging to the AAA family, and homo-hexameric ring architecture, typical of type II ATPases. VAT has been shown to act as an ATP-driven protein unfoldase. The N-domain is involved in substrate binding and is alone competent in the folding of permissive substrates<sup>[15]</sup>.

The role of VAT-N is dispensable in the full unfoldase activity of the VAT complex, which requires only the ATP-driven D1-D2 module. VAT-N thus clearly controls access of substrates to the D1-D2 unfoldase machine, although the mechanisms of this control remain controversial. Various proposals, including the “entropic brush” mechanism<sup>[16]</sup>, in which VAT-N has the role of removing unwanted substrates from the main D1-D2 binding site, must be considered and evaluated. The solution structure of VAT-N<sup>[17]</sup> revealed two equally sized sub-domains VAT-Nn and VAT-Nc, arranged into a kidney-shaped overall structure. The relative orientation of the two sub-domains and the possibility of inter-domain flexibility are important issues in determining the function of VAT-N. In particular, the opening of the cleft between sub-domains to expose the hydrophobic surface between them has been proposed as a functional mechanism<sup>[17]</sup>. An alternative proposal is that the binding site is located in a cleft between the loops which characterize the surface of both sub-domains<sup>[17]</sup>.

Chapter 5 discusses results from the relaxation analysis of VAT-N which were carried out to probe the possibility of large amplitude motions between sub-domains. Dipolar nuclear magnetic spin relaxation of protonated heteronuclei, such as  $^{15}\text{N}$ , is mediated by overall rotational tumbling of the molecule and by internal motions of the N-H bond vector<sup>[18]</sup> and has become widely accepted tool for characterization of their global and internal motions. Consequently, measurement of  $^{15}\text{N}$  spin relaxation parameters, primarily the spin-lattice and spin-spin relaxation rate constants and the steady state  $^{15}\text{N}$ - $^1\text{H}$  nuclear Overhauser effect (NOE), are powerful techniques for experimental investigation of dynamics in biological macromolecules<sup>[19]</sup>. Backbone  $^{15}\text{N}$  relaxation data can be analyzed in the model-free framework<sup>[20, 21]</sup>.

Exploration of the natural substrates for VAT-N has also been attempted and is described in chapter 6. One of the most important applications of biomolecular NMR spectroscopy is in the identification of the binding between a substrate and a ligand molecule. The ligand can either be a small organic molecule, an oligopeptide or even a protein. NMR spectroscopy not

only can detect the binding but it can locate the binding site, estimates the strength of binding, elucidate the structure of the bound state and changes in the dynamics [22-25]. NMR has become a first hand tool for discovering protein substrates and thus facilitates obtaining a complete picture of the functional mechanism of a macromolecule. These methodologies are also discussed in chapter 6.

Chapter 7 discusses results from an NMR study of DOTATOC, a somatostatin analogue that has been established recently for *in vivo* diagnosis and targeted therapy of somatostatin receptor-positive tumors. DOTATOC consists of a disulfide-bridged octapeptide, D-Phe<sup>1</sup>-Cys<sup>2</sup>-Tyr<sup>3</sup>-D-Trp<sup>4</sup>-Lys<sup>5</sup>-Thr<sup>6</sup>-Cys<sup>7</sup>-Thr<sup>8</sup>-ol, connected to a metal chelator DOTA (1, 4, 7, 10-tetraazacyclododecane-1, 4, 7, 10-tetraacetic acid). The structural investigations of Ga<sup>III</sup> and Y<sup>III</sup> chelated DOTATOC by <sup>1</sup>H and <sup>13</sup>C solution NMR and molecular modeling are shown.

## 2. Applications of Modern NMR Spectroscopy to Biological Systems

Nuclei possess an overall spin and therefore a spin angular momentum, characterized by the nuclear spin quantum number,  $I$ , which is quantized both in magnitude and orientation (a vector quantity). For spins with  $I = 1/2$  two orientation states exist, namely,  $m_z = +1/2$  and  $m_z = -1/2$ . For an isolated spin in the absence of external magnetic fields these orientations are of equal energy. In the presence of the external magnetic field ( $B_0$ ), however, polarization of the states takes place with the energy difference given by

$$E = -\hbar\gamma \cdot B_0 \quad [2-1],$$

where  $\gamma$  is the magnetogyric ratio leading to a population difference governed by a standard Boltzmann distribution. NMR is a spectroscopic method that exploits this energy and population difference. The transition between these two energy levels (often referred as Zeeman energy levels) can be achieved by application of radio frequency according to the Larmor condition <sup>[26]</sup>,

$$\omega_0 = -\gamma \cdot B_0 \quad [2-2],$$

where  $\omega_0$  is the precession frequency. The magnetogyric ratio is an intrinsic property and has distinct values for each nucleus, leading to well separated frequencies for different nuclei of interest.

The NMR absorptive signal is a response to perturbation of the net magnetization by relaxation processes and spin interactions.

### 2.1. Relaxation Mechanisms in NMR

Excited magnetization returns back to its original low-energy state *via* various relaxation processes. These relaxation processes are always chosen in such a way that it has minimum enthalpy and maximum entropy. In NMR spectroscopy, relaxation works mainly by two processes, spin-lattice relaxation ( $T_1$ ), spin-spin relaxation ( $T_2$ ).

The spin-lattice relaxation rate constant ( $R_1=1/T_1$ ) describes the recovery of the longitudinal magnetization to the thermal equilibrium or return of the population of the energy levels of the spin systems to the equilibrium Boltzmann distribution. The spin-spin relaxation ( $R_2=1/T_2$ ) rate constant describes the decay of the transverse magnetization to zero, or, the decay of transverse single quantum coherence (*vide infra*).

Additionally, relaxation mechanisms like dipole-dipole relaxation and chemical shift anisotropy (CSA) scale these processes.

The dipole-dipole relaxation mechanism is an important relaxation mechanism and gives rise to nuclear Overhauser effect (NOE). If both the nuclei undergoing this type of relaxation mechanism have similar sign of their respective magnetogyric ratios then this leads to an enhancement in the intensity otherwise a negative enhancement (lose in the intensity) can be observed. The sign of the NOE enhancement also depends on the product of the magnetic field and the rotational correlation time of the molecule. Chemical shift anisotropy also provides a gateway for the relaxation process in NMR spectroscopy (*vide infra*).

## 2.2. Spin Interactions in NMR Spectroscopy

NMR spectroscopy is dominated by the Zeeman interaction (which is the largest interaction due to large  $B_0$  field, *cf.* Eq [2-1]), while other interactions originating in the molecule are perturbations on the Zeeman interaction.

In a diamagnetic system other interactions are:

- chemical shielding interactions,
- electron coupled scalar interactions, and
- homo- and heteronuclear dipolar interaction.\*

All the NMR spin interactions can be denoted in the form of an operator. This operator is described as a Hamiltonian operator,  $\mathcal{H}$ . The internal spin interaction, in each case, behaves as a tensor of rank two and the interaction tensor is fixed in the molecular frame, and in the so-called Principal Axis System (PAS), the tensor is diagonal in its matrix representation.

### 2.2.1. Chemical Shift and Chemical Shift Anisotropy (CSA)

The electron distribution around the nucleus distorts the magnitude and the direction of the applied magnetic field,  $B_0$ , to produce an effective field that varies from one chemical environment to another. This chemical shielding result in a modification of the Zeeman energy levels called the chemical shift. The chemical shielding Hamiltonian is also expressed as,

---

\* For nuclei with integer and integer-half spin quantum number, additionally, informative quadrupolar interactions exist. Further discussion of quadrupolar interactions is beyond the scope of this work.

$$\mathcal{H}_{CSA} = \frac{\gamma_I \Delta\sigma B_0}{3} \sqrt{1 + \frac{\Delta\eta^2}{3}} \quad [2-3],$$

where  $\Delta\sigma$  is the shielding anisotropy. The chemical shielding is a symmetric second rank tensor completely characterized by the three principal elements  $\sigma_{11}$ ,  $\sigma_{22}$  and  $\sigma_{33}$ , and the orientation (given by three Euler angles) which are the diagonal elements of the tensor in the molecule fixed axis system or PAS. The chemical shielding interaction is characterized by the isotropic value ( $\sigma_{iso}$ ), the asymmetry parameter ( $\eta$ ) and the shielding anisotropy ( $\Delta\sigma$ ) which are given as:

$$\sigma_{iso} = \frac{1}{3} \text{Tr}(\sigma) = \frac{(\sigma_{11} + \sigma_{22} + \sigma_{33})}{3} \quad [2-4],$$

$$\Delta\sigma = \sigma_{33} - \frac{1}{2}(\sigma_{11} + \sigma_{22}) \quad [2-5],$$

and

$$\eta = \frac{(\sigma_{22} - \sigma_{11})}{(\sigma_{33} - \sigma_{iso})} \quad [2-6].$$

In solution NMR spectroscopy, the Brownian motion of the molecule leads to an isotropic averaging of the chemical shielding tensor to the scalar value  $\sigma_{iso}$ , Eq [2-4], which is the commonly known chemical shift and expressed in the *ppm* scale. Chemical shift is caused by electron current induced by  $B_0$  and can therefore be measured relative to  $B_0$  *i.e.* in *ppm*.

Partial or full alignment of a molecule with respect to  $B_0$  leads to an orientation dependent change in chemical shift, termed chemical shift anisotropy (CSA). Reintroduction of CSA takes place in partially aligned samples although its magnitude is very small compared to the pure solid crystalline sample. However, it is necessary to consider such contributions when achieving unambiguous resonance assignment under partial alignment. In a completely anisotropic case, orientational dependence between  $B_0$  and chemical shielding tensor (in a principal axis system) leads to resonance frequencies to give rise to an asymmetric powder pattern which is also referred as Pake pattern<sup>[27]</sup>.

CSA is also mentioned as a relaxation mechanism in the previous section. The maximum CSA for a particular nucleus is of the order of the chemical shift range for the nucleus, and therefore CSA is an important relaxation mechanism only for nuclei with a wide chemical shift range (such as  $^{15}\text{N}$  and  $^{13}\text{C}$ ). CSA contributions to  $R_1$  and  $R_2$  relaxation rate constants have a quadratic dependence on the applied magnetic field strength. Thus, use of higher magnetic fields does not always increase the achievable signal-to-noise ratio as much as expected theoretically.

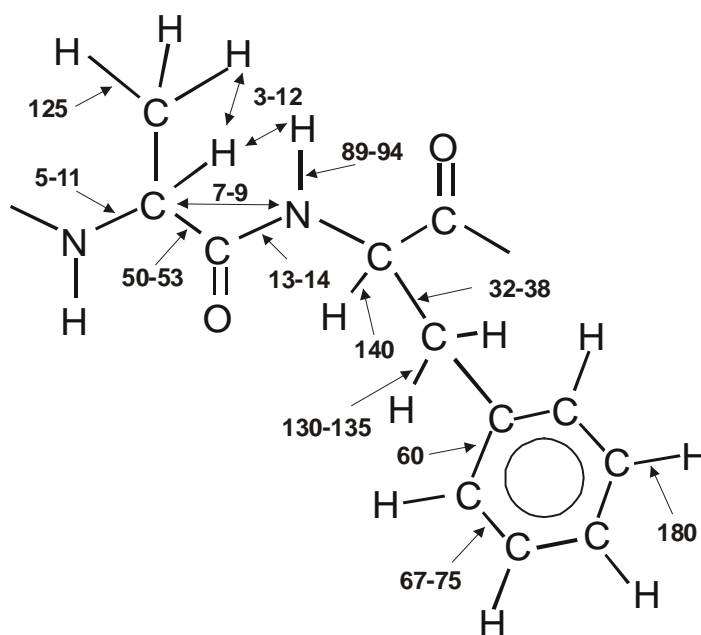
### 2.2.2. Scalar Coupling Interactions

The spin-spin coupling between two adjacent nuclei can occur through space or *via* spin delocalization through chemical bonds (mediated by the electrons forming the chemical bonds between the nuclei) [28]. The former gives rise to dipole-dipole interactions, while the latter gives rise to the scalar coupling denoted by a symbol  $J$ .

$J$ -coupling interaction is characterized by the tensor  $\mathbf{J}$  which has a finite trace. Molecular tumbling in liquid yields the isotropic average of  $J$ , *i.e.*,  $\mathbf{J}_{\text{iso}}$ . The Hamiltonian for  $J$  interaction is

$$\mathcal{H}_J = \mathbf{I} \cdot 2\pi \cdot \mathbf{J} \cdot \mathbf{S} \quad [2-7],$$

The scalar coupling interaction manifests in liquids as spin multiplets and provides  $J$ -coupling as another important parameter for spectral assignments, interpretation and as an aid for the structural elucidation. Figure 2-1 shows the scalar couplings [Hz] commonly found in an amino acid.



**Figure 2-1:** Scalar couplings [Hz] typically observed in a peptide. The sign of the coupling depends on the geometry as well as on the sign combination of gyromagnetic ratio of nuclei and is ignored here.

Table 2-1 gives a direct correlation of the scalar coupling and the secondary structure elements found in proteins.

**Table 2-1:** Characteristic  ${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha}}$ -coupling constants [Hz] in secondary structure elements of proteins <sup>[29, 30]</sup>

Secondary Structure Elements	${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha}}$	Secondary Structure Elements	${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha}}$
Helices	< 6	$\alpha$ -helices	3.9
$\beta$ -sheets	> 8	Parallel $\beta$ -sheets	9.7
Random coil	6-8	Anti-parallel $\beta$ -sheets	8.9

Scalar interactions also open the door to coherent magnetization transfer, which is the basis of most of the recently developed multidimensional NMR experiments. A discussion of these aspects appears in the latter part of this chapter.

### 2.2.3. Dipolar Interactions

For a pair of interacting nuclei **I** and **S**, the dipolar Hamiltonian can be expressed as

$$\mathcal{H}_D = -\frac{\gamma_I \gamma_S \mu_0 \hbar}{8\pi} \frac{1}{R_{IS}^3} (3\cos^2\theta - 1) 2I_z S_z \quad [2-8],$$

where  $\gamma_I$  and  $\gamma_S$  are the nuclear magnetogyric ratios of the designated spins,  $R_{IS}$  is the internuclear distance,  $\theta_{IS}$  is the angle between internuclear vector  $r_{IS}$  and the static magnetic field  $B_0$  directed along the  $z$  axis in the laboratory frame and  $I_z S_z$  is the spin operator function (*vide infra*). The dipolar interaction has the important property that  $D$  is axially symmetric and traceless. The dependence of dipolar interaction on the magnetogyric ratio has important consequences for abundant spins (*e. g.*  ${}^1\text{H}$ ,  ${}^{19}\text{F}$ ). The dependence of dipolar interactions on the  $r_{IS}$  and orientation  $\theta_{IS}$  implies that one can derive structural information (in terms of bond lengths and bond angles) from the dipolar interactions. Here one may distinguish homo-nuclear or hetero-nuclear dipolar interactions, depending upon whether the nuclear spin interacts within the same type or different types of nuclei, respectively.

The orientational dependence of the previously discussed scalar coupling (J interaction) is  $P_2(\cos\theta)$  (also called Legendre Polynomial) and is equal to  $(3\cos^2\theta - 1)$ , where  $\theta$  is the angle between the vector connecting **I** and **S** and  $B_0$ . Clearly, the  $P_2(\cos\theta)$  dependence makes the J interaction inseparable from dipolar interactions in solids or in the NMR of partially aligned molecules. Both  $D$  and  $J$  are field-independent and their separation can only be accomplished by a knowledge of one over the other.



In the next two chapters we will see how this can be achieved in the realm of solution NMR spectroscopy.

### 2.3. Some Important Concepts in NMR Spectroscopy

#### 2.3.1. Density Matrix and Product Operator Formalism

Most of NMR theory does not always need a sophisticated quantum mechanical background, and some of the NMR phenomena can be often expressed with equal elegance by both classical and quantum mechanical methods (because the NMR equation involves only classical terms, see Eq. [2-2]). Therefore, interactions of the spins, followed by a single 90° pulse or a train of such pulses, can be visualized by density matrix formalism (based on quantum mechanics)<sup>[31, 32]</sup> as well as by vector model analysis (based on classical mechanics)<sup>[33]</sup>. Classical methods, nevertheless, are inadequate to describe more complicated spin interactions. Quantum mechanical methods *via* the density matrix provide a complete description of the state of a spin system, however, the requisite matrix calculations quickly become cumbersome as the number of spins and eigenstates increases. A simple compromise, involving Cartesian spin operators and their conversions in the weak coupling limit, was developed by Sørensen *et al.*, namely, product operator formalism<sup>[34]</sup>. This approach will be used in this thesis for the description of various NMR experiments.

#### 2.3.2. Pulse Fourier-Transform NMR

The original discovery of NMR was a continuous wave (CW) absorption and induction technique and has almost exclusively been replaced by the more elegant and experimentally advantageous pulse Fourier Transform (FT) technique. The Fourier Transform (FT) technique was introduced to NMR by Ernst and Anderson in 1966<sup>[35]</sup>, leading to explosive developments in pulse NMR methodology, instrumentation and practice. The pulse FT method involves the application of short radio frequency pulses to the nuclear spins with an immediate capture of the time-domain response. The frequency domain spectrum  $F(\omega)$  can be recovered from the experimentally detected time-domain signal  $f(t)$  *via* Fourier transform, as the time and the frequency domains in the pulse NMR experiment are mathematically related by the FT relationship as,

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{i\omega t} dt \quad [2-9].$$

The time domain signal  $f(t)$  is a complex quantity and the FT embedded in Eq. [2-9] is a complex operation therefore the measured signal must also be complex. By a judicious combination of quadrature components and FT operations, the time domain signal, can be manipulated to produce a sign discriminated absorption and dispersion mode frequency spectrum from the pulse NMR experiment. It is customary to display and inspect only the absorption spectrum.

Experimental considerations in pulse FT-NMR first require the analog time domain signal to be sampled in a two-channel quadrature detector and then converted into a digital signal using an analog to digital converter (ADC). The digital sampling of the analog signal must satisfy the Nyquist sampling criterion<sup>[36]</sup>. The FT of this digital signal is carried out numerically on the spectrometer computer or on a remote workstation using the Cooley-Tukey algorithm (Fast Fourier Transform - FFT)<sup>[37]</sup> which requires the transform data size to be  $2^n$  complex data points.

### 2.3.3. *Two Dimensional NMR Spectroscopy*

One of the most important developments in pulse FT-NMR spectroscopy is the introduction of two-dimensional Fourier Transform NMR (2D FT-NMR) spectroscopy. The first experimental 2D NMR was performed in the laboratory of Prof. R. R. Ernst in mid-seventies<sup>[38]\*</sup>. Since then 2D, 3D and 4D NMR spectroscopy have become powerful tools for the structural elucidation of complex molecules in solution, especially proteins.

In 2D NMR spectroscopy, the 1D pulse sequence is extended to include a second time dimension. The total experiment is divided into four periods: preparation, evolution, mixing and detection. The preparation period leads to creation of a non-equilibrium state of the spin system by the application of suitable pulses. During the evolution period, the spin system is allowed to evolve under the influence of a suitable tailored Hamiltonian. The evolution period is incremented, providing an additional time period  $t_1$ , so as to collect an adequate sampling of data points. The mixing period corresponds to coherent or incoherent transfer of magnetization. The detection period is the same as in the 1D experiment, with the time domain signal detected in quadrature. The 2D experiment thus has two time domains with the

---

\* Prof. R. R. Ernst was awarded with the Nobel Prize in Chemistry in 1991 for his contributions to the development of the methodology of high resolution NMR spectroscopy. He is recognized as the father of the NMR spectroscopic technique.

NMR signal a function of two time variables  $t_1$  and  $t_2$ ,  $f(t_1, t_2)$ , which upon double Fourier transformation yields a two dimensional NMR spectrum.

To obtain pure phases (or quadrature detection) in 2D NMR both real and imaginary components must be acquired. Such information can be gained by using one of three popular methods, namely, States<sup>[39]</sup>, TPPI<sup>[40]</sup> and echo-antiecho<sup>[41]</sup>.

#### 2.3.4. Coherence and Coherence Transfer

Coherence is a relationship between two states across a single nuclear transition, or multiple states for multiple transitions<sup>[32, 42]</sup>. A diagonal matrix element of the density operator,  $\rho_{jj} = c_j c_j^*$ , is a real and positive number that corresponds to the population of the state described by the basis function  $|j\rangle$ . An off-diagonal element of the density operator,  $\rho_{jk}$ , represents coherence between eigen-states  $|j\rangle$  and  $|k\rangle$ , in the sense that the time-dependent phase properties of the various members of the ensemble are correlated with respect to  $|j\rangle$  and  $|k\rangle$ . Coherences can be classified by their coherence order  $p$ , which are various values of  $\Delta m$  (change in the spin angular-momentum quantum number):  $0, \pm 1, \pm 2$  etc. Those matrix elements that denote  $\Delta m = \pm 1$  are called *single quantum coherence*, those that denote  $\Delta m = \pm 2$  *double quantum coherence* and that denoting  $\Delta m = 0$  *zero-quantum coherence*.

The density operator after the radio frequency pulse is said to represent a coherent superposition between two states which is referred to as *coherence*. It describes correlation of quantum-mechanical phase among a number of systems (separate nuclei) that persist even after the r. f. field is removed. Coherence is a phenomenon associated with an NMR transition but is not a transition and coherence does not change the populations of the spin states.

An example of the coherence transfer pathways, occurring in the COSY experiment, is shown in figure 2-2. It is important to note that  $p = -1$  is detected in the receiver. Hence, other coherence pathways are not selected.



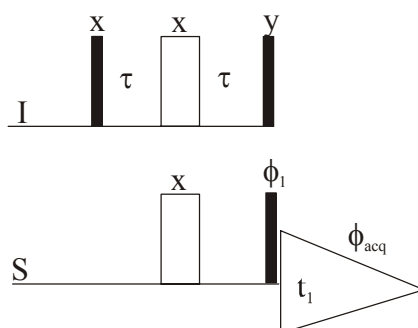
**Figure 2-2:** A pulse sequence of 2D-COSY experiment with P (left) and N (right) type coherence transfer pathways.

## 2.4. Experiments Necessary for Biomolecular NMR

### 2.4.1. The INEPT Experiment

The INEPT (Insensitive Nuclei Enhancement by Polarization Transfer) is a 1D equivalent experiment of the H-X correlation (in 2D manner) which was first brought into practice by Maudslay and Ernst in 1977 [43]. The INEPT [5] is widely used method which allows transfer of magnetization in a coherent and non-selective way from spin I to spin S and *vice versa* (conventionally, I is the sensitive spin and the S is the insensitive spin).

The sequence begins with the excitation of all I spins, which then evolve under the effects of chemical shift of the I spin and heteronuclear coupling to the S spin. After a period of  $1/4J_{IS}$ , the  $180^\circ$  pulse refocuses the chemical shift evolution (and the field inhomogeneity) during the second  $1/4J_{IS}$  period. The simultaneous application of a  $180^\circ$  pulse on the S nuclei ensures the evolution of heteronuclear coupling (counter-precessing relative to the proton evolution). After a total evolution period of  $1/2J_{IS}$ , a  $90^\circ$  pulse along the  $y$  axis for the I spin leaves magnetization along the  $z$ . The  $90^\circ$  pulse along  $x$  axis for the S spin converts  $S_z$  into the observable magnetization (anti-phase doublet).



**Figure 2-3:** The original INEPT sequence [5]. The narrow and the wide rectangular bars represent  $90^\circ$  and  $180^\circ$  pulse, respectively. The  $x$  and  $y$  denotes the direction of the pulse.  $\phi_1$  represents phase of the pulse which needs to be cycled and  $\phi_{acq}$  represents the receiver phase.

The product operator analysis of the INEPT sequence can be given as:

$$I_z \xrightarrow{\frac{\pi}{2}I_x - t - \pi(I_x + S_x) - t} -I_y \cos(2\pi J_{IS}t) + 2I_x S_z \sin(2\pi J_{IS}t) \xrightarrow{\frac{\pi}{2}(I_y + S_x)} 2I_z S_y \quad [2-10].$$

The cosine term in the above equation becomes zero, while the sine term retains (becomes one). The sensitivity gain by the INEPT sequence can be given by Eq. [2-11]

$$I_{\text{INEPT}} = I_{\text{conventional}} \left| \frac{\gamma_I}{\gamma_S} \right| \quad [2-11].$$

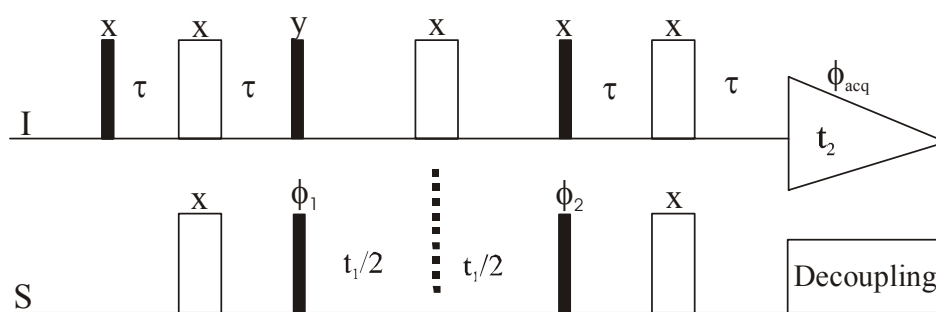
$I_{\text{INEPT}}$  (signal intensity) gains are directly related to the gyromagnetic ratio (and are absolute in sign), therefore make a notable gain in the intensity compared to the gains from the I-S NOE. In case of a  $^1\text{H}$ - $^{15}\text{N}$  spin pair one can gain up to 10 times in intensity using the INEPT transfers. Additionally, indirect detection of S nuclei (*i.e.* detection on I spin, which is used now on) can enhance the gain up to  $(\gamma_I/\gamma_S)^{3/2}$  times the conventional detection. The high magnetogyric ratio of the proton, and its nearly 100 % natural abundance and ubiquity makes direct proton observation more favourable in modern NMR spectroscopy.

In a high molecular weight protein or a protein complex (*ca.* 50 kDa molecular weight or more), the INEPT transfer suffers from enhanced relaxation losses during the evolution time (the maximal efficiency of transfer between the operators  $I_z$  and  $2I_zS_z$  depends only on the scalar coupling constant  $J$  and the net auto-correlated and cross-correlated relaxation rates of spin  $I$ ). To overcome this problem several other polarization transfer schemes have been proposed in recent years. These involve CRIPT<sup>[6]</sup>, CRINEPT<sup>[7]</sup> and the analytically derived optimally-controlled CROP sequence<sup>[8]</sup>. The utility of these modifications in polarization transfer along with specific labeling schemes (such as perdeuteration) have been demonstrated in the studies of the GroEL-GroES complex of 900 kDa<sup>[44]</sup>.

#### 2.4.2. The HSQC Experiment

The HSQC<sup>[9, 10]</sup> (Homonuclear Single Quantum Coherence), is a routinely used experiment in biomolecular NMR spectroscopy. It correlates the chemical shift of the proton with that of its attached heavy atom (I-S pair). This information can be very useful, particularly for recognizing whether a protein is folded and intact. It also forms the basis for nearly all multinuclear 3D spectra. The basic pulse sequence of HSQC is simple and consists of INEPT for transferring I spin magnetization to the S, where it is left to evolve during  $t_1$  time. This magnetization is transferred back to the I spins *via* a reverse INEPT step and then detected, as shown in figure 2-4.

Variants of the HSQC experiment are commonly seen in NMR literature today<sup>[45]</sup>. This includes gradient HSQC, and sensitivity enhancement by double INEPT during the reverse INEPT period<sup>[46, 47]</sup>. Apart from this, to reduce the intensity losses due to relaxation during the evolution time, concatenation of an  $J$  evolution period and  $t_1$  evolution period is possible, which is known as constant-time (CT)<sup>[48-50]</sup> and semi-constant time<sup>[51, 52]</sup>.



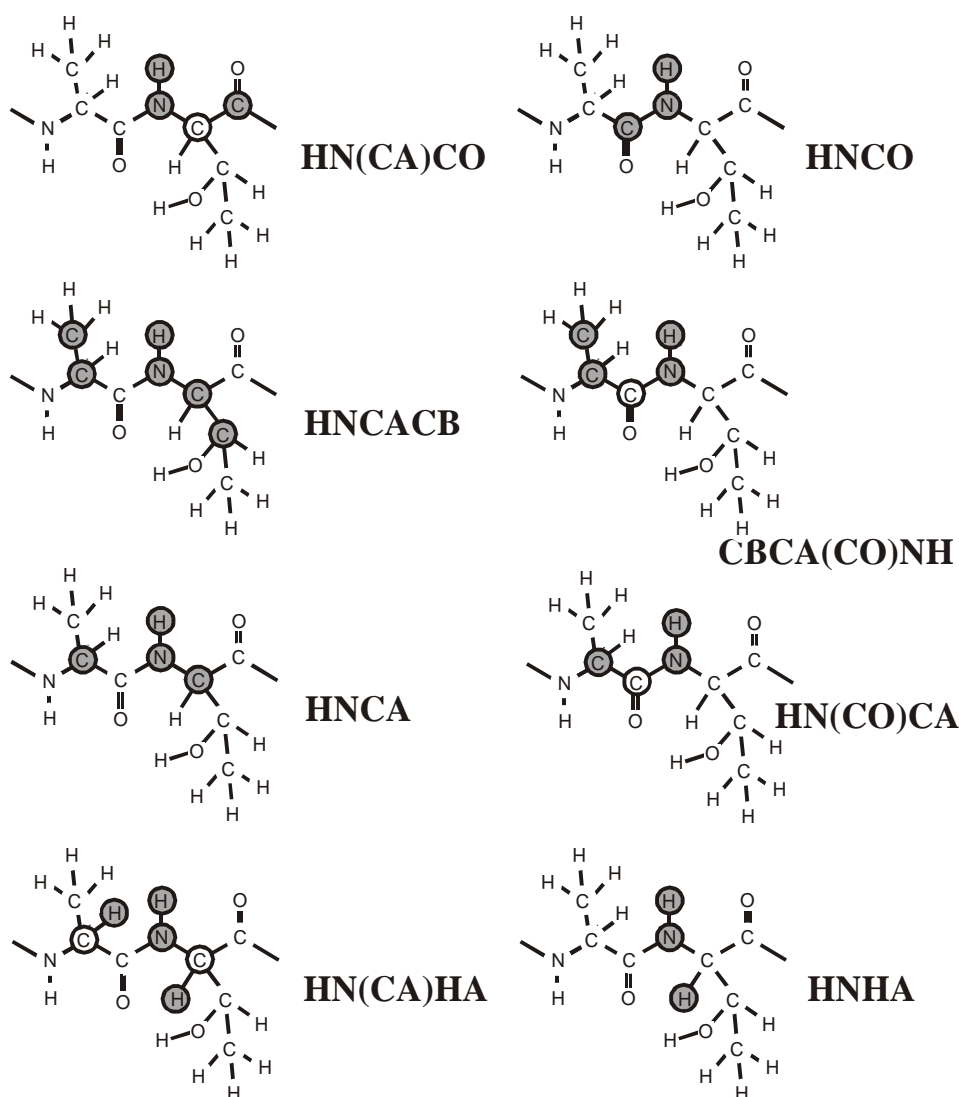
**Figure 2-4:** Pulse sequences for the fundamental HSQC sequence. The delay  $\tau$  is set to  $1/(4J_{IS})$ . The basic phase cycling is  $\phi_1 = x, -x, x, -x$ ,  $\phi_2 = x, x, -x, -x$ , and receiver  $\phi_{\text{acq}} = x, -x, -x, x$ <sup>[47]</sup>. Spin decoupling does not allow evolution of I-S coupling during the acquisition time period and is normally achieved by a combination of composite pulses.

Another method was developed subsequently for the heteronuclear correlation, which was named HMQC<sup>[53]</sup>. The distinction between these two proton-detected heteronuclear correlation techniques is that IS-spin coherence is stored as multi-quantum (HMQC) or single-quantum (HSQC), during the  $t_1$  evolution period. The HMQC approach is more robust and can be optimized for the double quantum-zero quantum relaxation occurring in the transverse plane<sup>[54]</sup>.

#### 2.4.3. Sequence Specific Assignments

Sequence-specific assignments have developed over last 15 years, due to the early efforts of Wüthrich and co-workers<sup>[30]</sup> and have undergone many fruitful modifications. Sequence-specific assignment yields the backbone and the side chain assignment strategy of a protein in a systematic way by making use of covalent connectivities. For example, sequence-specific assignment practiced today for the backbone assignment experiments, involving magnetization transfers from the amide nitrogen, to a carbon at the  $\alpha$ -position and to the carbonyl carbon.

The necessary condition of sequence specific assignment is a uniformly  $^{13}\text{C}$  and  $^{15}\text{N}$  labeled protein sample (commonly denoted as: U- $^{13}\text{C}$ - $^{15}\text{N}$ ) which can be achieved easily (though expensively) by expressing the protein in a bacterial host organism (usually *E. coli*) grown in media where  $^{13}\text{C}_6$ -Glucose and  $^{15}\text{NH}_4\text{Cl}$  are the only carbon and nitrogen sources. The introduction of  $^{13}\text{C}$  and  $^{15}\text{N}$  isotope labeling of NMR samples overcomes the low natural abundances of these nuclei in NMR.



**Figure 2-5:** Peptide backbone connectivities and the sequence specific assignment experiment based on them. Typically experiments are designed in following basic steps which involve magnetization transfer from protons to nitrogen via INEPT and then to the carbon, evolution of magnetization under tailored Hamiltonian ( $t_1$  period), which is transferred back to nitrogen for a second evolution time ( $t_2$  period) and later to the protons for detection ( $t_3$  period).

Experiments developed based on sequence specific assignments are shown in figure 2-5 and can be studied in more detail in a review by Sattler *et al.* [55].

#### 2.4.4. Structural Constraints in Protein Structural Calculation

Sequence specific pulse schemes are utilized for generating a complete list of resonances for each residue. Afterwards, combination of all or some structural restraints originating from

NOEs, J-couplings, chemical shifts, H-bond information and residual dipolar couplings, are used for structure calculation.

Initially, structure determination by NMR utilized the 2D-homonuclear NOESY experiment which was sufficient to give structural restraints for proteins up to 70 residues<sup>[30]</sup>. Isotope labeling in high molecular weight proteins provided the gateway not only for the assignment strategy but for the evolution of heteronuclear edited NOESY experiment<sup>[56, 57]</sup>. One such heteronuclear edited NOESY experiment was proposed in a 3D manner and became popular with the name HSQC-NOESY<sup>[58, 59]</sup>. This experiment provided correlations between the NH amide spin pair and all the other protons closer than about 5 Å. For observation of the side-chain/side-chain contacts an HSQC-NOESY experiment was proposed where evolution of NOE could be observed<sup>[58, 59]</sup>. Problems of extensive signal overlap in the protons can be solved using a combination of NNH-<sup>[60]</sup>, CCH-, NCH-, and CNH-NOESY<sup>[61]</sup> experiments that exploit the large spectral dispersion of the heteronuclei. Around the same time a 4D CNH-NOESY<sup>[62]</sup> was also proposed which is less in use because of the time investment involved.

The direct relation of  ${}^3J_{\text{H}^{\text{NH}}\alpha}$  to the secondary structural element is stated in the early section of this chapter (table 2-1). Secondary structure elements of proteins or peptides are defined by H-bonding between the residues. The backbone torsion angles ( $\phi$ - and  $\psi$ -angles) are restricted only to certain values such that the H-bonding should be possible for the formation of the secondary structure elements in a protein or in a peptide<sup>[63]</sup>. Several experiments allow the measurement of the  $\phi$ -angle value<sup>[64, 65]</sup>. Whereas measurement of  $\psi$ -angles is limited by the presence of the oxygen and nitrogen bound to the C'.

The local spatial arrangements of frequently found conformations in peptides and proteins are called secondary structure elements and can be estimated once the chemical shift assignment of backbone resonances is completed. With the primary effort of the group of Sykes, pioneered by K. Wüthrich, statistical lists were established to define random coil chemical shifts<sup>[66, 67]</sup>, also known as primary shifts. The chemical shift difference between the experimental value and its random coil value is called secondary shifts. This secondary shift information is used for identifying regions of secondary structure.

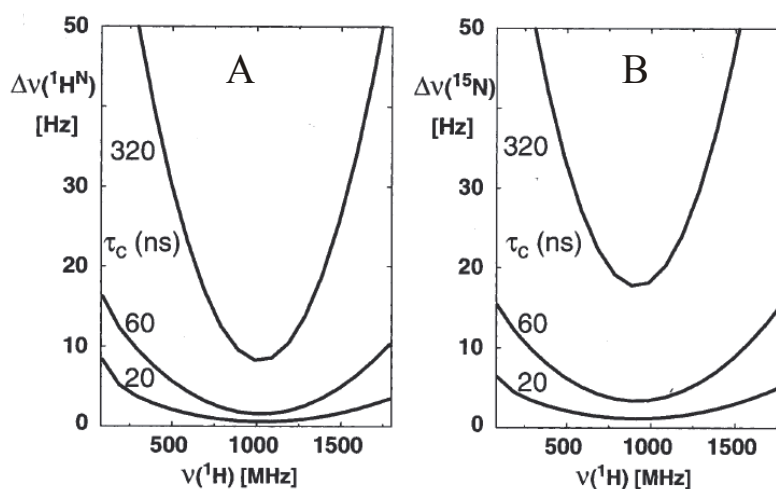
Solvent accessibility and hydrogen bonding can be characterized from hydrogen exchange rate measurements between labile protons (generally, backbone and side-chain amide protons) and the solvent (typically water)<sup>[68]</sup>. Hydrogen exchange rates can be, for example, measured with a MEXICO (Measurement of EXchange rates in Isotopically labeled COmpounds)



experiment<sup>[69]</sup>. H-bond information can be accomplished from a long-range HNCO experiment<sup>[70]</sup>.

#### 2.4.5. The TROSY Experiment

For large molecular systems, transverse relaxation *via* dipole-dipole coupling and chemical shift anisotropy leads to an overall increase in signal linewidth, and a corresponding decrease in spectral resolution (figure 2-6).



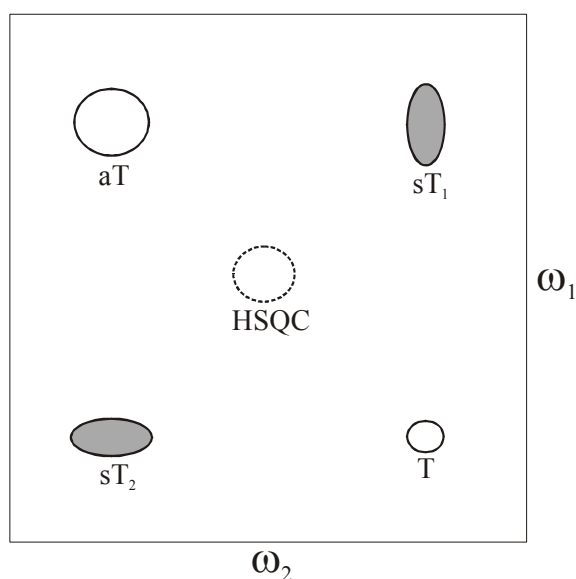
**Figure 2-6:** Frequency dependence from 100–1800 MHz of the full resonance line width at half height for amide groups in TROSY experiments calculated for three correlation times of  $\tau_c = 20, 60$  and  $320$  ns, which represent spherical proteins with molecular weights of 50, 150 and 800 kDa. (A)  $^1\text{HN}$  linewidth. (B)  $^{15}\text{N}$  linewidth. The calculation uses axial symmetric CSA tensor of  $^{15}\text{N} = 155$  ppm. and  $^1\text{HN} = 15$  ppm, and the angle between the principal tensor axis and the N–H bond was assumed to be  $15^\circ$  for  $^{15}\text{N}$  and  $10^\circ$  for  $^1\text{HN}$ ;  $l(\text{N–H}) = 0.104$  nm; effects of long-range dipole-dipole couplings with spins outside of the  $^{15}\text{N}$ - $^1\text{H}$  moiety were not considered. Figure reproduced from<sup>[71]</sup>.

It has been recognized that at very high magnetic field strengths, dipole-dipole (DD) and chemical shift anisotropy (CSA) interactions in a  $^{15}\text{N}$ - $^1\text{H}$  pair can be utilized to obtain sharp line widths for very large proteins or protein complexes. An important pulse sequence called TROSY (transverse relaxation optimized spectroscopy) has been developed<sup>[11, 72]</sup>. TROSY takes advantage of mutual cancellation of CSA and DD relaxation effects at high fields.

TROSY is basically a heteronuclear correlation experiment (particularly for  $^{15}\text{N}$ - $^1\text{H}$  spin pairs) in which the proton magnetization is first transferred to  $^{15}\text{N}$ , then evolves during  $t_1$  under

differential relaxation mechanisms of the  $^{15}\text{N}$  doublet due to CSA ( $^{15}\text{N}$ ) and dipole-dipole interaction ( $^{15}\text{N}$ - $^1\text{H}$ ). Magnetization is then transferred back to the proton prior to detection under differential line broadening of the proton doublet due to CSA ( $^1\text{H}$ ) and dipole-dipole interaction ( $^{15}\text{N}$ - $^1\text{H}$ ). In TROSY experiments, decoupling is not used, and J-coupled peaks are resolved.

When monitoring the  $^{15}\text{N}$  decoupled spectra, two peaks would be seen as there are two possible orientations of the bound hydrogen (*spin up* or *spin down* states). When the hydrogen nucleus is in the *spin up* state, the dipole-dipole coupling between the  $^{15}\text{N}$  and  $^1\text{H}$  will lead to a local  $^1\text{H}$  field which always has the same directionality as the CSA contribution. Conversely, when  $^1\text{H}$  is in the *spin down* state, the local  $^1\text{H}$  field always has directionality opposite that of the CSA contribution. This means that in the *spin down* state, the DD coupling of the system effectively reduces the chemical shift anisotropy. Since the chemical shift anisotropy is directly proportional to the square of the external magnetic field, it is possible to adjust the external field to a level at which the DD coupling and CSA exactly cancel each other (which occurs at 1.1 GHz).



**Figure 2-7:** Nomenclature and representation of the  $^{15}\text{N}$ - $^1\text{H}$  TROSY multiplet pattern. The slowest relaxing component, the TROSY peak, is marked with a T, the two semi-TROSY peaks in  $\omega_1$  ( $sT_1$ ) and  $\omega_2$  ( $sT_2$ ), respectively, as well as the so-called “anti-TROSY peak” labeled aT are also depicted in figure. In a decoupled HSQC, the central peak (ascribed as HSQC) appears as a superposition of fast and slow relaxing components, thus being prone to rather fast relaxation.

The resulting cross peak is a multiplet of four peaks, each having different width and relaxation rate in the  $\omega_1$  and  $\omega_2$  dimensions. In contrast, these four multiplets (arising from two different line widths for each N and H) for each amide proton are superimposed in the HSQC spectra due to decoupling in  $F_1$  and  $F_2$ . In TROSY spectra, among these four multiplets, only the one which is not affected by line broadening due to DD and CSA is selected by application of appropriate phase cycling. These four multiplets are shown in figure 2-7.

Several methodologies for obtaining TROSY spectra free of errors and artifacts have been developed in recent years<sup>[73, 74]</sup>. Use of TROSY elements in pulse programs for sequence specific assignment has now become part of routine NMR.

### 3. Residual Dipolar Couplings: Introduction and Theory

The Hamiltonian solution NMR of dipolar spins is mainly dominated by Zeeman, chemical shift and scalar couplings terms. CSA interactions are negligible at most working fields and dipolar interactions are averaged to zero in isotropic solution due to molecular tumbling. In contrast, solid state NMR shows large dipolar interactions, which are often larger than the average line width of the NMR resonance and thus it is practically very difficult to observe the resolution common in solution NMR. Techniques like Magic Angle Spinning (MAS)<sup>[75]</sup>, during which the sample is spun along the magic angle ( $54.7^\circ$ ), help to average out these interaction. For example,  $^1\text{H}$  resonances of polystyrene (which has intense dipolar interactions because of its rigidity) have line width of 25 kHz (under MAS at  $\sim 10$  kHz) whereas the static  $^1\text{H}$  spectrum of natural rubber (which has less intense dipolar interaction because of its mobility) shows partial resolution of the  $\text{CH}_2$  and  $\text{CH}_3$  groups, which are fully resolved under MAS even at a spinning speed of 500 Hz.

Since dipolar interactions are averaged out in solution NMR, spectral simplicity can be gained compared to the solid state NMR but dipolar interaction information is lost. Dipolar interactions are valuable as they are distance dependent and could thus provide restraints for structure calculation. Measurement of dipolar couplings in solution has therefore been attempted several times in the history of NMR.

Residual Dipolar Couplings (RDCs) are the dipolar couplings obtained in solution NMR by a tunable and tailored way, maintaining adequate spectral resolution. In this chapter, we will discuss the theoretical foundation necessary to understand concepts involved in the realm of RDCs (part of this work has been already published<sup>[14]</sup>).

#### 3.1. Historical Background and Development of RDCs

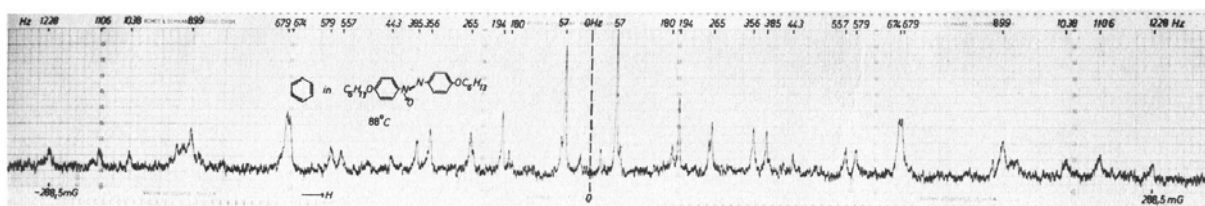
##### 3.1.1. First Observation of Dipolar Couplings in Solution

In 1963, the anisotropic dipolar interactions in high resolution NMR were reintroduced by Saupe and coworkers by dissolving benzene in nematic\* solvents (4,4'-bis(hexyloxy)azoxybenzene)<sup>[76]</sup>. The  $^1\text{H}$  spectrum of benzene no longer displayed a single peak, but rather, was a complex spectrum of more than 50 lines. At the same time, the

---

\*nematic: thread in Greek, initially used to describe rod-like solvent molecules.

resolution of the solute spectrum was retained, and any signals from the nematic solvent disappeared in the background which can be seen in figure 3-1. The intramolecular dipole-dipole interactions enhanced the complexity of the benzene spectrum, nevertheless, a high resolution spectrum was retained due to the reduction of intermolecular dipole-dipole interactions (compared to the solid state) by rapid translational diffusion. This opened a new era of liquid crystal NMR which compromises both high resolution NMR and solid state NMR [77].



**Figure 3-1:** Proton NMR of benzene in nematic solvent 4,4'-bis(hexyloxy)azobenzene. Proton spectrum of benzene consists of many resonances due to reintroduction of anisotropic dipolar interactions [76].

For the structural interpretation of residual dipolar couplings, it was necessary to interpret the average angular dependence of residual dipolar couplings given by the quantity  $\langle 3(\cos^2 \theta - 1)/2 \rangle$ , where  $\theta$  is the angle between the internuclear vector connecting the coupled nuclei and the external magnetic field  $B_0$ , and the angle brackets denote averaging due to molecular reorientation. To extract structural information from this equation one would require complete knowledge of the distribution function governing molecular orientation. Since this angular dependence is a second rank spherical harmonic, the relevant part of the probability distribution could be expressed as a linear combination of just the five elements of second rank spherical harmonics [78-80]. Hence, measurement of five or more suitably independent residual dipolar couplings in a known rigid element would permit extraction of the structural information. In the following sections, we will see that these five spherical harmonics are directly related to the five elements of the alignment tensor, and are the basis for the order matrix approach to extract structural information.

Though the theoretical foundation was laid on the early work of Saupe and Englert in 1964 [79], application of liquid crystal techniques for the measurement of anisotropic interactions in macromolecules remained challenging. Particularly, in larger molecular systems spectra became complex due to additional hundreds of resonances. The very first applications of residual dipolar couplings for structural analysis of macromolecules

materialized from the direct alignment of solute molecules at high magnetic fields<sup>[81, 82]</sup> and not from the alignment by liquid crystalline media.

### 3.1.2. *Alignment of Molecules by External Magnetic Field*

The external magnetic field  $B_0$  induces orientation to molecules which have high magnetic susceptibility anisotropies. The size of induced magnetic moments in such molecules, and therefore, the energy of interaction with the magnetic field, would vary with orientation and produce non-isotropic distributions. Lohman and MacLean<sup>[83]</sup> observed magnetic alignment for the first time in the form of quadrupolar splitting\* for  $^2\text{H}$  in the aligned benzene- $d_6$ . The observation of residual dipolar couplings under direct field-induced orientation awaited the technical developments for higher available fields, mainly because of the weak, non-cooperative nature of the orientation caused by the magnetic field. Orientation induced by the magnetic field leads to a dipolar splitting that scales quadratically with the field.

The first demonstration of measurable residual dipolar coupling came from Bothner-By and co-workers, where the paramagnetic system bis[*o*-toluyltris(pyrazolyl)borato]cobalt(II) ( $\text{Co}(\text{TTPB})_2$ ), was aligned in the magnetic field. The alignment achieved in this case was almost an order of magnitude larger than in diamagnetic systems<sup>[84]</sup>. Furthermore, they were able to measure quadrupolar and residual dipolar couplings in porphyrin and nucleic acid systems, where anisotropy in susceptibility is diamagnetic in origin<sup>[85, 86]</sup>.

The first observation of residual dipolar coupling by direct field induced orientation to a protein came after the availability of a  $^{15}\text{N}$  labeled protein and higher magnetic fields. The measured residual dipolar contributions to the scalar one bond  $^{15}\text{N}$ - $^1\text{H}$  couplings was only 2-5 Hz, even in a 750 MHz (*ca.* 17 T) spectrometer<sup>[87]</sup>. Nevertheless, their agreement with the values predicted from the X-ray derived geometries was sufficient to demonstrate structural utility in macromolecules.

### 3.1.3. *Alignment by External Alignment Media*

In practice, not many macromolecules have large magnetic anisotropies, making the level of alignment small and limiting the number of residual dipolar coupling measurements that can be made with reproducibility and low errors. This stumbling block was removed very recently

---

\* Quadrupolar splittings display the same  $(3(\cos^2 \theta - 1)/2)$  dependence as residual dipolar couplings, but are larger in magnitude.

by the use of a dilute liquid crystalline medium, where a ten-fold increase in macromolecular alignment (relative to the paramagnetic alignment) could be achieved without any sacrifice in the spectral resolution<sup>[12]</sup>. The medium used was a dilute ‘bicelle’ medium, which is based on an aqueous dispersion of lipid bilayer disks<sup>[88, 89]</sup>. This medium proved compatible with proteins and other biomolecules, and is amenable to adjustment for ideal levels of alignment. This discovery was a significant step that not only improved the compromise between alignment magnitude and spectral resolution, but also permitted measurement of residual dipolar couplings in a much broader range of systems. In their pioneering work, Bax and co-workers have shown that the residual dipolar coupling contributions to  $^{15}\text{N}$ - $^1\text{H}$  splittings measured in Ubiquitin were as large as 20 Hz, and could be measured with a precision of approximately 0.2 Hz. This approach of introducing RDCs has allowed the determination of an internuclear vector orientation with impressive accuracy, ranging between 0.5 and 5 degrees.

Based on these developments, residual dipolar couplings (RDC) have found a wide range of applications in high resolution NMR of biomolecules in the liquid state in recent years. Today, with the rapid development of the alignment media, any kind of macromolecule can be aligned irrespective of its surface and physical properties. In the next chapter, various alignment media and their utilities will be discussed.

### 3.2. The Concept of the Alignment Tensor

The next sections present an intuitive introduction to the alignment tensor and an elementary derivation of key equations. The fundamental question of how to calculate the expected residual dipolar coupling constant for a homonuclear (*e. g.*  $^1\text{H}$ - $^1\text{H}$ ) or heteronuclear (*e. g.*  $^{15}\text{N}$ - $^1\text{H}$ ) spin pair is discussed. This turns out to be a surprisingly simple calculation if one knows the orientation and the three principal components of the so-called *alignment tensor*. This *alignment tensor* is a key concept and the understanding of the physical meaning of the alignment tensor is crucial in understanding residual dipolar couplings.

Commonly found derivations for the *alignment tensor* use mathematically elegant, but not very intuitive approaches based on spherical harmonics, their addition theorems, Legendre polynomials, Wigner rotation matrices, and a confusing number of angles between various axes<sup>[13, 90]</sup>. These methods lead to difficulties in fully understanding the physical meaning of the alignment tensor.

The following development is a streamlined geometric approach, similar to the original derivation by Saupe<sup>[76, 79]</sup>, based on the Cartesian representation of vectors. Except for the most basic rules of matrix and vector multiplication, only elementary mathematics is needed

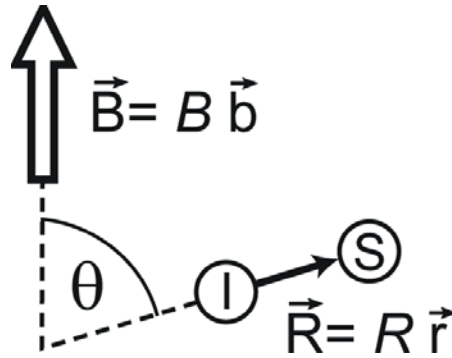
to derive the *alignment tensor*. Understanding of the *alignment tensor* is achieved in the later part using the explanation of the concept of the related *probability tensor*. Numerical examples and illustrating figures are used to convey the physical meaning of these tensors. Various expressions for the residual dipolar coupling constants commonly found in the literature are also derived from the presented key results.

### 3.2.1. Static Dipolar Coupling Hamiltonian

Let us consider two spins  $I$  and  $S$  with an internuclear vector  $\vec{R}$  (figure 3-2). This vector can be expressed in the form

$$\vec{R} = R \vec{r} = R \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} \quad [3-1],$$

where  $R$  is the distance between the two nuclei and  $\vec{r}$  is a unit vector pointing in the direction of  $\vec{R}$ .



**Figure 3-2:** Definition of the angle  $\theta$  between the internuclear vector  $\vec{R}$  (connecting spins  $I$  and  $S$ ) and the magnetic field vector  $\vec{B}$ . The unit vectors  $\vec{r}$  and  $\vec{b}$  point in the direction of  $\vec{R}$  and  $\vec{B}$ , respectively.

Similarly, the vector representing the external magnetic field  $\vec{B}$  can be expressed in the form

$$\vec{B} = B \vec{b} = B \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} \quad [3-2],$$

where  $B$  is the magnitude of the static magnetic field, and  $\vec{b}$  is a unit vector pointing in the direction of the magnetic field. In the lab frame  $(x^L, y^L, z^L)$ , where by convention the magnetic field points along the  $z^L$  axis, the (truncated) dipolar coupling Hamiltonian has the form <sup>[32]</sup>,



$$\mathcal{H}_D = 2\pi D \left\{ I_{z^L} S_{z^L} - \frac{1}{2} I_{x^L} S_{x^L} - \frac{1}{2} I_{y^L} S_{y^L} \right\} \quad [3-3].$$

If the spins I and S are heteronuclear, the second and third term in the bracket can be neglected, resulting in the simpler weak dipolar coupling Hamiltonian

$$\mathcal{H}_D = 2\pi D I_{z^L} S_{z^L} \quad [3-4],$$

(which has the same form as the weak heteronuclear J-coupling Hamiltonian). In both cases, the dipolar coupling constant (which in the weak coupling limit corresponds directly to the experimentally observed line splittings in units of Hz) <sup>[32]</sup> is:

$$D = \frac{\kappa}{R^3} \left( \cos^2 \theta - \frac{1}{3} \right) \quad [3-5],$$

where  $\theta$  is the angle between the internuclear vector and the magnetic field (figure 3-2).

The term,

$$\kappa = -\frac{3}{8\pi^2} \gamma_I \gamma_S \mu_0 \hbar \quad [3-6],$$

depends only on physical constants: the gyromagnetic ratios  $\gamma_I$  and  $\gamma_S$  of spin I and S respectively, the Planck constant  $\hbar = h/2\pi$ , and the permeability of vacuum  $\mu_0$  <sup>[32]</sup>. *E. g.*, for  $^1\text{H}$ - $^1\text{H}$ ,  $^{13}\text{C}$ - $^1\text{H}$  and  $^{15}\text{N}$ - $^1\text{H}$  spin pairs,  $\kappa = -360.3 \text{ kHz } \text{\AA}^3$ ,  $-90.6 \text{ kHz } \text{\AA}^3$  and  $36.5 \text{ kHz } \text{\AA}^3$ , respectively. The maximum possible value of  $\cos^2 \theta$  is 1 (for  $\theta = 0$  or  $\pi$ ), and hence, according to Eq. [3-5], the maximum possible dipolar coupling constant is

$$D_{\max} = \kappa / R^3 (1 - 1/3) = (2/3) \kappa / R^3 \quad [3-7],$$

which corresponds, *e. g.*, to 21.7 kHz for a  $^{15}\text{N}$ - $^1\text{H}$  spin pair with distance  $R = 1.04 \text{ \AA}$ .

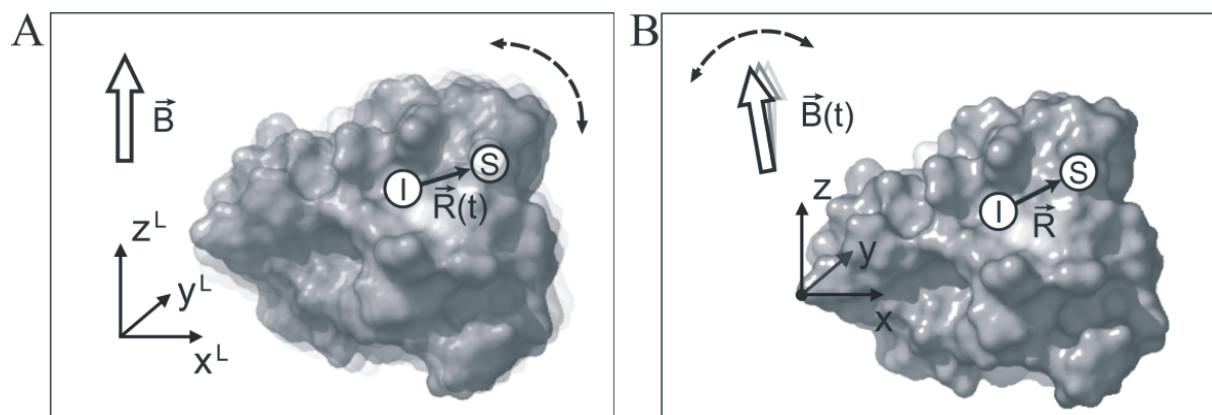
Remembering that the scalar product between two unit vectors is identical to the cosine of the angle  $\theta$  between the two vectors, the term  $\cos \theta$  (Eq. [3-5]) can always be expressed in the form

$$\cos \theta = \bar{\mathbf{b}}^T \bar{\mathbf{r}} \quad [3-8].$$

Here,  $\bar{\mathbf{b}}^T$  is a row vector (the transpose of the column vector  $\bar{\mathbf{b}}$ ) which allows us to write the scalar product of the two vectors as a usual matrix product between the 1x3 matrix  $\bar{\mathbf{b}}^T$  and the 3x1 matrix  $\bar{\mathbf{r}}$  (*vide infra*).

### 3.2.2. Time Dependent and Average Dipolar Coupling Hamiltonian

Now let us consider the two spins I and S to be part of a molecule in solution. The magnetic field vector  $\bar{\mathbf{B}}$  is constant (pointing along the  $z^L$  axis), in the laboratory frame, but the internuclear vector  $\bar{\mathbf{R}}$  is now time-dependent (figure 3-3 (A)).



**Figure 3-3:** Effect of molecular tumbling of a rigid molecule as seen (Panel A) from the lab frame of reference (with axes  $x^L$ ,  $y^L$ ,  $z^L$ ) and (Panel B) from an arbitrary molecular frame of reference (with axes  $x$ ,  $y$ ,  $z$ ). In the lab frame (Panel A), the magnetic field  $\vec{B}$  is constant and points by definition along the  $z^L$  axis, whereas the internuclear vector  $\vec{R}$  keeps changing its direction. In a molecular frame (Panel B), the situation is reversed: here, any given internuclear vector is constant, whereas the orientation of the magnetic field is time-dependent.

For simplicity, let us assume that the molecule is rigid (no internal dynamics and constant distance  $R$ ), such that the time-dependence of  $\vec{R}$  is solely due to the rotational tumbling motion of the molecule. Hence, the term  $\cos\theta$  (and as a result also the dipolar coupling constant  $D$  and the dipolar coupling Hamiltonian) is time-dependent. For proteins, the rotational correlation time is in the order of nanoseconds and on the time-scale of the NMR experiment, only the time-averaged dipolar Hamiltonian  $\overline{\mathcal{H}_D}$  gives rise to splittings in the spectrum (relaxation effects caused by the fluctuations of the dipolar Hamiltonian will not be considered here). The time-averaged dipolar coupling constant

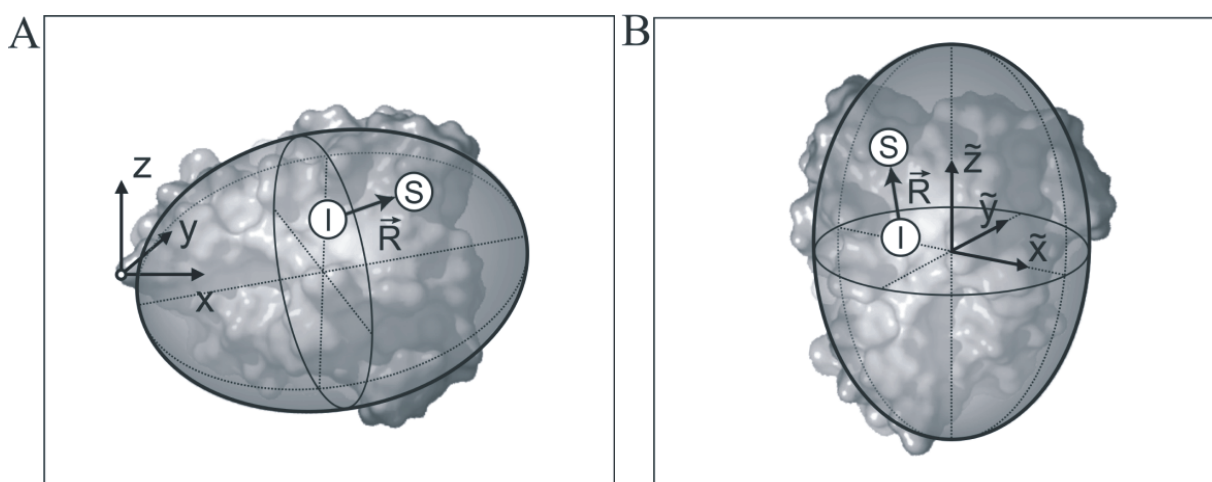
$$\overline{D} = \frac{\kappa}{R^3} \left( \overline{\cos^2 \theta} - \frac{1}{3} \right) \quad [3-9],$$

represents the so-called residual dipolar coupling constant, which depends on the average alignment of the molecule.

### 3.2.3. Outline of the Key Results

The goal of the further discussion is to derive a general approach for the calculation of  $\overline{D}$  for any pair of spins if the “alignment properties” of the molecule are known. Before going into the formal derivation, a brief outline of the steps and the final result is given. First, let us

move from the lab frame  $(x^L, y^L, z^L)$  (figure 3-3 (A)) to a frame of reference  $(x, y, z)$  that is fixed to the molecule. In this frame of reference, the term  $\overline{\cos^2 \theta}$  can be conveniently expressed with the help of a probability tensor  $P$ , which is a second order approximation of the orientational probability distribution of the direction of the external magnetic field in the molecule-fixed frame of reference<sup>[13, 32]</sup>. This probability tensor  $P$  can be represented by an ellipsoid (figure 3-4 (A)) with a fixed orientation in the chosen molecular frame  $(x, y, z)$ . The principal values  $P_{\tilde{x}}$ ,  $P_{\tilde{y}}$  and  $P_{\tilde{z}}$  of the probability tensor (*i.e.*, the lengths of the half axes of the probability ellipsoid) are the probabilities of finding the magnetic field along the corresponding principal axes of the probability ellipsoid, and hence  $P_{\tilde{x}} + P_{\tilde{y}} + P_{\tilde{z}} = 1$ .



**Figure 3-4:** The molecule, a given internuclear vector  $\vec{R}$  and the probability ellipsoid (a graphical representation of the probability tensor  $P$ , cf. Eq. [3-23]) are shown (Panel A) in an arbitrarily chosen molecular frame (cf. figure 3-3 (B)) and (Panel B) in the special coordinate system (with axes  $\tilde{x}$ ,  $\tilde{y}$ ,  $\tilde{z}$ ) defined by the principal axes of the probability ellipsoid.

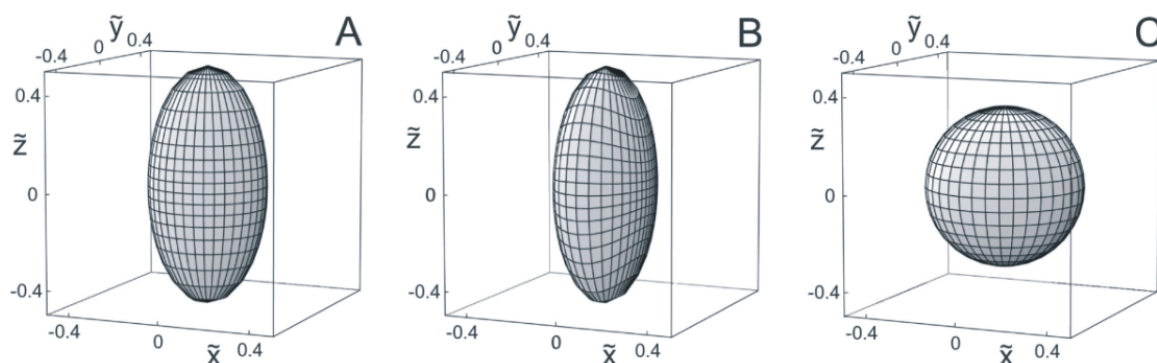
For example, for an isotropically reorienting molecule,  $P_{\tilde{x}} = P_{\tilde{y}} = P_{\tilde{z}} = 1/3$ , and the probability ellipsoid is reduced to a sphere (figure 3-5 (C)). On the other hand, if a molecule is fully aligned,  $P_{\tilde{x}} = P_{\tilde{y}} = 0$  and  $P_{\tilde{z}} = 1$  (by convention, the principal elements are ordered with increasing magnitude), *i.e.*, the probability tensor is reduced to a single line in the direction of the magnetic field.

In general, the principal axes of the probability ellipsoid define a special molecule-fixed axis system  $(\tilde{x}, \tilde{y}, \tilde{z})$ , in which the calculation of residual dipolar coupling constants is

especially simple (figure 3-4 (B)): If one knows the three Cartesian components  $r_{\tilde{x}}$ ,  $r_{\tilde{y}}$  and  $r_{\tilde{z}}$  of any given internuclear unit vector  $\tilde{r}$  in this principal axis system, the term  $\overline{\cos^2 \theta}$  in Eq. [3-8] is simply given by

$$\overline{\cos^2 \theta} = P_{\tilde{x}} r_{\tilde{x}}^2 + P_{\tilde{y}} r_{\tilde{y}}^2 + P_{\tilde{z}} r_{\tilde{z}}^2 \quad [3-10].$$

If this simple equation (derived below) is inserted into Eq. [3-9], the residual coupling constant can be predicted for any arbitrary spin pair in a molecule, as long as the orientation and principal values of the probability tensor are known.



**Figure 3-5:** Examples of three characteristic probability ellipsoids (graphical representations of the probability tensor  $P$ , cf. Eq. [3-23]) as seen from the principal axis system with axes  $\tilde{x}$ ,  $\tilde{y}$ ,  $\tilde{z}$  (cf. figure 3-4 (B)). Panel A shows an axially symmetric probability ellipsoid with  $P_{\tilde{x}} = P_{\tilde{y}} = 0.25$  and  $P_{\tilde{z}} = 0.5$  (Panel A). Panel B depicts a rhombic probability ellipsoid with  $P_{\tilde{x}} = 0.2$ ,  $P_{\tilde{y}} = 0.3$  and  $P_{\tilde{z}} = 0.5$ . Panel C shows an isotropic probability ellipsoid with  $P_{\tilde{x}} = P_{\tilde{y}} = P_{\tilde{z}} = 1/3$ .

With this key result, one can calculate everything and one could stop here, except that residual dipolar coupling constants are commonly not expressed in terms of the introduced probability tensor  $P$  (corresponding in general to a real symmetric 3x3 matrix with trace 1) but in terms of its traceless part (its “resolvent”)  $\mathbf{P} - 1/3 \mathbf{1}$ , which is called the alignment tensor  $\mathbf{A}$  [12].

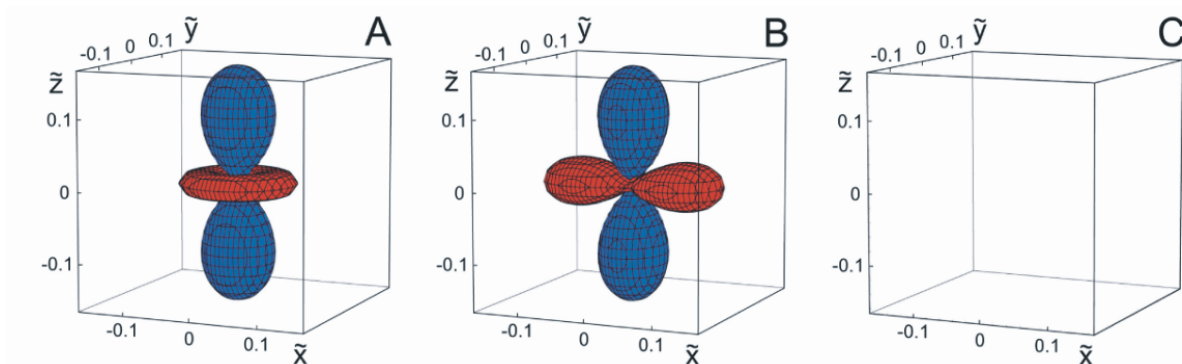
$$\mathbf{A} = \mathbf{P} - \frac{1}{3} \mathbf{1} \quad [3-11].$$

The three principal components  $A_{\tilde{x}}$ ,  $A_{\tilde{y}}$  and  $A_{\tilde{z}}$  of the alignment tensor  $\mathbf{A}$  are simply given by,

$$A_{\tilde{x}} = P_{\tilde{x}} - \frac{1}{3}, \quad A_{\tilde{y}} = P_{\tilde{y}} - \frac{1}{3} \quad \text{and} \quad A_{\tilde{z}} = P_{\tilde{z}} - \frac{1}{3} \quad [3-12],$$

and the principal axes of A and P are identical.

Note that in contrast to the probability tensor P (figure 3-4 and figure 3-5), the alignment tensor A cannot be represented as an ellipsoid, because one or two of the principal components  $A_{\tilde{x}}$ ,  $A_{\tilde{y}}$ , and  $A_{\tilde{z}}$  of the alignment tensor are negative if any of the three components is nonzero due to  $A_{\tilde{x}} + A_{\tilde{y}} + A_{\tilde{z}} = 0$ . Alternative graphical representations of the alignment tensor are shown in figure 3-6 and figure 3-7 (*vide infra*).



**Figure 3-6:** Graphical representations of the alignment tensors (Panel A) which correspond to the three probability tensors shown in figure 3-5 (A-C). The principal components of the alignment tensor are (A)  $A_{\tilde{x}} = A_{\tilde{y}} = 0.25 - 1/3 = -1/12$ ,  $A_{\tilde{z}} = 0.5 - 1/3 = -1/6$ , (B)  $A_{\tilde{x}} = 0.2 - 1/3 = -2/15$ ,  $A_{\tilde{y}} = 0.3 - 1/3 = -1/30$ ,  $A_{\tilde{z}} = 0.5 - 1/3 = 1/6$  and (C)  $A_{\tilde{x}} = A_{\tilde{y}} = A_{\tilde{z}} = 1/3 - 1/3 = 0$ . The plots show the surfaces where  $(\bar{\mathbf{r}}^T \mathbf{A} \bar{\mathbf{r}}) / R^3 = 1 \text{ \AA}^{-3}$  (blue surface) or  $-1 \text{ \AA}^{-3}$  (red surface) if the  $\tilde{x}$ ,  $\tilde{y}$  and  $\tilde{z}$  axes are labeled in units of  $\text{\AA}$ .

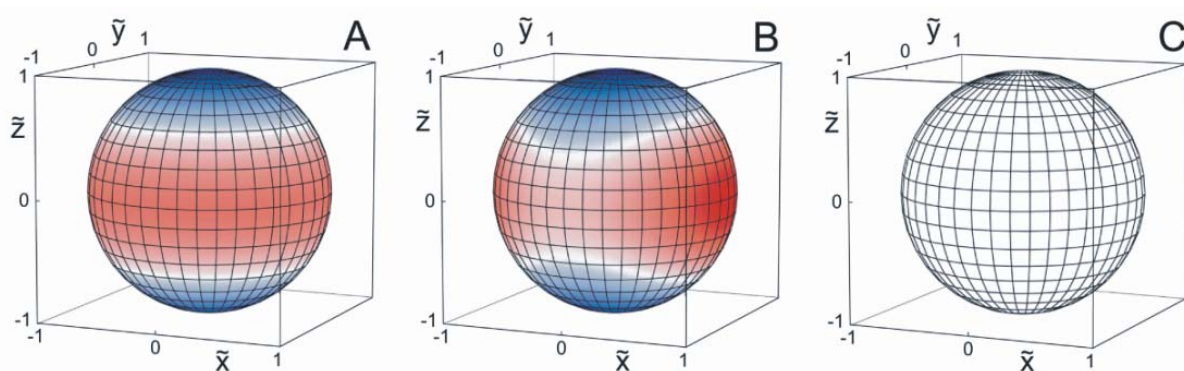
In terms of the principal components of the alignment tensor, the term  $(\overline{\cos^2 \theta} - 1/3)$  in the equation for the residual dipolar coupling constant (Eq. [3-9]) can be expressed as

$$\left( \overline{\cos^2 \theta} - \frac{1}{3} \right) = A_{\tilde{x}} r_{\tilde{x}}^2 + A_{\tilde{y}} r_{\tilde{y}}^2 + A_{\tilde{z}} r_{\tilde{z}}^2 \quad [3-13].$$

If this equation is inserted into Eq. [3-9], it is again possible to predict the residual coupling constant for any arbitrary spin pair in a molecule, provided that the orientation and principal values of the alignment tensor are known.

Conversely, the alignment tensor A (or the probability tensor P) can be determined if a sufficient number of experimental dipolar coupling constants are measured for a given molecule<sup>[91]</sup>. As will be shown below, the alignment tensor A (and the probability tensor P) is

characterized by five independent parameters. Therefore, at least five dipolar coupling constants need to be measured in order to determine the five unknown parameters<sup>[91]</sup>. In many cases, it is also possible to accurately predict the alignment tensor  $A$ <sup>[92]</sup> or the probability tensor  $P$  for a given molecule in a given liquid crystalline solvent, and hence to predict the expected dipolar coupling constants for a proposed molecular structure from first principles.



**Figure 3-7:** For the three cases shown in figure 3-5 and figure 3-6 with (A)  $A_{\bar{x}} = A_{\bar{y}} = -1/12$ ,  $A_{\bar{z}} = -1/6$ , (B)  $A_{\bar{x}} = -2/15$ ,  $A_{\bar{y}} = -1/30$ ,  $A_{\bar{z}} = 1/6$  and (C)  $A_{\bar{x}} = A_{\bar{y}} = A_{\bar{z}} = 0$  the scaling factor  $(\overline{\cos^2 \theta} - 1/3)$  is color-coded on a unit sphere as a function of the orientation of the internuclear vector  $\bar{R}$  (white: vanishing scaling factor, blue: positive scaling factor, red: negative scaling factor).

### 3.3. Derivation of the Probability and Alignment Tensors

The definition of  $\cos \theta$  via the scalar product of the unit vectors  $\bar{b}$  and  $\bar{r}$  (cf. Eq. [3-8]) is valid in any frame of reference. Hence,  $\cos \theta$  can be expressed in the molecular frame as a function of the components of the unit vectors  $\bar{b}$  and  $\bar{r}$ , which point in the (varying) direction of the magnetic field  $\bar{B}$  and of the (constant) internuclear vector  $\bar{R}$  (cf. Eq [3-1] and [3-2], respectively):

$$\begin{aligned} \cos \theta &= \bar{b}^T \cdot \bar{r} = \begin{pmatrix} b_x(t) & b_y(t) & b_z(t) \end{pmatrix} \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} \\ &= b_x(t)r_x + b_y(t)r_y + b_z(t)r_z \end{aligned} \quad [3-14],$$

where  $(t)$  represents time-dependent form (in the molecule-fixed frame) and

$$\begin{aligned}
\cos^2 \theta &= (b_x(t)r_x + b_y(t)r_y + b_z(t)r_z)^2 \\
&= b_x^2(t)r_x^2 + b_x(t)b_y(t)r_x r_y + b_x(t)b_z(t)r_x r_z \\
&\quad + b_y(t)b_x(t)r_y r_x + b_y^2(t)r_y^2 + b_y(t)b_z(t)r_y r_z \\
&\quad + b_z(t)b_x(t)r_z r_x + b_z(t)b_y(t)r_z r_y + b_z^2(t)r_z^2
\end{aligned} \tag{3-15}$$

Note that Eq. [3-15] can also be expressed in the form

$$\cos^2 \theta = \begin{pmatrix} r_x & r_y & r_z \end{pmatrix} \begin{pmatrix} b_x^2(t) & b_x(t)b_y(t) & b_x(t)b_z(t) \\ b_x(t)b_y(t) & b_y^2(t) & b_y(t)b_z(t) \\ b_x(t)b_z(t) & b_y(t)b_z(t) & b_z^2(t) \end{pmatrix} \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} \tag{3-16}$$

Hence, the time average of  $\cos^2 \theta$  is given by

$$\overline{\cos^2 \theta} = \begin{pmatrix} r_x & r_y & r_z \end{pmatrix} \begin{pmatrix} \overline{b_x^2(t)} & \overline{b_x(t)b_y(t)} & \overline{b_x(t)b_z(t)} \\ \overline{b_x(t)b_y(t)} & \overline{b_y^2(t)} & \overline{b_y(t)b_z(t)} \\ \overline{b_x(t)b_z(t)} & \overline{b_y(t)b_z(t)} & \overline{b_z^2(t)} \end{pmatrix} \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} = \bar{\mathbf{r}}^T \mathbf{P} \bar{\mathbf{r}} \tag{3-17}$$

The matrix

$$\mathbf{P} = \begin{pmatrix} \overline{b_x^2(t)} & \overline{b_x(t)b_y(t)} & \overline{b_x(t)b_z(t)} \\ \overline{b_x(t)b_y(t)} & \overline{b_y^2(t)} & \overline{b_y(t)b_z(t)} \\ \overline{b_x(t)b_z(t)} & \overline{b_y(t)b_z(t)} & \overline{b_z^2(t)} \end{pmatrix} \tag{3-18}$$

is called the probability matrix<sup>[14]</sup>. For a known probability matrix  $\mathbf{P}$ , the residual dipolar coupling constant (Eq. [3-9]) is given by

$$\bar{D} = \frac{\kappa}{R^3} \left( \bar{\mathbf{r}}^T \mathbf{P} \bar{\mathbf{r}} - \frac{1}{3} \right) \tag{3-19}$$

The matrix  $\mathbf{P}$  is real, symmetric, and has a trace of 1 because

$$\text{tr}\{\mathbf{P}\} = P_{xx} + P_{yy} + P_{zz} = \overline{b_x^2(t)} + \overline{b_y^2(t)} + \overline{b_z^2(t)} = \overline{(b_x^2(t) + b_y^2(t) + b_z^2(t))} = 1 \tag{3-20}$$

since by definition,  $\bar{\mathbf{b}}$  is a unit vector, and hence,  $b_x^2(t) + b_y^2(t) + b_z^2(t) = 1$  for all times  $t$ .

Therefore,  $\mathbf{P}$  is fully specified by only five independent parameters. The matrix  $\mathbf{P}$  can be represented graphically as an ellipsoid (figure 3-4 and figure 3-5). The three principal axes  $\tilde{x}$ ,  $\tilde{y}$  and  $\tilde{z}$  of this ellipsoid are defined by the three eigenvectors of the matrix  $\mathbf{P}$  and the lengths of the three half axes are defined by the eigenvalues  $P_{\tilde{x}}$ ,  $P_{\tilde{y}}$  and  $P_{\tilde{z}}$  (figure 3-4 (A)).

In the special frame of reference defined by this principal axis system (figure 3-4 (B)), the matrix  $\mathbf{P}$  is diagonal:

$$\mathbf{P} = \begin{pmatrix} P_{\tilde{x}} & 0 & 0 \\ 0 & P_{\tilde{y}} & 0 \\ 0 & 0 & P_{\tilde{z}} \end{pmatrix} \quad [3-21].$$

In this case the eigenvalues (principal values)  $P_{\tilde{x}} = \overline{b_x^2}$ ,  $P_{\tilde{y}} = \overline{b_y^2}$  and  $P_{\tilde{z}} = \overline{b_z^2}$  are the probabilities to find the magnetic field along the principal axes  $\tilde{x}$ ,  $\tilde{y}$  and  $\tilde{z}$ , respectively. Therefore let us call P simply the probability tensor. (Rigorously, P corresponds to the sum of the zero and second order term of a spherical harmonics expansion of the probability distribution function describing the orientation of a reference vector relative to a rigid body [80, 93]).

In the principal axis system, Eq. [3-19] for the calculation of the residual dipolar coupling reduces simply to

$$\overline{D} = \frac{\kappa}{R^3} \left( P_{\tilde{x}} r_{\tilde{x}}^2 + P_{\tilde{y}} r_{\tilde{y}}^2 + P_{\tilde{z}} r_{\tilde{z}}^2 - \frac{1}{3} \right) \quad [3-22].$$

For example, in the static case,  $\vec{b} = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix}$  is constant, and hence,

$$\mathbf{P} = \begin{pmatrix} b_x^2 & b_x b_y & b_x b_z \\ b_x b_y & b_y^2 & b_y b_z \\ b_x b_z & b_y b_z & b_z^2 \end{pmatrix}. \quad [3-23].$$

The matrix has a much simpler form in the principal axis frame ( $\tilde{x}$ ,  $\tilde{y}$ ,  $\tilde{z}$ ) where the  $\tilde{z}$  axis is parallel to the vector  $\vec{b}$ . In this reference frame,

$$\vec{b} = \begin{pmatrix} b_{\tilde{x}} \\ b_{\tilde{y}} \\ b_{\tilde{z}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } \mathbf{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad [3-24].$$

In this case, the probability ellipsoid is reduced to a line along the  $\tilde{z}$  axis and the dipolar coupling constant is

$$\overline{D} = D = \frac{\kappa}{R^3} \left( r_{\tilde{z}}^2 - \frac{1}{3} \right) \quad [3-25].$$

For a completely isotropically reorienting molecule, the averages  $\overline{b_x(t)b_y(t)}$ ,  $\overline{b_x(t)b_z(t)}$ ,  $\overline{b_y(t)b_z(t)}$  are zero, and  $P_{\tilde{x}} = P_{\tilde{y}} = P_{\tilde{z}} = 1/3$ , *i.e.*, the probability matrix



$$\mathbf{P} = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix} \quad [3-26],$$

is diagonal in any molecule-fixed frame of reference. Hence, there is an equal probability of 1/3 of the magnetic field direction pointing along all three axes of reference. The corresponding probability ellipsoid is a sphere with radius 1/3 (figure 3-5 (C) ), and the residual dipolar coupling constant is

$$\bar{D} = \frac{\kappa}{R^3} \left( \frac{1}{3} (r_{\tilde{x}}^2 + r_{\tilde{y}}^2 + r_{\tilde{z}}^2) - \frac{1}{3} \right) = 0 \quad [3-27].$$

Figure 3-5 A shows an example of an axially symmetric probability ellipsoid with the principal values  $P_{\tilde{x}} = P_{\tilde{y}} = 0.25$  and  $P_{\tilde{z}} = 0.5$ . Figure 3-5 B shows an example without axial symmetry where  $P_{\tilde{x}} = 0.2$ ,  $P_{\tilde{y}} = 0.3$  and  $P_{\tilde{z}} = 0.5$ . Note that the lack of axial symmetry simply means that there are two different probabilities  $P_{\tilde{x}} \neq P_{\tilde{y}}$  of the magnetic field pointing along the principal axes  $\tilde{x}$  and  $\tilde{y}$  of the molecule-fixed probability tensor. However, this does by no means imply that in the lab frame there are different probabilities for the molecule to be aligned along the  $x^L$  or  $y^L$  direction. For example in the case shown in figure 3-5 B,  $P_{\tilde{x}} = 0.2$ ,  $P_{\tilde{y}} = 0.3$  and  $P_{\tilde{z}} = 0.5$  are the probabilities that the  $\tilde{x}$ ,  $\tilde{y}$  and  $\tilde{z}$  axes are aligned parallel to  $B_0$ .

In the NMR literature, it is not customary to consider the probability tensor  $\mathbf{P}$  (which can be nicely depicted as an ellipsoid), but to use its traceless part which is called the alignment tensor

$$\mathbf{A} = \mathbf{P} - \frac{1}{3} \mathbf{1} \quad [3-28].$$

If one multiplies  $\mathbf{A}$  from the left with the unit row vector  $\bar{\mathbf{r}}^T$  and from the right with the column vector  $\bar{\mathbf{r}}$  and using Eq. [3-17] and Eq. [3-28], one gets,

$$\begin{aligned} \bar{\mathbf{r}}^T \mathbf{A} \bar{\mathbf{r}} &= \bar{\mathbf{r}}^T \left( \mathbf{P} - \frac{1}{3} \mathbf{1} \right) \bar{\mathbf{r}} \\ &= \bar{\mathbf{r}}^T \mathbf{P} \bar{\mathbf{r}} - \frac{1}{3} \bar{\mathbf{r}}^T \bar{\mathbf{r}} \\ &= \overline{\cos^2 \theta} - \frac{1}{3} \end{aligned} \quad [3-29],$$

which can also be used to calculate the residual dipolar coupling constant in Eq. [3-19]:

$$\bar{D} = \frac{\kappa}{R^3} (\bar{\mathbf{r}}^T \mathbf{A} \bar{\mathbf{r}}) \quad [3-30].$$

$\mathbf{P}$  and  $\mathbf{A}$  have the same principal axis system ( $\tilde{x}$ ,  $\tilde{y}$ ,  $\tilde{z}$ ) (except for a possible reordering of the axis labels if the convention is used that  $|P_{\tilde{x}}| \leq |P_{\tilde{y}}| \leq |P_{\tilde{z}}|$  and  $|A_{\tilde{x}}| \leq |A_{\tilde{y}}| \leq |A_{\tilde{z}}|$ ), and the principal values are related by

$$A_{\tilde{x}} = P_{\tilde{x}} - \frac{1}{3}, \quad A_{\tilde{y}} = P_{\tilde{y}} - \frac{1}{3} \quad \text{and} \quad A_{\tilde{z}} = P_{\tilde{z}} - \frac{1}{3} \quad [3-31],$$

with  $A_{\tilde{x}} + A_{\tilde{y}} + A_{\tilde{z}} = 0$ .

In the principal axis system

$$\left( \overline{\cos^2 \theta} - \frac{1}{3} \right) = A_{\tilde{x}} r_{\tilde{x}}^2 + A_{\tilde{y}} r_{\tilde{y}}^2 + A_{\tilde{z}} r_{\tilde{z}}^2 \quad [3-32],$$

and hence, the residual dipolar coupling constant is given by

$$\bar{D} = \frac{\kappa}{R^3} (A_{\tilde{x}} r_{\tilde{x}}^2 + A_{\tilde{y}} r_{\tilde{y}}^2 + A_{\tilde{z}} r_{\tilde{z}}^2) \quad [3-33].$$

The alignment tensor cannot be represented as an ellipsoid, because at least one of the principal values is always negative if  $\mathbf{A} \neq 0$ . In figure 3-6, a graphical representation of the  $\mathbf{A}$  tensors is shown which correspond to the  $\mathbf{P}$  tensors shown in figure 3-5. The plots show the surfaces where the term  $\left| \frac{\bar{\mathbf{r}}^T \mathbf{A} \bar{\mathbf{r}}}{R^3} \right|$  is constant. Hence, if spin I is assumed to be located at the origin, the plots show the possible locations of spin S for which the residual dipolar coupling constant has the same magnitude. For the case of an isotropically reorienting molecule (spherical probability tensor), the residual dipolar coupling is always zero, and no such surface exists.

The dependence of the scaling factor  $(\overline{\cos^2 \theta} - 1/3)$  on the orientation of the internuclear vector is sometimes shown by the color of a unit sphere. For the three cases shown in figure 3-6 and, the corresponding color coded surface representations of the alignment tensors are shown in figure 3-7. The color represents the scaling factor (white: 0, blue: positive, red: negative) of a residual dipolar coupling constant if spin I is located at the origin and spin S is moved over the surface, *i.e.*, assuming a constant internuclear distance.

For example, in the axially symmetric case shown in figure 3-7 A with  $A_{\tilde{x}} = A_{\tilde{y}} = -1/12$  and

$A_{\tilde{z}} = 1/6$ , the scaling factor  $(\overline{\cos^2 \theta} - 1/3)$  is zero if the  $\tilde{z}$ -component of the internuclear vector is  $r_{\tilde{z}} = \sqrt{1/3}$ , which is a straightforward result if Eq. [3-32] is set to zero and using  $r_{\tilde{x}}^2 + r_{\tilde{y}}^2 = 1 - r_{\tilde{z}}^2$ . This corresponds to an angle of  $\vartheta = \arccos \sqrt{1/3} = 54.74^\circ$  (the magic angle)

between the internuclear vector and the  $\tilde{z}$ -axis. For the case shown in figure 3-7 (B) with  $A_{\tilde{x}} = -2/15$ ,  $A_{\tilde{y}} = -1/30$  and  $A_{\tilde{z}} = 1/6$ , the polar angle  $\vartheta$ , when the scaling factor is zero, depends also on the azimuthal angle  $\varphi$  between the  $\tilde{x}$ -axis and the projection of  $\bar{r}$  on the  $\tilde{x}/\tilde{y}$  plane. For example, in the  $\tilde{x}/\tilde{z}$  plane, the scaling factor is zero if  $r_{\tilde{z}} = 2/3$  (corresponding to  $\vartheta = \arccos(2/3) = 48.19^\circ$ ), and in the  $\tilde{y}/\tilde{z}$  plane, the scaling factor is zero if  $r_{\tilde{z}} = \sqrt{1/6}$  ( $\vartheta = \arccos\sqrt{1/6} = 65.91^\circ$ ). In the isotropic case shown in figure 3-7 (C), the scaling factor  $(\overline{\cos^2 \theta} - 1/3)$  is zero for all orientations of the internuclear vector  $\bar{R}$ .

### 3.4. Expression of Dipolar Coupling Constant in Various Formats

In this section, the key equations (Eq. [3-22] and Eq. [3-33]) for the calculation of the residual dipolar coupling constant  $\bar{D}$  are re-expressed in various forms commonly found in the literature. If the unit vector  $\bar{r}$  is defined in terms of the polar coordinates  $\vartheta$  and  $\varphi$  in the principal axis system of the alignment tensor **A**, then

$$\bar{r} = \begin{pmatrix} r_{\tilde{x}} \\ r_{\tilde{y}} \\ r_{\tilde{z}} \end{pmatrix} = \begin{pmatrix} \sin \vartheta \cos \varphi \\ \sin \vartheta \sin \varphi \\ \cos \vartheta \end{pmatrix} \quad [3-34],$$

and hence (according to Eq. [3-32]):

$$\left( \overline{\cos^2 \theta} - \frac{1}{3} \right) = A_{\tilde{x}} \sin^2 \vartheta \cos^2 \varphi + A_{\tilde{y}} \sin^2 \vartheta \sin^2 \varphi + A_{\tilde{z}} \cos^2 \vartheta \quad [3-35].$$

This can be simplified by noting that  $\cos^2 \varphi = (1 + \cos 2\varphi)/2$  and  $\sin^2 \varphi = (1 - \cos 2\varphi)/2$ :

$$\begin{aligned} \left( \overline{\cos^2 \theta} - \frac{1}{3} \right) &= \frac{A_{\tilde{x}}}{2} \sin^2 \vartheta + \frac{A_{\tilde{x}}}{2} \sin^2 \vartheta \cos 2\varphi + \frac{A_{\tilde{y}}}{2} \sin^2 \vartheta - \frac{A_{\tilde{y}}}{2} \sin^2 \vartheta \cos 2\varphi + A_{\tilde{z}} \cos^2 \vartheta \\ &= \frac{A_{\tilde{x}} + A_{\tilde{y}}}{2} \sin^2 \vartheta + \frac{A_{\tilde{x}} - A_{\tilde{y}}}{2} \sin^2 \vartheta \cos 2\varphi + A_{\tilde{z}} \cos^2 \vartheta \end{aligned} \quad [3-36].$$

Since **A** is a traceless matrix,  $A_{\tilde{x}} + A_{\tilde{y}} = -A_{\tilde{z}}$ , and Eq. [3-36] can be rewritten as

$$\left( \overline{\cos^2 \theta} - \frac{1}{3} \right) = A_{\tilde{z}} \left( \cos^2 \vartheta - \frac{\sin^2 \vartheta}{2} \right) + \frac{A_{\tilde{x}} - A_{\tilde{y}}}{2} \sin^2 \vartheta \cos 2\varphi \quad [3-37].$$

The pre-factor of  $A_{\tilde{z}}$  can be further simplified by using the relation  $\sin^2 \vartheta = 1 - \cos^2 \vartheta$ :

$$\begin{aligned} \cos^2 \vartheta - \frac{\sin^2 \vartheta}{2} &= \cos^2 \vartheta - \frac{(1 - \cos^2 \vartheta)}{2} \\ &= \frac{1}{2} (3 \cos^2 \vartheta - 1) \end{aligned} \quad [3-38].$$

Thus,

$$\left(\overline{\cos^2 \theta} - \frac{1}{3}\right) = \frac{A_z}{2} (3 \cos^2 \vartheta - 1) + \frac{A_x - A_y}{2} \sin^2 \vartheta \cos 2\varphi \quad [3-39].$$

Eq. [3-39] can alternatively be expressed in terms of the principal values  $S_x$ ,  $S_y$  and  $S_z$  of the Saupe matrix (or order matrix)  $S$ , which is simply the alignment matrix  $A$  scaled by a factor of  $3/2$ , if the optical axis of the liquid crystal is collinear with the direction of the magnetic field <sup>[76, 79]</sup>:

$$S = \frac{3}{2} A \quad [3-40].$$

Hence,

$$\left(\overline{\cos^2 \theta} - \frac{1}{3}\right) = \frac{1}{3} \left\{ S_z (3 \cos^2 \vartheta - 1) + (S_x - S_y) \sin^2 \vartheta \cos 2\varphi \right\} \quad [3-41].$$

Often, the axial component  $A_a$  of the alignment tensor is defined as <sup>[12]</sup>

$$A_a = \frac{3}{2} A_z = S_z \quad [3-42],$$

and the rhombic component  $A_r$  of the alignment tensor is defined as

$$A_r = A_x - A_y = \frac{2}{3} (S_x - S_y) \quad [3-43].$$

With these definitions, Eqs. [3-39] and [3-41] can be expressed as

$$\left(\overline{\cos^2 \theta} - \frac{1}{3}\right) = \frac{1}{3} \left\{ A_a (3 \cos^2 \vartheta - 1) + \frac{3}{2} A_r \sin^2 \vartheta \cos 2\varphi \right\} \quad [3-44],$$

which in turn is often written as

$$\begin{aligned} \left(\overline{\cos^2 \theta} - \frac{1}{3}\right) &= \frac{A_a}{3} \left\{ (3 \cos^2 \vartheta - 1) + \frac{3}{2} R \sin^2 \vartheta \cos 2\varphi \right\} \\ &= \frac{A_a}{3} \left\{ (3 \cos^2 \vartheta - 1) + \eta \sin^2 \vartheta \cos 2\varphi \right\} \end{aligned} \quad [3-45],$$

where

$$R = \frac{A_r}{A_a} \quad [3-46],$$

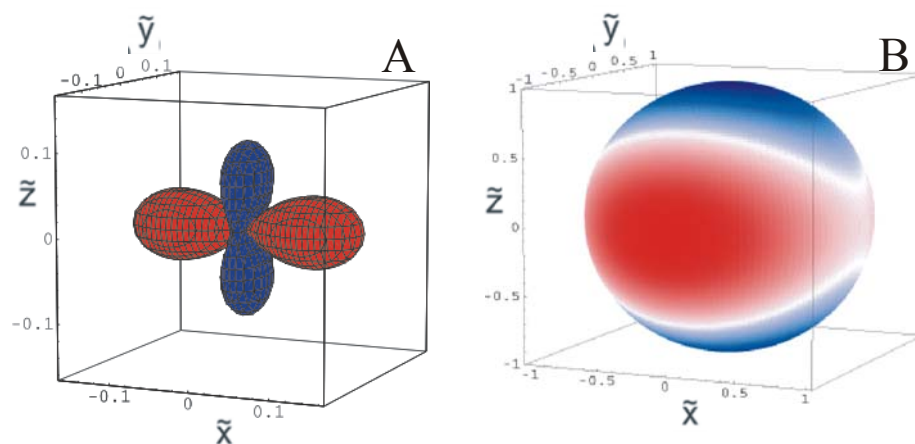
is called the rhombicity of the alignment tensor and

$$\eta = \frac{A_x - A_y}{A_z} = \frac{S_x - S_y}{S_z} = \frac{3}{2} R \quad [3-47],$$

is called the asymmetry parameter which describes the deviation from axially symmetric ordering <sup>[13]</sup>.

### 3.4.1. Representation of a Highly Rhombic Tensor

An alignment tensor is called highly rhombic if the main contributions come from the  $A_{\tilde{y}}$  and  $A_{\tilde{z}}$  terms. For example, if an alignment tensor has three components such as  $A_{\tilde{x}} = 0.1$ ,  $A_{\tilde{y}} = 0.4$  and  $A_{\tilde{z}} = 0.5$ , then it is a highly rhombic tensor. Such a tensor is presented in figure 3-8.



**Figure 3-8:** Graphical representations of the highly rhombic alignment tensors (A) which correspond to the probability tensor shown in figure 3-5 (B). The principal components of the alignment tensor are  $A_{\tilde{x}} = 0.1 - 1/3 = -7/30$ ,  $A_{\tilde{y}} = 0.4 - 1/3 = -3/15$ ,  $A_{\tilde{z}} = 0.5 - 1/3 = 1/6$ .

In panel (B), the scaling factor  $(\overline{\cos^2 \theta} - 1/3)$  is color-coded on a unit sphere as a function of the orientation of the internuclear vector  $\tilde{R}$  (white: vanishing scaling factor, blue: positive scaling factor, red: negative scaling factor).

### 3.5. Alignment Tensor in the Presence of Internal Motion

The discussion up to this point was made under the assumption of a rigid molecule tumbling in solution. In the presence of internal motions the derivation of residual dipolar couplings becomes more complicated [13, 20, 94]. Provided the alignment process is not affected by intramolecular motion, the analysis is relatively straightforward. If the internal motion of the internuclear vector  $\tilde{r}$  is axially symmetric with respect to the average orientation  $\tilde{r}_{av}$ , the dipolar coupling expected for this average orientation is scaled by a factor  $\lambda$ , which is identical to a generalized order parameter  $S$  ( $0 \leq S \leq 1$ ) [94]. The latter corresponds mathematically to the spin relaxation order parameter [20, 21], but exhibits a sensitivity to

motions extending to the millisecond time scale<sup>[13, 94]</sup>. This leads to the following equation of the residual dipolar coupling constant:

$$\bar{D} = S \frac{\kappa}{3} \frac{A_a}{R^3} \left\{ (3 \cos^2 \vartheta - 1) + \eta \sin^2 \vartheta \cos 2\varphi \right\} \quad [3-48].$$

This expression is often rewritten using the maximum dipolar coupling  $D_{\max} = (2/3)\kappa/R^3$  (cf. Eq. [3-7]) or the so called magnitude of the residual dipolar coupling tensor<sup>[90]</sup>:

$$D_a = D_{\max} A_a / 2 \quad [3-49].$$

Therefore,

$$\begin{aligned} \bar{D} &= S D_a \left\{ (3 \cos^2 \vartheta - 1) + \eta \sin^2 \vartheta \cos 2\varphi \right\} \\ &= S \frac{D_{\max}}{2} A_a \left\{ (3 \cos^2 \vartheta - 1) + \eta \sin^2 \vartheta \cos 2\varphi \right\} \\ &= S D_{\max} A_a \left\{ P_2(\cos \vartheta) + \frac{\eta}{2} \sin^2 \vartheta \cos 2\varphi \right\} \end{aligned} \quad [3-50],$$

where  $P_2(x) = (3 \cos^2 x - 1)/2$  is the second-order Legendre polynomial.

### 3.6. Generalised Degree of Order

To conclude, the concepts of the generalized degree of order (GDO) of a given alignment tensor  $\mathbf{A}$ <sup>[95]</sup> and the generalized angle between two different alignment tensors  $\mathbf{A}^{(1)}$  and  $\mathbf{A}^{(2)}$ <sup>[96]</sup> will be introduced using the results obtained in the previous section.

In complete analogy to the scalar product between two real vectors, the scalar product between two real matrices (e. g. two alignment matrices  $\mathbf{A}^{(1)}$  and  $\mathbf{A}^{(2)}$ ) is defined as

$$\langle \mathbf{A}^{(1)} | \mathbf{A}^{(2)} \rangle = \sum_{i,j} A_{ij}^{(1)} A_{ij}^{(2)} \quad [3-51],$$

and the norm  $|\mathbf{A}|$  of the real matrix  $\mathbf{A}$  is given by

$$|\mathbf{A}| = \sqrt{\langle \mathbf{A} | \mathbf{A} \rangle} = \sqrt{\sum_{i,j} A_{ij}^2} \quad [3-52].$$

The maximum order is found for the static case, where the probability tensor  $\mathbf{P}_{\max}$  is given by Eq. [3-24] in the principal axis system. The corresponding maximum alignment tensor  $\mathbf{A}_{\max} = \mathbf{P}_{\max} - 1/3 \mathbf{1}$  has the form

$$\mathbf{A}_{\max} = \begin{pmatrix} -1/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 2/3 \end{pmatrix} \quad [3-53].$$

The norm of  $\mathbf{A}_{\max}$  is given by

$$|\mathbf{A}_{\max}| = \sqrt{\frac{1}{9} + \frac{1}{9} + \frac{4}{9}} = \sqrt{\frac{2}{3}} \quad [3-54].$$

The generalized degree of order (GDO) of a given order matrix  $\mathbf{A}$  can be defined as

$$\text{GDO} = \frac{|\mathbf{A}|}{|\mathbf{A}_{\max}|} = \sqrt{\frac{3}{2}} |\mathbf{A}| \quad [3-55].$$

In terms of the Saupe matrix  $\mathbf{S} = 3/2 \mathbf{A}$  (*cf.* Eq. [3-40]), this can be written as <sup>[13, 95]</sup>,

$$\text{GDO} = \sqrt{\frac{2}{3}} |\mathbf{S}| \quad [3-56].$$

In literature, the symbol “ $\mathcal{G}$ ” is often used for the GDO but will not be used here in order to avoid confusion with the polar angle  $\mathcal{G}$  defined in Eq. [3-34].

The GDO is independent of the molecular-fixed frame, in which the alignment tensor  $\mathbf{A}$  is expressed. In the principal axis system only the diagonal elements of  $\mathbf{A}$  are nonzero and Eq. [3-55] simplifies to

$$\text{GDO} = \sqrt{\frac{3}{2}} \sqrt{A_x^2 + A_y^2 + A_z^2} \quad [3-57].$$

For axially symmetric alignment tensors ( $A_x = A_y = -A_z/2$ ) this simplifies further to <sup>[95]</sup>:

$$\begin{aligned} \text{GDO} &= \sqrt{\frac{3}{2} \left( \frac{1}{4} A_x^2 + \frac{1}{4} A_x^2 + A_z^2 \right)} = \frac{3}{2} \sqrt{A_z^2} \\ &= \frac{3}{2} |A_z| = |S_z| \end{aligned} \quad [3-58].$$

With the help of the scalar product, a generalized angle  $\beta$  between two alignment tensors  $\mathbf{A}^{(1)}$  and  $\mathbf{A}^{(2)}$  can be defined, which corresponds *e. g.* to two different alignment media.

If the matrix representations of  $\mathbf{A}^{(1)}$  and  $\mathbf{A}^{(2)}$  are given in a common molecular frame of reference, the cosine of the generalized angle  $\beta$  between these alignment tensors can be defined as their normalized scalar product <sup>[96]</sup>:

$$\cos \beta = \frac{\langle \mathbf{A}^{(1)} | \mathbf{A}^{(2)} \rangle}{|\mathbf{A}^{(1)}| |\mathbf{A}^{(2)}|} \quad [3-59].$$

### 3.7. Conclusion

A non-isotropic orientational distribution of a molecule can be created by creating anisotropy in the solution either by magnetic field or by addition of external alignment media. This gives rise to non-zero averaged dipolar coupling. For using these couplings as structural constraint, familiarity with the concept of the alignment tensor is necessary.

An intuitive introduction to the alignment tensor and an elementary derivation of key equations was accomplished in this chapter. Vital concepts like the probability tensor and alignment tensor were discussed. Various formats, often used in the literature, were also derived in a simple approach based on the Cartesian representation of vectors. This approach was extended to derive dipolar coupling constant equation in the time dependent case as well as in the presence of two alignment media.



## **4. Practical Aspects of Residual Dipolar Couplings: Subdomain Orientation in VAT-N**

Frontier work in the biomolecular NMR spectroscopy, in recent years, was mainly done in the area of finding/optimizing alignment media, developing practical methods for measuring and analyzing RDCs, and utilizing RDCs for subsequent structure calculations. This chapter contains some of these developments and demonstrates an application of RDCs, showing their ability to precisely define the subdomain orientation in a multi-domain protein (VAT-N: 20.5 kDa).

### **4.1. Alignment Media**

The choice of the alignment medium is always critical before actually starting RDC measurements. It mainly depends upon various factors (mostly physical properties) such as pH and temperature compatibility, surface charges, stability, solubility, affinity to the protein under study *etc.* Some of the alignment media and their properties are listed in table 4-1. This information can be helpful in particular to choose one of these alignment media suited for biomolecules under consideration.

Amongst them, bicelles (composed of DMPC-DHPC), phages (bacteriophage Pf1) and polyacrylamide gels have fetched more attention and were used more intensively compared to other alignment media because of their efficacy and availability. Therefore, a discussion related to these alignment media will follow.

**Table 4-1:** Media used to align molecules and measure residual dipolar coupling.

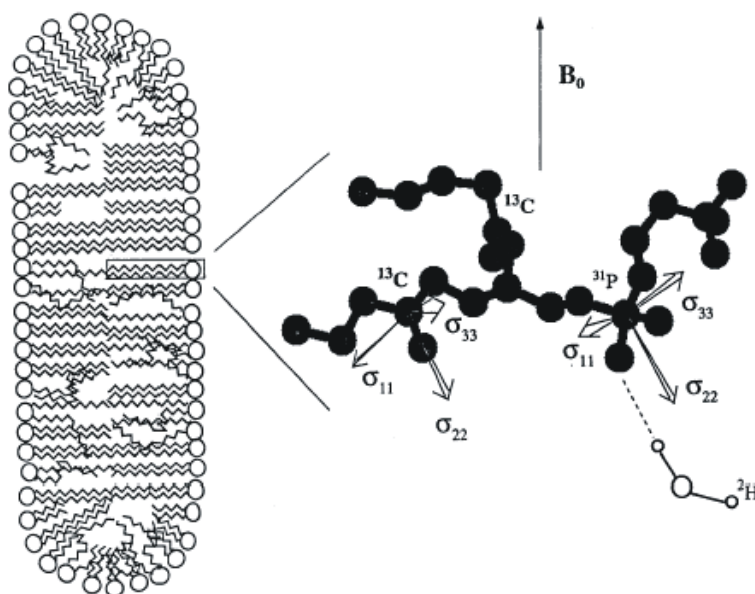
Medium	Orientation/shape/Temp.[°C]	Features	Major Applications	Ref.
DMPC:DHPC	Perpendicular/disc/27–45	Other lipids can be substituted	Proteins, nucleic acids, carbohydrates	[12, 97]
DMPC:DHPC:CTAB	Perpendicular/disc/27–42	For positively charged proteins	Proteins, nucleic acids, carbohydrates	[98]
DMPC:DHPC:SDS	Perpendicular/disc/27–42	For negatively charged proteins	Proteins, nucleic acids, carbohydrates	[98]
DMPC:DHPC:DMPX	Perpendicular/disc/35–40	For negatively charged proteins	Membrane peptide	[99]
DMPC:CHAPSO	Perpendicular/disc/30–40	Zwitterions	Proteins, glycolipids	[100]
DIODPC:CHAPSO	Perpendicular/disc/10–55	Stable over wide range of pH (pH 1.0 till pH 6.0), only acidic conditions, hydrolysis resistant due to ether linkages	Proteins	[101]
Pfl Phage	Parallel/rod-like/5–60	Very easy to work, wide temperature and concentration range	Proteins, nucleic acids, carbohydrates	[102, 103]
Purple membrane	Parallel/disc-like/< 70	Wide range of temperature	Proteins, peptides, low concentration	[96]
DLPC-CHAPSO	Perpendicular/disc/7–50	Wide range of temperature	Proteins	[104]
DMPC:DHPC + Ln <sup>3+</sup>	Parallel/ disc/35–90	Changes direction of the orientation	Proteins	[105]

Medium	Orientation/shape/Temp.[°C]	Features	Major Applications	Ref.
DBPC:DHPC	Parallel/disc/8–40	Biphenyl group	Carbohydrates	[106]
Rod shaped viruses	Parallel/rod-like/5–60	For hydrophobic patched proteins, wide temperature and concentration range	Proteins, nucleic acids, carbohydrates	[107]
Strained Polyacrylamide gel	Mechanical/gel/5–45	Easy recovery of macromolecule but difficult to align	Proteins	[108]
<i>N</i> -alkyl-poly(ethylene glycol)/ <i>n</i> -alkyl alcohol + glucopone/ <i>n</i> -hexanol mixture	Perpendicular/lamellar/0–40	Uncharged alignment medium, no effect of pH change, salt concentration and high protein concentration, stable, inert	Proteins, nucleic acids	[109]
Charged Polyacrylamide gel (Acrylamide/acrylate copolymer)	Mechanical/gel/5–45	Charged acrylate results different alignment tensor than Polyacrylamide gel	Proteins	[110]
Cetylpyridinium bromide/ <i>n</i> -hexanol/sodium bromide	Parallel/lamellar/0–70	Tolerant to buffer concentration, temperature and protein concentration, robust and versatile, sensitive to salt concentration	Proteins, carbohydrates	[111]
Cellulose crystallite	Perpendicular/??/37	Readily producible, no interaction with protein, stable for solution conditions	Proteins.	[112]
Filamentous bacteriophage fd	Parallel/rod-like/5–60	pH sensitive, pH modulates alignment tensor	Proteins, nucleic acids	[113]

Medium	Orientation/shape/Temp.[°C]	Features	Major Applications	Ref.
Polyacrylamide-stabilized Pfl phage	Mechanical/gel/??	Robust	Proteins, nucleic acids	[114]
Vanadium pentoxide or mineral liquid crystal	Parallel/ribbons/20-??	pH < 3, negative charge	Carbohydrates	[115]
Polystyrene gel	Mechanical/gel/solvent dependent	Robust, easily prepared, inert over wide range of conditions	Peptides, organic molecules	[116]

#### 4.1.1. DMPC-DHPC Bicelles

Dimyristoylphosphatidylcholine (DMPC) and dihexanoylphosphatidylcholine (DHPC) lipids, when mixed together in an aqueous solution, in 3:1 respective molar proportion, forms disk shaped objects of an average thickness of 41 Å (figure 4-1).



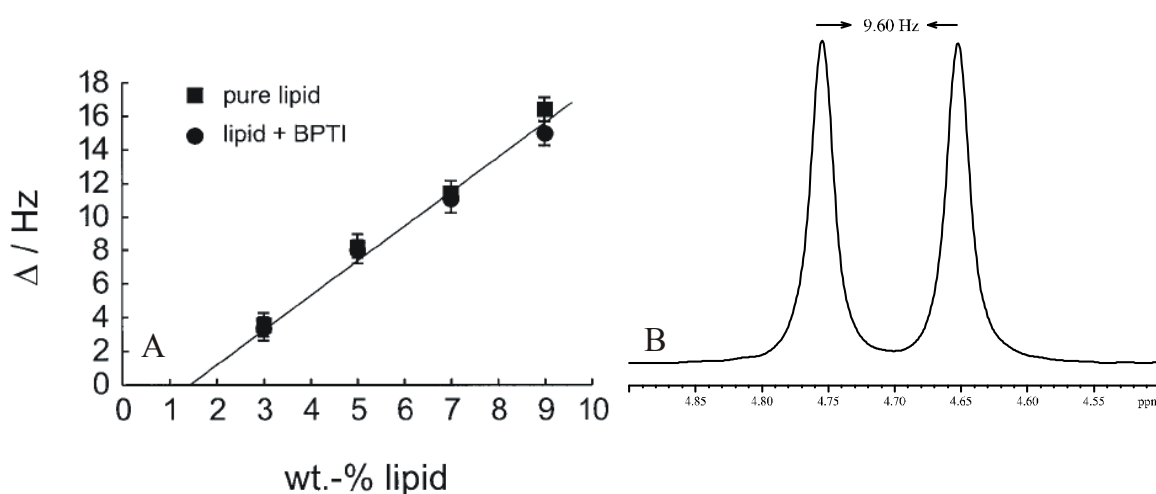
**Figure 4-1:** The disk shaped assemblies composed of DMPC and DHPC form bicelles. Open circles denote the phospho-diester backbone, while long chain is denoted by criss-cross line representing carbons (DMPC: 14 carbons, DHPC: 6 carbons). DMPC makes up the bulk of the plane of the disc, whereas DHPC stabilizes the edges. An enlarged structure of DMPC is shown along with the  $^{13}\text{C}$  and  $^{31}\text{P}$  CSA tensors and their orientation with respect to the external magnetic field  $B_0$ . DMPC and DHPC have the same phosphate backbone with different length of aliphatic chain and therefore DMPC is more lipophilic than DHPC.

Formation of disks is very similar to micelle formation except that these disks have two layers and therefore are called “bicelles” (**B**ilayered **micelles**). These disks align themselves, when placed in an external magnetic field, along the direction of  $B_0$  and adopt a lamellar liquid crystalline phase. Aligned bicelles cause hindrance to the isotropic Brownian motion of the solvent and the solute molecule, making the motional averaging anisotropic.

Deuterium, a quadrupolar nucleus, gives rise to a doublet pattern in an anisotropic medium.  $^2\text{H}$  magnetization under quadrupolar coupling oscillates harmonically with the precession under Zeeman or CSA. As a consequence, one component of the time signal is zero for the

time signal evolving under pure quadrupolar interactions. This can be explained in terms of two counter-rotating vectors, which would correspond to the two lines (transitions). This is interpreted as a quadrupolar splitting of  $^2\text{H}$  in an anisotropic medium <sup>[117]</sup>.

The lipid concentration and the deuterium splitting are in linear relationship above a threshold concentration of 1.5 wt % (figure 4-2 A) <sup>[118]</sup>. Therefore, the lipid concentration in the aligned solution and bicelle formation can be monitored by the quadrupolar splitting of the solvent  $^2\text{H}$  signal (figure 4-2 (B)) which is used for achieving the field-frequency lock condition and often present in 10% concentration.



**Figure 4-2:** The spectrum of  $^2\text{H}$  shows a quadrupolar splitting (in the form of a doublet) in an anisotropic medium. The anisotropy was created by the addition of bicelles that align in the external magnetic field. (A) shows the dependence  $^2\text{H}$  quadrupolar splitting in  $\text{D}_2\text{O}$  on bicelle wt % <sup>[118]</sup>. (B)  $^2\text{H}$  quadrupolar splitting in  $\text{D}_2\text{O}$  observed in the bicelles solutions prepared in our laboratory and which were used for further experimental work (cf. text). The presence of two well resolved and equally intense signals (a doublet) suggests that the sample is homogenous.

Neutral behaviour of the above mentioned phospholipids over a wide pH range makes them applicable to both positively and negatively charged biomolecules. However, electrostatic interactions between protein and the bicelles can be tuned to some degree by the addition of small amounts (10 % of DHPC) of charged amphiphiles. This includes positively charged CTAB (cetyltrimethyl-ammonium bromide) <sup>[98]</sup>, negatively charged SDS (sodium dodecyl sulfate) <sup>[98]</sup>, and negatively charged DMPX (3,7-dimethyl-1-propargylxanthine) <sup>[99]</sup>.

Addition of low amounts of the bicelles to an isotropic protein sample does not affect the translational diffusion coefficient of protein <sup>[119]</sup> unless some affinity between the bicelles and protein exists.

Bicelles are known to undergo hydrolysis within 2-3 weeks of preparation (at pH 6.0-7.0) because of their phospho-diester backbone. Therefore, it is necessary to prepare them freshly.

#### 4.1.2. Experimental Procedure for the Preparation of Bicelles

Here is a procedure which was frequently used in our laboratory for the bicelle preparation (15% w/v) in an Eppendorf tube (*Eppendorf AG, Hamburg, Germany*). DMPC, DHPC and CTAB (*Avanti Polar Lipids Inc., Alabaster, AL, USA*) were obtained and kept at 243 K. A buffer solution was prepared from 80 mM potassium phosphate in H<sub>2</sub>O at pH 5.9 with 120 mM NaCl, and 5 mM NaN<sub>3</sub> and will be referred to as "buffer" here afterwards. DMPC (77 mg, 90 μmol) was suspended in 200 μL of buffer. DHPC (17 mg, 30 μmol), in 1/3 molar proportion of DMPC, was dissolved in 100 μL buffer. CTAB (1.7 mg, 4 μm) was dissolved with 50 μL of buffer. All the suspensions and solutions were vortexed independently. DMPC was still suspended in 200 μL due to its high lipophilicity whereas DHPC and CTAB were easily soluble. The DHPC aqueous solution was then added to the DMPC suspension and the empty eppendorf was washed with 100 μL buffer solution twice. It should be noted here that the bicelle formation did not took place when DMPC suspension was added to the aqueous DHPC solution. This mixture was vortexed for 3 min. and kept on ice for 5 min. This was repeated 3 times. Subsequently, the CTAB solution was added to it and the empty eppendorf was washed with the remaining 50 μL of buffer. This solution was allowed to stay at 310 K for half an hour, after vortexing for a minute. Soon after, the solution-suspension was vortexed for 1 minute and kept on ice for 15 min. This cycle of warming and cooling was repeated until the solution become homogenous and obtained a viscous phase above 297 K.

Soon after, bicelles were kept at 276 K for 18 hrs. The standing duration of 18 hrs is critical for bicelle formation. Bicelles did not form completely in the magnet at 6 and 12 hrs of standing time which resulted in an isotropic <sup>2</sup>H signal along with quadrupolar splitted <sup>2</sup>H resonance.

The NMR sample was prepared in a shigemi tube (*Shigemi Co., Ltd, Tokyo, Japan*), by mixing 100 μL bicelle stock, 180 μL protein solution and 20 μL D<sub>2</sub>O. The shigemi tube was transferred from the ice bath into the preheated (313 K) magnet such that the sample experiences a fast temperature transition. This quick transition is necessary for DMPC to

undergo a phase transition at 298 K which facilitates formation of bicelles. A slow phase transition does not allow proper bicelle formation.

#### 4.1.3. Filamentous Phage Pfl

Bacteriophage Pfl have been used most extensively for the biomolecule alignment<sup>[102]</sup> mainly because it is very easy to use and can be directly added to the sample. It gains advantage over bicelles as no meticulous experimental preparation is needed.

It consists of a single stranded circular DNA genome packaged in coat protein (*ca.* 1:1 ratio) which forms a rod (*ca.* 60 Å in diameter and *ca.* 20,000 Å in length). This rod orients with its long axis parallel to the field. Phages are extraordinarily stable to different conditions in the solution, and unlike the bicelles, the ordered phase exists over a wide temperature range (278 -350 K). Moreover, the phage fully aligns over a very wide range of phage concentration. Pfl appears particularly well suited for studies of nucleic acids. Since Pfl is negatively charged at physiological pH (pI = 4.0), negatively charged nucleic acid molecules will not bind to the phage particle, thereby preventing unfavorably high levels of alignment or sample aggregation<sup>[102]</sup>. The high negative charge and inability to vary the charge makes phages less applicable to highly positively charged macromolecules. However, in a systematic study, low concentrations (*ca.* 2-4 mg/mL) of phage were shown to align positively charged proteins<sup>[120]</sup>.

#### 4.1.4. Polyacrylamide Gel

Cross-linked polyacrylamide swells in an aqueous solution to form a gel. These gels are elastic, neutral, hydrophilic and chemically inert. Cross-linking provides cavities in which macromolecules can diffuse. Staining these cavities (*i.e.*, the gel) vertically or horizontally creates restricted motion and thus anisotropy<sup>[108, 121]</sup>. This property can be exploited to use these gels for measuring RDCs.

Initially, polymerization of the gel is carried out in the presence of a cross-linking agent (*e. g.* *N,N'*-methylenebisacrylamide), an initiator and in the absence of protein. The gel can be later dried for storage purpose. At the time of use, the gel is allowed to swell in the solution containing protein. However, this approach suffers largely from slow diffusion of the protein into the tiny cavities of the gel. Another approach utilizes polymerization of an acrylamide sample additionally in the presence of the protein sample. The presence of the protein in the same solution during polymerization helps avoiding the problem of possible slow diffusion of the protein into tiny cavities.



The staining, in either case, can be produced by using the piston from the shigemi tube or by allowing the gel to expand along the vertical axis. A commercial device for inserting partially swollen gel into an NMR tube is meanwhile available <sup>[122]</sup>.

## 4.2. NMR Experiments for Measuring RDCs

The NMR methods for measuring coupling constants can be divided into two general categories: frequency resolved methods where separation of peak centers is measured in a frequency domain <sup>[123]</sup>, and intensity based (J-modulated) experiments, where the coupling is extracted from the resonance intensity <sup>[64, 124-126]</sup>. The principle underlying quantitative J-type experiments is to pass the observed signal through a period in which the intensity is modulated by a known function of the spin-spin coupling. As the modulation by the coupling assumes a sinusoidal form, at least 6-10 points (*i.e.*, experiments) are needed so as to acquire reasonably good data, which makes J-modulated coupling measurement a time demanding technique. The frequency resolved and intensity based methods are somewhat complementary because of different sources of systematic error.

Three major approaches based on a frequency resolved method are discussed below. Nevertheless, the attainable accuracy and precision, with the use of any of these methods of measurement, highly depends on signal to noise ratio, line shape, number of additional passive couplings, and the method employed for extracting peak positions. It should also be noted that this section treats only measurement of <sup>15</sup>N-<sup>1</sup>H one-bond scalar and dipolar coupling. Methodologies for measuring other couplings can be looked up elsewhere <sup>[127-129]</sup>.

### 4.2.1. Coupled HSQC

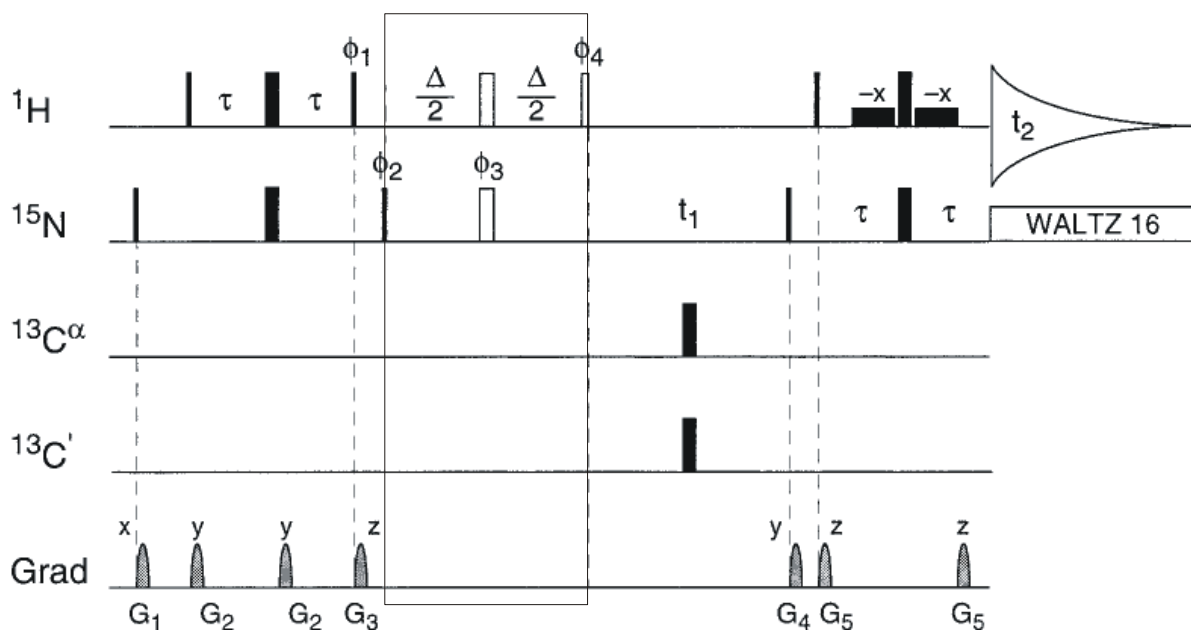
A simple method to record a heteronuclear coupling is by using a modified HSQC sequence (*cf. section 2.4.2*), where the spin decoupling pulse is omitted during the  $t_1$  period. This allows evolution of the coupling during the  $t_1$  period and J or J±D can be measured as the *in-phase* splitting in the indirect F<sub>1</sub> dimension after spectral processing. The measurement of splittings in the indirect dimension (<sup>15</sup>N) is straightforward and often preferred over measurement in the direct dimension (<sup>1</sup>H) because of longer <sup>15</sup>N spin-spin relaxation (T<sub>2</sub>). However, this approach leads to spectral crowding even with moderately medium sized proteins (< 9-10 kDa).

### 4.2.2. <sup>15</sup>N-<sup>1</sup>H-IPAP-HSQC

Many techniques have been introduced to avoid the spectral crowding problem from the coupled HSQC, where separate collection of each component of the doublet in two different

spectra is done, *e. g.* spin-state selective excitation ( $S^3E$ ) experiments<sup>[130-133]</sup>. Conversely, the clean separation of doublet components using these techniques is often quite sensitive to the size of the one-bond couplings<sup>[134]</sup> which is variable particularly in the presence of additional contributions from dipolar couplings.

An approach that is less sensitive to the variations in the couplings has been introduced to achieve spin-state separated spectra. This method is based on the collection of two spectra: in one of them, the coupling evolves *in-phase* (IP) and in the other the coupling evolves *anti-phase* (AP)<sup>[134]</sup>. Addition and subtraction of these spectra yields two new spectra, each containing one of the doublet components. A pulse sequence of  $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC is shown in figure 4-3 (a working pulse program in the BRUKER pulse program format is given in section 9.1.1). This pulse sequence is designed specifically for the measurement of directly bonded  $^{15}\text{N}$ - $^1\text{H}$  couplings in doubly labeled proteins<sup>[134]</sup>.



**Figure 4-3:**  $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC experiment can be used to measure  $^1J_{\pm D_{\text{NH}}}$  couplings<sup>[134]</sup>. Narrow and wide pulses correspond to  $90^\circ$  and  $180^\circ$  pulses respectively. The sequence element " $\Delta/2$ - $180^\circ(^{15}\text{N}-^1\text{H})$ - $\Delta/2$ - $90^\circ\phi_4$ " (box) is only used for generating the anti-phase spectrum (AP) and is omitted for generating the in-phase (IP) spectrum. Phase cycling:  $\phi_1 = -y, y$ ;  $\phi_2 = 2(x), 2(-x)$  for IP;  $\phi_2 = 2(-y), 2(y)$  for AP;  $\phi_3 = 4(x), 4(y), 4(-x), 4(-y)$ ;  $\phi_4 = 8(x), 8(-x)$ ; Receiver =  $x, 2(-x), x$  for IP; Receiver =  $x, 2(-x), x, -x, 2(x), -x$  for AP. Quadrature detection in the  $t_1$  dimension is obtained by altering  $\phi_2$  (IP) or  $\phi_2$  and  $\phi_3$  simultaneously (AP) in the States-TPPI manner;  $\Delta = 5.3$  ms.

The generation of the *in-phase* sub-spectra is simple and exactly identical to the regular  $^1\text{H}$ -coupled  $^{15}\text{N}$ - $^1\text{H}$  HSQC which generates *in-phase* doublets in the  $F_1$  dimension.

For a  $^{15}\text{N}$  spin ( $S$ ) at an angular offset frequency  $\Omega$ , coupled to its amide proton ( $I$ ), *anti-phase*  $S$  spin magnetization,  $2I_z S_y$ , is generated at the end of the first INEPT transfer which later evolves according to,

$$\begin{aligned} 2I_z S_y &\xrightarrow{t_1} \cos(\Omega t_1) \cos(\pi J_{NH} t_1) 2I_z S_y - \sin(\Omega t_1) \cos(\pi J_{NH} t_1) 2I_z S_x + (\text{IP}) \\ &\xrightarrow{90_x(I,S)} -\cos(\Omega t_1) \cos(\pi J_{NH} t_1) 2I_y S_z + (\text{MQ}) \\ &\xrightarrow{(2J_{HN})^{-1}} \cos(\Omega t_1) \cos(\pi J_{NH} t_1) I_x \end{aligned} \quad [4-1],$$

where (IP) refers to the in-phase term present at the end of the  $t_1$  evolution period, which is not converted into observable I-spin magnetization, and (MQ) is unobservable two-spin coherence.

In order to obtain the *anti-phase* component of the coupling, it is necessary to introduce a refocusing period prior to  $t_1$  evolution. This is accomplished using an " $\Delta/2$ - $180^\circ(I,S)$ - $\Delta/2$ - $90^\circ\phi_4$ " sequence element,

$$2I_z S_y \xrightarrow{\Delta/2-180_x(I,S)-\Delta/2 90_x(I)} -\sin(\pi J_{NH} \Delta) 2S_y + \cos(\pi J_{NH} \Delta) 2I_z S_y \quad [4-2].$$

The cosine term in the above equation is cancelled by phase cycling, which is not followed by the receiver. The  $S_y$  magnetization evolves into *anti-phase* magnetization during  $t_1$ ,

$$\begin{aligned} -\sin(\pi J_{NH} \Delta) 2S_y &\xrightarrow{t_1} -\cos(\omega t_1) \sin(\pi J_{NH} t_1) \sin(\pi J_{NH} \Delta) 2I_z S_x \\ &\quad + \sin(\omega t_1) \sin(\pi J_{NH} t_1) \sin(\pi J_{NH} \Delta) 2I_z S_y + (\text{IP}) \end{aligned} \quad [4-3],$$

where  $I_z S_x$  is now modulated *anti-phase* by the coupling. The reverse INEPT transfers magnetization back to protons and yields observable I magnetization, which is modulated by  $\sin(\pi J_{NH} t_1) e^{-i(\omega t_1)}$ .

After Fourier transformation, addition and subtraction of the two signals yields individual spectra for each component of the doublet<sup>[134]</sup>. The IPAP method has also been implemented in a variety of triple resonance NMR experiments for the detection of other couplings, *e.g.* N-CO<sup>[134]</sup>.

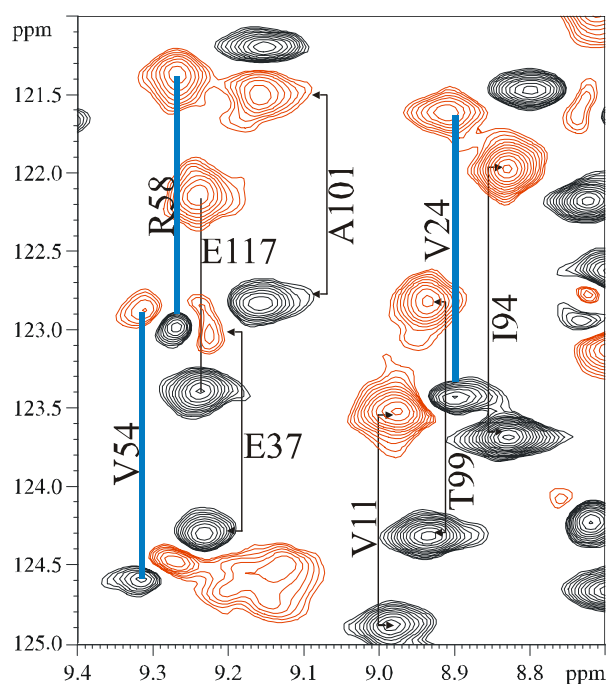
#### 4.2.3. TROSY-SemiTROSY (Tr-SmTr)

Spin-state selective excitations are utilizing in TROSY experiment (*cf. section 2.4.5*). Appropriate use of the phase cycling for the selection of TROSY or semiTROSY signals can be achieved in two different experiments, leading to selection of either signal of the splitting<sup>[135]</sup>. Both spectra can be analyzed to extract the low frequency and high frequency signal. The difference in the frequencies, for a specific residue, corresponds to the coupling.

Since the TROSY signal is generated by the differential canceling of the CSA and the dipolar interaction, it has the narrowest possible linewidth. Therefore, the source of the error can be mainly the semi-TROSY signal.

#### 4.2.4. Comparison of the $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC and TROSY-SemiTROSY Approach

A destructive source of error in the IPAP method is the interference of the doublet pattern of two or more residues. For example, in an AP sub-spectrum, interference can arise due to overlap of the positive signal of a residue with the negative signal of the other. Since both the sub-spectra have almost twice the resonances, IPAP-HSQC is very prone to have such artifacts. Figure 4-4 B, a section of *anti-phase* sub-spectra, shows interference of positive signal of R58 with the negative signal of E37 and also negative signal of V54. Figure also illustrates interference of the positive signal of V24 with the negative signal of V11. Therefore, in the small section shown as an example, an error-free estimation of the dipolar couplings for R58, V54, V24 and V11 is impossible.



**Figure 4-4:** The IPAP-HSQC method may have some artifacts. A connecting line between two signals denotes corresponding IP or AP doublets. A section of anti-phase sub-spectra showing artifacts arising due to interference of one of the doublet component of a residue with the other (black: positive and red: negative signal). Residues experiencing such errors are shown by thick blue lines. This artifact hampers measurement of the accurate coupling for both the residues.

Artifacts arising in such manner can be only circumvented by varying the magnitude of the alignment which can be cumbersome.

In section 9.1.2, a comparison of couplings obtained by both, IPAP-HSQC and Tr-SmTr, approaches have been shown for a protein, VAT-N. According to this analysis, 74 % (43 out of 58) residues exhibited a good agreement (within respective error limits) and rest was in disagreement (26 % (15 out of 59)) couplings. These residues were looked again systematically in the respective spectra and the J and the D couplings were re-extracted. Most of the disagreements (13 out of 15, *ca.* 84 % erroneous couplings) were caused by the presence of one of an artifact in the IPAP-HSQC experiments.

Therefore, Tr-SmTr approach to measure couplings can be potentially useful, provided enough care is taken in obtaining good spectral resolution.

### 4.3. Calculation of the Alignment Tensor

After the extraction of the dipolar couplings, the alignment tensor can be calculated *a priori* to the structure calculation or refinement with RDCs. Knowledge of the fit of the predicted and the experimental alignment tensor can give a direct measure of the correctness and the quality of the structure of biomolecule under discussion. The most popular programs to achieve this task are MODULE<sup>[136]</sup> and PALES<sup>[92]</sup>. Both use a singular value decomposition algorithm<sup>[91]</sup> based on matrix manipulation for the calculation of predicted values of dipolar couplings from the input structure. Thus, both softwares need to have a reasonably well-defined starting structure for the calculation of predicted values of the alignment tensor. In case of non-availability of starting structure, a histogram based approach can be used to calculate components of the alignment tensor<sup>[137]</sup>.

Residual Dipolar Coupling constants define the quality of the structures by the deviation between experimental RDCs and predicted RDCs from the structural model, measured in terms of the  $\chi^2$  and/or  $Q$  value. The  $\chi^2$  and  $Q$  are defined in Eq. [4-4] and [4-5].

$$\chi^2 = \frac{\sum (D_{\text{calc}} - D_{\text{exp}})^2}{\sigma_{\text{exp}}^2} \quad [4-4],$$

where  $\sigma_{\text{exp}}$  denotes the experimental error and,

$$Q = \frac{\sum_j (D_j^{\text{exp}} - D_j^{\text{calc}})^2}{\sum_j (D_j^{\text{exp}})^2} \quad [4-5].$$

$Q$  is defined as the ratio of the mean square deviation between observed and calculated couplings and the mean square of the observed couplings<sup>[138]</sup>.

NMR structures calculated with RDCs typically exhibit  $Q$  values between 0.05 and 0.3, whereas a non-RDC derived structure may have  $Q$  values within 0.3 and 0.8 [139].

While MODULE interprets errors in the fit only by  $\chi^2$ , PALES utilizes both approaches (*i.e.*,  $\chi^2$  and  $Q$ ). MODULE is designed to be very user-friendly, while PALES offers wide options for the calculation of the alignment tensor. It should be noted that the *axial component*,  $A_a$ , of the alignment tensor generated by PALES is already divided by two (*i.e.*, half) compared to the value generated by MODULE. Therefore, PALES facilitates the direct multiplication of  $A_a$  and  $D_{\max}$  for obtaining the *magnitude* of the alignment tensor,  $D_a$  (Eq. [3-49]).

#### 4.4. Structure Calculation

A break-through step in the routine use of RDCs has been its incorporation in the structure calculation algorithm. CNS [140] and its derivation XPLOR-NIH [141] incorporate RDCs as a structural restraint.

CNS refines a NMR restrained structure using a simulated annealing approach in which an ensemble of molecule is heated to very high temperature, where it loses practically all physical interactions and contains a very high degree of freedom. These molecules are allowed to cool, soon afterwards, in extremely slow steps under restraints obtained by NMR experiments (*cf. chapter 2 for discussion on other restraints*). Molecules would slowly fall into various energy minima on the potential energy surface.

During the course of cooling, restraints like NOEs, J-couplings, chemical shift information, H-bonds are utilized in the first place for obtaining an appropriate global fold. RDCs are employed in the later stages with more preference and are used only to fine-tune the structure. In the absence of motional averaging, a single residual dipolar coupling measurement restricts the orientation of an internuclear vector to two cones of orientations subtended by the angle  $\theta$  relative to the magnetic field ( $B_0$ ).

Residual dipolar couplings are incorporated into the structure calculation by means of a penalty function [142, 143], which is generated by summing, for each measured residual dipolar contribution, the weighted ( $W$ ) square of the difference between the experimental splitting and the calculated splitting for a molecular structure,

$$E_{\text{DD}} = W \sum_i (D_{i,\text{calc}} - D_{i,\text{exp}})^2 \quad [4-6].$$

This penalty function (or pseudo-energy) is added to normal molecular and NOE distance constraint energies, and a search for a minimum energy structure is conducted using a

simulated annealing protocol<sup>[144-146]</sup>. A significant advantage of the simulated annealing approach is that it is easy to add residual dipolar pseudo-energy terms to a molecular dynamics force field along with other pseudo-energy terms for NOE and scalar coupling constraints.

An input file of RDCs, in the CNS format, is given in appendix (*cf. section 9.1.3*). Additionally, a guess value of the alignment tensor is needed for the structure calculation. This includes, DFS (Depth First Search),  $D_a$  (magnitude of the alignment) and the rhombicity ( $R$ ).

When more than one spin-pair generated dipolar couplings are used for the structure calculation, tensorial components of the other spin-pairs need to be scaled to the tensor components generated from the H-N spin pair.

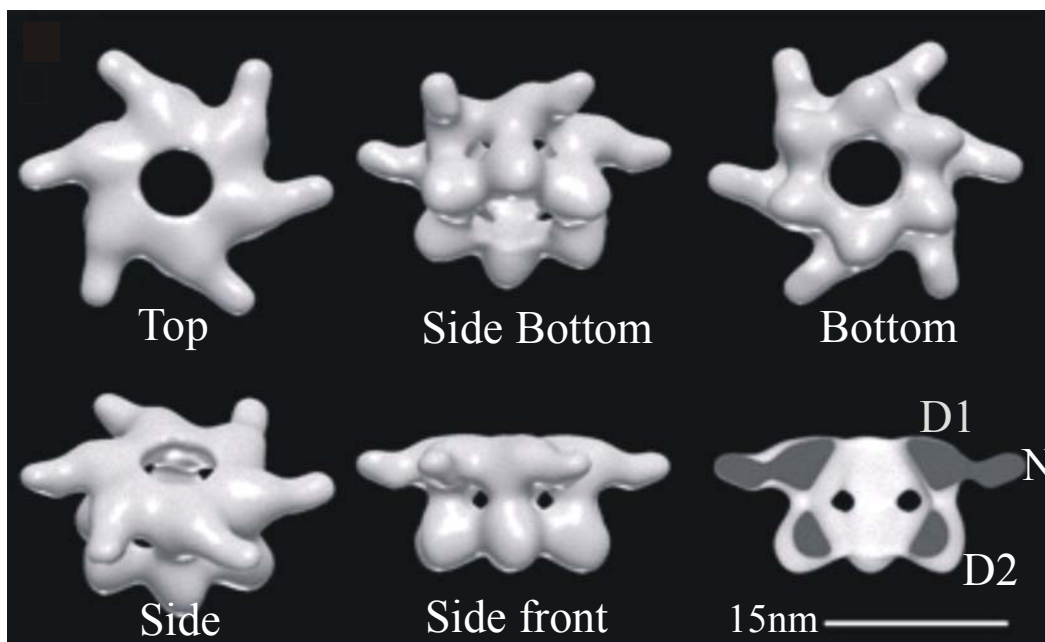
#### **4.5. Application of RDCs: Determination of Subdomain Orientation of VAT-N**

##### *4.5.1. Introduction to the VAT Complex*

Proteins of the AAA (ATPases associated with different cellular activities) family are involved in a large number of cellular processes, including membrane fusion, organelle biogenesis, protein degradation and cell cycle regulation<sup>[147]</sup>. They are characterized by a common motif that is defined by a sequence of 230–250 amino acids. It includes the Walker type A and B cassettes, which are important for ATP binding and hydrolysis, and other regions of similarity unique to AAA proteins<sup>[147]</sup>.

One extensively studied AAA-ATPases is mammalian p97 (first termed VCP, for valosin-containing protein<sup>[148]</sup> and its highly conserved homologues are identified in *Saccharomyces cerevisiae* (Cdc48p)<sup>[149, 150]</sup>, *Xenopus laevis*<sup>[151]</sup>, *Thermoplasma acidophilum* (VAT)<sup>[152]</sup> and many other organisms.

VAT (Valosine-containing protein-like ATPases of *Thermoplasma acidophilum*) displays a tripartite domain structure, N-D1-D2, and homohexameric ring architecture. It has been shown to act as an ATP-driven protein unfoldase.



**Figure 4-5:** The hexameric VAT assembly consists of a tripartite domain structure of N-D1-D2 of molecular weight 520 kDa. The N-domain (monomeric unit 20.5 kDa) is believed to be the substrate recognition domain. Equal sized D1 and D2 domains are responsible for ATP binding and substrate hydrolysis. Figure is generated from an EM image<sup>[15]</sup>.

#### 4.5.2. Solution Structure and the Subdomain Orientation in VAT-N

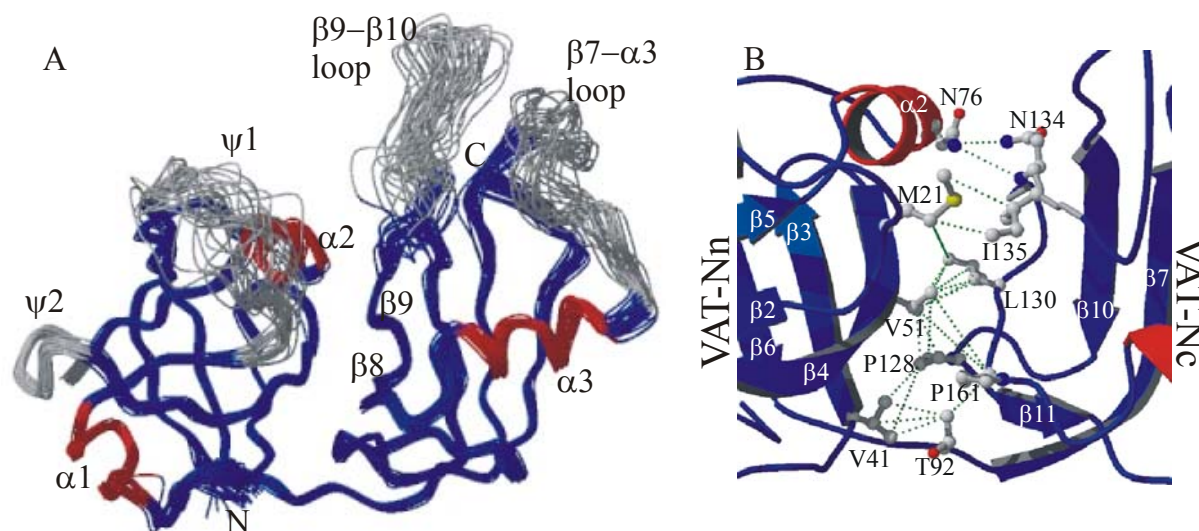
The N-terminal domain of VAT, VAT-N (20.5 kDa), is believed to take part in the substrate binding and might be capable in the folding of permissive substrates<sup>[15]</sup>.

The full unfoldase activity of VAT complex requires only ATP driven D1-D2 modules. Therefore, the role of the N-domain remains dispensable in the N-D1-D2 assembly. Therefore, it is believed that VAT-N controls access of substrate to the D1-D2 unfoldase machine, although the mechanisms of this control remain controversial. Various proposals, including the “entropic brush” mechanism<sup>[16]</sup>, where VAT-N has the role of removing unwanted substrates from the main D1-D2 binding site, must be considered and evaluated.

The solution structure of VAT-N<sup>[17]</sup> had been determined in our laboratory previously using mainly NOEs, H-bond, and scalar coupling information. The presence of two equally sized sub-domains, namely, VAT-Nn and VAT-Nc, was revealed from the structural studies. These two sub-domains are arranged into a kidney-shaped rather than a dumbbell-shaped overall structure, with a cleft between sub-domains formed on the concave side (figure 4-6 (A) ). The definition of the relative subdomain orientation relies on 28 unambiguous subdomain NOE connectivities, whereas about 2000 long, middle and short range NOE connectivities were found to define the structure within sub-domains (figure 4-6 (B) ).



The relative orientation of the two sub-domains and the possibility of inter-domain flexibility are important issues in determining the function of VAT-N. In particular, the opening of the cleft between the sub-domains to expose the hydrophobic surface between them has been proposed as a functional mechanism<sup>[17]</sup>. An alternative proposal is that the binding site is located in a cleft between the loops which characterize the surface of both sub-domains<sup>[17]</sup>.



**Figure 4-6:** (A) Ensemble of 20 best structures defining the tertiary structure as well as the sub-domain orientation of VAT-N. Overall kidney shaped structure of VAT-N can be seen. Secondary structure elements and the loop regions are marked on the structure. (B) Inter sub-domain NOE connectivities are demonstrated in figure. Relatively fewer NOEs define the relative subdomain orientation<sup>[17]</sup>.

Therefore, we utilized potential of RDCs for the determination of more precise subdomain orientation of VAT-N. The following sections deal with the experimental part and the results obtained from RDC studies on VAT-N.

#### 4.5.3. Experimental Section

U-<sup>[15</sup>N] VAT-N sample was produced and purified in the group of Prof. Baumeister, MPI of Biochemistry, Martinsried, according to previously described procedure<sup>[15]</sup>.

Samples of 0.7 mM uniformed <sup>15</sup>N-labeled VAT-N were prepared in 80 mM phosphate buffer at pH 5.9, 120 mM NaCl, and containing 5 mM NaN<sub>3</sub> and 10% D<sub>2</sub>O.

Filamentous phages Pf1 was obtained from ASLA (*Asla Biotech, Latvia*) and titrated with <sup>15</sup>N VAT-N *as is* in the proportion of 8 mg/mL.

DMPC, DHPC and CTAB were purchased as dry powders commercially (*Avanti Polar Lipids, Inc. Alabaster, AL*) and were used without further purification. Partial alignment was achieved by diluting the isotropic protein sample into a liquid-crystalline bicelle medium in 2:1 proportion. The bicelles comprised of DMPC, DHPC, and CTAB in 3.0:1.0:0.1 molar proportions respectively, and prepared in the same buffer prepared for protein sample.

The  $^2\text{H}$  quadrupolar splitting (shown in figure 4-2 (B) ) of 9.6 Hz, at 313 K, corresponds to the bicelle concentration of *ca.* 5 % w/v<sup>[118]</sup> in the protein solution.

NMR experiments were performed on a BRUKER spectrometer operating at a proton precessional frequency of 600.13 MHz (14.1 T) with triple resonance TXI-5 mm probe with gradient pulse facility.

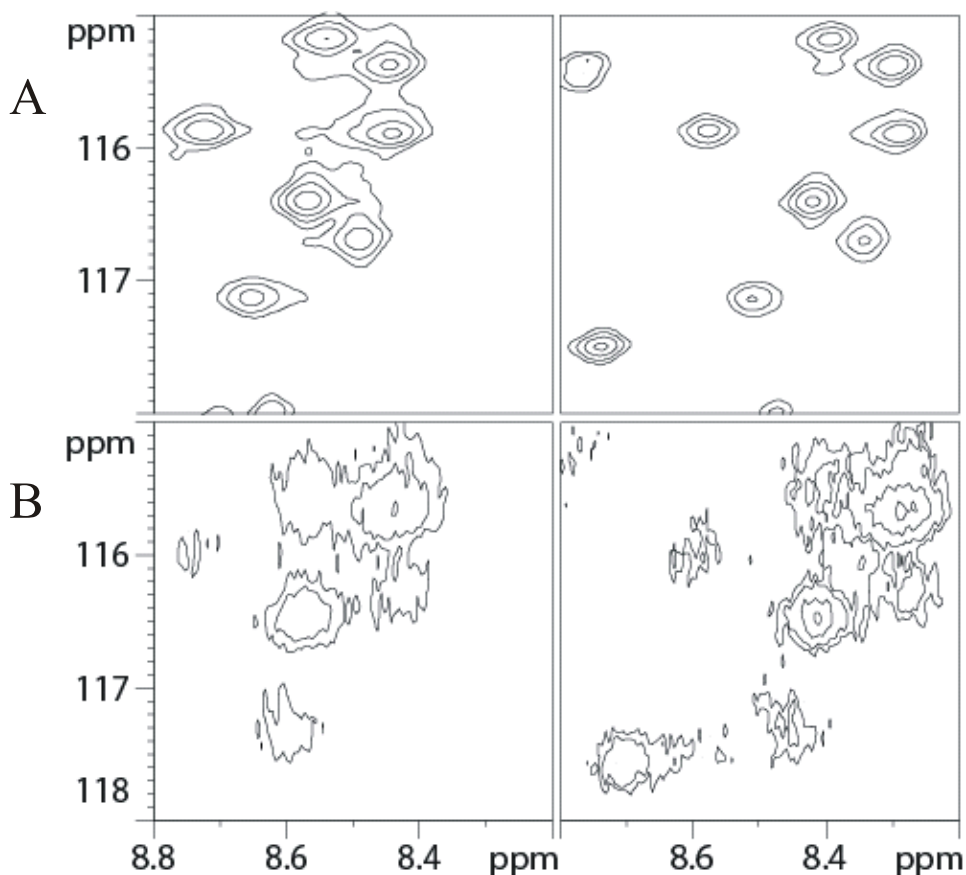
RDC measurements were carried out at 313 K in the bicelles and at 320 K in the phages. The H-NH scalar and dipolar couplings were measured under isotropic and partially aligned conditions using 2D-IPAP (*In-Phase Anti-Phase*)  $^{15}\text{N}$ - $^1\text{H}$  HSQC experiments<sup>[134]</sup> and with TROSY sequence, choosing TROSY and Semi-TROSY signals (Tr-SmTr) as coupling partners<sup>[135]</sup>. Coupling measurements in the  $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC were done in the  $F_1$  dimension while selection of the semi-TROSY signals in the TROSY-SemiTROSY approach was done in the  $F_2$  dimension. Residual NH dipolar couplings ( $^1D_{\text{H-NH}}$ ) were extracted by subtracting the  $^1J_{\text{H-NH}}$  scalar coupling constant, measured using the isotropic sample, from the  $^1J_{\text{H-NH}} \pm ^1D_{\text{H-NH}}$  values obtained using the liquid-crystalline bicelle sample. Uncertainties in  $^1D_{\text{H-NH}}$  were estimated to be 2, 3 or 4 Hz depending on the degree of line broadening, spectral resolution and the experiment of choice. Calculation of the alignment tensor from the observed dipolar couplings was achieved by MODULE<sup>[136]</sup> and PALES<sup>[92]</sup>. The error in the fit is measured as  $\chi^2$  and  $Q$ <sup>[138]</sup>.

The RDC refined structure of VAT-N was calculated using home made extension to XPLOR-NIH<sup>[141]</sup>. All other constraints which were used for the original structure calculations<sup>[17]</sup> *e. g.* short, medium and long-range NOEs, H-bond information, scalar couplings *etc.* were used in addition to RDCs. The penalty factor for RDCs is weighed to one (owing to 100 % priority to RDCs over all other restraints). Backbone RMSD of the RDC refined and non-refined structure was calculated by superimposing backbone atoms of both the structure in the program INSIGHT (*Biosym/MSI, San Diego*).

#### 4.5.4. Alignment of VAT-N with Phages

As stated in the previous sections, Pfl phage is readily available, widely studied and easy to use. Therefore, phages were utilized for achieving partial alignment of VAT-N.

Filamentous phage (Pf1) was added to a  $^{15}\text{N}$ -labeled sample of VAT-N in 8 mg/mL proportion. Tr-SmTr spectra were obtained on this sample. Huge line broadening of all the resonances of VAT-N was observed. In the sample with VAT-N and phages, the line width for  $^{15}\text{N}$  resonances (in F1) is *ca.* 55 Hz, which is almost four times higher compared to the linewidth in the free VAT-N sample (*ca.* 15 Hz). This effect can be seen in figure 4-7.



**Figure 4-7:** A spectral region of Semi-TROSY signals (left) and corresponding TROSY signals (right) can be seen for the isotropic VAT-N sample (A) and the sample titrated with 8 mg/mL bacteriophage Pf1 (B). Huge linewidth caused due to the non-specific interaction between negatively charged phages and positively charged VAT-N leading to possible binding between them causing increased correlation time. Therefore, use of phages as an alignment medium for VAT-N failed.

Unusual line broadening for all the resonances can only be explained by the presence of an electrostatic interaction between partially positive surface patch of VAT-N and negatively charged phages. Opposite charges of the protein and the alignment medium might cause a non-specific binding between them. This can give rise to the increase in the rotational correlation time of protein and therefore broadening of signals. Effects originating from the

CSA mechanism (present in an partially or fully aligned state) would have perturbed chemical shifts of signals. However, the chemical shifts of signals were not shifted in the anisotropic VAT-N. This indicates that Pf1 does not align VAT-N. Therefore, measurement of RDCs was practically not feasible for VAT-N in phages.

#### 4.5.5. *Effect of Temperature Change (from 320 K to 313 K) on VAT-N*

Due to failure of phages to align VAT-N, the use of other alignment media such as DMPC:DHPC:CTAB bicelles system was considered. Nevertheless, DMPC-DHPC bicelles are known to be stable in the temperature range of 308-314 K<sup>[97]</sup>, forcing us to reduce the measurement temperature to 314 K. Recalling that the structural studies of VAT-N were performed at 320 K, it was necessary to check the intactness of the secondary as well as global fold of VAT-N at the lowered temperature. A change in the temperature may cause changes in the secondary structure elements and therefore global structure of the protein.

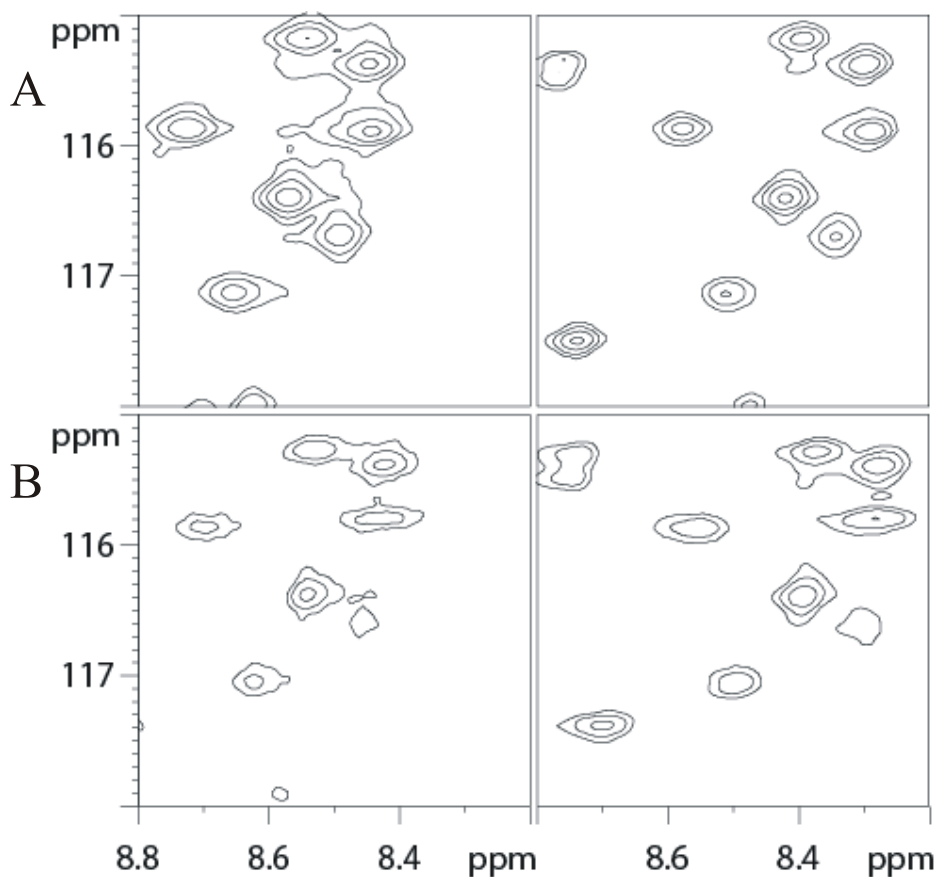
Tracing of the resonances from a series of <sup>15</sup>N-<sup>1</sup>H HSQC spectrum with an interval of 2 K in the temperature range 322 K- 310 K were performed. Chemical shifts of the HSQC cross peaks were calibrated by using 3-(trimethylsilyl) propionic acid Na salt (TSPA) as an external standard. None of the peaks in <sup>15</sup>N-<sup>1</sup>H HSQC spectra, were seen to be unaffected by the change of temperature. In conclusion, overall structure of VAT-N remains intact.

#### 4.5.6. *Alignment of VAT-N in Bicelles*

A stable anisotropic phase was obtained by addition of bicelles to VAT-N sample (figure 4-8). The <sup>2</sup>H quadrupolar splitting for this batch of bicelle preparation is shown in figure 4-2 B and a bicelle concentration of *ca.* 5 % w/v in the solution could be derived.

The following point is worth considering before getting into formal calculation and the analysis of the alignment tensor. If the NOE-derived sub-domain orientation is already well defined, a unique alignment tensor would be sufficient to define the vector orientations of the residues belonging to both the domains at once. However, a disagreement in the sub-domain orientation would lead to two different alignment tensors for each sub-domain.

In VAT-N, out of 184 residues (72 non-proline secondary structure elements residues), 56 unambiguous couplings were extracted. For 16 residues spectral overlap did not yield an error/artifact-free J or D value. Residues belonging to the flexible part (such as loop regions) showed an averaged RDC value (because averaging takes place due to at the magic angle leading to zero contribution from D) and therefore taken out of the analysis.

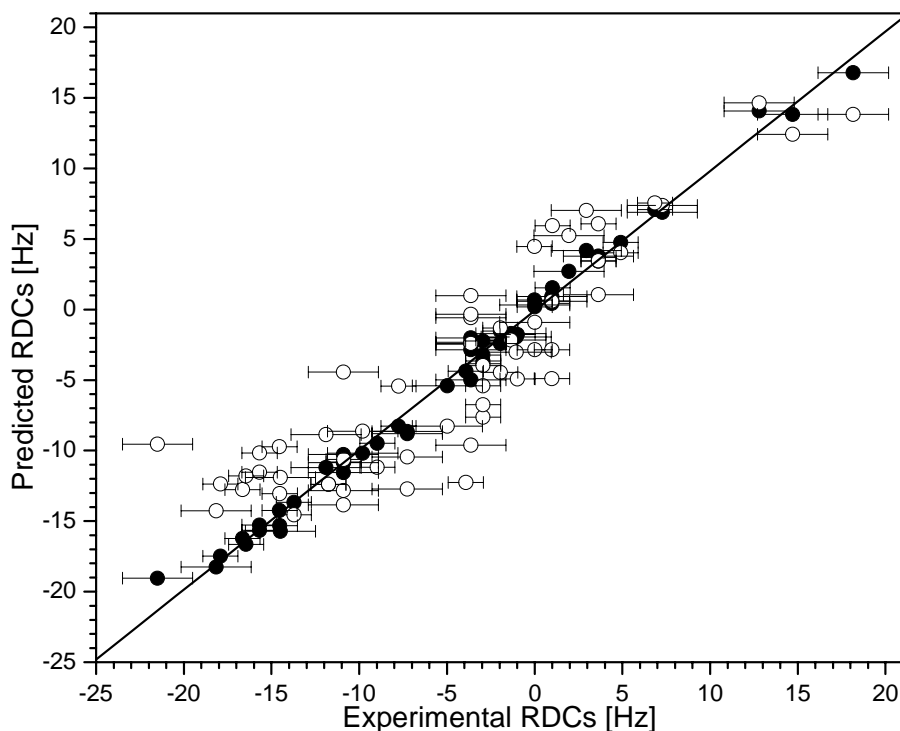


**Figure 4-8:** A spectral region of Semi-TROSY cross peaks (left) and corresponding TROSY crosspeaks (right) can be seen for the isotropic VAT-N sample (A) and the sample with 5 % w/v DMPC:DHPC:CTAB (3.0:1.0:0.1) bicelles (B). Very good spectral resolution in the latter case allowed measurement of RDCs.

RDC data was obtained from extracting couplings as discussed in the previous sections.

It was fitted to the non-RDC refined structure of VAT-N<sup>[17]</sup>. Figure 4-9 shows a correlation values predicted from the fit and the observed values of RDCs for the non-RDC refined structure. The agreement between the experimental and the predicted RDCs was very poor and is reflected in the  $\chi^2 = 197.71$  and  $Q = 0.401$ . The alignment tensor for this fit resulted with  $D_a = 8.830$  Hz, and  $R = 0.235$ . Due to high errors on the  $\chi^2$  and  $Q$ , it was evident that a single alignment tensor is not sufficient for the correct definition of the sub-domain orientation. Therefore, prediction of the alignment tensor was done separately for both sub-domains, *i.e.*, VAT-Nn and VAT-Nc. Fitting RDC data of only VAT-Nn sub-domain (30 couplings) resulted in an alignment tensor:  $D_a = 9.444$  Hz, and  $R = 0.215$  with  $\chi^2 = 109.882$  and  $Q = 0.380$ , and for VAT-Nc (26 couplings):  $D_a = 7.388$  Hz, and  $R = 0.275$  with  $\chi^2 = 68.440$  and  $Q = 0.381$ .

It should be noted that the tensorial components are moderately different for both sub-domains. Thus, the sub-domain orientation derived in a non-RDC structure contain flaws. Nevertheless, the moderate correspondence between them rules out the possibility of a significantly different orientation of the two sub-domains (such as dumbbell shaped orientation). These results indicate that the local geometry of the residues constituting these sub-domains can be defined with better accuracy.



**Figure 4-9:** *Fit of the predicted and the experimental RDCs for the non-RDC refined structure of VAT-N (open circles) and for the RDC-refined structure (filled circles). A relatively bad correlation of the non-RDC refined structure underlines need for the better definition of the sub-domain orientation and the local geometry. Error bars indicate errors in the experimental values.*

Therefore, an XPLOR calculation was performed, which included RDC data along with conventionally obtained restraints.

As expected a better agreement between the predicted and experimental RDCs was obtained for the RDC refined structure (figure 4-9).

A calculation of the alignment tensor for RDC refined structure resulted into following components and errors:

for VAT-N:  $D_a = 10.151 \text{ Hz}$ , and  $R = 0.255$  with  $\chi^2 = 8.863$  and  $Q = 0.090$ ,

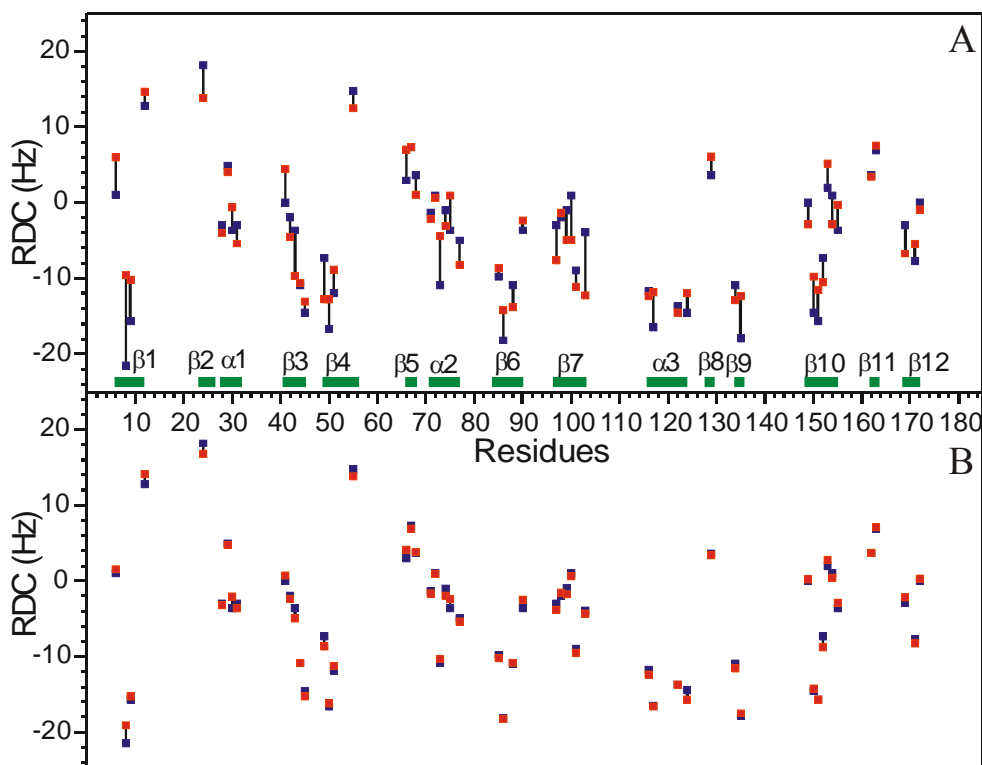
for VAT-Nn:  $D_a = 10.450 \text{ Hz}$ , and  $R = 0.265$  with  $\chi^2 = 6.102$  and  $Q = 0.084$ , and

for VAT-Nc:  $D_a = 9.786 \text{ Hz}$ , and  $R = 0.245$  with  $\chi^2 = 1.181$  and  $Q = 0.062$ .

Very low  $Q$  and reduction in the overall  $\chi^2$  indicates the proper definition of each N-H vector with respect to the alignment tensor.

#### 4.5.7. Subdomain Orientation from the Residual Dipolar Coupling

The subdomain orientation in VAT-N is defined by NOE connectivities between the  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$ , and  $\alpha_2$  secondary structure elements belonging to VAT-Nn and  $\beta_8$ ,  $\beta_9$ , and  $\beta_{11}$  belonging to VAT-Nc (figure 4-6(B)).



**Figure 4-10:** Differences between the predicted (red) and the experimental RDCs (blue) for non-RDC refined structure (A) and for RDC refined structure (B). The poor correlation in (A) for region involved in sub-domain NOE-connectivity i.e.,  $\beta_2$ ,  $\beta_3$ ,  $\beta_4$ ,  $\alpha_2$ ,  $\beta_8$ ,  $\beta_9$ , and  $\beta_{11}$  can be clearly seen. Effects seen in other regions i.e.,  $\beta_1$ ,  $\beta_6$ ,  $\beta_7$ ,  $\beta_{10}$  and  $\alpha_3$  are caused due to their structural involvement with the residues constituting sub-domain NOEs. The differences between the predicted and the experimental RDCs have disappeared due to better definition of the sub-domain orientation and other region (B).

A relatively poor correlation of predicted and experimental values of RDCs of the individual residues in these secondary structure elements was observed (figure 4-10). These differences in the RDCs are a very clear indication of the necessity for a better definition of the local geometry of these secondary structure elements.

Additionally, a bad correlation was also observed for the other secondary structure elements, *e. g.*  $\beta$ 1,  $\beta$ 6,  $\beta$ 7,  $\beta$ 10 and  $\alpha$ 3. The cause of these violations could be the spatial vicinity of these residues with the residues defining subdomain orientation. It should be noted that these residues take part in the formation of the secondary structural with the residues defining subdomain orientation.

The experimental RDCs were mostly in disagreement with the predicted RDCs for residues which were actually involved in the subdomain NOE connectivities. The poor correlation of the experimental and the predicted RDCs was also seen for the residues preceding and following to residues involved in the subdomain NOE contacts. Additionally, local geometry of the non-RDC-refined structure can be defined with little more accuracy as seen from the slight differences in the predicted and experimental RDCs for other residues.

The RDC-refined structure possesses a better compatible definition of the sub-domain orientation. All the violations in the predicted and experimental RDCs stated for the non-RDC-refined structure do not exist in the RDC-refined structure. The RDC-refined structure also proposes an overall kidney shape.

#### 4.5.8. *The RDC Refined Structure*

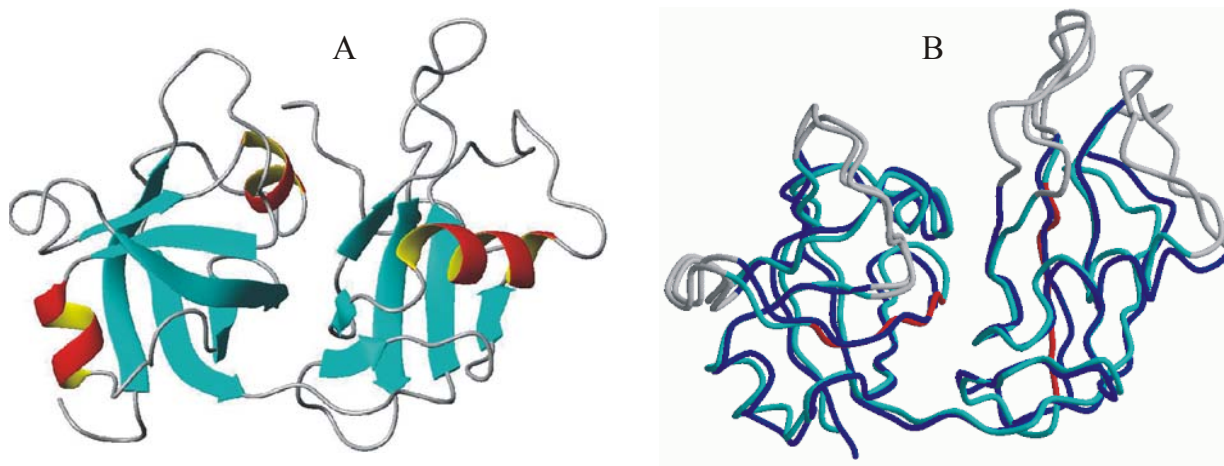
The RDC refined structure simultaneously satisfies both existing inter-domain NOEs and additional RDCs. It is very interesting to note that the RDC refined structure resulted in only two medium NOE violations out of 1814 input NOEs, *i.e.*, 107ASP-HN:106LYS-HA and 26LEU-HA:67VAL-HA. The NOE violating residues, respectively, belong to the  $\beta$ 7- $\alpha$ 3 loop (VAT-Nc) and to the  $\psi$ 1 loop (VAT-Nn). Their spatial position is away from the residues defining sub-domain orientation.

The RDC refined structure is mainly more compact than the non-RDC refined structure. The definition of some local geometries is also more accurate in the new structure *e. g.* the  $\alpha$ 1 helix is more clearly defined. The relative position of preceding and succeeding residues for both  $\beta$ 9- $\beta$ 10 and  $\beta$ 7- $\alpha$ 3 loops in VAT-Nc is more compact and accurate in the new structure.

The components of the alignment tensor derived for both sub-domains are very similar and the RDC-refined structure simultaneously satisfies both the NOEs and RDCs. This point is significant for the interpretation of backbone motion of the protein and the inter-domain



motion. Large-scale inter-domain motion, even on very slow timescales, would be expected to result in averaged components of the alignment tensor for the domain as a whole. RDCs resulting from such a tensor would be difficult to resolve with the observed inter-domain NOEs, which define one of the most “closed” inter-domain orientations possible.



**Figure 4-10:** *RDC refined structure (A) and a superposition (B) of the non-RDC refined structure (blue) with the RDC refined structure (cyan). The red region in (B) shows the violated area and the region mainly responsible for the sub-domain orientation of VAT-N. A backbone RMSD of 1.2 Å was observed for this overlay.*

The slight differences in the components of the alignment tensors for VAT-Nn and VAT-Nc in the RDC-refined structure might be due to a small amplitude motion. It can be also concluded that the large-amplitude motions do not exist, therefore, implying rigid nature of the protein. The complete analysis of the internal dynamics of VAT-N will be discussed in the next chapter.

#### 4.6. Conclusion

Residual dipolar couplings provide an access to probe the quality of structures obtained by the conventional restraints. This can be particularly useful for cases where the number and the quality of the conventionally obtained restraints are limited. RDCs can be introduced by partial alignment of the protein in the presence of external alignment media. The availability of many alignment media can help in choosing a suitable alignment system. Software developments like MODULE and PALES facilitate easy analysis of the RDC data. RDC-based and conventionally obtained restraints can be incorporated in XPLOR and CNS for structure calculations.

VAT-N possesses two equally sized sub-domains and their orientation is a crucial factor for the functional aspect of the protein. Sub-domain orientation of VAT-N was poorly defined due to few inter-domain NOEs. RDCs proved to be helpful for the validation of the sub-domain orientation obtained from conventional restraints. Secondary structure elements involved in the subdomain NOEs are defined more precisely in the RDC-refined structure without violating NOEs. As expected the RDC-refined structure also proposes a kidney shaped molecule. Analysis of the alignment tensors obtained from the fitting two sub-domains differently suggests that the protein is mainly rigid and might have low amplitude sub-domain motions.

## 5. Backbone Relaxation and Internal Dynamics of VAT-N

It is well known that proteins are highly dynamic systems covering a wide range of amplitudes and time-scales ranging from picoseconds to hours. The motions may also correlate with protein function, *e. g.*, enzyme action<sup>[153]</sup>. NMR spectroscopy offers advantage over all the other techniques by enabling detailed characterization of sequence specific local and global dynamical properties of proteins in aqueous solution. Heteronuclear <sup>15</sup>N and <sup>13</sup>C relaxation studies have been used extensively during the last few years to characterize the backbone/sidechain dynamics and motional properties of many protein molecules<sup>[154-157]</sup>. The measurement of <sup>15</sup>N or <sup>13</sup>C relaxation rates is particularly useful for obtaining dynamics information, since the relaxation of these nuclei is governed mainly by the dipolar interaction with directly bound protons and to a much smaller extent by the chemical shift anisotropy mechanism<sup>[158]</sup>. However, CSA relaxation mechanism contributes equally at relatively higher magnetic field strengths (such as 800 and 900 MHz).

Relaxation is principally governed by rotational diffusion of a molecule in the solvent, which is associated with a rotational correlation time,  $\tau_m$ . Both rotational and translational diffusion mainly depend on the size and the shape of the molecule and the viscosity of the solvent and can be described by independent tensors, if the coupling between the two is neglected. Diffusion is a tensorial property, ascribed by tensor  $D$  and represented by three diagonal elements,  $D_{xx}$ ,  $D_{yy}$  and  $D_{zz}$ . Depending on the values of these tensor elements, the diffusion tensor represents an isotropic ( $D_{xx} \cong D_{yy} \cong D_{zz}$ ), axially symmetric ( $D_{xx} \cong D_{yy} < D_{zz}$ ) or completely anisotropic ( $D_{xx} < D_{yy} < D_{zz}$ ) tumbling.

In this chapter, an introduction to the measurement and analysis of backbone relaxation is given along with its application for calculation of the internal dynamics of VAT-N.

### 5.1. Spin Relaxation and Motions in a Protein

Investigations of dynamical processes by high-resolution solution-state NMR spectroscopy can be categorized on the basis of the correlation times for experimentally accessible motional processes: laboratory frame nuclear spin relaxation measurements ( $T_2$  and  $T_1$ ) sensitive to picosecond to nanosecond (ps-ns) time scales, and rotating frame nuclear spin relaxation measurements ( $T_{1\rho}$ ) sensitive to microsecond to millisecond ( $\mu$ s-ms) time scales. The relaxation times are often expressed in terms of a rate, which is reciprocal of the time ( $s^{-1}$ ).

In addition to the above relaxation rate constants, cross-correlation or interference between different relaxation mechanisms (such as dipolar interactions) can also provide unique information, particularly if the relative geometric relationships between the interactions are fixed. Since the relaxation pathway of a  $^{15}\text{N}$  spin mainly depends on the dipolar interactions with directly bound protons, the heteronuclear steady-state NOE may be exploited as it is experimentally accessible for its contribution to the relaxation processes. The  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE is defined as:  $\text{NOE} = I_{\text{sat}} / I_{\text{eq}}$ , where  $I_{\text{sat}}$  and  $I_{\text{eq}}$  are the intensities of a signals in the spectra collected with and without proton saturation, respectively. The  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE studies distinguish unstructured and partially flexible parts (such as surface exposed loops) from the folded core.

The dynamics demonstrated by various structural elements in a protein are different. Most often unstructured terminal residues and the loop regions exhibit large amplitude dynamics. In contrast, secondary structural elements hardly show any flexibility mainly due to structural restraints. However, short structural elements (*i.e.*, elements formed by few residues) may undergo a conformational exchange between structured and unstructured state. Such an exchange phenomenon is often seen with short helices. On the other hand, elements like long  $\alpha$ -helices (8-10 residues or more) and  $\beta$ -strands forming long  $\beta$ -sheets are frequently seen as highly rigid elements.

## 5.2. Spectral Density Function

The spectral density function is simply the Fourier transform of a correlation function  $c(t)$ , and a correlation function establishes a correlation between a parameter at time  $t$  and at some time later  $(t + \tau)$ , *e. g.* it correlates the isotropic tumbling of a molecule with time, and its rate constant for the decay is in fact the rotational correlation time,  $\tau_m$ . The rotational correlation time is the average time for the molecule to rotate by one radian.

The movement of the NH bond axis is characterized by the spectral density function  $J(\omega)$ , which is related to three parameters that describe the relaxation of the  $^{15}\text{N}$  spin: the longitudinal relaxation rate ( $R_1$ ), the transverse relaxation rate ( $R_2$ ), and the steady-state NOE enhancement (NOE) <sup>[159]</sup>. The relaxation parameters of  $^{15}\text{N}$  are related to  $J(\omega)$ , at five different frequencies (*i.e.*,  $J(0)$ ,  $J(\omega_N)$ ,  $J(\omega_H)$ ,  $J(\omega_H + \omega_N)$  and  $J(\omega_H - \omega_N)$ ) by the following equations.

$$R_1 = \frac{1}{4} d^2 \{ J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N) \} + c^2 J(\omega_N) \quad [5-1],$$

$$R_2 = \frac{1}{8}d^2 \{4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)\} + \frac{c^2}{6} \{4J(0) + 3J(\omega_N)\} + R_{ex} \quad [5-2],$$

and,

$$\text{NOE} = \frac{d^2}{4R_1} \frac{\gamma_H}{\gamma_N} \{6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)\} + 1 \quad [5-3],$$

in which  $d$  and  $c$  are:

$$d = \frac{\mu_0}{4\pi} \gamma_H \gamma_N \frac{h}{2\pi} \langle r_{\text{NH}}^{-3} \rangle \quad [5-4],$$

and

$$c = \omega_N (\sigma_{\parallel} - \sigma_{\perp}) / \sqrt{3} \quad [5-5],$$

where  $\mu_0$  is the permeability of the free space,  $\gamma_H$  and  $\gamma_N$  is the gyromagnetic ratio of  $^1\text{H}$  and  $^{15}\text{N}$ , respectively,  $\omega_H$  and  $\omega_N$  are the Larmor frequencies of  $^1\text{H}$  and  $^{15}\text{N}$  respectively,  $r_{\text{NH}}$  is the N–H bond length (angular bracket denotes averaging of the bond length over time due to vibrational motions) and  $J(\omega_i)$  are the spectral densities at the angular frequency  $\omega_i$ . An axially symmetric chemical shift tensor has been assumed for  $^{15}\text{N}$  with  $\sigma_{\parallel} - \sigma_{\perp} = -160 \text{ ppm}^{[160]}$  (considering helical polypeptide chain). However, a value of  $-172 \text{ ppm}$  is also in use and has been shown to be more appropriate than the conventional value ( $-160 \text{ ppm}$ ) in some cases<sup>[161-164]</sup>.

$R_{ex}$  has been included in Eq. [5-2] to accommodate chemical exchange and other pseudo-first-order processes that contribute to the decay of transverse magnetization<sup>[165]</sup>. The  $R_{ex}$  term in Eq. [5-2] represents line broadening due to chemical exchange and/or conformational averaging on a time scale slower ( $\mu\text{s}$ - $\text{ms}$ ) than the overall rotational correlation time,  $\tau_m$ .

### 5.3. Model-free Formalism

A quantitative interpretation of the relaxation data of these commonly available three relaxation parameters can be achieved in terms of dynamical variables. A commonly used approach for this purpose is the model-free formalism pioneered by Lipari and Szabo<sup>[20, 21]</sup> and extended by Clore and coworkers<sup>[166]</sup>. Here dynamical variables include the order parameters ( $S^2$ ), the internal correlation time ( $\tau_i$ ), the global rotational correlation time ( $\tau_m$ ), and conformational or chemical exchange rates ( $R_{ex}$ ). Model-free analysis assumes that the

internal and global motions are independent and provides the amplitudes ( $S^2$ ) and the effective correlation times of the internal motions of a protein.

Motions represented by the generalized order parameter are often referred to as dynamics on the ps to ns time scale. The order parameter specifies the degree of spatial restriction of the NH bond. Assuming that the motion of the NH bond can be described by diffusion on a cone of semiangle  $\theta$ ,  $S^2$  is given by:

$$S^2 = \frac{\cos^2\theta (1 + \cos\theta)^2}{4} \quad [5-6].$$

Thus,  $S^2 = 0$  ( $\theta = \pi/2$ ), for internal motions, and  $S^2 = 1$  ( $\theta = 0$ ), in the absence of motion.

The spectral density function  $J(\omega)$  is modeled separately in the model-free formalism, depending upon whether the rotational diffusion tensor is isotropic or anisotropic.

In the former case, as per the modification by Clore and co-workers<sup>[167]</sup>, when the internal motions of the NH bond are considered to occur fast on two significantly different time scales characterized by two effective correlation times,  $\tau_f$  and  $\tau_s$ , with  $\tau_f \ll \tau_s \ll \tau_m$ <sup>[167]</sup>,

$$J(\omega) = \frac{2}{5} \left[ \frac{S^2 \tau_m}{1 + (\omega \tau_m)^2} + \frac{(1 - S_f^2) \tau'_f}{1 + (\omega \tau'_f)^2} + \frac{(S_f^2 - S^2) \tau'_s}{1 + (\omega \tau'_s)^2} \right] \quad [5-7],$$

where

$$\frac{1}{\tau'_f} = \frac{1}{\tau_f} + \frac{1}{\tau_m} \quad [5-8],$$

and

$$\frac{1}{\tau'_s} = \frac{1}{\tau_s} + \frac{1}{\tau_m} \quad [5-9],$$

where  $S^2 = S_f^2 S_s^2$  is the square of the generalized order parameter characterizing the amplitude of internal motions of each NH bond, and  $S_f^2$  and  $S_s^2$  are the squares of the order parameters for internal motions on the fast and slow time scales, respectively. The model-free spectral density function in Eq. [5-7] assumes that the overall tumbling motion of the molecule is isotropic. The isotropic correlation time  $\tau_m$  is related to  $D$  by the equation:  $\tau_m = (6D)^{-1}$ .

For the situation when the rotational diffusion tensor is anisotropic, more complicated expressions have been described<sup>[168-171]</sup>. However, for the case of an axially symmetric tensor ( $D_x = D_y = D_\perp; D_z = D_\parallel$ ), simplifications occur and the spectral density function is approximated, for the situations when the internal motions are much faster than the overall tumbling rate<sup>[169]</sup>, as:

$$J(\omega) = \frac{2}{5} \left[ S^2 \sum_{k=1}^3 \frac{A_k \tau_k}{1 + (1 + (\omega \tau_k)^2)} + \frac{(1 - S^2) \tau}{1 + (\omega \tau)^2} \right] \quad [5-10],$$

with  $A_1 = (1.5 \cos^2 \alpha - 0.5)^2$ ,  $A_2 = 3 \sin^2 \alpha \cos^2 \alpha$ , and  $A_3 = 0.75 \sin^4 \alpha$ , where  $\alpha$  is the angle between the NH bond vector and the unique axis of the principal frame of the diffusion tensor,  $\tau_1 = (6D_{\perp})^{-1}$ ,  $\tau_2 = (6D_{\parallel} + 5D_{\perp})^{-1}$ ,  $\tau_3 = (4D_{\parallel} + 2D_{\perp})^{-1}$ , and  $\tau^{-1} = 6D + \tau_1^{-1}$ , where  $D = \frac{1}{3} D_{\parallel} + \frac{2}{3} D_{\perp}$  (which is 1/3 the trace of the diffusion tensor),  $D_{\parallel}$  and  $D_{\perp}$  are the components of the axially symmetric diffusion tensor, parallel and perpendicular to the axis of symmetry, respectively. Their ratio is a measure of the diffusion anisotropy.

### 5.3.1. Dynamical Model Selection and Parameter Estimation

Maximum of six free parameters may be required for the fitting of the experimental data using Eq. [5-7], *i.e.*, the five parameters in Eq. [5-7] and  $R_{ex}$  in Eq. [5-2]. To achieve this, five simpler dynamical models derived from Eq. [5-7] are used.

Each model contains an overall rotational correlation time, a maximum of three internal motional parameters and at most a single internal time scale parameter, either  $\tau_f$  or  $\tau_s$ . For convenience in the following, the internal time parameter will be referred to as  $\tau_i$ . With this notation, the five models consisted of the following subsets of the extended model-free parameters:

Model 1:  $S^2$ ;

Model 2:  $S^2, \tau_i = \tau_f$ ;

Model 3:  $S^2, R_{ex}$ ;

Model 4:  $S^2, \tau_i = \tau_f, R_{ex}$ ; and

Model 5:  $S_f^2, S^2, \tau_i = \tau_s$ .

Model 1 is obtained by assuming that  $S_s^2 = 1$  and  $\tau_f \rightarrow 0$  and is applicable if motions on the slow time scale are not present or negligible and motions on the fast time scale are very fast ( $< 20$  ps). Model 2 is obtained by assuming that  $S_s^2 = 1$  and is applicable if motions on the slow time scale are not existing or negligible. Model 2 is the original formulation of Lipari and Szabo<sup>[20, 21]</sup>. Models 3 and 4 are derived from models 1 and 2, respectively, by including a non-zero chemical exchange contribution,  $R_{ex}$ , in these model. For models 1 through 4,  $S^2 = S_f^2$  is assumed. Model 5 is obtained by assuming only that  $\tau_f \rightarrow 0$  and it includes a

very fast and a slower internal motion. The form of the spectral density function for model 5 is isomorphous with an approximate spectral density function incorporating anisotropic rotational diffusion<sup>[20, 21, 172]</sup>. Model 5 is an extension to the original Lipari-Szabo formalism and was originally proposed to describe backbone dynamics of certain residues in loops that undergo fast librational motions as well as slower motion due to dihedral transition<sup>[167]</sup>.

#### 5.4. Estimation of Correlation Time and Diffusion Tensor

Since the relaxation analysis mainly depends only on three measured parameters, the model-free analysis, based on six dynamical parameters, can be underdetermined. Therefore, data analysis is based on strict statistical analysis to avoid over-interpretation. This can be achieved by the application of the simplest model and iterative addition of dynamic parameters until the improvement in the fit is no longer statistically significant.

For achieving efficient statistical analysis in the model-free parameter optimization, a tentative value of the diffusion tensor can be given as a starting point. Such a tentative guess value of the diffusion tensor and the correlation time can be estimated by the following approaches.

##### 5.4.1. Hydrodynamic Calculations

The hydrodynamic behaviour of arbitrarily shaped rigid particles can be modeled and computed using models, composed of spherical frictional elements<sup>[173, 174]</sup>. A computer program based on these principles is HYDRONMR<sup>[175]</sup>, where a primary hydrodynamical model or a shell model is constructed first by replacing each non-hydrogen atom by a spherical element whose radius is often referred to as the atomic element radius (AER). The lower bound of this radius is typically given by the van der Waals radius of the respective atoms. However, the AER may be varied due to hydration between 2-5 Å. The hydrodynamic simulation is performed by filling beads of decreasing size, and the results are extrapolated for infinitely small bead size.

Since proteins often exhibit unstructured loops and hydrophilic residues at the surface, results purely based on hydrodynamic calculation are prone to be erroneous. For example, a comparison of the predicted molecular correlation time and the diffusion tensor for VAT-N can be seen in table 5-1. For the RDC-refined structure of VAT-N, a temperature of 320 K and viscosity of solvent (H<sub>2</sub>O) at this temperature 0.005 Poise were used in these simulations. The hydrodynamic calculation predicts longer correlation time with increasing bead size whereas the diffusion anisotropy decreases (table 5-1).



**Table 5-1:** Hydrodynamic calculation for VAT-N with varying Atomic Element Radius (AER) in Å.

AER [Å]	2.4	2.8	3.2	3.6	4.0	4.4	4.8
Correlation Time [ns] <sup>a</sup>	6.85	7.07	7.38	7.88	8.35	8.71	9.17
D <sub>ratio</sub> <sup>b</sup>	1.382	1.371	1.351	1.335	1.333	1.315	1.311

<sup>a</sup> rotational correlation time obtained from the relation  $\tau_m = (6D)^{-1}$ ,

<sup>b</sup>  $D_{\text{ratio}} = 2D_{zz} / (D_{xx} + D_{yy})$

It was later found that the actual rotational correlation time of VAT-N corresponds to the value predicted by HYDRONMR at 4.8 Å AER (*vide infra*). However, actual diffusion anisotropy ( $D_{\text{ratio}} = 1.14$ ) is in disagreement with the predicted diffusion anisotropy. This disagreement is easily explained by the presence of flexible unstructured loops on the surface of VAT-N which are considered as fixed beads and simulated as rigid structure by HYDRONMR.

#### 5.4.2. $R_2/R_1$ Ratio

The overall rotational dynamics of the quasi-rigid structures can be expressed in terms of a single quantity, the correlation time, which can be derived from the ratio of <sup>15</sup>N longitudinal and transversal relaxation times. If no three-dimensional structure of the molecule under investigation is available, or if the molecule is known to have a low degree of rotational anisotropy, then the overall rotational correlation time,  $\tau_m$ , can be estimated from a mean value of  $R_2/R_1$  by solving the equation<sup>[176]</sup>:

$$\frac{R_2}{R_1} = \frac{4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N) + (c^2 / 3d^2) \{4J(0) + 3J(\omega_N)\}}{2J(\omega_H - \omega_N) + 6J(\omega_N) + 12J(\omega_H + \omega_N) + 2(c^2 / 3d^2)J(\omega_N)} \quad [5-11],$$

in which,

$$J(\omega) = \frac{2}{5} S^2 \left[ \frac{\tau_m}{1 + (\omega\tau_m)^2} \right] \quad [5-12],$$

is obtained from Eq. [5-7] assuming that the internal motions are limited (large  $S^2$ ) and fast ( $\tau_i < 10$ ps). Eq. [5-11] is independent of  $S^2$  and only depends on  $\tau_m$ .

An estimation of the magnitude of the anisotropy of the diffusion tensor can be also approximated by  $R_2/R_1$ <sup>[176]</sup>. Residues with large-amplitude internal motions and undergoing

conformational exchange must be excluded from this estimation. While the former can be detected from a low NOE value, latter are selected by the following condition:

$$\left( \frac{\langle T_2 \rangle - T_{2,n}}{\langle T_2 \rangle} - \frac{\langle T_1 \rangle - T_{1,n}}{\langle T_1 \rangle} \right) > 1.5 \times \text{SD} \quad [5-13],$$

where,  $\langle T_2 \rangle$  is the average value of  $T_2$ ,  $T_{2,n}$  is the value of  $T_2$  of residue  $n$ , and the SD is the standard deviation of the distribution of the value in brackets over all residues.

A determination of the diffusion tensors for spherical and axial symmetric tumbling can be achieved on the basis of Eq. [5-11]. Mainly two approaches have been developed, of which one utilizes a four dimensional grid search on the normalized error function and has been proposed by Tjandra and co-workers<sup>[169]</sup>. Another approach uses a quadratic representation of the relaxation data for the calculation of spherical, axially-symmetric, and fully anisotropic tumbling<sup>[177, 178]</sup>.

### 5.5. Analysis of the Relaxation Data

The analysis of relaxation data in the model-free framework for the model selection and for the estimation of the correlation time can be achieved by programs like MODELFREE 4.1 (A. G. Palmer, Columbia University). Relaxation analysis in MODELFREE 4.1<sup>[179]</sup> is done in the steps described in section 5.4. However, it suffers from many complicated input preparations as well as several manual interferences, therefore is limited to few experts. Meanwhile, FAST-MODELFREE (Facial Analysis and Statistical Testing – MODELFREE)<sup>[180]</sup>, a program which interfaces MODELFREE 4.1, has been developed to make the relaxation analysis accessible to a broader audience, particularly to chemists and to biologists.

In MODELFREE 4.1, the first stage of analysis is the selection of the best model for each residue by fitting the experimental data to the different models separately and selecting the one which requires a minimum statistically significant amount of parameters. After selecting the best model in this manner,  $\tau_m$  is optimized along with the other model parameters using the grid search method. Errors for the model selection in MODELFREE 4.1 are estimated by following approaches.

#### 5.5.1. Error Estimation on the Model Selection

All optimization in the grid search method involves minimization of the  $\chi^2$  function<sup>[179]</sup>:

$$\chi^2 = \sum_i^n \Gamma_i = \sum_i^n \sum_j^{m_j} (E_{ij} - S_{ij})^2 / \sigma_{ij}^2 \quad [5-14],$$

where the index  $i$  refers to an amide  $^{15}\text{N}$  site with  $n$  being the total number of sites, and  $\Gamma_i$  is the summed-squared error (SSE) for site  $i$ .  $m_j$  represents the number of experimentally determined relaxation parameters for the  $i$ th site.  $E_{ij}$ ,  $S_{ij}$ , and  $\sigma_{ij}$ , respectively, are experimental relaxation parameters, simulated relaxation parameters (predicted), and the experimental uncertainty in the  $j$ th relaxation parameter. The criterion used for acceptance of proposed models are 95 % confidence tests comparing the experimental  $\chi^2$  to  $\chi^2$  distributions based on simulated datasets from Monte-Carlo sampling of Gaussian distributions.

To judge the statistical significance an additional parameter  $F$  is calculated.  $F$  is defined as,

$$F = \frac{(N - n)\{\chi_m^2 - \chi_n^2\}}{(n - m)\chi_n^2} \quad [5-15],$$

for the comparison of models fitting  $N$  variables with  $m$  and  $n$  parameters.

## 5.6. Experimental Measurement of Relaxation Data

The experimental methods for measuring the  $^{15}\text{N}$  relaxation times are very well established<sup>[176]</sup>. These experiments have a ‘relaxation period’,  $T$ , in addition to the four basic periods of a 2D experiments (*cf. section 2.3.3*) and which is often incorporated between preparation and evolution period.

The experiment starts with the proton to nitrogen magnetization transfer *via* an INEPT sequence. The magnetization present after the preparation period provides the initial condition for the relaxation period. The relaxation-encoded frequency-labeled transverse proton magnetization is recorded during the  $t_2$  acquisition period. The relaxation rate constant measured in a given experiment depends on the initial magnetization and on any manipulations of the magnetization during  $T$ . In most experiments, the relaxation period,  $T$ , is increased parametrically in a time-series of 2D NMR spectra or is increased in interleaved fashion within a single 2D experiment (pseudo-3D experiment).

The relaxation information is encoded in the intensity of the resonance signal. Therefore, intensities for a resonance signal are extracted from each 2D frequency domain spectrum collected at various relaxation times  $T$ . These intensities are fitted to a single exponential decay function. Errors on the  $R_1$  and  $R_2$  rates are often estimated differently and discussed in the later part of this chapter.

The  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE experiment is recorded in an interleaved fashion, in which the first experiment is carried out with saturation and the second without saturation of proton

magnetization. The sensitivity of the heteronuclear NOE experiment is inherently low because the pulse sequence starts with the equilibrium  $^{15}\text{N}$  magnetization <sup>[181]</sup>.

Residues undergoing motions with large amplitude can easily be identified by low NOE values ( $< 0.65$ ). Unimolecular chemical reactions give rise to chemical exchange contribution  $R_{ex}$  to the spin-spin relaxation rate  $R_2$  determined by CPMG spin-echo sequences <sup>[182]</sup>.

### 5.6.1. Error Estimation on the Relaxation Data

Relaxation measurements are quite sensitive to peripheral parameters like temperature, pulse imperfections, unspecific oligomerization states of protein *etc.* A variation in these parameters can introduce a significant error in the measurement of the relaxation data. Realistic error estimation on the relaxation data is necessary to eliminate misinterpretation of the data. Therefore, measurement of relaxation data is practiced with certain norms which include measurement of minimum eight to ten data points of the complete exponential decay, duplication of complete measurement or at least few data points *etc.* Relaxation measurements are often performed twice on the same sample, several months apart for better estimation of the errors <sup>[169]</sup>.

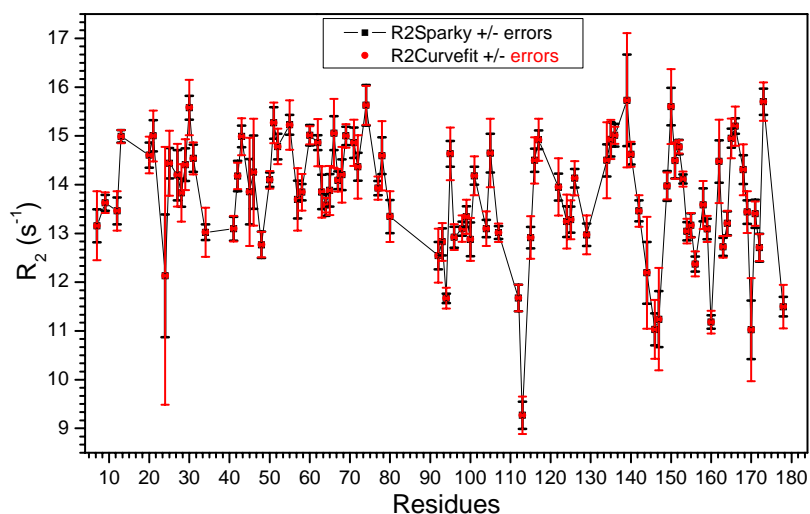
Error estimation is done differently with different programs, for example, SPARKY <sup>[183]</sup> and DASHA <sup>[184]</sup> use a base plane noise estimation procedure which reports error on the fit of experimentally determined rates to the estimated rate by stimulating Gaussian noise (using Monte-Carlo fitting) in the signal intensities. Therefore, these two programs are good where no duplication of data is available. In contrast, CURVEFIT <sup>[185]</sup> utilizes several duplicate time points for the error estimation. It uses calculation of the average deviation and a Jackknife fitting algorithm (rather than Monte-Carlo) for simulating data.

Figure 5-1 shows a comparison of the relaxation rates and the error estimation obtained by SPARKY and CURVEFIT on a dataset. The relaxation data was obtained on a 0.3 mM VAT-N sample (320 K) at 600 MHz spectrometer. Other sample conditions can be found in *section 5.7.1.*

The relaxation rate and error determination were achieved by fitting eight data points to a single exponential in SPARKY, while additionally two duplicated data points were used in CURVEFIT. It is worth mentioning that the deviation in the duplicated data point was not statistically significant and thus the choice of the duplicated data points did not influence the result.

Rates determined by SPARKY and CURVEFIT are exactly identical whereas errors estimated by SPARKY are approximately half compared to those estimated by CURVEFIT. Rates and errors

obtained from both the sources were used further for a MODELFREE 4.1 calculation. For SPARKY generated relaxation rates and the errors, an improper model selection was observed, *i.e.*, most of the residues were ascribed to model 3, which represents a conformational exchange model. In contrast, CURVEFIT generated relaxation data and errors were able to choose proper model selection, *i.e.*, most of the residues belonging to the rigid regions were selected in model 1. The proper model selection in the latter case was an outcome of more realistic estimation of errors.



**Figure 5-1:** SPARKY (black) and CURVEFIT (red) generated relaxation rates and errors on relaxation data obtained on VAT-N. Error estimation in CURVEFIT is approximately two times higher than SPARKY whereas the relaxation rate estimation is identical in both cases.

In conclusion, the measurement of at least few duplicate data points for  $T_1$  and  $T_2$  along with carrying out the error estimation procedure by CURVEFIT can provide a good starting point for the relaxation analysis in the model-free framework using the program MODELFREE 4.1.

### 5.7. Relaxation Measurement for VAT-N

The role of VAT as an energy-dependent unfoldase suggested that ATP hydrolysis may cause major changes in the location of the peptide-binding sites, thus exerting mechanical force on the bound polypeptide<sup>[17]</sup>. Such global changes in the position of the amino-terminal domains have been observed in the studies of the NSF, a protein belonging to AAA family and having similar N-D1-D2 hexameric architecture<sup>[186]</sup>.

Coles *et al.* proposed two kinds of hinge motions possible from the solution structure of VAT-N, one of the amino-terminal domains relative to the ATPases ring and the other between the two VAT-N sub-domains<sup>[17]</sup>. Both types of motions could lead to the gradual unfolding of a bound polypeptide. Therefore, the dynamical behavior of VAT-N is necessary to be exploited in order to envisage the functional role.

The apical domain of GroEL has been shown to be implicated in the unfolding of bound polypeptides through mechanical force<sup>[187]</sup> and has some striking similarities to VAT-N. Both domains are located at the upper, outer rim of barrel-shaped complexes that are involved in chaperone activities, and both domains can catalyze the refolding of permissive substrates\*. Nevertheless, the three-dimensional folds and the nature of the surface of GroEL and VAT-N are not similar, *i.e.*, the apical domain of GroEL uses an exposed hydrophobic surface for substrate binding whereas the putative binding cleft of VAT-N is charged.

The RDC-refined structure of VAT-N shows that sub-domains are fixed with respect to each other in solution and the molecule shows an overall rigid kidney shape. However, a complete picture of dynamical behaviour of any protein can not be obtained only by the analysis of RDCs generated for only one bond vector<sup>[139]</sup>. At least five different alignment media for one bond vector or five different bond vectors per residue in one alignment medium are needed to study backbone dynamics based on RDC studies<sup>[188]</sup>. The latter approach is expensive as it necessitates uniformly doubly labeled sample while the former could not be utilized because alignment of a highly charged protein in many alignment medium is difficult. In an attempt, it was shown that VAT-N could not be aligned by filamentous phage Pf1 (*cf. section 4.5.4*).

At the same time, the dynamical studies of VAT-N are necessary to be carried out since they provide insight into the sub-domain motions and therefore might shed some light on the speculation of the hinge motion proposed earlier. Therefore, a study of backbone dynamics of VAT-N in the model-free framework by measuring <sup>15</sup>N backbone relaxation rates has been done. Experimental details and the results obtained from this analysis are discussed in the next sections.

### 5.7.1. Experimental Section

Uniformly <sup>15</sup>N labeled VAT-N sample was produced and purified in the group of Prof. Baumeister, MPI of Biochemistry, Martinsried as described previously<sup>[15]</sup>.

---

\* substrates that do not require ATP for refolding.

Samples of 1.1, 0.7, 0.3, 0.2 and 0.1 mM uniformly  $^{15}\text{N}$ -labeled VAT-N were prepared in 80 mM phosphate buffer at pH 5.9, 120 mM NaCl, containing 5 mM  $\text{NaN}_3$  and 10%  $\text{D}_2\text{O}$ .

$^{15}\text{N}$  relaxation measurements ( $R_1$ ,  $R_2$ , and  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE) were carried out on a 320 K at 600 MHz (14.1 T) spectrometer equipped with a cryo probe and at 900 MHz (21.1 T) spectrometer equipped with a TXI-probe.

Ten different mixing times were recorded for both  $R_1$  and  $R_2$  experiments with 5 s and 2 s recycle delay, respectively. The pulse schemes used were fully interleaved modifications of experiments described earlier<sup>[176]</sup>.  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE spectra of VAT-N were recorded with and without proton saturation during the relaxation delay. A recycle delay of 5 s was used for the spectrum recorded in the absence of proton saturation, whereas a 2 s recycle delay followed by a 3 s period of proton saturation was used with the NOE experiment.  $^1\text{H}$  saturation was achieved with a series of  $120^\circ$  proton pulses at 5 ms intervals<sup>[181]</sup>.

Peak intensities were extracted using the relaxation fitting algorithm in SPARKY<sup>[183]</sup>. A script “SPARKY2RATE” (Patric Loria, Yale University) was used to convert rates into an input file for CURVEFIT (A. G. Palmer, Columbia University). A first initial guess of the molecular rotational diffusion tensor was obtained from the  $R_2/R_1$  ratios of individual residues using the programs R2R1\_TM (A. G. Palmer, Columbia University) and QUADRIC DIFFUSION (A. G. Palmer, Columbia University) and PDB coordinate files obtained from RDC-refined structure. Highly mobile residues or residues with relaxation contributions from chemical exchange were excluded from this estimation using the criteria described in Eq. [5-13].

The model-free analysis of the relaxation data was performed with MODELFREE 4.1 (A. G. Palmer, Columbia University) interfaced with FAST-MODELFREE<sup>[180]</sup>. As stated previously, FAST-MODELFREE automatically performs the rigorous statistical testing protocol for the assignment of the model function for each individual residue<sup>[179]</sup>. Rigid body hydrodynamic modeling of the diffusion tensor and relaxation rates was performed with the program HYDRONMR<sup>[175]</sup> using the previously mentioned structures and an atomic bead radius of 4.8 Å.

### 5.7.2. *Unspecific Oligomerization of VAT-N*

The model-free analysis of the relaxation data can be erroneous if the exact oligomeric/monomeric (micro-crystalline aggregation) state of the protein is not known. Therefore, it is important, prior to any analysis, to evaluate the exact nature of the protein under investigation.

Initial relaxation experiments on VAT-N were carried out at 1.1 mM concentration which had been also used for the structural studies. It was considered that VAT-N remains basically monomeric at this concentration based on the measurement of the translational diffusion coefficient<sup>[189]</sup>. However, at this concentration the average transverse relaxation rate ( $R_2$ ) for VAT-N was *ca.* 20 s<sup>-1</sup> (corresponding  $T_2 = 50$  ms) compared to the predicted average value of 12.5 s<sup>-1</sup> (corresponding  $T_2 = 80$  ms). The latter value is predicted by HYDRONMR and is in agreement with transverse relaxation rates experimentally found in similarly sized proteins. This information implies that VAT-N has equilibrium of monomeric and oligomeric states at this concentration. Possible oligomerization of VAT-N at 1.1 mM concentration was supported by the first estimate of the molecular rotational correlation time  $\tau_m \cong 17.20$  ns obtained from the individual <sup>15</sup>N  $R_2/R_1$  ratio. The unspecific oligomerization of VAT-N might have caused by the vivid charge distribution on the surface.

**Table 5-2:** Relaxation rates and estimated correlation times of VAT-N at various concentrations.

Concentrations [mM]	1.1	0.7	0.3	0.2	0.1	0.07
Relaxation Parameters						
Averaged $R_2$ [s <sup>-1</sup> ]	19.23 ± 0.52	16.94 ± 0.47	13.38 ± 0.48	12.65 ± 0.28	12.34 ± 43	12.20 ± 0.55
Averaged $R_1$ [s <sup>-1</sup> ]	1.00 ± 0.07	1.05 ± 0.07	1.28 ± 0.06	1.33 ± 0.06	1.36 ± 0.05	1.38 ± 0.09
Correlation time $\tau_m$ [ns] <sup>a</sup>	17.20 ± 0.06	15.14 ± 0.05	9.72 ± 0.04	9.24 ± 0.01	8.87 ± 0.04	8.85 ± 0.08

<sup>a</sup> rotational correlation time obtained from Quadratic Diffusion (A. G. Palmer, Columbia University).

Further relaxation measurements were carried out at lower concentrations of VAT-N. Table 5-2 shows relaxation rates measured for various concentrations VAT-N (at 600 MHz) and first estimations of the correlation time from  $R_2/R_1$ . The residues exhibiting low heteronuclear NOE values (< 0.65) and the residues not satisfying condition in Eq. [5-13] were taken out of this analysis.

At and below a concentration of 0.2 mM, VAT-N primarily remains in the monomeric state as evident from the estimated correlation time (table 5-2). Therefore, the model-free relaxation analysis was carried out on the experimental data acquired at a concentration of 0.2 mM.

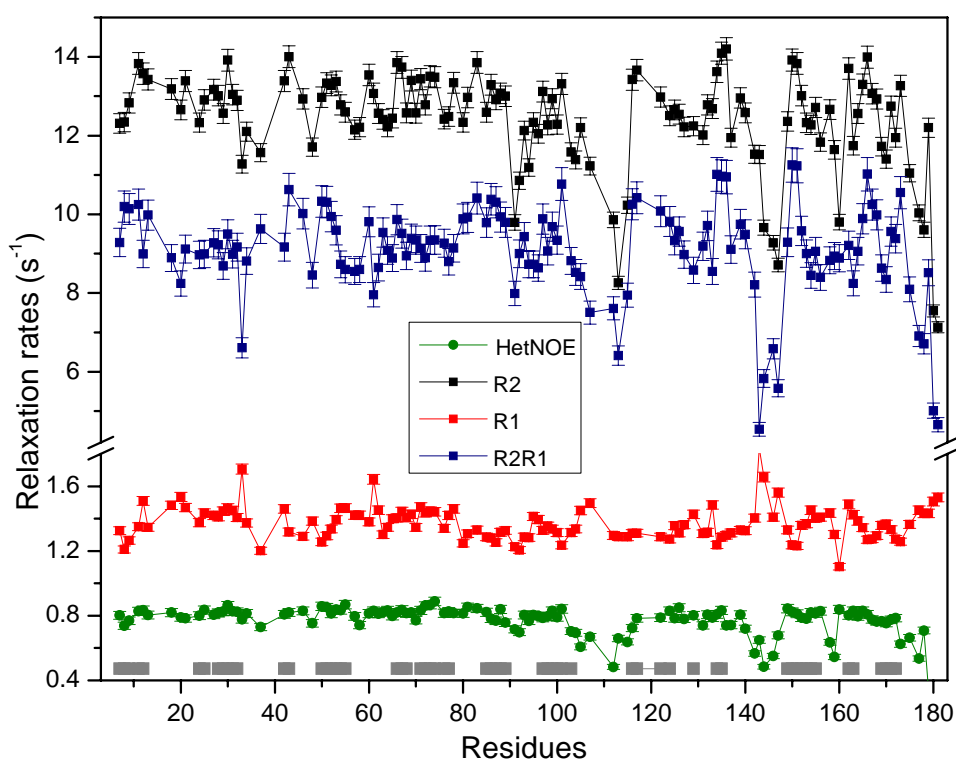


The  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE spectra for 0.1 and 0.07 mM samples exhibited relatively poor signal intensities compared to that of the 0.2 mM sample, and the signal intensities for the former samples could only be enhanced at the expense of spectrometer time (approximately four additional days for one sample). Therefore, analysis of the relaxation data acquired at other concentrations was not performed.

It is worth mentioning that the 0.2 mM sample showed larger  $R_2$  values (*ca.*  $27\text{ s}^{-1}$ ,  $T_2 = 36\text{ ms}$ ) after 3 months of storage at 276 K. This observation made it clear that VAT-N remains monomeric at very low concentration only for a short time and has hampered our attempts to estimate errors from the relaxation data acquired several months apart.

### 5.8. Residue Specific Relaxation Rate Analysis

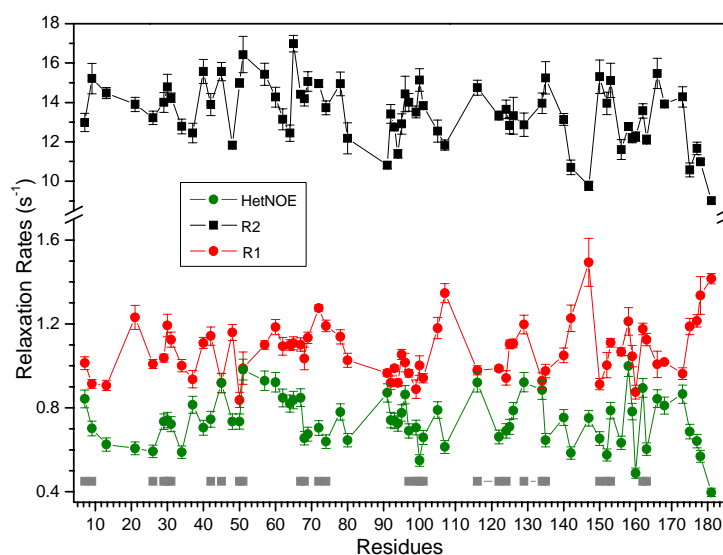
A plot of residue specific relaxation rates, their ratio and the  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE for VAT-N is given in figure 5-2.



**Figure 5-2:** Residue-specific  $^{15}\text{N}$  relaxation rates  $R_2$  (black),  $R_1$  (red) and their ratio  $R_2/R_1$  (blue) for VAT-N at 600 MHz. The  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE is shown in green. Secondary structure elements are shown in grey. Bars indicate errors estimated by CURVEFIT (A. G. Palmer, Columbia University). Highly dynamic loop regions show a sudden drop in  $R_2$  and  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE.

This plot is often sufficient for identifying highly dynamic regions well before accomplishing a complete model-free analysis. These regions can be located by a moderately large sudden drop in the  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE,  $R_2$  and  $R_2/R_1$  values.

A sudden drop in these values in figure 5-2 is found for the unstructured loops in VAT-Nc ( $\beta_9$ - $\beta_{10}$  and  $\beta_7$ - $\alpha_3$ , *cf.* figure 4-6) distinguishing them from rest of the protein residues and demonstrate their flexible nature. However, the  $\psi_1$  (residues 14-22) and the  $\psi_2$  (residues 57-65) loops in VAT-Nn are rigid relative to the other loops located in the sequence between residues 35-41, 102-116, 136-149, and 155-162. Additionally, the domain linking region formed by residues 90-97, linking VAT-Nn to VAT-Nc, is rigid in the solution. This is an important information, in particular, in the presence of the large amplitude motions between sub-domains, this region would be more flexible in contrast to the current observation. Terminal residues also show a drop in the  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE,  $R_2$  and  $R_2/R_1$  values corresponding to their flexible nature, exactly as expected.



**Figure 5-3:** Residue-specific relaxation rates  $R_2$  (red),  $R_1$  (black), and the  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE (green) measured for VAT-N at 900 MHz. Secondary structure elements are shown in grey Bars indicate errors estimated by CURVEFIT (A. G. Palmer, Columbia University). Very poor  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE data did not allow for further treatment of the relaxation rates.

A similar behavior of the residue specific relaxation rates can be seen in the data obtained at 900 MHz for the 0.1 mM sample of VAT-N (figure 5-3). Due to the increase in the field strength  $R_2$  values are increased and  $R_1$  values are decreased compared to 600 MHz. However, the  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE data obtained on this sample showed a very poor sensitivity owing to the reasons stated earlier, as well as due to use of a normal probe.

Therefore, this relaxation data set was not used further for the model-free analysis. Nevertheless, the ratio of the relaxation rates ( $R_2/R_1$ ) can be calculated from this data. Information derived from this ratio is discussed in the next part of this chapter.

It should be noted that residue specific relaxation rate analysis only provides a qualitative picture of the local dynamics and should be interpreted only as a very first inference.

### 5.8.1. Estimation of the Diffusion Tensor and the Model-free Calculation

The relaxation data obtained on 0.2 mM sample of VAT-N at 600 MHz were further used for estimating a diffusion tensor. As discussed in previous sections, such estimation can be accomplished using the ratio of the relaxation rates, *i.e.*,  $R_2/R_1$ . The average value obtained for the ratio of the relaxation rates was 9.35. Prediction of the tensor and the correlation time was done using the QUADRIC DIFFUSION approach<sup>[177, 178]</sup>. The estimated diffusion tensor was axially-symmetric based, on the F test value of 17.4 for the axially symmetric, 0.39 for the fully anisotropic, with a diffusion anisotropy,  $D_{\text{ratio}} = 1.19 \pm 0.01$  and a molecular correlation time of  $\tau_m = 9.24 \pm 0.017$  ns. Estimated values of the diffusion tensor and the correlation time are in agreement with the values predicted by HYDRONMR at 4.8 Å AER. These values were used as input tensor for running a MODELFREE 4.1 calculation.

The model-free calculations were run for the RDC-refined structure of VAT-N with three different inputs for relaxation data sets for:

- 1) All residues belonging to both sub-domains, 2) Only VAT-Nn sub-domain, and
- 3) Only VAT-Nc sub-domain.

In the anisotropic tumbling, if the optimized components of the diffusion tensor are matched in all the three cases mentioned above, very likely no motion between sub-domains of VAT-N can occur.

Results obtained from the MODELFREE 4.1 calculations are listed in table 5-3. The model-free optimized tensor components of VAT-N are very resembling to the tensor components optimized for VAT-Nn and VAT-Nc though marginal differences exists. This point is significant suggesting that the inter-domain motions are almost absent.

A large amplitude motion even on a very slow time scale between two sub-domains (such as wobbling motion) would have resulted in substantially different components of the diffusion tensors for both sub-domains and the overall tensorial components fitting to the complete protein would be an average of it. The components of the diffusion tensor representing full length VAT-N have very high similarity in the tensorial components of VAT-Nn and VAT-Nc. Therefore no large amplitude motion between sub-domains of VAT-N can exist.

**Table 5-3:** Diffusion tensor analysis for VAT-N (concentration: 0.2 mM)

Tensor <sup>a</sup>		$\tau_m$ [ns] <sup>b</sup>			$D_{\text{ratio}}$ <sup>c</sup>			$\chi^2$			F		
		VAT-N	Nn	Nc	VAT-N	Nn	Nc	VAT-N	Nn	Nc	VAT-N	Nn	Nc
Estimated Diffusion Tensor <sup>d</sup>	isotropic	9.28 ± 0.01	9.25 ± 0.03	9.33 ± 0.04	--	--	--	877.38	307.02	562.84	--	--	--
	axial symmetric	9.24 ± 0.01	9.28 ± 0.01	9.18 ± 0.03	1.19 ± 0.01	1.15 ± 0.01	1.23 ± 0.01	535.12	181.30	237.68	17.4	10.4	15.0
	fully anisotropic	9.24 ± 0.02	9.28 ± 0.01	9.19 ± 0.03	1.19 ± 0.01	1.15 ± 0.01	1.23 ± 0.01	529.92	178.74	231.16	0.39	0.30	0.43
HYDRONMR <sup>e</sup>		9.169	---	---	1.313	---	---	---	---	---	---	---	---
optimized <sup>f</sup>		9.15 ± 0.01	9.18 ± 0.02	9.03 ± 0.03	1.14 ± 0.01	1.09 ± 0.01	1.20 ± 0.01	---	---	---	---	---	---

<sup>a</sup> anisotropy of the diffusion tensor,

<sup>b</sup> rotational correlation time obtained from the relation  $\tau_m = (6D_{\text{iso}})^{-1}$ ,

<sup>c</sup> ratio of the elements of the diffusion tensor: for axially symmetric case:  $D_{\text{ratio}} = D_{\parallel} / D_{\perp}$ , for fully anisotropic case and for HYDRONMR  $D_{\text{ratio}} = 2D_{\text{zz}} / (D_{\text{xx}} + D_{\text{yy}})$ ,

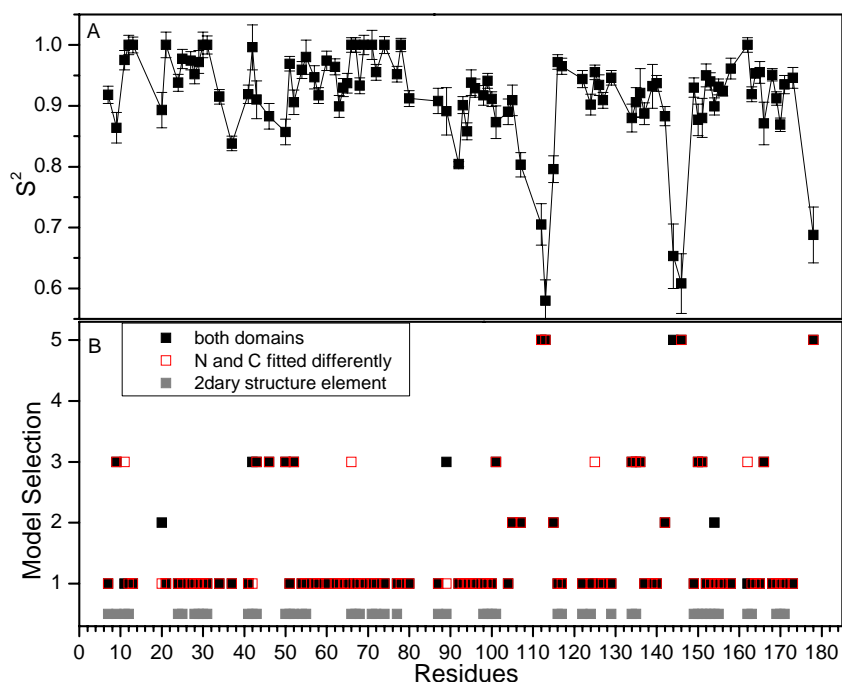
<sup>d</sup> the diffusion tensor and the rotational correlation time were estimated using program QUADRIC DIFFUSION,

<sup>e</sup> results from hydrodynamic calculations using HYDRONMR,

<sup>f</sup> model-free results optimized using MODELFREE 4.1.

### 5.8.2. Residue Specific Model and Squared Order Parameter Selection

Model-free parameters such as the squared order parameter and the dynamical model selection were extracted from the MODELFREE 4.1 output files. A plot of the squared order parameter and residues is given in figure 5-4 (A). The dynamic model selection of the residues is given in figure 5-4 (B).

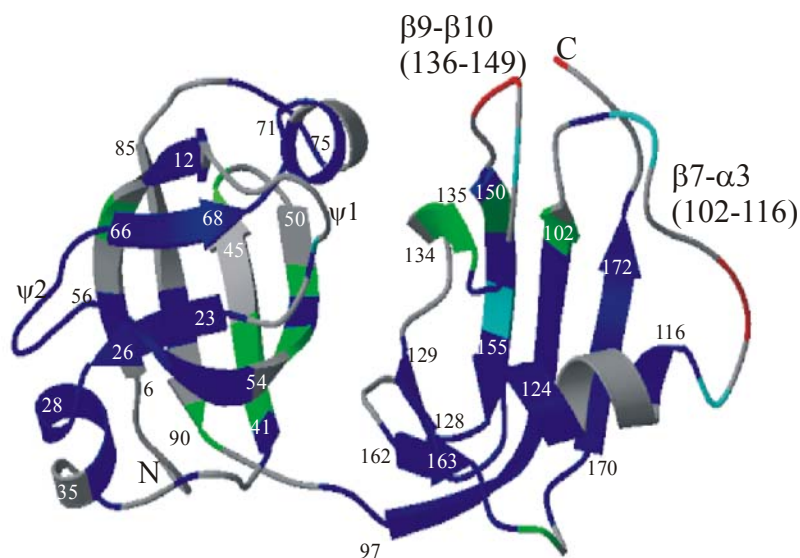


**Figure 5-4:** Residue specific squared order parameter (A) and the dynamical model selection (B) as selected by MODELFREE 4.1 (A. G. Palmer, Columbia University). The secondary structure elements are shown in grey. See text for the definition of the models 1-5. Highly populated model 1 represents a rigid backbone of the protein, whereas residues selected in model 3 might undergo chemical exchange.

The two minima obtained in  $S^2$  correspond to highly flexible surface loops in VAT-Nc ( $\beta 7$ - $\alpha 3$  and  $\beta 9$ - $\beta 10$ ). For the residues belonging to these loops, model 5 is chosen indicating that the motions can be described on two time scales, fast and slow. Other residues in the protein exhibit more or less a uniform squared order parameter which is also reflected in the fact that they are selected in model 1. This indicates that the motions experienced by these residues can be described by a single time scale. The plot in figure 5-4 (A) indicates that the residues belonging to the  $\psi 1$ ,  $\psi 2$  and other loop regions (*cf.* figure 4-6) along with the linker

region are rigid since they exhibit squared order parameter approaching to a value of one. Most of these residues are also fitted to model 1.

The calculation of the squared order parameters and model selection was not influenced by fitting two sub-domains independently. Model 1 is a highly populated and has been chosen for most of the residues belonging to the secondary structure elements. At the same time, several residues were fitted to model 3 (second highly populated model). The dynamic models are mapped on the structure and are shown in figure 5-5.



**Figure 5-5:** Map of model selection residues on the structure, grey: no data available; blue: model 1; cyan: model 2; green: model 3; and red: model 5. Selection of model 3 for the residues at the interface of the sub-domain indicates that this region can undergo a conformational exchange.

Selection of model 3 for certain residues can be now explained on the basis of figure 5-5. The residues taking part in the NOE connectivities and residing at the sub-domain interface are mainly belonging to model 3. Other residues selected in model 3 are associated with the sub-domain interface residues *via* spatial connectivities (*e. g.* adjacent  $\beta$ -strands) and hence can be ignored while analyzing inter-domain motion.

Model 3 includes a chemical exchange term and its frequent occurrence at the interface of the sub-domain interface indicates that a very small amplitude motion may exist between the two sub-domains. It is worth to note that these were mainly violating residues when the RDC data was fitted to the non-RDC-refined structure (*cf. section 4.5.7*). At the same time, it was already shown that the large amplitude motion do not exist based on the tensorial component analysis.

### 5.8.3. Small Amplitude Motions between sub-domains of VAT-N

Concrete information about such small amplitude motion between two sub-domains can be gained from measuring  $^{15}\text{N}$  relaxation at different magnetic field [190]. If the relaxation data analysis acquired at the second field provides a different diffusion tensor and rotational correlation time compared to the relaxation data obtained at the former magnetic field, then the inter-domain motion exists. Nevertheless, the differences in the tensorial components and the rotational correlation time give qualitative information.

Since our attempts to analyze the relaxation data acquired at 900 MHz in the model-free framework were restricted by the poor quality of the  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE signals, the diffusion tensor and the rotational correlation time were estimated by the ratio of the relaxation rates. For the exclusion of the residues based on the low NOE values,  $^{15}\text{N}$ - $^1\text{H}$  heteronuclear NOE values obtained at 600 MHz were used. Additionally, the condition posed by Eq. [5-13] was applied.

An axially symmetric diffusion tensor with  $D_{\text{ratio}} = 1.27 \pm 0.01$  and  $\tau_m = 7.44 \pm 0.015$  ns was obtained for this analysis. This value is different from the value  $D_{\text{ratio}} = 1.11 \pm 0.03$  and  $\tau_m = 8.87 \pm 0.040$  ns estimated on the same sample (concentration: 0.1 mM) at 600 MHz. This observation gives a clear indication that a small amplitude motion exists between two sub-domains of VAT-N. However, it should be understood that the magnitude of the dynamics shown by sub-domains is less significant compared to the postulation of Coles *et al.* as well as compared to the dynamics of the apical domain of GroEL.

Owing to the presence of the groove between sub-domains such small motion can exist, and its biological importance can only be analyzed in the presence of a ligand. Additionally, the peptide linker connecting VAT-N to the D1 domain might also provide the flexibility necessary to relocate VAT-N for ligand binding. Therefore, investigations on the functional role of VAT-N in VAT assembly firmly necessitate similar relaxation studies of VAT-N in the N-D1 mutants of VAT.

## 5.9. Conclusions

A detailed characterization of sequence specific local and global dynamical properties of proteins in aqueous solution can be accomplished by NMR spectroscopy. This information is accessible by NMR relaxation processes.  $^{15}\text{N}$  relaxation rates can be measured by well-established techniques. Analysis of the protein relaxation data can be done in the framework of the model-free analysis. The first estimation of the diffusion tensor is necessary for starting

model-free analysis and can be obtained from the ratio of the relaxation rates. Further optimization of the diffusion parameters in the model-free framework gives direct information on residue specific motions as well as global motions. An application of dynamics studies derived from NMR spin relaxation to VAT-N has been shown in this chapter. Opening of the cleft between sub-domains had been proposed as one of the possibilities for accommodating substrates. Such a mechanism would necessitate large amplitude motional changes between sub-domains. Relaxation data acquired on VAT-N has been studied to address this question and to confirm the existence of such a mechanism.

The analysis of the relaxation data in the model-free framework suggests an axially symmetric diffusion tensor for both sub-domains as well as for full length VAT-N. The tensorial components fitted to two sub-domains separately and to complete VAT-N do not show any differences. This is a strong indication of a correct sub-domain orientation, as well as an absence of large amplitude motion at any timescale. However, the relaxation analysis shows that the residues involved in sub-domain NOE contacts undergo a conformational exchange. Therefore, a small amplitude motion between the two sub-domains cannot be ruled out. Relaxation data acquired at 900 MHz also supports a small amplitude motion between the two sub-domains. At the same time, it should be noted that the protein is overall rigid and does not undergo substantially larger motions.



## 6. Substrate Binding Studies of VAT-N

NMR spectroscopy has proven to be very useful for the identification of the binding between a substrate and/or a ligand molecule. Additionally, NMR spectroscopy can also provide specific information about the strength of the binding [22-25]. An introduction to some of these techniques and substrate binding studies of VAT-N will be discussed in this chapter.

### 6.1. Ligand screening, the Nature of the Binding and Location of the Binding Site

Ligand binding by NMR often offers a choice to observe binding either on the ligand or on the substrate resonances (*e. g.* protein).

#### 6.1.1. Primary Approaches

Substrate binding studies by NMR spectroscopy has developed as a first hand tool for finding out the functional mechanism of a protein. Therefore, several NMR methodologies have been evolved in recent years to explore the substrate binding. A short introduction to such methods is given in the following part of this section.

One dimensional spectroscopic methods like STD (Saturation Transfer Difference) [191] which relies on 1D proton NMR spectra and therefore does not require isotope labeling of specific nuclei. This is a very fast and convenient method and utilizes saturating magnetization of the protein and transferring it to the ligand (and *vice versa*). Saturation-transfer difference (STD) NMR spectroscopy exploits chemical exchange and spin diffusion to label ligands with magnetization (or saturation) from a protein. If a ligand shows two different signals because of a slow exchange between the bound state and the unbound state a transfer of saturation is possible between the free and the bound state. By irradiating signals of the free ligand, the signals of the bound ligand may be identified. This technique can be easily used for homo-nuclear spectroscopy, especially proton NMR experiments, to obtain well-resolved spectra of the ligand alone.

Observation of the chemical shift perturbations of the methyl  $^{13}\text{C}$  resonances upon ligand addition was also proposed recently. It benefits from the fact that the side chain methyl groups are abundant, spectra can be acquired fast, and the carbon chemical shifts are more dispersed resulting in sharp resolved resonances.

Alternatively, translational and rotational diffusion of a ligand bound to a protein can be also studied. A bound ligand will exhibit a slower translational or rotational diffusion coefficient than for the free ligand.

It should be noted that these approaches report the binding but do not contain any specific information about the location of the binding site and the mechanism of the binding.

‘SAR by NMR’<sup>[192]</sup> (structure–activity relationships by NMR) has become a very useful tool for detecting binding. It utilizes the chemical shift mapping method and therefore can give insight into the binding site.

### 6.1.2. *Chemical Shift Mapping*

Chemical shift mapping helps to locate an exact binding site. Information on the backbone resonance assignment and the protein structure (or a homology based model) for a protein is a prerequisite for chemical shift mapping. Ligand binding causes change in the electronic environment of the residues which are in the vicinity of the ligand. This changed environment induces change in the chemical shift for these residues in the <sup>15</sup>N/<sup>13</sup>C-<sup>1</sup>H HSQC (or HMQC) experiments.

If the ligand binds relatively weak, (fast exchange), addition of increasing concentrations of the ligand will lead to progressive shifts of the resonances, such that each amide peak can be followed from its position in the free protein to its position in the bound complex. For the tight binding (slow exchange), affected residues will be characterized by the disappearance of the peak from the free protein and the appearance of a peak from the complex. In either event, it is possible to identify from the spectrum all the amide groups whose environment is affected by ligand binding. These will include groups both in residues that make contact with the ligand and in residues that are affected indirectly by ligand-induced changes in the protein structure (allosteric effects). If the shift changes are mapped onto the protein structure, a clear surface patch of affected residues is generally observed, and this indicates the location of the binding site.

### 6.1.3. *Distance Measurements between the Ligand and the Substrate*

Much more precise identification of binding sites, in terms of distances between atoms of the protein and those of the bound ligand are provided by intermolecular NOEs. It is worth considering here that this approach fails if the binding is weak because of  $r^{-6}$  dependence of NOE. The distance measurement can be achieved using edited experiments where substrate and the ligand are differentially labeled and then the NOE is measured using conventional 3D NOESY experiments, introduced in the first chapter. Because these methods yield inter-atomic distances, they can be used not only to locate the binding site but also to ‘dock’ the ligand into a known or modeled protein structure for obtaining a structure of the complex.

## 6.2. Substrate Binding Studies of VAT-N

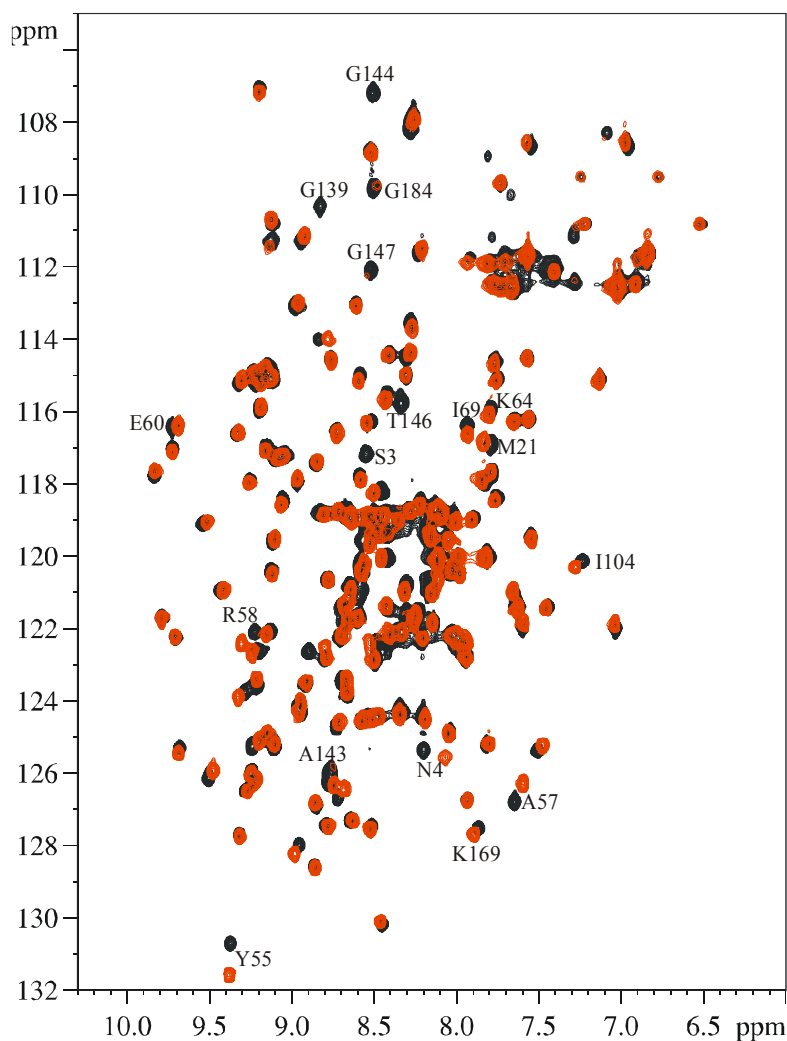
As discussed in the previous two chapters, VAT-N is the N-terminal domain of VAT protein which is a hexameric assembly and acts as a molecular machine in the eukaryotic cells. The role of VAT-N was thought to bind to the unstructured terminal residues of a substrate protein *via* its concave surface groove. It was also proposed that after binding, VAT-N leads the substrate protein into the D1 domain hexamer. Then, the hydrolysis of the substrate protein would take place in the D1 domain. The hinge connecting the hexameric assembly of the N and D1 domains would provide necessary flexibility for this mechanism. To investigate this hypothesis, it is necessary to look for the natural substrates of VAT-N. Exploring natural substrate would clearly demonstrate the role of the N-domain in the VAT assembly.

We have titrated  $^{15}\text{N}$  labeled VAT-N with a peptide SsrA, 8.5 kDa Ubiquitin, 8 kDa Barstar and 23.5 kDa casein. Following sections will deal with the experimental conditions, choice of the specific substrate as well as results from these studies.

### 6.2.1. Assignment of VAT-N Amide Resonances at pH 6.7

As seen in the previous chapters, all the structural and dynamic studies on VAT-N were performed at a pH 5.9. However, three major difficulties existed in using pH 5.9 for pursuing binding studies. Barstar has a pI at pH 6.0. Many proteins, including Barstar, tend to aggregate near their pI value. Thus, the binding studies of VAT-N and Barstar were not feasible at pH 5.9. At the same time, casein and Ubiquitin have favorable pH ranges around 7.0.

Therefore, all the titration experiments were performed at pH 6.7. However, it is known that a small change in the pH can cause changes in the secondary structure elements and thus in the global structure of the protein. Hence, the  $^{15}\text{N}$ - $^1\text{H}$  HSQC of VAT-N at pH 6.7 was recorded and compared the proton and the nitrogen shifts with the spectra recorded at pH 5.9. The overlay of spectrum recorded at a pH of 5.9 and a pH 6.7 is shown in figure 6-1.



**Figure 6-1:** An overlay of  $^{15}\text{N}$ - $^1\text{H}$  HSQC of VAT-N at pH 5.9 (black) and 6.7 (red). No major changes in the chemical shift were observed for all residues, indicating that the secondary structure elements and the global fold of the protein have not changed.

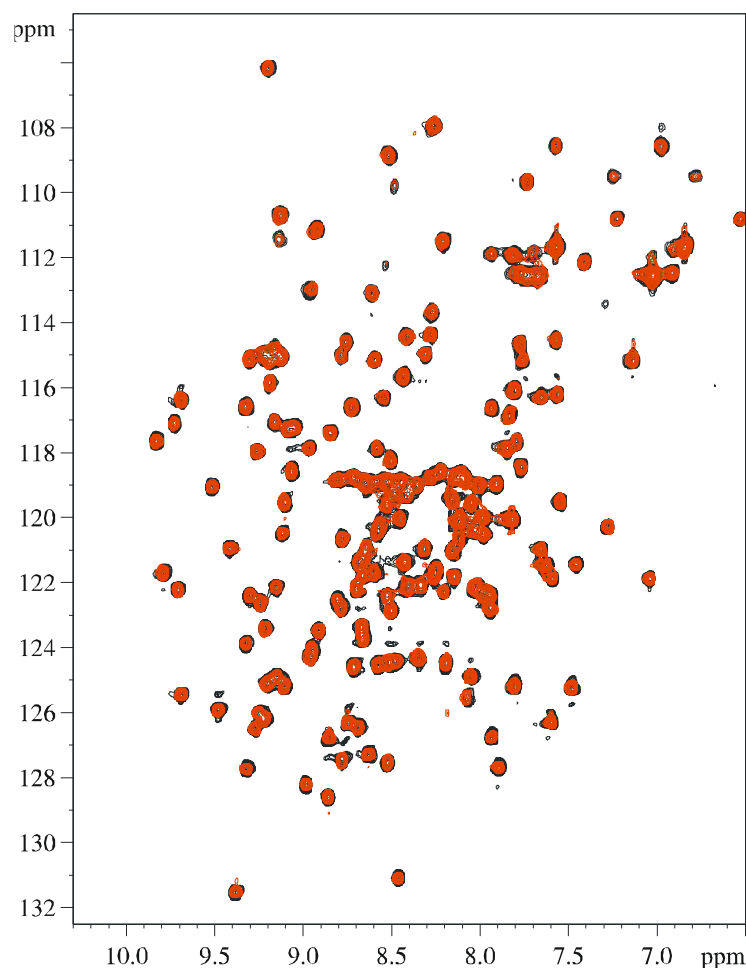
Resonances showing comparatively large changes in chemical shifts were (above 25 Hz in any dimension) N4, E13, M21, V24, V54, Y55, A57, I104 and N134. It was unambiguous to assign the shifted resonances because of their vicinity to the parent resonance. A closer inspection reveals that these residues were distributed over the complete sequence and they mainly consist of either the starting or the ending residue of a secondary structure element. A pH change has induced a small change in the local environment of these residues which caused these shift. Apart from the differences in chemical shifts of certain residues, it was also noted that some of the resonances disappear completely. These disappearing resonances mainly belong to the residues of the  $\beta$ 9- $\beta$ 10 loop in VAT-Nc (residues: A143-G147, *cf.* figure 4-6) and to the flexible terminus (residues: S3, N4, G6 and G184). The disappearance of these resonances is very well possible since the water exchange of the flexible loop is accelerated at

higher pH. In conclusion, the overall structure of VAT-N remains intact as evidenced from the high resemblance in the HSQC pattern.

### 6.2.2. *VAT-N:SsrA*

Florescence in the GFP (Green Florescence Protein), is caused by the presence of the chromophore, resulting from the spontaneous cyclization and oxidation of the sequence Ser65-Tyr66-Gly67. The native protein fold is required for both formation of the chromophore and fluorescence emission <sup>[193]</sup>.

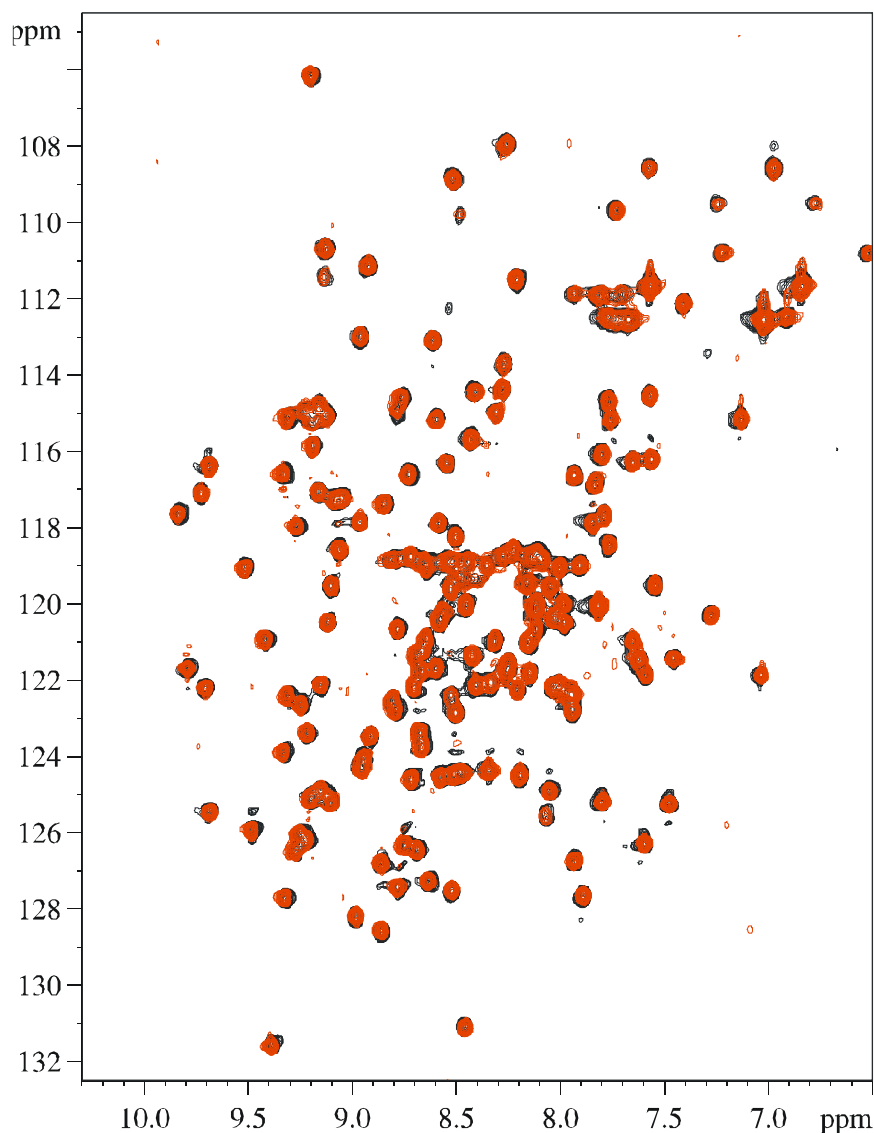
It was observed that VAT acts as unfoldase for SsrA tagged GFP (acronym: GFP-SsrA). SsrA is an unstructured peptide tag consisting of 19 amino acids and fused at the C-terminal end of GFP. Similar studies on the wild type GFP and VAT yielded reduced affinity. Therefore, it was thought that in GFP-SsrA, SsrA first binds to VAT-N which then feeds it to the D1-D2 domain hexamer. Since GFP is attached covalently to the SsrA tag, it is guided to D1, hence, an increased binding was seen. Based on this observation, the titration of VAT-N with the SsrA tag was carried out to monitor the changes in the <sup>15</sup>N-<sup>1</sup>H HSQC spectrum during each step of the titration. The sample conditions were made uniform so as to avoid artifacts resulting from the non-similar sample conditions. <sup>15</sup>N-<sup>1</sup>H HSQC experiments were carried out under identical experimental conditions for each step of titration at 320 K. Titration of <sup>15</sup>N labeled VAT-N and unlabeled SsrA was done in the increasing order of molar ratio of SsrA (*i.e.*, SsrA versus VAT-N ratios of 0.5, 1.0, 2.0 and 3.0 were used). No changes in the <sup>15</sup>N-<sup>1</sup>H HSQC chemical shifts or the appearance of the resonances for VAT-N were observed (figure 6-2) in all the stages of titration. This indicates that no binding interaction between VAT-N and SsrA exists. Therefore, the proposed role of VAT-N in GFP-SsrA binding might be different.



**Figure 6-2:** An overlay of  $^{15}\text{N}$ - $^1\text{H}$  HSQC of the free VAT-N (black) and with a three fold molar excess of SsrA titrated VAT-N (red). No change in the chemical shifts or non-appearance of new signals for all the residues indicates that the SsrA tag does not bind to VAT-N.

### 6.2.3. VAT-N: Ubiquitin

Ubiquitin is one of the highly studied proteins by NMR mainly because of its availability, stability and small size. In our context, the N-terminus of p97-D1-D2 protein complex degrades substrate proteins in Ubiquitin dependent pathway<sup>[194]</sup>. p97-D1-D2 complex is found in the higher eukaryotic cells. It is a well-known AAA protein and has a very resembling structure to VAT<sup>[195]</sup>. Therefore, it was thought that Ubiquitin might bind to the N-domain of VAT, VAT-N. A series of titration of VAT-N with Ubiquitin was carried out in order to have final concentration of Ubiquitin in the molar excess range (0.5, 1.0, 2.0 and 3.0 molar equivalence of unlabeled Ubiquitin). The  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectra were recorded for 320 K and 329 K. Inclusion of the latter temperature was due to the fact that Ubiquitin assumes a partially unfolded state. An overlay of the spectra is shown in figure 6-3.



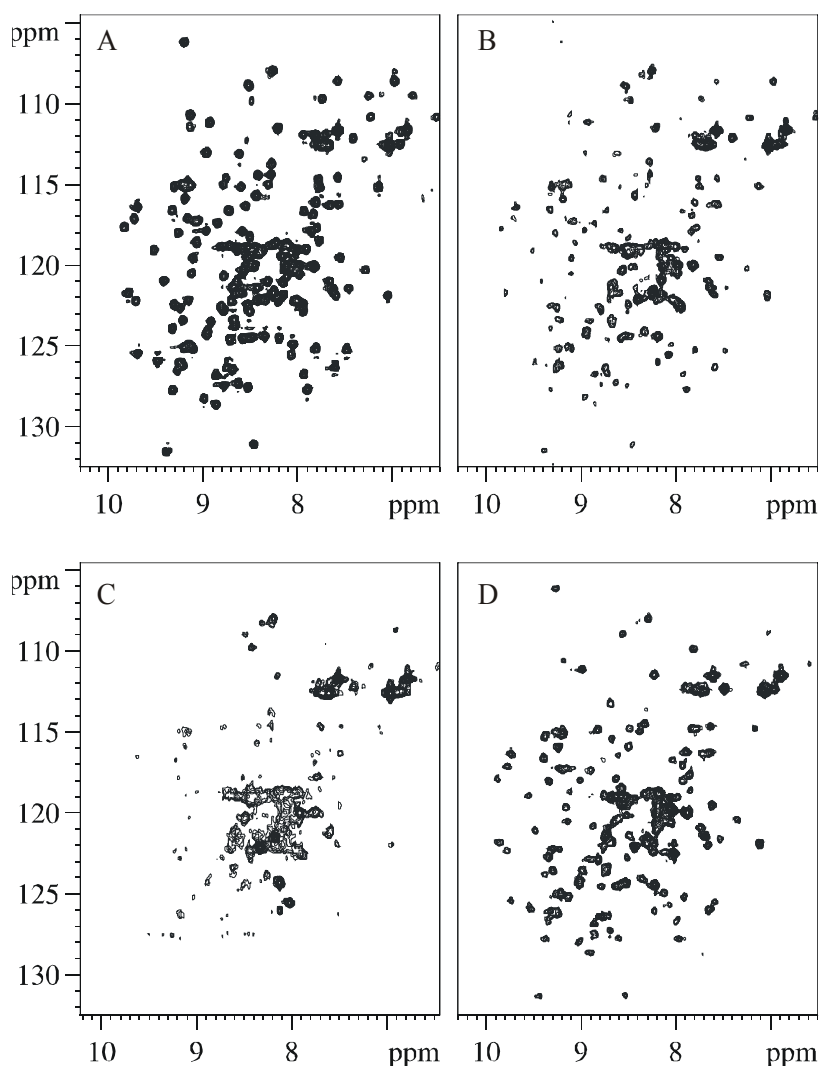
**Figure 6-3:** An overlay of  $^{15}\text{N}$ - $^1\text{H}$  HSQC of the free VAT-N (black) and with a two fold molar excess of Ubiquitin titrated to VAT-N (red). No change in the chemical shifts or non-appearance of the new signals for all the residues indicates that Ubiquitin does not bind to VAT-N. The appearance of the overlay did not change at elevated temperature (329 K).

No chemical shift perturbation or appearance of new resonances, at both temperatures, concludes that there is no binding between VAT-N and Ubiquitin.

#### 6.2.4. VAT-N:casein

Casein is a viscous protein (mainly unstructured) of 23.5 kDa and is derived from bovine milk. Casein is known to stimulate ATPases activity in ClpB and therefore may be regarded as substrate [196]. ClpB is a homologous protein of VAT (belong to the AAA family) and has the N-D1-D2 domain structure. In an exploration for finding a natural substrate of VAT-N,

$^{15}\text{N}$  labeled VAT-N was titrated with 0.5, 1.0 and 2.0 equimolar unlabeled casein. The titrations were carried out at 320 K and 329 K (figure 6-4).



**Figure 6-4:**  $^{15}\text{N}$ - $^1\text{H}$  HSQC of VAT-N titrated casein. VAT-N:casein molar concentrations are in the order, A=1:0, B=1:1 and C= 1:2, respectively, at 320 K. Unusual line widths of most of the residues were seen with increasing concentration, such that in C hardly any peak can be seen. When the temperature of C is increased to 329 K, (D) re-emergence of these resonances was seen (though the line width was more than twice). This observation implies that the oligomerization of VAT-N was induced by casein.

With the addition of 0.5 equivalence of casein (25  $\mu\text{L}$  of 1 mM), at 320 K, resonances of VAT-N were seen to broaden and reached to approximately twice the original linewidth. Broadening of resonances was linear with increasing concentration of titrated casein and at 2.0 equimolar concentration, resonances were too broad to be observed except those belonging to the flexible unstructured free loops. This behaviour suggests that oligomerization



of VAT-N was triggered by casein. This oligomerization pronounces slower correlation time and therefore resonances belonging to the rigid part of the protein disappear due to increased relaxation rate.

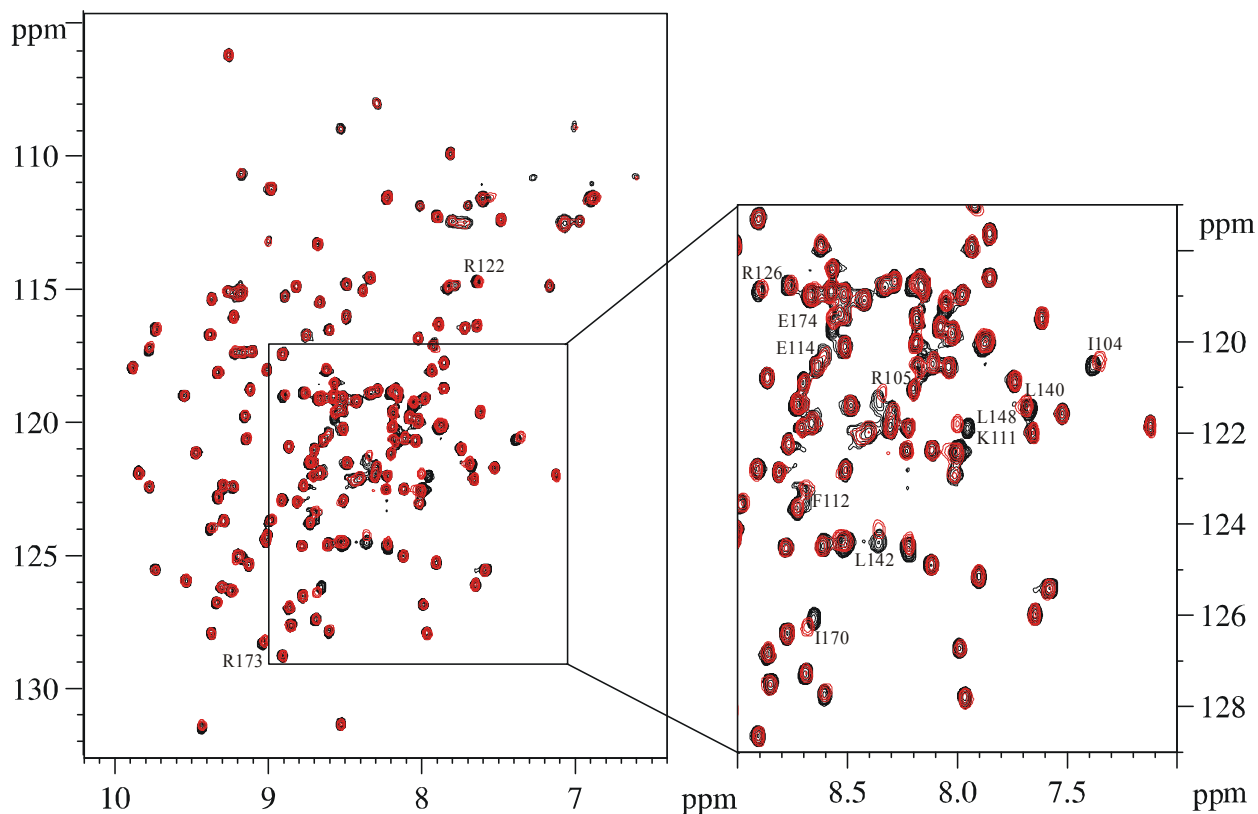
A  $^{15}\text{N}$ - $^1\text{H}$  HSQC comparison of the free and casein titrated VAT-N, at 329 K, gave exactly identical spectrum, although in the latter case the averaged line width of the resonances was almost two times higher. This confirmed that there was no chemical shift perturbation or appearance of new resonances. At the same time, it should be also noted that the increased linewidth in VAT-N was caused due to oligomerization and not by any conformational exchange. Line broadening due to the conformational exchange would not be equally pronounced for each residue. It is hardly plausible that the whole molecule undergoes chemical exchange.

Though the binding studies between VAT-N and casein look attractive at the first sight, it has no functional importance due to its non-specific nature.

#### 6.2.5. *VAT-N:Barstar*

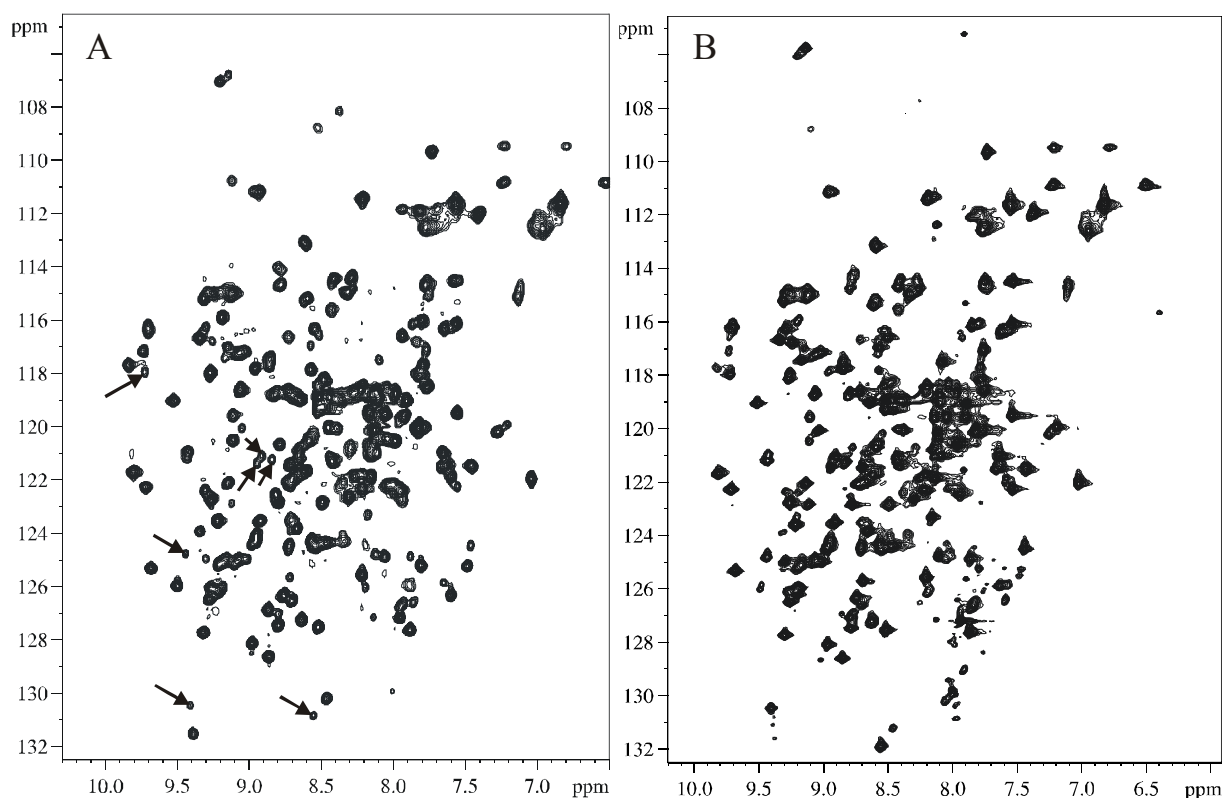
The selection of Barstar, as a substrate for VAT-N, resulted from a biochemical study which showed that thermal precipitation of Barstar was slowed down in the presence of VAT-N. A series of titration of VAT-N with Barstar was carried out in the similar way as was done for earlier three cases.

Upon immediate titration of the first batch of Barstar (0.5 molar equivalence) changes in the chemical shift for the certain residues were observed. These residues are labeled in figure 6-5. Surface mapping reveals these residues belong to the flexible loop  $\beta 7$ - $\alpha 3$  in VAT-Nc (residues: I104-F112) and to the C-terminal residues (residues: I170-E174). Two other residues, R122 ( $\alpha 3$ ) and R126 ( $\alpha 3$ - $\beta 8$  loop), were affected mainly due to the allosteric effects. The first and the last residue of the  $\beta 9$ - $\beta 10$  loop also show chemical shift perturbation. Since this whole loop has already disappeared due to the change in the pH (*vide supra*), any direct evidence of its perturbation in presence of Barstar could not be achieved. Therefore, it manifests that Barstar weakly binds to VAT-N and the binding mainly takes place at the free surface loops present in VAT-Nc.



**Figure 6-5:** Chemical shift perturbations seen immediately upon addition of 0.25 equimolar Barstar to VAT-N sample. Residues showing different chemical environment due to ligand addition mainly belong to the loop region in the C-terminal subdomain, VAT-Nc.

The chemical shift perturbations were expected to be seen more pronounced in the next step of the titration. However, the appearance of new resonances was observed. Initially, newly appearing resonances were weak and gained intensity when the concentration of Barstar reached to 2 molar equivalence of VAT-N (figure 6-6).

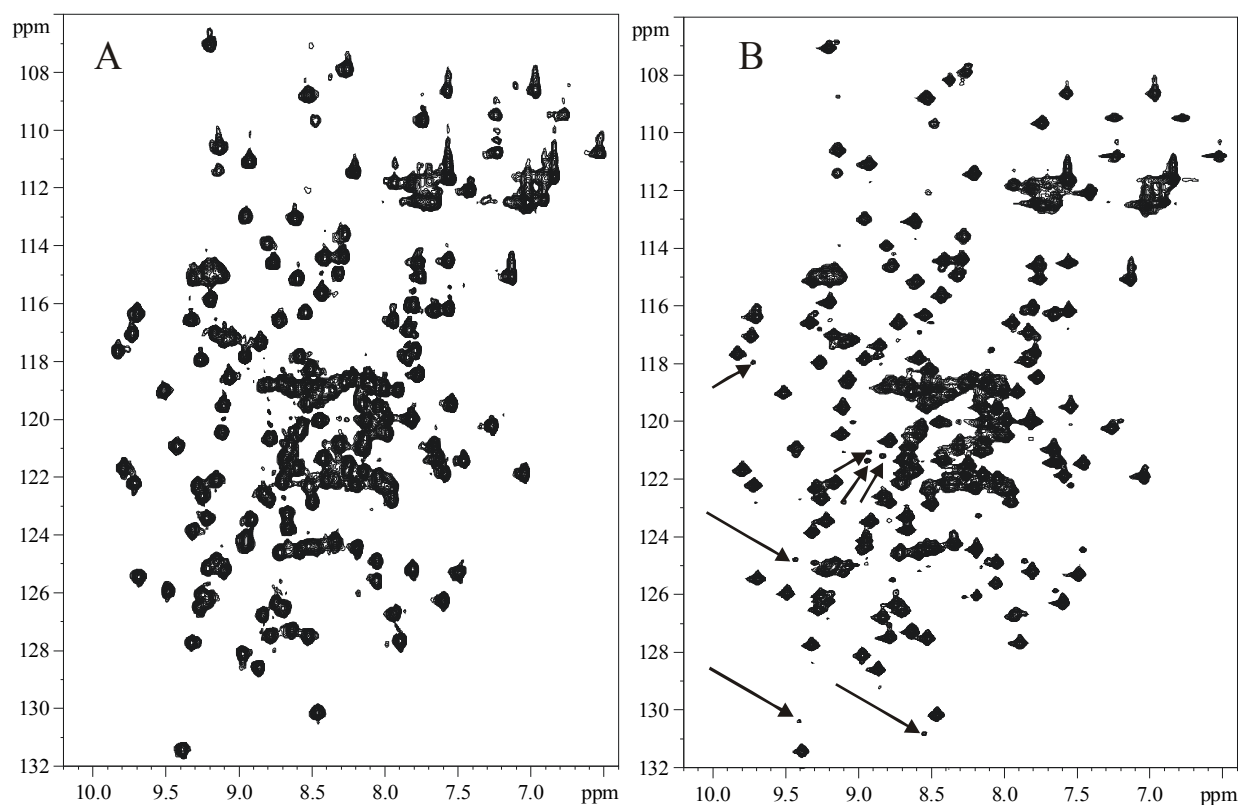


**Figure 6-6:** *Molar excess titration of Barstar to VAT-N results in the appearance of new signals (marked with arrows), which is typical for a tightly bound complex (contradictory to the results in figure 6-5). The equilibrium between the free and the bound state shifts slowly towards the bound state as evidenced from the changed ratio of old signals to new signals (after 24 hrs.: 80:20 (A), after 6 months: 05:95 (B) ).*

Newly appearing resonances were ascribed to the Barstar bound state of VAT-N. These resonances became stronger over a period, indicating that the conformational equilibrium is slowly shifting from free VAT-N to the bound VAT-N. An estimation of the average ratio between the free and the bound state was 80:20, respectively, after 24 hrs of molar excess titration of Barstar. It turned out that this ratio changes to 10:90 over a period of 6 months which can be seen in figure 6-6 (the sample was stored at 276 K meanwhile).

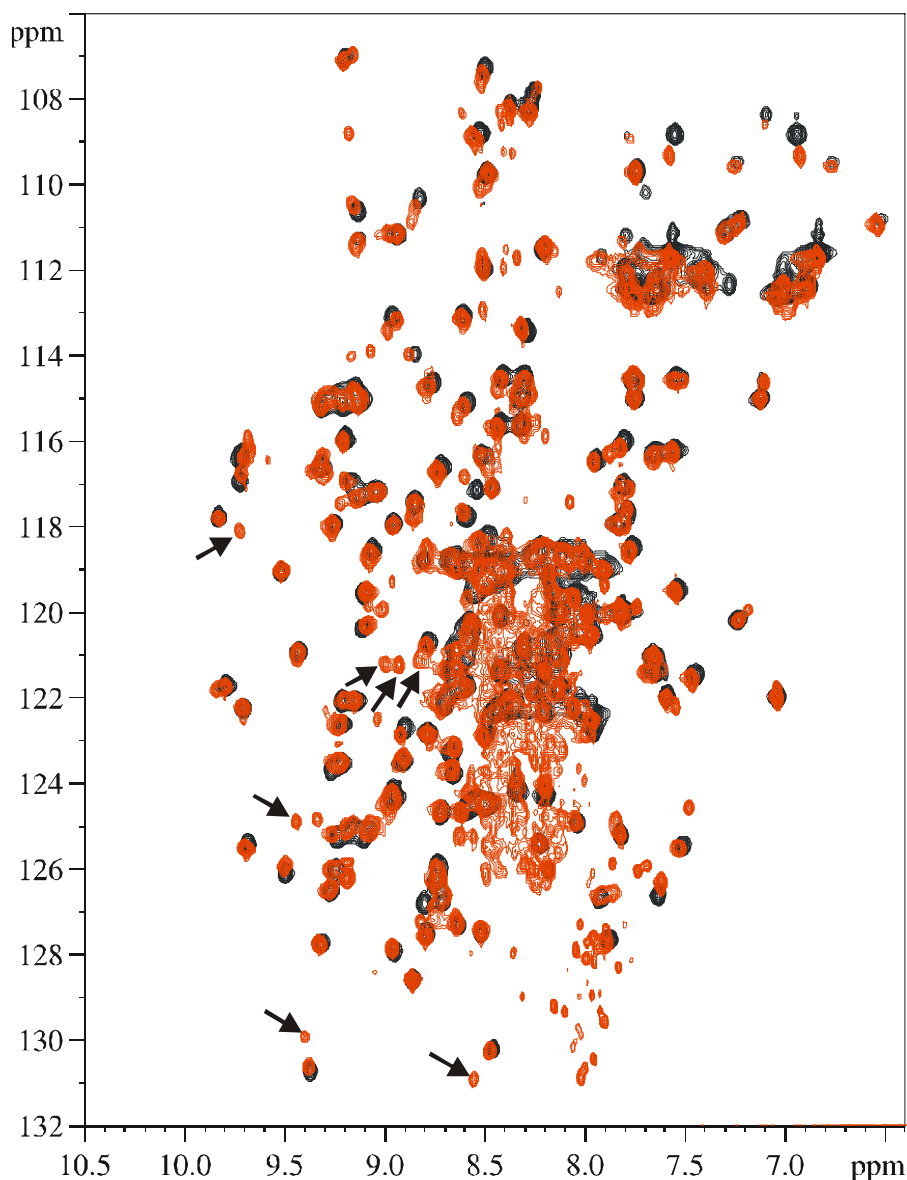
Based on above observation, it was obvious that Barstar binds to VAT-N weakly at low concentrations and later relocates to form a tight complex. The newly appearing resonances were mapped on the surface. They belong to the rigid  $\psi$ 2 loop in the N-subdomain, VAT-Nn. The relocation of Barstar seems unrealistic as the rigid  $\psi$ 2 loop in VAT-Nn and the  $\beta$ 7- $\alpha$ 3 and  $\beta$ 9- $\beta$ 10 loops belonging to VAT-Nc are quite far away in their spatial location. (approximately 45 Å). Barstar, being a small molecule, it is less likely that this kind of mechanism would work, unless Barstar unfolds, before the second effect takes place.

Quite simultaneously, we observed that these newly appearing resonances (appearing in the Barstar titrated VAT-N sample) do occur in the 6 months old clean samples of VAT-N which is shown in figure 6-7 and figure 6-8 (the sample was stored at 276 K during this time). The ratio between the original VAT-N resonances and the newly appearing resonances was approximately 95:5, respectively. The new resonances in VAT-N might correspond to the partially degraded VAT-N.



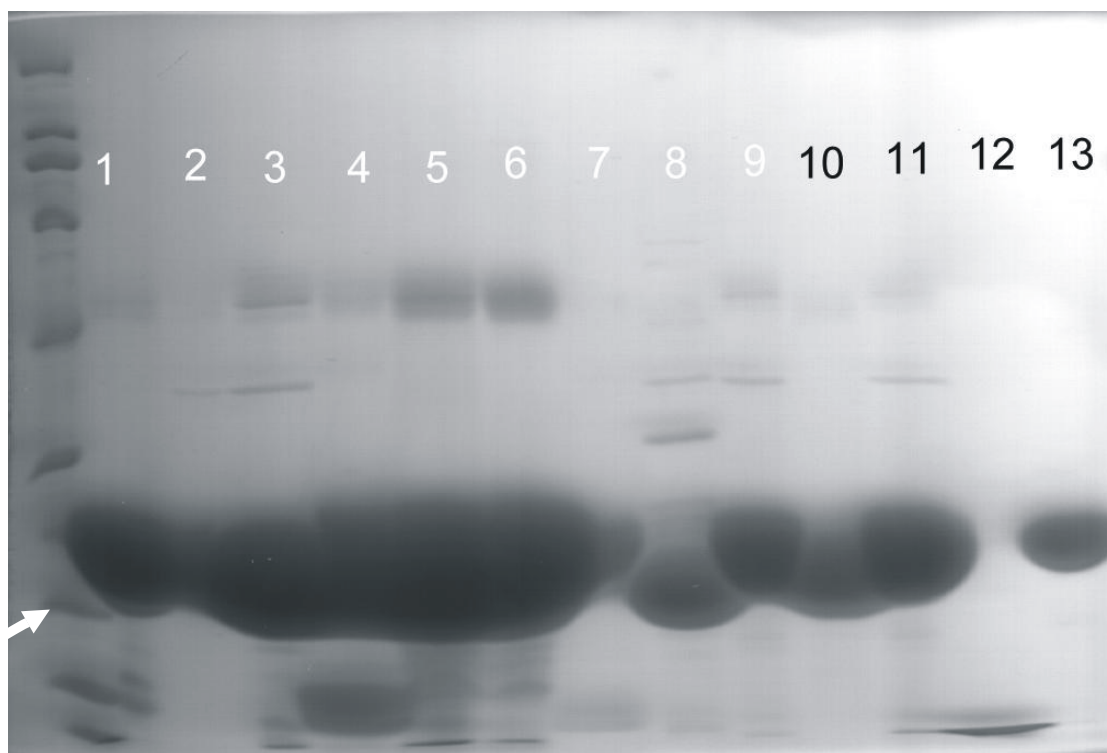
**Figure 6-7:** (A)  $^{15}\text{N}$ - $^1\text{H}$  HSQC of the freshly prepared VAT-N sample. (B) HSQC of the same sample after 6 months during which it was stored at 276 K. Newly appearing signals (marked with arrows) resemble with those found in Barstar titrated VAT-N sample.

In the view of these experimental results, it can be concluded that Barstar triggers degradation of VAT-N and accelerates it.



**Figure 6-8:**  $^{15}\text{N}$ - $^1\text{H}$  HSQC of the freshly prepared VAT-N sample (A) and a sample more than 4 years old (B), both at pH 5.9. Newly appearing signals (marked with arrows) resemble with those found in Barstar titrated VAT-N and are more pronounced compared to figure 6-7.

These various samples of VAT-N were loaded on a native as well as SDS gel to detect possible degradation in the form of new low molecular weight bands. These results of the SDS gel are shown in figure 6-9. Column 3 represents old sample of clean VAT-N stored for 6 months at 276 K. Existence of a small band of a low MW polypeptide chain along 6 kDa marker line indicates partial degradation of VAT-N. Similar band can be seen in a sample which was titrated with Barstar. Interestingly, no protease can be present in the sample of neat Barstar as Barstar gives a single sharp band along 9 kDa. Therefore, the accelerated degradation of VAT-N can only be explained by a certain biological role of Barstar.



**Figure 6-9:** A SDS gel ran on various samples of VAT-N. Left hand side band is for the marker (200, 116, 97, 66, 45, 31, 21, 14, 6.5 kDa). The white arrow represents the monomeric weight of VAT-N. Column 3 represents VAT-N sample that was kept for 6 months at pH 6.7 and at 276 K. Columns 4-7 represents very old samples of VAT-N, columns 8-10 are batched derived from sample 3, Column 11 is VAT-N:Barstar sample, column 12 is titrated neat Barstar, column 13 is freshly prepared unlabeled VAT-N. Similar patches in the column 3 and 11, particularly, around 14 kDa and 6.5 kDa might be due to partial degradation of VAT-N.

### 6.3. Conclusions

Binding studies by NMR provide a primary tool for the location of the binding partner of the biomolecules. These studies are mainly attainable by observing chemical shift perturbations or by observing newly appearing resonances in the  $^{15}\text{N}$ - $^1\text{H}$  HSQC spectrum upon addition of ligand. In an attempt to find natural substrate for VAT-N, we have carried out equimolar and molar excess titrations of VAT-N were carried out with SsrA, Ubiquitin, casein and Barstar. The titration of SsrA and Ubiquitin with VAT-N indicates no binding between VAT-N and substrates. Casein induces oligomerization of VAT-N as became evident from the increased linewidth of all the resonances belonging to the structured residues. Initial low concentration titration of VAT-N and Barstar reveals that the binding takes place at the  $\beta 7$ - $\alpha 3$  and  $\beta 9$ - $\beta 10$  loops belonging to VAT-Nc. With increased concentration of Barstar, surprisingly, additional resonances arise which were also found in a partially degraded sample of titration-free

VAT-N sample. Therefore, it can be concluded that Barstar accelerates degradation of VAT-N. Thus, the exploration for a natural substrate for VAT-N, driven by homology search, failed. Intensive attempts are necessary to find out the binding role of VAT-N.

## 7. Investigation of the Structural Differences in Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC

### 7.1. Somatotropin Release Inhibiting Factor (SRIF)

Somatostains, which are also known as Somatotropin-release inhibiting factors (SRIFs), form a family of cyclo-peptides that are mainly produced by normal endocrine, gastrointestinal, immune and neuronal cells, as well as by certain tumors<sup>[197-199]</sup>. Somatostains peptide contains 14 and 28 amino acids and is generated as C-terminal products of prosomatostatin. Exploratory clinical trials of natural SRIF-14 were carried out for the treatment of a range of conditions, including: diabetes type I and II; hypersecretory tumor, growth hormone-secreting pituitary adenomas, gastrinomas, insulinomas, glucagonomas and vapomas *etc.*<sup>[200, 201]</sup>.

Although these studies indicated that SRIF-14 is efficacious in certain conditions, they also showed that its full therapeutic potential cannot be exploited *in vivo* owing to its rapid proteolytic degradation (plasma half-life of 3 min)<sup>[202]</sup>. The therapeutic limitation of a shorter plasma half-life can be overcome either by a synthesis of short-chain (reduced size) stable analogues with receptor specific selectivity or by achieving metabolic stability by chemical modification, involving the incorporation of D-amino acids or *N*-methylated amino acids.

This led to a search for the synthetic analogues. Synthetic peptides mainly comprising octreotide, lanreotide, and vapreotide (figure 7-1), were successfully used in clinical applications<sup>[203]</sup>. The basic features of these SRIF based peptides are the cyclization by a cysteine disulfide bridge, causing restricted conformational flexibility, and the introduction of D-Trp<sup>8</sup>-Lys<sup>9</sup> in the somatostatin sequence in the *i+1* and *i+2* positions of the  $\beta$ -turn, respectively.



Structure	Name
$  \begin{array}{cccccccc}  & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\  \text{H-Ala-Gly-Cys-Lys-Asn-Phe-Phe-Trp} \\  & & &   & & & & &   \\  & & & \text{OH-Cys-Ser-Thr-Phe-Thr-Lys}  \end{array}  $	<b>Somatostatin 14</b> SST 14
$  \begin{array}{cccccccc}  \text{H-Ser-Ala-Asn-Ser-Asn-Pro-Ala-Met-Ala-Pro} \\    \\  \text{Arg-Lsy-Ala-Gly-Cys-Lys-Asn-Phe-Phe-Trp} \\    \\  \text{OH-Cys-Ser-Thr-Phe-Thr-Lys}  \end{array}  $	<b>Somatostatin 28</b> SST 28
$  \begin{array}{ccc}  \text{Pro-Phe-D-Trp} \\    \quad   \\  \text{Phe-Thr-Lys}  \end{array}  $	<b>Veber-Peptide</b>
$  \begin{array}{cccc}  \text{H-D-Phe-Cys-Phe-D-Trp} \\    \quad   \quad   \\  \text{Thr(ol)-Cys-Thr-Lys}  \end{array}  $	<b>Octreotide (OC)</b> Sandostatin
$  \begin{array}{ccc}  \text{H-D-Phe-Cys-Tyr-D-Trp} \\    \quad   \\  \text{Thr(ol)-Cys-Thr-Lys}  \end{array}  $	<b>Tyr<sup>3</sup>-Octreotide (TOC)</b>
$  \begin{array}{ccc}  \text{H-D-Phe-Cys-Phe-D-Trp} \\    \quad   \\  \text{H}_2\text{N-Trp-Cys-Val-Lys}  \end{array}  $	<b>Vapreotide</b>
$  \begin{array}{ccc}  \text{H-D-}\beta\text{Nal-Cys-Phe-D-Trp} \\    \quad   \\  \text{HO-Thr-Cys-Thr-Lys}  \end{array}  $	<b>Lanreotide</b>
$  \begin{array}{ccc}  \text{H-D-Phe-Cys-Tyr-D-Trp} \\    \quad   \\  \text{NH}_2\text{-Thr-Cys-Lys}  \end{array}  $	<b>TT-232</b>

**Figure 7-1:** Structures of somatostatin 14 and 28 and some commonly used examples of somatostatin analogues like octreotide, vapreotide and lanreotide etc.

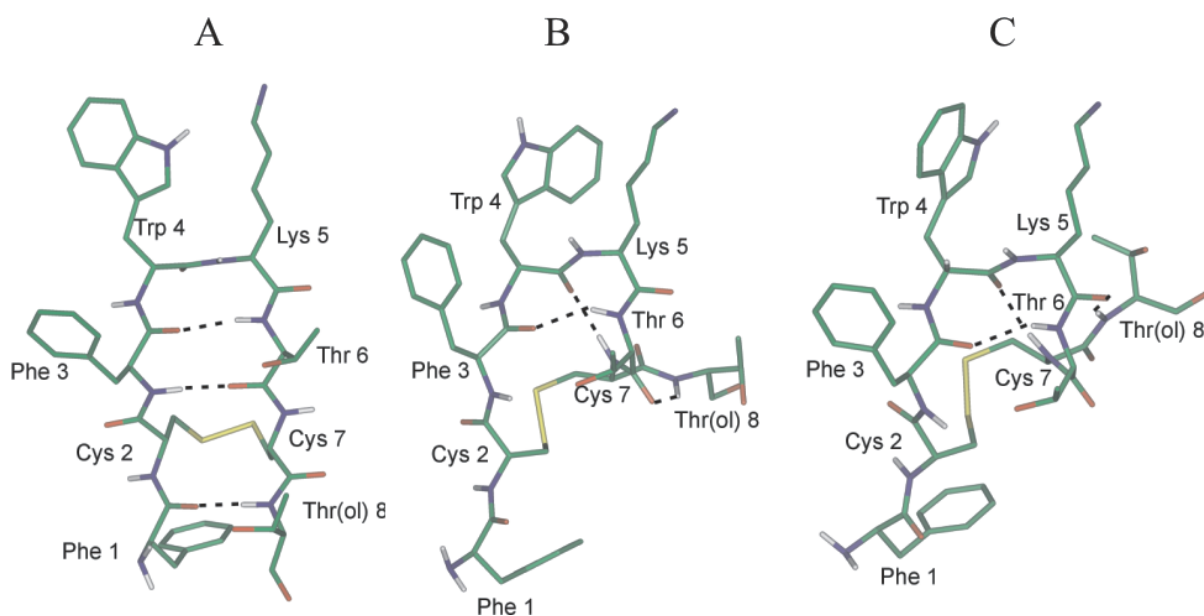
### 7.1.1. Receptor Subtypes

The discovery of the five SRIF receptor subtypes, SSTR<sub>1-5</sub>, in the early 1990s triggered in-depth research into their binding properties and coupling to multiple signaling pathways. Clinically used SRIF analogues, such as octreotide and lanreotide, were found to preferentially bind to the SSTR<sub>2</sub> receptor subtype. All five SRIF receptor subtypes bind their natural ligands SRIF-14 and SRIF-28 in nano-molar affinity whereas synthetic analogues bind with high affinity to a particular receptor subtype (micro-molar affinity)<sup>[204]</sup>. Signal transduction through SRIF receptors is complex and involves binding of somatostatins, SRIF analogues to various SRIF receptor subtypes. Binding of these ligands to SRIF receptor

induces G-protein activation and signaling through various pathways. Activation of several key enzymes occurs as a consequence. Modulation of several proteins occurs along with changes in the intracellular level of calcium and potassium ions [205].

### 7.1.2. Structural Investigations of Somatostatin Analogues

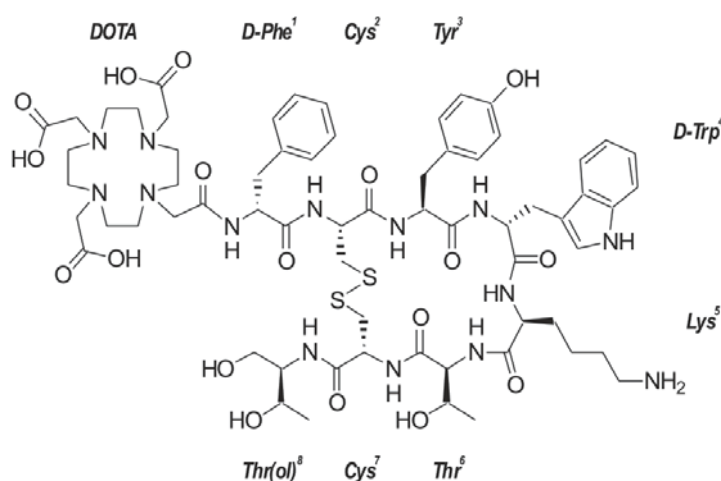
The varied medicinal applications of somatostatin have made it the object of numerous structural investigations using different techniques like NMR spectroscopy and X-ray diffraction [206]. The early NMR studies of van Binst *et al.* in water and in DMSO-*d*<sub>6</sub> solution [207-209] indicated that the octreotide adopts a predominant antiparallel  $\beta$ -sheet conformation characterized by a type II'  $\beta$ -like turn across residues D-Trp<sup>4</sup> and Lys<sup>5</sup>. Similar conformations were obtained in the water-methanol solvent system [210]. In a recent study, Melacini *et al.* [211] have used NOE restraints for molecular dynamics calculations. Violations of NOE distance and  $^3J_{\text{H}^{\text{N}}\text{H}^{\alpha}}$  dihedral angle restraints showed that the NMR data on octreotide could not be explained by a single conformation. Instead, Melacini *et al.* found equilibrium between antiparallel  $\beta$ -sheet structures and conformations in which the C-terminal residues fold into a  $3_{10}$ -helix-like array or a similar helical ensemble. Similar structures were also found in an X-ray diffraction study [206]. These structures are shown in figure 7-2.



**Figure 7-2:** Backbone structures of the octreotide derived from X-ray crystallography. Conformations found are (A) antiparallel  $\beta$ -sheet, (B and C)  $3_{10}$ -helix (CSD: YICMUS, PDB: 1SOC, 2SOC) [206]. The different arrangement of Cys<sup>7</sup> and Thr<sup>8</sup>(ol) leads to differences in the (B) and (C) structure. Similar conformations are also found in the NMR studies.

### 7.1.3. Radio Labeled Analogues

Apart from the peptidic and non-peptidic agonists and antagonists, radio labeled analogues of regulatory peptides have been used recently for the contrast enhanced diagnosis and radiotherapy of primary tumors and their metastases. [203, 212]. Further modifications for radio-labeling of these peptides were used successfully for the *in vivo* localization of SRIF receptor-positive tumors [213]. The isotopes <sup>111</sup>In, <sup>99m</sup>Tc, <sup>186/188</sup>Re, <sup>66/67/68</sup>Ga, <sup>177</sup>Lu and <sup>90</sup>Y were amongst the radionuclides which were tested for this purpose. The primary structure of such peptides connected to a metal ion complex is shown in figure 7-3 for the most widely used macrocyclic chelator DOTA (1, 4, 7, 10-tetraazacyclododecane-1, 4, 7, 10-tetraacetic acid). DOTA usually confers very high kinetic and thermodynamic stability to its metal complexes. <sup>111</sup>In<sup>III</sup>-DOTATOC (*i.e.*, [<sup>111</sup>In<sup>III</sup>-DOTA, Tyr<sup>3</sup>]-octreotide) and <sup>90</sup>Y<sup>III</sup>-DOTATOC have been shown to be excellent targeting and therapeutic agents in animal models and in patients [212, 214, 215]. The peptidic part of the M<sup>III</sup>-DOTATOC compounds consists of [Tyr<sup>3</sup>]-octreotide.



**Figure 7-3:** Primary structure of DOTATOC. The peptidic part consists of D-Phe<sup>1</sup>-Tyr<sup>3</sup>-octreotide, the DOTA chelator is attached to the N terminus via an amide bond. In this study a metal ion (Gallium, Yttrium or Europium, not shown here) is complexed by the four DOTA nitrogens and additional carboxyl oxygens, depending on the ionic radius of the metal ion (*cf.* figures 7-11 & 7-12).

In several studies the properties of DOTATOC labeled with <sup>67</sup>Ga, <sup>111</sup>In and <sup>90</sup>Y were investigated *in vitro* and *in vivo*. Specifically, the IC<sub>50</sub> value of Ga<sup>III</sup>-DOTATOC, measured in a (subtype SSTR<sub>2</sub>) receptor binding assay with [<sup>125</sup>I]-[Leu<sup>8</sup>, D-Trp<sup>22</sup>, Trp<sup>25</sup>]-somatostatin-28 as a radioligand, was about five times higher than the value of Y<sup>III</sup>-DOTATOC [216]. In

addition, biodistribution data in an AR4-2J bearing nude mouse model showed differences for the two radiopeptides, with a more than two times higher tumor uptake for <sup>67</sup>Ga<sup>III</sup>-DOTATOC. Moreover, the kidney uptake of <sup>67</sup>Ga<sup>III</sup>-DOTATOC was significantly lower than the one for <sup>90</sup>Y<sup>III</sup>-DOTATOC. The very good performance of <sup>67</sup>Ga<sup>III</sup>-DOTATOC *in vitro* and in the animal model prompted different groups to study <sup>68</sup>Ga<sup>III</sup>-DOTATOC as a PET tracer<sup>[217, 218]</sup>. Gallium-68 is especially attractive, since it has a 68 min half-life time and is generator produced, with a very favourable 280 d half-life time of the parent isotope <sup>68</sup>Ge. The human data look indeed very promising and parallel the preclinical results.

#### 7.1.4. Scope of the Present Work

The reasons for the significant differences between Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC are still not fully understood. The metal ion dependence for kidney uptake may originate from geometrical differences within the metal<sup>III</sup>-DOTA complex, affecting their biophysical properties. In all previously mentioned structural investigations, the focus of structure determination was solely on the peptide sequence, without the metal chelator attached. Thus, structural studies of the explicit metal-DOTA-peptides could form a basis for understanding the intricate *in vivo* behavior of the different metal-DOTATOC combinations.

This chapter discusses structural results based on <sup>1</sup>H- and <sup>13</sup>C -NMR data of the Ga<sup>III</sup> and Y<sup>III</sup>-complexes of DOTATOC in aqueous solution. While the peptidic parts of Ga<sup>III</sup>-DOTATOC and Y<sup>III</sup>-DOTATOC exhibit similar solution conformations, *i.e.*, a fast equilibrium of a <sub>310</sub>-helical- and a β-sheet-like structure, the specific metal coordination geometry in Y<sup>III</sup>-DOTA-D-Phe<sup>1</sup> causes an additional slow *cis-trans* isomerisation about the DOTA-D-Phe<sup>1</sup> amide bond. Additionally, NMR studies on Eu<sup>III</sup>-DOTATOC were carried out. Due to paramagnetic Eu<sup>III</sup>, we observe hyperfine shifts, which make spectral assignment and structure calculation almost an impossible task.

## 7.2. Experimental Conditions

Samples of DOTATOC were obtained from the laboratory of Prof. Mäcke, Basel and were synthesized according to previously published procedures<sup>[216]</sup>.

Metal<sup>III</sup>-DOTATOC samples were prepared in 90:10 H<sub>2</sub>O:D<sub>2</sub>O solvent. All NMR experiments were performed on a BRUKER 600 MHz NMR spectrometer, equipped with a triple resonance probe with gradient pulse facility and temperature control unit. Optimum resolution in the 1D spectrum was the criterion for the choice of the temperature in each case. Additional measurements at 275 K were performed for Y<sup>III</sup>-DOTATOC because of the narrower

linewidth of the D-Phe<sup>1</sup> amide signals. Exact sample conditions which were used for Ga<sup>III</sup>, Y<sup>III</sup> and Eu<sup>III</sup>-DOTATOC during NMR studies are given in table 7-1.

Table 7-1: Sample conditions of M<sup>III</sup>-DOTATOC during NMR studies

SRIF analogues	pH	Temperature [K]	Molar Mass [g/mol]	Concentration [mM]
Ga <sup>III</sup> -DOTATOC	6.0	290	1488.36	9.0
Y <sup>III</sup> -DOTATOC	6.0	275 and 290	1507.54	9.7
Eu <sup>III</sup> -DOTATOC	5.0	275	1570.60	9.3

Standard pulse programs were used for data acquisition, but occasional modifications were incorporated in order to suppress artifacts. A WATERGATE<sup>[219, 220]</sup> sequence was used in all NMR experiments for effective water signal suppression. The D<sub>2</sub>O signal was used throughout all experiments for achieving a field-frequency lock condition. All spectra were calibrated with 3-(trimethylsilyl) propionic acid sodium salt (TSPA) as an external standard at 0 ppm in the proton dimension, whereas carbon chemical shifts were calibrated indirectly<sup>[221-223]</sup>. Once proton shifts are calibrated, the heteronuclear chemical shifts can be directly calibrated by using the following equation.

$$\nu_0^X = \nu_0^H \frac{\gamma_X}{\gamma_H} \quad [7-1],$$

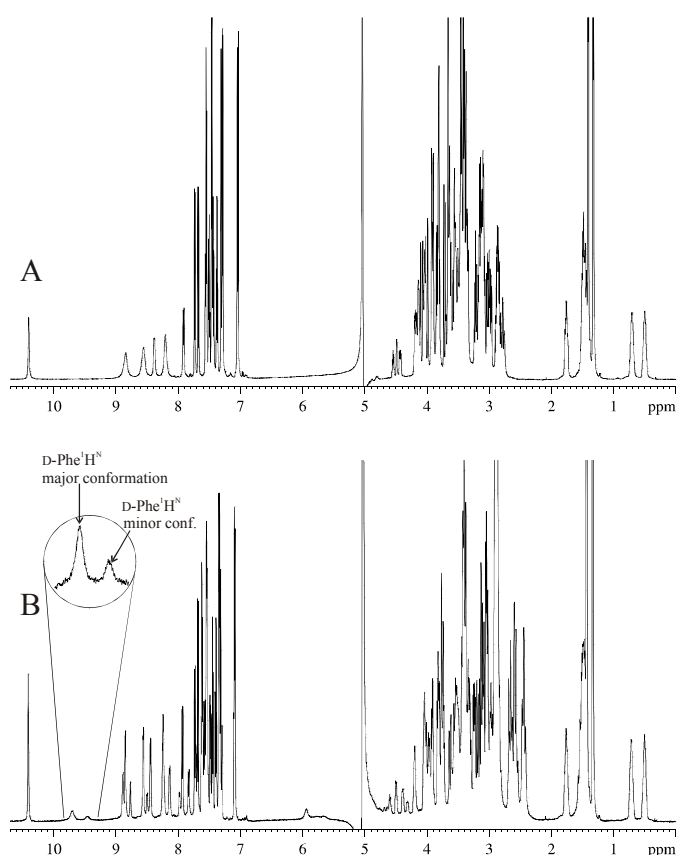
where  $\nu_0$  is the absolute frequency of 0 ppm for the nucleus (X: heteronuclei and H: proton) and  $\gamma$  is the gyromagnetic ratio of the respective nuclei. The values of  $\gamma_X/\gamma_H$  for external standard TSPA correspond to 0.25144954 (for <sup>13</sup>C) and 0.10132900 (for <sup>15</sup>N).

Homonuclear 2D NMR experiments like TOCSY<sup>[224]</sup>, DQF-COSY<sup>[225]</sup> and E.COSY<sup>[226]</sup> were used for <sup>1</sup>H chemical shifts assignment. <sup>13</sup>C chemical shifts were determined from heteronuclear 2D HSQC<sup>[227]</sup> and HMBC<sup>[228]</sup> experiments. Distance restraints were derived from 2D offset compensated ROESY<sup>[229]</sup> (80 ms mixing time) and NOESY<sup>[230]</sup> experiments with 100 ms NOE mixing time. Data processing and analysis were performed using Bruker XWINNMR software (*version 3.2*) with standard data processing tools and baseline correction.

### 7.3. Results and Discussion

#### 7.3.1. <sup>1</sup>H NMR and Spectral Assignments

The 1D proton NMR spectra of Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC are shown in figure 7-4. From a set of 2D homo- and heteronuclear NMR experiments (*vide supra*), all <sup>1</sup>H and <sup>13</sup>C resonances of Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC could be assigned, except for the highly symmetric DOTA parts, where no unambiguous chemical shift assignment was possible. In the case of Eu<sup>III</sup>-DOTATOC, paramagnetic Eu<sup>III</sup>- ion is causing a shift of resonances (*hyperfine shift*) and thus the spectral width of the 1D proton NMR of Eu<sup>III</sup>-DOTATOC ranges from +35 to -20 ppm (figure 7-5 (A)).

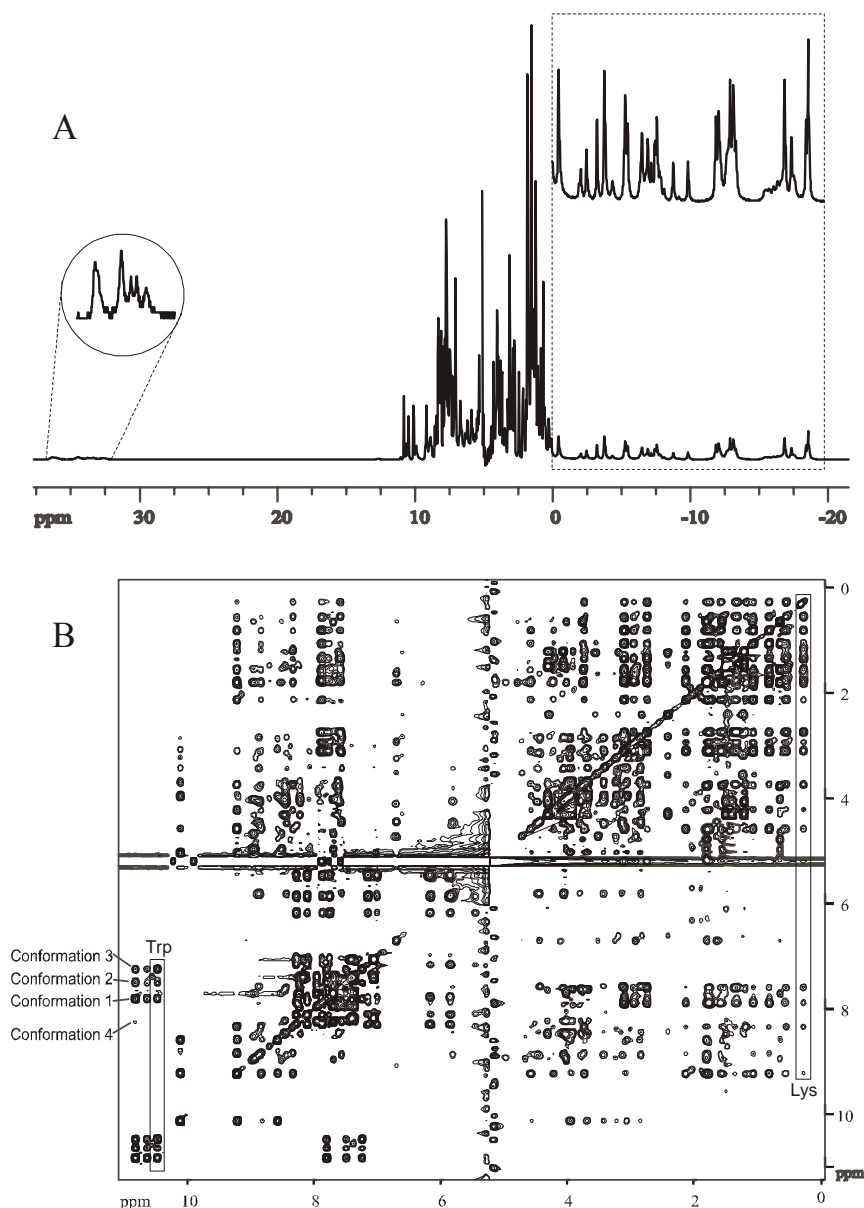


**Figure 7-4:** 1D proton NMR spectra of (A) Ga<sup>III</sup>-, and (B) Y<sup>III</sup>-DOTATOC. For Y<sup>III</sup>-DOTATOC, the D-Phe<sup>1</sup>H<sup>N</sup> resonance appears at ca. 9.5 ppm, a much higher value than expected for an amide proton. In addition, the spectrum shows the presence of two signal sets, most clearly for the D-Phe<sup>1</sup>H<sup>N</sup> resonance (insert).

#### 7.3.2. Characterization of Eu<sup>III</sup>-DOTATOC

Eu<sup>III</sup> has an outer shell electronic configuration [Xe] 4f<sup>4</sup> 6s<sup>2</sup>. The four unpaired electrons in the outermost *f* shell make Eu<sup>III</sup>-DOTATOC paramagnetic. Paramagnetic lanthanides induce

changes in the relaxations and/or the chemical shifts of protons on the vicinity<sup>[231]</sup> (in this case DOTATOC). With the low spin population in the outermost  $f$  orbital (e. g.  $f^2$ ), the Europium<sup>III</sup> ion influences mainly the chemical shift (causing hyperfine shifts).



**Figure 7-5:** 1D proton NMR spectra of Eu<sup>III</sup>-DOTATOC (A). Paramagnetic influence of Eu<sup>III</sup> ion causes hyperfine shifts. (B) 2D TOCSY spectra of Eu<sup>III</sup>-DOTATOC. Presence of at least four exchanging conformations can be seen in this case. Huge hyperfine shifts as well as many conformations in solution make further study of Eu<sup>III</sup>-DOTATOC practically impossible and less relevant with respect to its bioactivity.

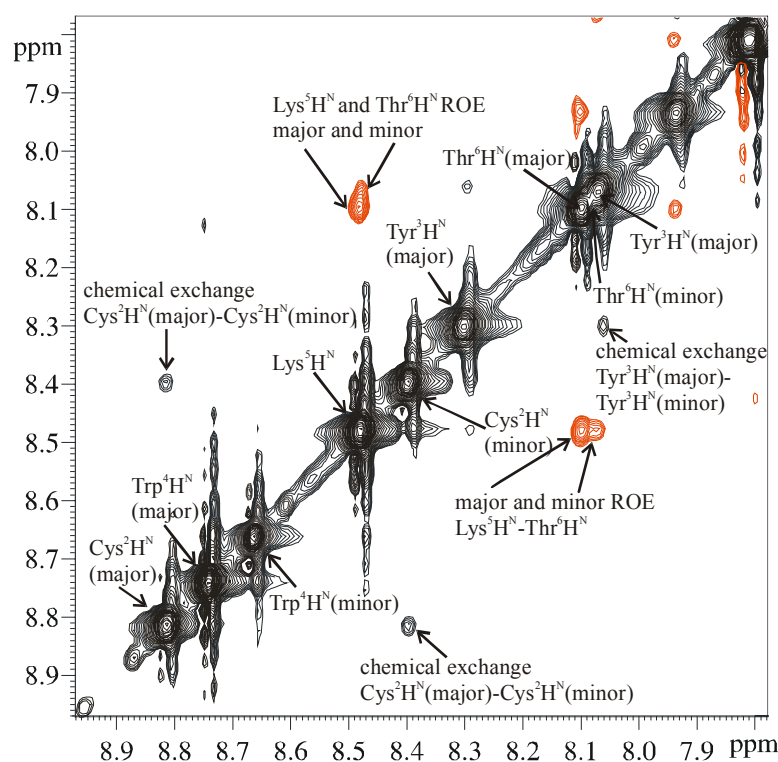
Whereas, the highly spin populated orbitals (e. g.,  $f^7$ ) contribute more to faster relaxation compared to the chemical shift changes<sup>[231]</sup>. We have seen these effects in the 1D proton NMR of Eu<sup>III</sup>-DOTATOC. Due to paramagnetic Eu<sup>III</sup> in the vicinity, all the proton chemical

shifts of the DOTATOC are dispersed from +35 to -20 ppm (figure 7-5 (A)). The spectral assignment of this compound thus became a challenging task.

Apart from that, the TOCSY correlation spectra (figure 7-5 (B)) suggest the presence of at least four conformations. In combination with the large hyperfine shifts, this fact strongly limited our attempts to further investigate the structure of Eu<sup>III</sup>-DOTATOC.

### 7.3.3. Characterization of Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC

The 1D proton spectrum of Ga<sup>III</sup>-DOTATOC exhibits just a single set of NMR signals (figure 7-4(A)). In contrast, the <sup>1</sup>H spectrum of Y<sup>III</sup>-DOTATOC shows a second signal set consisting of weaker resonance lines, most clearly observable for the downfield signal of D-Phe<sup>1</sup>-H<sup>N</sup> (figure 7-4(B)). At 290 K the ratio between the two signal sets is 67:33, determined by integration of several carefully deconvoluted resonances. In the ROESY spectrum of Y<sup>III</sup>-DOTATOC weak exchange cross-peaks can be observed between the two sets of signals, indicating the existence of slow exchange between them<sup>[232]</sup> as shown in figure 7-6.

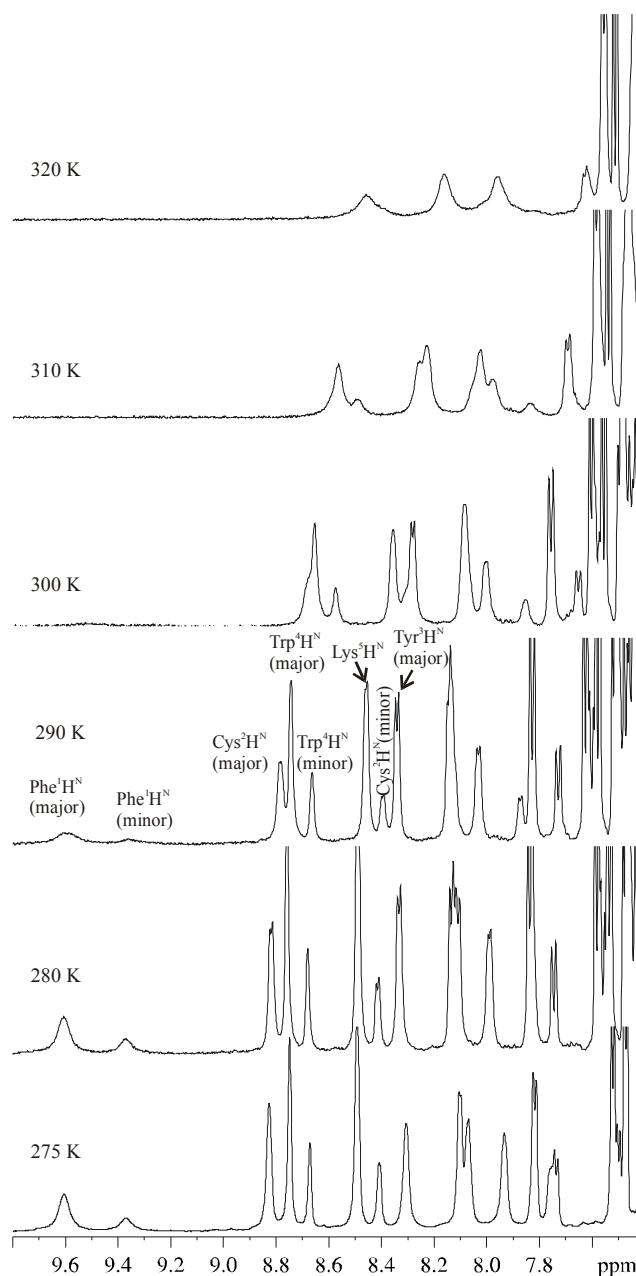


**Figure 7-6:** Contour plot of low field region of the 2D ROESY spectrum of Y<sup>III</sup>-DOTATOC (290 K). Positive exchange cross-peaks (black) and negative NOE cross-peaks (red) can be distinguished, demonstrating the existence of two slowly interchanging conformations (cf. annotations).



Therefore this double signal set clearly represents two different solution conformations for Y<sup>III</sup>-DOTATOC, slowly interconverting on the NMR time scale, which will henceforth be referred to as *major* and *minor* conformation denoting the more and less populated conformer, respectively.

The existence of such separate signals sets has not been reported in earlier NMR studies on similar DOTA model compounds<sup>[233]</sup>. The ratio of the two conformers was found to be temperature dependent (figure 7-7).



**Figure 7-7:** Temperature dependence of the <sup>1</sup>H NMR spectrum of Y<sup>III</sup>-DOTATOC. Upon temperature increase from 275 K to 320 K a line broadening due to proton exchange with the solvent can be observed. However, coalescence of the two conformations does not occur in <sup>13</sup>C-<sup>1</sup>H HSQC spectra up to 330 K (table 7-2).

At lower temperature (275 K), clear and distinct resonances (average proton linewidth  $\sim$  4-6 Hz) are observed for both conformers, with a ratio of 55:45. Upon temperature increase a broadening of the amide proton resonances is observed, with an average proton linewidth of  $\sim$  20 Hz at 310 K, due to an accelerated exchange of the H<sup>N</sup> protons with the solvent.

At the same time, it should be noted that no coalescence occurs between the various resonances of the *major* and *minor* conformers even at temperatures up to 330 K, as can be judged from well-resolved 2D <sup>13</sup>C-<sup>1</sup>H HSQC spectra (table 7-2).

**Table 7-2:** Chemical shift difference [Hz] and the coalescence for the *major* and *minor* conformation of Y<sup>III</sup>-DOTATOC evaluated from 1D proton NMR for H<sup>N</sup> signals and from <sup>13</sup>C-<sup>1</sup>H HSQC for H $\beta$  resonances of Cys<sup>2</sup>.

Temperature [K]	D-Phe <sup>1</sup> H <sup>N</sup> ( <i>major</i> )- D-Phe <sup>1</sup> H <sup>N</sup> ( <i>minor</i> )	Cys <sup>2</sup> H <sup>N</sup> ( <i>major</i> )- Cys <sup>2</sup> H <sup>N</sup> ( <i>minor</i> )	Cys <sup>2</sup> H $\beta^u$ ( <i>major</i> )- Cys <sup>2</sup> H $\beta^u$ ( <i>minor</i> )	Cys <sup>2</sup> H $\beta^d$ ( <i>major</i> )- Cys <sup>2</sup> H $\beta^d$ ( <i>minor</i> )
290	145	236	141	179
300	145	224	140	168
310	--	212	135	159
315	--	Not Available	130	151
320	--	203	128	141
330	--	--	111	132

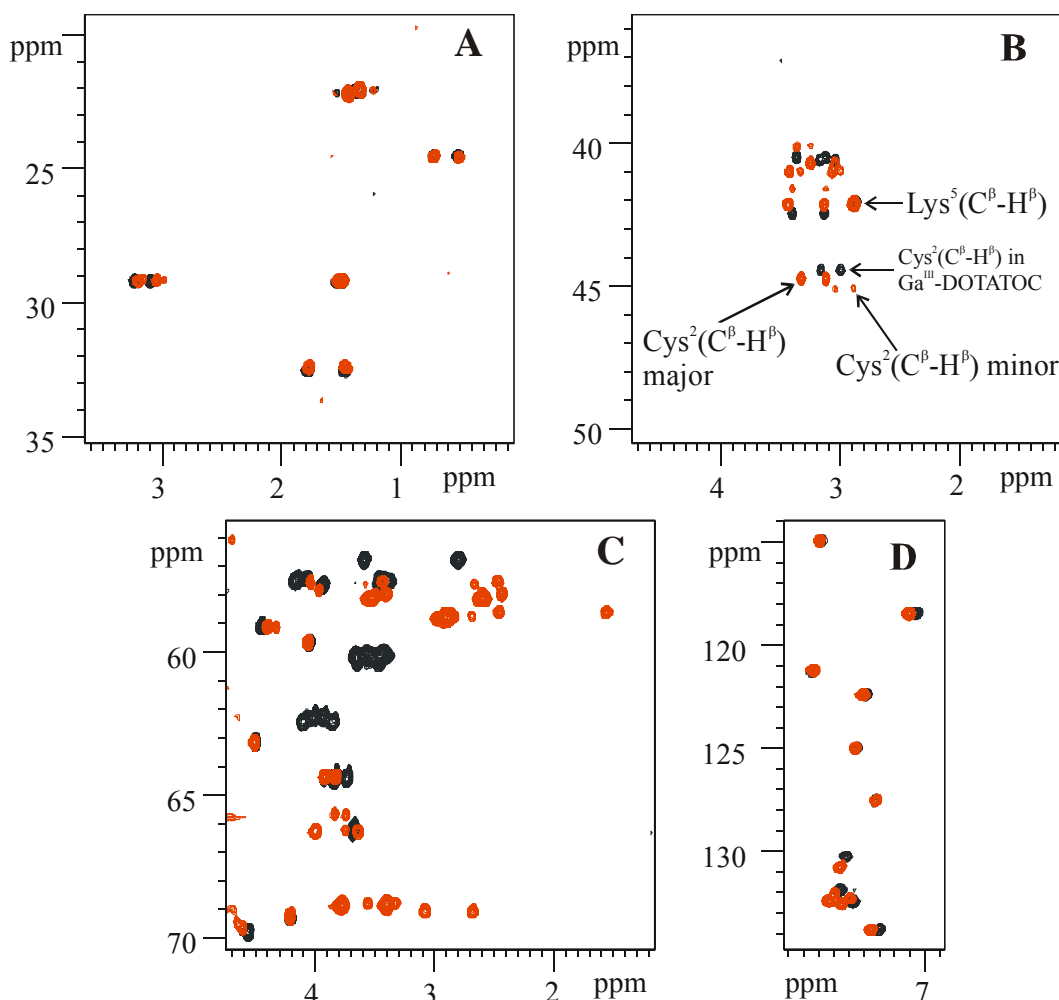
Dash indicates absence of data due to increased water exchange,

<sup>u</sup> and <sup>d</sup> corresponds to upfield and downfield on the chemical shift scale, respectively.

Compared to the other amide signals, the D-Phe<sup>1</sup>-H<sup>N</sup> protons also show an unusual line broadening, in addition to their pronounced downfield shift. Both features can be attributed to a complex formation between the carbonyl oxygen of the DOTA-D-Phe<sup>1</sup> peptide bond and the metal ion, as seen in the X-ray crystal structure of the model peptide Y<sup>III</sup>-DOTA-D-Phe-NH<sub>2</sub> [216] (figure 7-9). The association of the amide carbonyl with the metal ion causes the D-Phe<sup>1</sup> amide proton to resonate further downfield, together with an increase of its acidity and hence solvent exchange rate, resulting in a larger linewidth compared to other amide resonances in the peptide.

The <sup>1</sup>H and <sup>13</sup>C chemical shift differences between corresponding atoms of the two conformations are largest in the vicinity to DOTA (*cf. sections: 9.3.3-9.3.6*). Specifically, the <sup>1</sup>H chemical shift differences between the two conformers follow the order D-Phe<sup>1</sup>  $\approx$  Cys<sup>2</sup> > Tyr<sup>3</sup> > Cys<sup>7</sup> > D-Trp<sup>4</sup>, whereas they are practically absent for Lys<sup>5</sup> and Thr<sup>6</sup>. The <sup>13</sup>C shifts

behave in a similar way: the  $\alpha$  and  $\beta$  carbons of D-Phe<sup>1</sup>, Cys<sup>2</sup>, Tyr<sup>3</sup> and Cys<sup>7</sup> show two well-separated carbon chemical shifts indicating two different environments, whereas the other carbon atoms of the peptide part remain essentially unaffected. Interestingly, L-Thr(ol)<sup>8</sup>, though being close to D-Phe<sup>1</sup> in the primary structure of the peptide, nevertheless shows only one signal set, except for a small chemical shift difference for its amide proton.



**Figure 7-8:** Comparison of the,  $^{13}\text{C}$ - $^1\text{H}$  HSQC patterns of Ga<sup>III</sup>- (black) and Y<sup>III</sup>-DOTATOC (red). Panels (A) and (B) display the aliphatic region containing mainly CH<sub>3</sub> and CH<sub>2</sub> correlations. Panel (C) shows the C<sup>α</sup>-H<sup>α</sup> correlations and the crowded DOTA region, panel (D) depicts the aromatic region. Agreement between the chemical shifts of peptidic protons and carbons of Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC (panel (A), (B) and (D)) suggest that the peptide conformation is very similar for both compounds. Dispersed chemical shifts in the DOTA region (panel (C)) are indicative of the presence of the different conformations near this region. Presence of only 16 carbons for the DOTA region of Y<sup>III</sup>-DOTATOC reveals that the conformational exchange affects only to certain region of DOTA suggesting an amide cis-trans isomerisation across the linker, i.e., the (DOTA)CH<sub>2</sub>CO-D-Phe<sup>1</sup>H<sup>N</sup> bond.

Chemical shifts could not be unambiguously assigned for the DOTA moiety due to its high symmetry and resulting spectral complexity, therefore, chemical shift differences between the two conformers in the DOTA ligand could not be determined. However, one of the DOTA protons in the major conformation resonates at a characteristic upfield shift of 1.54 ppm (corresponding carbon shift: 58.49 ppm). This shift can only be explained if the proton is spatially close to the D-Phe<sup>1</sup> phenyl ring and located above the ring plane, thus being influenced by the anisotropic ring current. Indeed, an NOE cross-peak can be observed between this specific DOTA proton and the H $\delta$  and H $\epsilon$  protons of D-Phe<sup>1</sup>, as well as NOEs between aromatic protons of D-Phe<sup>1</sup> and other DOTA protons in the major conformation (figure 7-8).

#### 7.3.4. Identification of the Two Conformations of Y<sup>III</sup>-DOTATOC

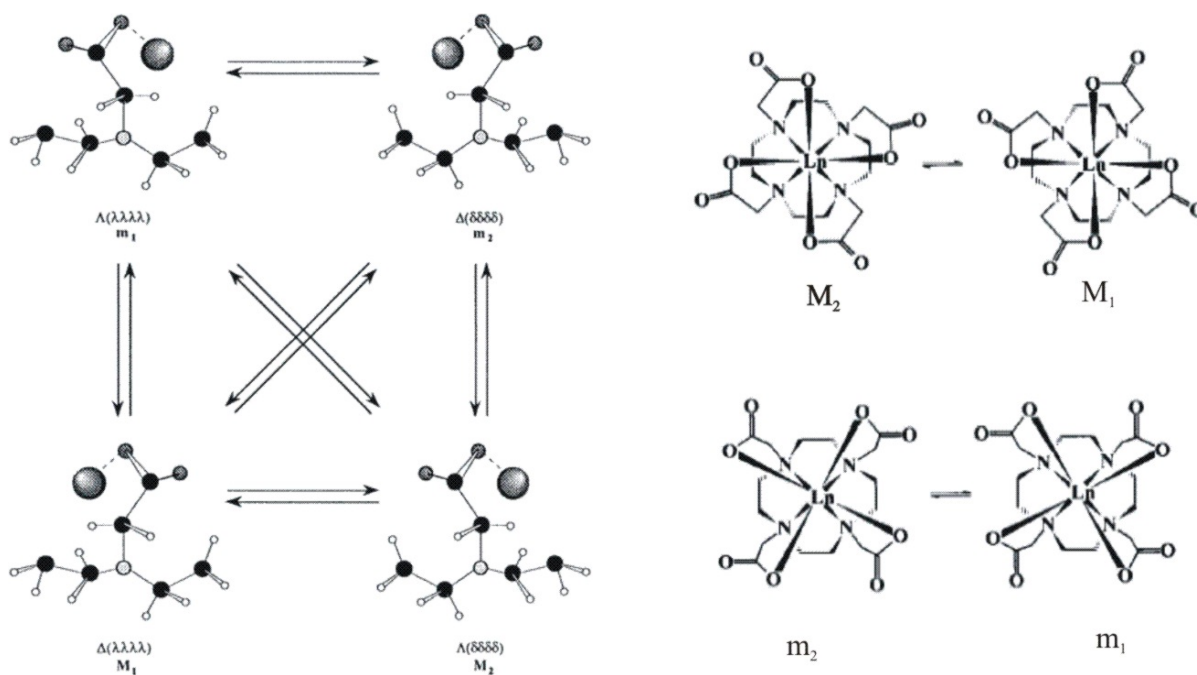
There are three possible explanations for the two conformations in the Y<sup>III</sup>-DOTATOC, namely,

- (1) Two slowly exchanging conformations of the peptide part,
- (2) The two well-known *m/M* diastereomeric conformations of the chelator often found in Ln<sup>III</sup>-DOTA complexes, or
- (3) A *cis-trans* isomerisation occurring at the amide bond in the linker (*i.e.*, between the carboxylic carbon of the acetate sidechain of Y<sup>III</sup>-DOTA and the amide nitrogen of D-Phe<sup>1</sup>).

A first clue can be derived from the observation that most of the carbon atoms of the peptide part of Y<sup>III</sup>-DOTATOC show only one single NMR signal. The chemical shifts of these atoms are practically identical with the corresponding carbons of the single signal set of Ga<sup>III</sup>-DOTATOC (figure 7-8 (A, B and D)). For the rest of the peptidic Y<sup>III</sup>-DOTATOC carbons, the chemical shift differences in the two conformations are quite small. Together with the very similar NOE pattern of the Ga<sup>III</sup>- and Y<sup>III</sup>-complexes, this practically excludes a conformational change of the peptide part (*cf. sections: 9.3.11-9.3.13*), and the second signal set must be caused by the DOTA or linker sections of the molecule.

DOTA lanthanide (III) complexes have already been extensively studied by various methods, and they are known to exhibit a square-antiprismatic geometry (eightfold co-ordination with four nitrogens and four oxygens around the lanthanide ion). The arrangement of the ethylene bridges and the positioning of the acetate sidechains give rise to four exchanging basic conformations, commonly denoted as  $m_1$ ,  $m_2$ ,  $M_1$  and  $M_2$  [234] and are shown in figure 7-9. Here,  $m_1$  and  $M_1$  (or  $m_2$  and  $M_2$ ) describe the different diastereomers, while  $m_1$  and  $m_2$  (or  $M_1$

and  $M_2$ ) constitute enantiomeric forms that normally cannot be distinguished by NMR. On the other hand, Ga<sup>III</sup>-DOTA-D-PheNH<sub>2</sub> shows a six-fold octahedral coordination geometry, with four nitrogens and two oxygens of the carboxylate arm complexing the central ion <sup>[216]</sup>.



**Figure 7-9:** Various conformations of the DOTA lanthanide (III) complex which are normally denoted as  $m_1$ ,  $m_2$ ,  $M_1$  and  $M_2$ . Different diastereomers are ascribed as  $m_1$  and  $M_1$  (or  $m_2$  and  $M_2$ ) while  $m_1$  and  $m_2$  (or  $M_1$  and  $M_2$ ) constitute the enantiomeric forms <sup>[234, 235]</sup>.

Y<sup>III</sup>-DOTA, as a pseudo-lanthanide complex, could be generally expected to show four conformations in solution ( $m_1$ ,  $m_2$ ,  $M_1$  and  $M_2$ ). However, it has been found that – due to its specific ionic radius – Y<sup>III</sup>-DOTA exhibits exclusively the  $M$  conformation in solution <sup>[234]</sup> (while it adopts only the  $m$  conformation in the crystalline form <sup>[216]</sup>). Nevertheless, it is principally conceivable that addition of the bulky peptide part in Y<sup>III</sup>-DOTATOC might influence the conformational equilibrium to give rise to a second DOTA conformation.

A comparison of the carbon chemical shifts of the DOTA region in the Ga<sup>III</sup> and Y<sup>III</sup>-DOTATOC should give further insight into the probability of such a conformational equilibrium. In an overlay of the <sup>13</sup>C-<sup>1</sup>H HSQC spectra of Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC (figure 7-8 (C)), the carbon chemical shifts of the DOTA part show a behavior similar to that observed for the peptide signals. The DOTA part of Ga<sup>III</sup>-DOTATOC shows 12 distinct carbon chemical shifts (corresponding to the 12 different proton-bearing carbon positions in the molecule), while for Y<sup>III</sup>-DOTATOC 16 carbon resonances could be identified. However, in case of the existence of two diastereomeric conformations ( $m$  and  $M$ ), 24 distinct carbon

chemical shifts should have been observed. If in our case this inter-conversion was slow enough to lead to split peptide signals, the effect on the DOTA signals should be even more pronounced, *i.e.*, two clearly separated signals would be expected for all DOTA signals – not just for four out of 12.

In addition, Aime *et al.* have studied the Lu<sup>III</sup>-DOTA complex by solution NMR [235]. They have reported distinctive carbon chemical shifts for the *m* and *M* conformations in the DOTA ring: 57.6 / 56.9 / 67.4 ppm for  $\text{NCCN} / \text{NCCN} / \text{NCCO}$  in the *M* form, and 55.9 / 50.9 / 61.5 ppm for the *m* conformer, respectively. In Y<sup>III</sup>-DOTATOC, the carbon chemical shifts for the DOTA signals occur at 54.08-56.32, 60.76 - 63.63, and 65.24 - 66.70 ppm at 275 K (no more degenerate due to the attached peptide). This is in good agreement with the carbon chemical shifts found for the *M* conformer of Lu<sup>III</sup>-DOTA, but incompatible with the values for the *m* form [235]. Lu<sup>III</sup> and Y<sup>III</sup> have very similar ionic radii (0.97 and 1.04 Å) and are both diamagnetic, hence, a direct comparison of the carbon chemical shifts should indeed be meaningful. Clearly, if Y<sup>III</sup>-DOTA would have adopted an *m* form as one of its conformations, the corresponding carbon chemical shifts should be pronouncedly shifted towards lower frequencies.

### 7.3.5. The Coalescence between the Conformations

For Lu<sup>III</sup>-DOTA it has also been reported that the coalescence of proton resonances occurs around 310 K, corresponding to an energy barrier of ~ 60 kJ/mol for the *m* / *M* transition [235]. If the same exchange between *m* and *M* was responsible for the second signal set in Y<sup>III</sup>-DOTATOC, then coalescence should occur in the same temperature range. However, in our measurements hardly any change was observed in the splitting of the <sup>1</sup>H and <sup>13</sup>C resonances of the two conformations over the whole temperature range up to 330 K (table 7-2) – a clear indication that the energy barrier is significantly higher for the conformational exchange observed in Y<sup>III</sup>-DOTATOC than known for the *m* / *M* transition. The calculation of the energy barrier from the coalescence temperature can be done using following equation [236].

$$\Delta G^\# = R \tau_c \left( 22.96 + \ln \frac{\tau_c}{\delta\nu} \right) \quad [7-2],$$

where  $\Delta G^\#$  is the Gibbs free energy (J mol<sup>-1</sup>), R is the gas constant (8.314 J K<sup>-1</sup> mol<sup>-1</sup>),  $\tau_c$  is the coalescence temperature and  $\delta\nu$  is the chemical shift difference (Hz) between the corresponding resonances of two different conformations at highest possible population

difference (*i.e.*, at the lowest possible temperature). At the logarithmic scale to the base 10, above equation becomes,

$$\Delta G^\ddagger = 19.14 \tau_c (9.97 + \log \frac{\tau_c}{\delta\nu}) \quad [7-3].$$

Our NMR data suggest a coalescence temperature of  $\geq 350$  K (with  $\delta\nu$  as 236 Hz for Cys<sup>2</sup>H<sup>N</sup>) for Y<sup>III</sup>-DOTATOC, corresponding to a lower limit for the energy barrier of 68 kJ/mol.

Alternatively, energy barrier can be also calculated from the integration of the diagonal and the chemical exchange crosspeaks observed in the ROESY (or EXSY) spectrum (figure 7-6). At 275 K, the compensated value for the peak volume of the diagonal signals for Cys<sup>2</sup>H<sup>N</sup> (*major*) is 703.98 and for Cys<sup>2</sup>H<sup>N</sup> (*minor*) are 536.18. Similarly, the cross peak volume for Cys<sup>2</sup>H<sup>N</sup> (*major*)-Cys<sup>2</sup>H<sup>N</sup> (*minor*) is 8.55 and for Cys<sup>2</sup>H<sup>N</sup> (*minor*)-Cys<sup>2</sup>H<sup>N</sup> (*major*) is 6.74. The rate constants can be approximated under the assumption that the ROESY mixing time was in the regime of initial build up of cross peak intensity. The rate constant, in such a case, can be given as <sup>[237]</sup>,

$$I_{ij(\tau_m)} \approx k_{ji} \cdot \tau_m \cdot M_j^0 \quad \text{for } i \neq j \quad [7-4],$$

where  $I_{ij(\tau_m)}$  is the average integrated cross-peak volume,  $k_{ji}$  is the unidirectional pseudo-first-order rate constant ( $s^{-1}$ ) from site  $j$  to  $i$ ,  $\tau_m$  is the mixing time used in ROESY (or EXSY experiment) and  $M_j^0$  corresponds to the equilibrium magnetization at  $\tau_m = 0$ .

For the two conformations in the Y<sup>III</sup>-DOTATOC, the pseudo-first-order rate constant corresponds to  $0.155 s^{-1}$  (with  $j$  as *minor* and  $i$  with *major* conformation and  $\tau_m = 80$  ms). The rate constant can lead to the energy barrier between the two conformations as <sup>[236]</sup>:

$$\Delta G^\ddagger = RT (23.76 - \ln \frac{k}{T}) \quad [7-5],$$

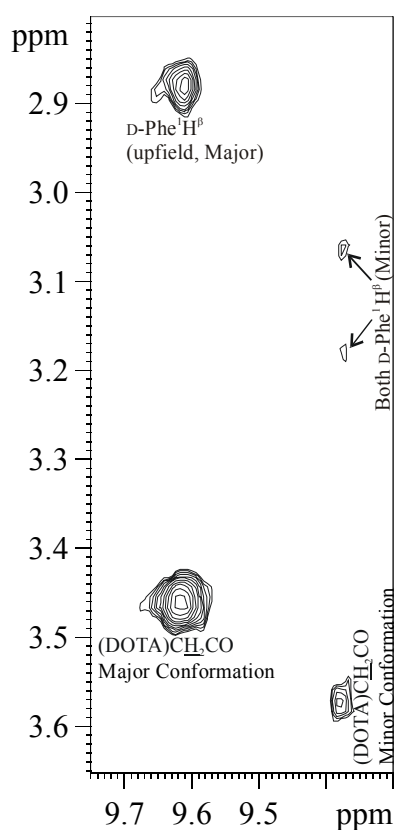
where  $k$  is the rate constant ( $s^{-1}$ ) and  $T$  is the temperature.

This results in a value for  $\Delta G^\ddagger$  of 71 kJ/mol for the conformational exchange occurring in the Y<sup>III</sup>-DOTATOC. These values are in good agreement with the approximately 72-80 kJ/mol expected for a peptide bond *cis-trans* isomerisation.

Based on all these facts, an amide *cis-trans* isomerisation across the linker, *i.e.*, the (DOTA)CH<sub>2</sub>CO-D-Phe<sup>1</sup>H<sup>N</sup> bond, seems the only possible explanation. This would also explain the observations that the NMR signals of D-Phe<sup>1</sup> are most affected by the conformational exchange, and that only four carbons in the DOTA part of Y<sup>III</sup>-DOTATOC resonate at two different frequencies.

7.3.6. *Cis-trans Isomerization in Y<sup>III</sup>-DOTATOC*

Upon closer inspection of the NMR data, a very weak NOE cross-peak between D-Phe<sup>1</sup>H<sup>N</sup> of the major conformation and the nearest CH<sub>2</sub> group of DOTA could be observed, indicating that the major conformation is *trans* configured. To confirm these findings, an additional ROESY spectrum was recorded at 275 K for better resolution and higher intensity of the D-Phe<sup>1</sup> amide signals. A set of 2D experiments (TOCSY, <sup>13</sup>C-<sup>1</sup>H HSQC, <sup>13</sup>C-<sup>1</sup>H HMBC) was run at this temperature to reassess all proton and carbon resonances.



**Figure 7-10:** NOE cross-peaks between acetate-CH<sub>2</sub> of DOTA and D-Phe<sup>1</sup>H<sup>N</sup> (275 K). Left row (at 9.60 ppm): major conformation of D-Phe<sup>1</sup>H<sup>N</sup>, right row (at 9.36 ppm): minor conformation of D-Phe<sup>1</sup>H<sup>N</sup>. The NOE pattern indicates that only the major conformation corresponds to a *trans* configured amide bond, showing a strong cross-peak with a DOTA proton (and one of its own β protons). In contrast, the *cis* conformation (minor) displays only a much weaker cross-peak to a DOTA signal (and both its β protons).

At 275 K, the above mentioned NOE correlation appears as a relatively strong cross-peak in the spectrum (figure 7-10). On the other hand, only a very weak cross-peak exists for the minor conformation. However, based on purely geometric considerations, in the *trans*



configuration D-Phe<sup>1</sup>H<sup>N</sup> should give rise to two NOE cross-peaks, one corresponding to an average distance of 2.5 Å (to the CH<sub>2</sub> group of the covalently attached acetate arm), the other with an average distance of 3.8 Å (to one of the CH<sub>2</sub> groups of the cyclen backbone of DOTA). The latter NOE is absent in the ROESY spectrum, probably due to its weaker nature and the still large linewidth of the D-Phe<sup>1</sup> amide. In a similar consideration, for the *cis* configuration two NOE cross-peaks should be observed corresponding to average distances of 3.6 Å and 4.5 Å (the first between D-Phe<sup>1</sup>H<sup>N</sup> and the CH<sub>2</sub> group of the covalently attached acetate arm, the latter between D-Phe<sup>1</sup>H<sup>N</sup> and one of the CH<sub>2</sub> groups of the cyclen backbone of DOTA).

If the conformational exchange was occurring in the DOTA part (*i.e.*, between the *m* and *M* form) and the amide bond was *trans* configured in both conformations, then two strong cross-peaks would be expected from the two D-Phe<sup>1</sup>H<sup>N</sup> resonances to the CH<sub>2</sub> group of the covalently attached acetate arm, with a distance of ~ 2.5 Å. Absence of this strong NOE in the minor signal set of Y<sup>III</sup>-DOTATOC again rules out the possibility of conformational exchange in the DOTA part.

Interestingly, no NOE cross-peak could be detected between the two DOTA protons at 3.44 ppm and 3.56 ppm (figure 7-10). This suggests that both belong to the same group in the two different conformations (although no exchange cross-peak could be observed). It seems plausible that the DOTA <sup>1</sup>H resonances at 3.44 ppm and 3.56 ppm belong to the CH<sub>2</sub> group of covalently attached acetate arm.

The conversion of the NOE intensities measured for the D-Phe<sup>1</sup>H<sup>N</sup> – DOTA cross-peaks into distances (after correcting for the appropriate population ratios) resulted in some discrepancy from the distances expected from the geometric considerations. The NOEs correspond to distances of 3.5 Å (2.5 Å) for the major conformation and 4.5 Å (3.6 Å) for the minor conformation (expected values in parentheses). However, the D-Phe<sup>1</sup>H<sup>N</sup> signals are much broader than all other <sup>1</sup>H resonances (linewidth major: 38 Hz, minor: 42 Hz, peptide amide protons: ~ 12 Hz at 275 K). Obviously, the large linewidth and hence the existence of significant alternative relaxation mechanisms could readily explain the reduced absolute NOE signal intensities for the D-Phe<sup>1</sup>H<sup>N</sup> resonances. Nevertheless, the observed large intensity differences between the NOE cross-peaks of the two conformations agree very well with a *cis-trans* isomerisation. The ROESY spectrum at 275 K also shows a correlation between D-Phe<sup>1</sup>H<sup>N</sup> and only one of the D-Phe<sup>1</sup>Hβ protons in the major conformation, whereas in the minor conformation, D-Phe<sup>1</sup>H<sup>N</sup> correlates with both Hβ protons, pointing to different sidechain conformations of D-Phe<sup>1</sup>.

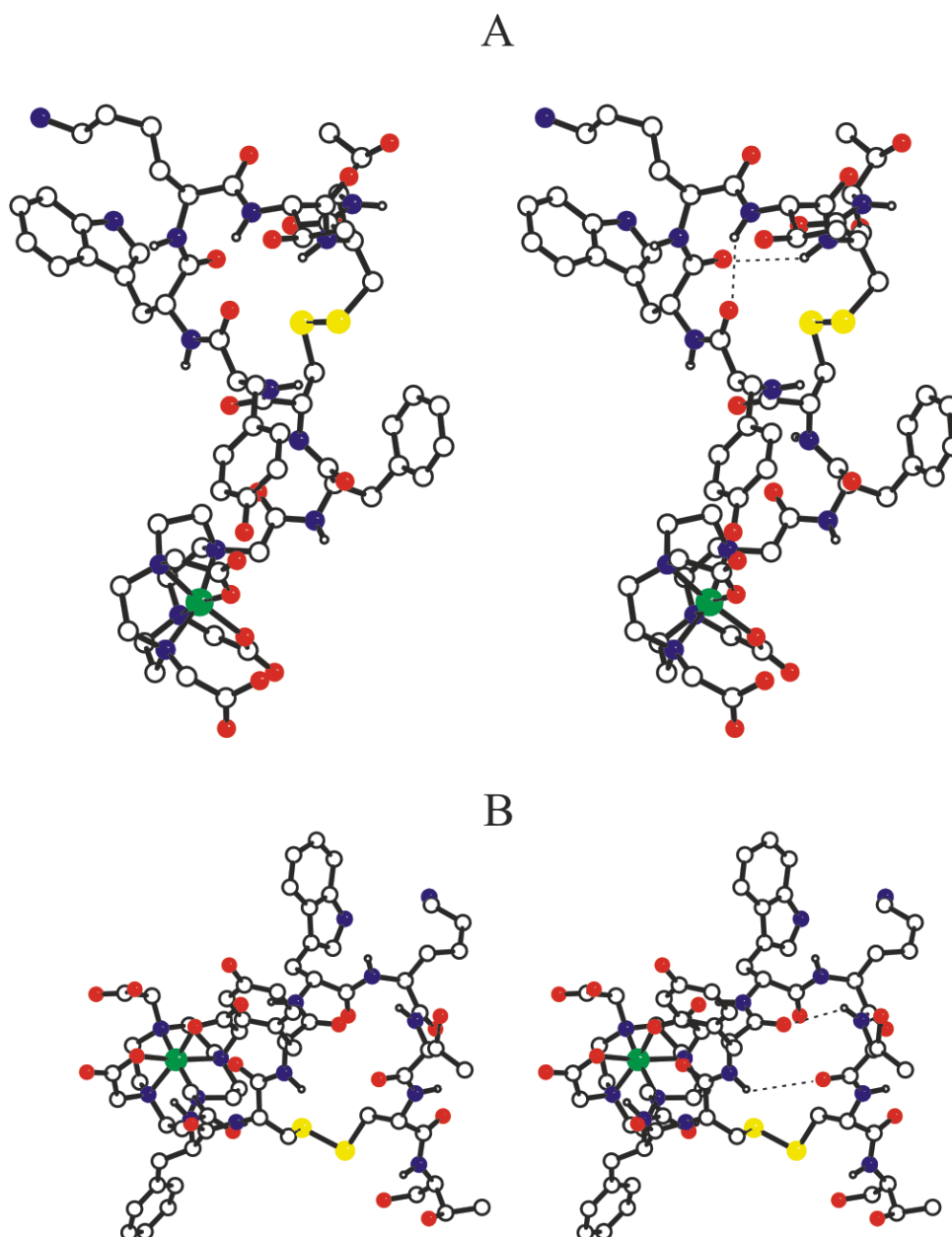
### 7.3.7. Structure Calculations and MD Simulations

Interproton distances were calculated from integration of the offset compensated cross-peaks<sup>[229]</sup> of the ROESY spectra. A tolerance of  $\pm 10\%$  was applied on these distances to derive lower and upper bounds as distance restraints for structure calculations. For the calculations of Ga<sup>III</sup>-DOTATOC, 64 such restraints were used. Due to the excessive overlap between the two signal sets of Y<sup>III</sup>-DOTATOC, only 27 and 28 restraints could be unambiguously extracted and used for calculations for the major and minor signal set, respectively. Since the <sup>1</sup>H signals of the DOTA chelator could not be assigned unambiguously, no restraints were included for this part for both Ga<sup>III</sup>-and Y<sup>III</sup>-DOTATOC.

Initial conformational searches were performed with distance geometry (DG) calculations with a modified version of DISGEO program by Mierke *et al.*<sup>[238-242]</sup>. Further refinement was done by molecular dynamics (MD) simulations with the DISCOVER program package (version 2.9.8)<sup>[242]</sup> with time averaged distance restraints protocols<sup>[243-246]</sup> in the form of an in-house written extension<sup>[247, 248]</sup>. In order to take the metal ions into account, all dynamics simulations were performed with the ESFF force field implemented in DISCOVER. To allow conformational transitions during the simulation runs, time averaged distance restraints protocols<sup>[243-246]</sup> were utilized in form of an in-house written extension for the DISCOVER package<sup>[247, 248]</sup>.

Due to the absence of H $\alpha$ (i)-H $\alpha$ (i+1) cross-peaks in the ROESY spectra, all peptide bonds (except for D-Phe<sup>1</sup>H<sup>N</sup> in the Y<sup>III</sup>-DOTATOC) were restricted to the *trans* configuration in all structure calculations. For the *major* and *minor* conformations of Y<sup>III</sup>-DOTATOC calculations were performed separately, with the amide bond between D-Phe<sup>1</sup>H<sup>N</sup> and the DOTA moiety set to *trans* or *cis*, respectively. Since the peptide parts of Ga<sup>III</sup>-DOTATOC and both conformations of Y<sup>III</sup>-DOTATOC consist of more than a single conformation in solution, the initial DG calculations (100 structures for each dataset) led to somewhat distorted structures. Both the antiparallel  $\beta$ -sheet structures and the  $3_{10}$ -helical structures proposed by Melacini *et al.*<sup>[211]</sup> were contained in the DG ensembles of all three NMR datasets (Ga<sup>III</sup>-DOTATOC, Y<sup>III</sup>-DOTATOC in the *minor* and *major* conformation). Therefore, from each dataset those sheet and helical structures fulfilling the experimental data best were chosen as starting structures for further MD simulations. With each starting structure, a restrained dynamics simulation of 500 picoseconds duration was performed with time averaged distance restraints protocols. The resulting trajectories were then clustered. The program NMRCLUST<sup>[249]</sup> was used to sort the frames of the dynamics trajectories into structural families. Since no experimental data had been available for the DOTA sections, superposition and clustering

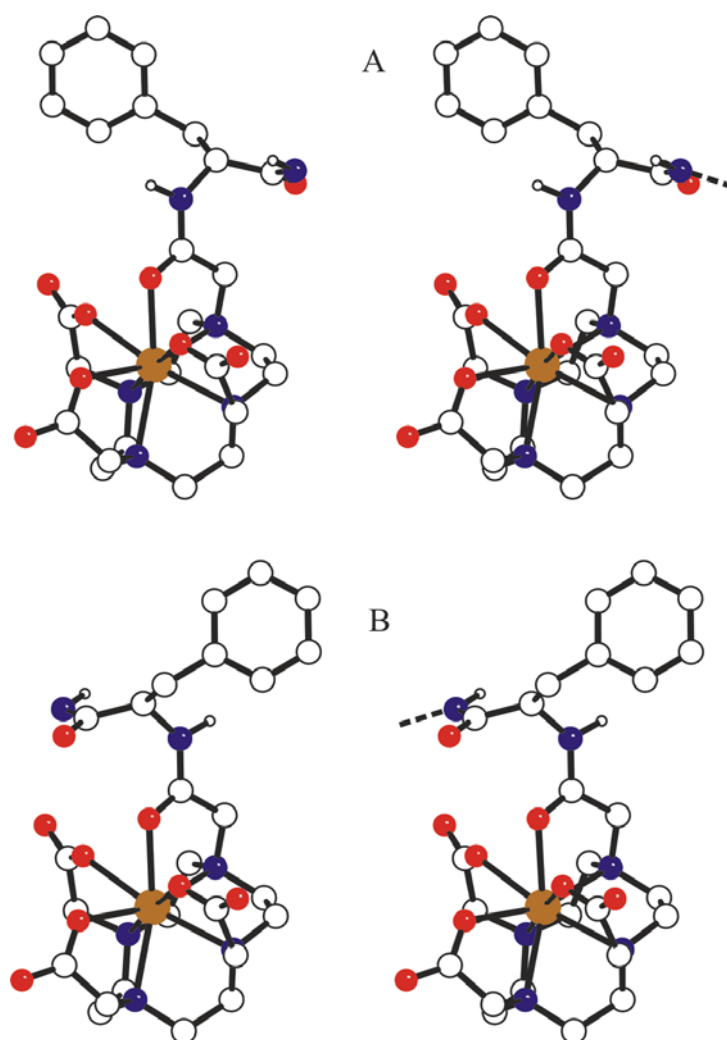
were based on the peptide cycle, *i.e.*, the backbone atoms of the fragment Cys<sup>2</sup>-Tyr<sup>3</sup>-Trp<sup>4</sup>-Lys<sup>5</sup>-Thr<sup>6</sup>-Cys<sup>7</sup> plus the disulfide bridge. The cluster analysis clearly revealed the highly flexible nature of the peptidic parts of Ga<sup>III</sup>- as well as Y<sup>III</sup>-DOTATOC. Both the sheet and helical conformations were represented in the trajectories of all three datasets (figure 7-11), in addition to a variety of other conformations. A thorough variation of all critical parameters of the time averaged distance restraints (exponential decay time  $\tau$ , the force constants of the restraints, and simulation time) did not change this finding.



**Figure 7-11:** Stereo views of representatives of the helical (A) and sheet (B) conformation of Ga<sup>III</sup>-DOTATOC, taken from the time-averaged MD trajectory. Similar peptide conformations are found in the case of Y<sup>III</sup>-DOTATOC.

The dihedral angle CSSC of the disulfide bridge, which can be either +90° or -90°, showed no influence on the structures obtained for both Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC.

In conclusion, the results of these simulations indicate that the metal ion does not have any detectable influence on the backbone structure of the peptide itself, only a minor shift of the conformational equilibrium between sheet and helical forms seems possible from the MD simulations. It has been shown that not only the orientation of the sidechains of Tyr<sup>3</sup>, Trp<sup>4</sup> and Lys<sup>5</sup>, but also that of D-Phe<sup>1</sup>, play an important role in the binding of the peptide to the somatostatin receptor [211]. Therefore, the significant differences in bioavailability between Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC must be due to the differences in the D-Phe<sup>1</sup> linker, *i.e.*, its inclusion in the metal coordination sphere and *cis-trans* isomerisation (figure 7-12) in the Y<sup>III</sup>-complex, in contrast to its essentially unrestricted (extended) conformation in Ga<sup>III</sup>-DOTATOC.



**Figure 7-12:** Stereo models showing the (A) *cis* and (B) *trans* forms of the amide bond between DOTA and D-Phe<sup>1</sup> in Y<sup>III</sup>-DOTATOC. Due to the participation of the amide carbonyl oxygen in the metal coordination sphere, steric interactions are comparable for both isomers.

#### 7.4. Conclusions

This <sup>1</sup>H and <sup>13</sup>C NMR study of the solution structure of Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC has shown that the peptidic parts of both compounds can be characterized by a fast equilibrium of two predominant conformations, displaying a helical and a sheet-like structure, as had been shown for octreotide alone. Specifically, the peptidic moieties of both NMR signal sets of Y<sup>III</sup>-DOTATOC show essentially the same helical and sheet-like contributions as the Ga<sup>III</sup> complex. An investigation into the nature of the two observable signal sets of Y<sup>III</sup>-DOTATOC by variable temperature NMR and various 2D NMR experiments confirmed a *cis-trans* isomerisation across the DOTA – peptide linker, *i.e.*, the (DOTA)CH<sub>2</sub>CO-D-Phe<sup>1</sup>H<sup>N</sup> amide bond. This phenomenon is caused by the incorporation of the carbonyl oxygen of this amide bond into the coordination sphere of the Y<sup>III</sup> ion. The resulting conformational differences at the D-Phe<sup>1</sup> residue represent the only structural cause for the significant differences in the biological activities *in vivo* of Ga<sup>III</sup>- and Y<sup>III</sup>-DOTATOC.

## 8. References

- [1] E. M. Purcell, H. C. Torrey, R. V. Pound, *Phys. Rev.* **1946**, *69*, 37-38.
- [2] F. Bloch, W. W. Hansen, M. Packard, *Phys. Rev.* **1946**, *69*, 127-127.
- [3] C. V. Raman, Krishnan, *Proc. Roy. Soc.* **1927**, *A115*, 549-554.
- [4] C. V. Raman, Krishnan, *Phil. Magn.* **1927**, *3*, 713-735.
- [5] G. A. Morris, R. Freeman, *J. Am. Chem. Soc.* **1979**, *101*, 760-762.
- [6] R. Brüschweiler, R. R. Ernst, *J. Chem. Phys.* **1992**, *96*, 1758-1766.
- [7] R. Riek, G. Wider, K. Pervushin, K. Wüthrich, *Proc. Natl. Acad. Sci.* **1999**, *96*, 4918-4923.
- [8] N. Khaneja, B. Luy, S. J. Glaser, *Proc. Natl. Acad. Sci.* **2003**, *100*, 13162-13166.
- [9] L. Müller, *J. Am. Chem. Soc.* **1979**, *101*, 4481- 4484.
- [10] G. Bodenhausen, D. J. Ruben, *Chem. Phys. Lett.* **1980**, *69*, 185-189.
- [11] K. Pervushin, R. Riek, G. Wider, K. Wüthrich, *Proc. Natl. Acad. Sci.* **1997**, *94*, 12366-12371.
- [12] N. Tjandra, A. Bax, *Science* **1997**, *278*, 1111-1114.
- [13] J. H. Prestegard, H. M. al-Hashimi, J. R. Tolman, *Q. Rev. Biophys.* **2000**, *33*, 371-424.
- [14] F. Kramer, M. V. Deshmukh, H. Kessler, S. J. Glaser, *Conc. Magn. Reson.* **2004**, *21 A (1)*, 10-21.
- [15] R. Golbik, A. N. Lupas, K. K. Koretke, W. Baumeister, J. Peters, *Biol. Chem.* **1999**, *380*, 1049-1062.
- [16] T. Ishikawa, M. R. Maurizi, A. C. Steven, *J. Struct. Biol.* **2004**, *146*, 180-188.
- [17] M. Coles, T. Diercks, J. Liermann, A. Groger, B. Rockel, W. Baumeister, K. K. Koretke, A. Lupas, J. Peters, H. Kessler, *Curr. Biol.* **1999**, *9*, 1158-1168.
- [18] R. E. London, in *Magn. Reson. Biol., Vol. 1* (Ed.: J. S. Cohen), Wiley, New York, **1980**, pp. 1-69.
- [19] G. Wagner, *Curr. Opin. Struc. Biol.* **1993**, *3*, 748-753.
- [20] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, *104*, 4546-4559.
- [21] G. Lipari, A. Szabo, *J. Am. Chem. Soc.* **1982**, *104*, 4559-4570.
- [22] B. J. Stockman, *Prog. Nucl. Mag. Res. Sp.* **1998**, *33*, 109-151.
- [23] G. C. Roberts, *Drug Discov. Today* **2000**, *5*, 230-240.
- [24] H. Kessler, M. Heller, G. Gemmecker, T. Diercks, E. Planker, M. Coles, *Ernst Schering Res. Found. Workshop* **2003**, 59-85.
- [25] B. Meyer, T. Peters, *Angew. Chem. Int. Ed.* **2003**, *42*, 864-890.
- [26] E. Andrew, *Nuclear Magnetic Resonance*, The University Press, Cambridge, **1956**.

- [27] B. C. Gerstein, C. R. Dybrowski, *Transient techniques in NMR of solids*, Academic Press, Orlando, FL, **1985**.
- [28] N. F. Ramsey, E. M. Purcell, *Phys. Rev.* **1952**, *85*, 143-144.
- [29] V. F. Bystrov, *Prog. Nucl. Mag. Res. Sp.* **1976**, *10*, 41-81.
- [30] K. Wüthrich, *NMR of proteins and nucleic acids*, Wiley-Interscience Publication, John Wiley and Sons, New York, **1986**.
- [31] K. Blum, *Density Matrix Theory and Applications*, Plenum Press, New York, **1981**.
- [32] R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of nuclear magnetic resonance in one and two dimensions*, Oxford University Press, New York, **1987**.
- [33] A. Derome, *Modern NMR techniques for chemistry research*, The Pergamon Press, Oxford, **1987**.
- [34] O. W. Sørensen, G. W. Eich, M. H. Levitt, G. Bodenhausen, R. R. Ernst, *Prog. Nucl. Mag. Res. Sp.* **1983**, *16*, 163-192.
- [35] R. R. Ernst, W. A. Anderson, *Rev. Sci. Instrum.* **1966**, *37*, 93-102.
- [36] D. Shaw, *Fourier transform NMR spectroscopy*, 2nd Edition ed., Elsevier, Amsterdam, **1984**.
- [37] J. W. Cooley, J. W. Tukey, *Math. Comput.* **1965**, *19*, 297-&.
- [38] W. P. Aue, E. Bartholdi, R. R. Ernst, *J. chem. Phys.* **1976**, *64*, 2229-2246.
- [39] D. J. States, R. A. Haberkorn, D. J. Ruben, *J. Magn. Reson.* **1982**, *48*, 286-292.
- [40] D. Marion, K. Wüthrich, *Biochem. Bioph. Res. Co.* **1983**, *113*, 967-974.
- [41] A. L. Davis, J. Keeler, E. D. Laue, D. Moskau, *J. Magn. Reson.* **1992**, *98*, 207-216.
- [42] A. D. Bain, *J. Magn. Reson.* **1984**, *56*, 418-427.
- [43] A. A. Maudsley, R. R. Ernst, *Chem. Phys. Lett.* **1977**, *50*, 368-372.
- [44] J. Fiaux, E. B. Bertelsen, A. L. Horwich, K. Wüthrich, *Nature* **2002**, *418*, 207-211.
- [45] P. K. Mandal, A. Majumdar, *Conc. Magn. Reson.* **2004**, *20A*, 1-23.
- [46] A. G. Palmer, J. Cavanagh, P. E. Wright, M. Rance, *J. Magn. Reson.* **1991**, *93*, 151-170.
- [47] J. Cavanagh, A. G. Palmer, P. E. Wright, M. Rance, *J. Magn. Reson.* **1991**, *91*, 429-436.
- [48] H. Kessler, C. Griesinger, J. Zarbock, H. R. Loosli, *J. Magn. Reson.* **1984**, *57*, 331-336.
- [49] H. Kessler, C. Griesinger, G. Zimmermann, *Magn. Reson. Chem.* **1987**, *25*, 579-583.
- [50] G. W. Vuister, A. Bax, *J. Magn. Reson.* **1992**, *98*, 428-435.
- [51] T. M. Logan, E. T. Olejniczak, R. X. Xu, S. W. Fesik, *J. Biomol. NMR* **1993**, *3*, 225-231.

- [52] S. Grzesiek, J. Anglister, A. Bax, *J. Magn. Reson.* **1993**, *101B*, 114-119.
- [53] A. Bax, R. H. Griffey, B. L. Hawkins, *J. Magn. Reson.* **1983**, *55*, 301-315.
- [54] V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, L. E. Kay, *J. Am. Chem. Soc.* **2003**, *125*, 10420-10428.
- [55] M. Sattler, J. Schleucher, C. Griesinger, *Prog. Nucl. Mag. Res. Sp.* **1999**, *34*, 93-158.
- [56] C. Griesinger, O. W. Sørensen, R. R. Ernst, *J. Magn. Reson.* **1987**, *73*, 574-579.
- [57] C. Griesinger, O. W. Sørensen, R. R. Ernst, *J. Am. Chem. Soc.* **1987**, *109*, 7227-7228.
- [58] S. W. Fesik, E. R. P. Zuiderweg, *J. Magn. Reson.* **1988**, *78*, 588-593.
- [59] D. R. Muhandiram, N. A. Farrow, G. Y. Xu, S. H. Smallcombe, L. E. Kay, *J. Magn. Reson.* **1993**, *102B*, 317-321.
- [60] O. Zhang, J. D. Forman-Kay, *Biochemistry* **1993**, *36*, 3959-3970.
- [61] T. Diercks, M. Coles, H. Kessler, *J. Biomol. NMR* **1999**, *15*, 177-180.
- [62] S. Grzesiek, P. Wingfield, S. Stahl, J. D. Kaufman, A. Bax, *J. Am. Chem. Soc.* **1995**, *117*, 9594-9595.
- [63] S. Spera, A. Bax, *J. Am. Chem. Soc.* **1991**, *113*, 5490-5492.
- [64] G. W. Vuister, A. Bax, *J. Am. Chem. Soc.* **1993**, *115*, 7772-7777.
- [65] M. Wittekind, L. Mueller, *J. Magn. Reson* **1993**, *101*, 201-205.
- [66] D. S. Wishart, B. D. Sykes, F. M. Richards, *Biochemistry* **1992**, *31*, 1647-1651.
- [67] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biomol. NMR* **1995**, *5*, 67-81.
- [68] S. Linse, O. Teleman, T. Drakenberg, *Biochemistry* **1990**, *29*, 5925-5934.
- [69] G. Gemmecker, W. Jahnke, H. Kessler, *J. Am. Chem. Soc.* **1993**, *115*, 11620-11621.
- [70] M. Ikura, L. E. Kay, A. Bax, *Biochemistry* **1990**, *29*, 4659-4667.
- [71] K. Wüthrich, *Nat. Struct. Biol.* **1998**, *5(Suppl.)*, 492-495.
- [72] K. Pervushin, *Q. Rev. Biophys.* **2000**, *33*, 161-197.
- [73] T. Schulte-Herbrüggen, O. W. Sørensen, *J. Magn. Reson.* **2000**, *144*, 123-128.
- [74] C. Kojima, M. Kainosho, *J. Magn. Reson.* **2000**, *143*, 417-422.
- [75] E. R. Andrew, A. Bradbury, R. G. Eades, *Nature* **1959**, *183*, 1802-1803.
- [76] A. Saupe, G. Englert, *Phys. Rev. Lett.* **1963**, *11*, 462-464.
- [77] I. W. Emsley, J. C. Lindon, *NMR spectroscopy using liquid crystal solvents*, Pergamon Press, Oxford, **1975**.
- [78] A. D. Buckingham, J. A. Pople, *T. Faraday Soc.* **1963**, *59*, 2421.
- [79] E. Englert, S. A., *Z. Naturforsch. Pt.* **1964**, *19A*, 161-171.
- [80] L. C. Snyder, *J. Chem. Phys.* **1965**, *43*, 4041-4050.



- [81] E. W. Bastiaan, C. Maclean, P. C. M. Van Zijl, A. A. Bothner-By, *A. Rep. NMR Spec.* **1987**, *19*, 35-77.
- [82] E. W. Bastiaan, C. MacLean, *NMR - Basic Princ. Prog.* **1990**, *25*, 17-43.
- [83] J. A. B. Lohman, C. Maclean, *Chem. Phy.* **1978**, *35*, 269-274.
- [84] A. A. Bothner-By, P. J. Domaille, C. Gayathri, *J. Am. Chem. Soc.* **1981**, *103*, 5602-5603.
- [85] M. A. Lisicki, P. K. Mishra, A. A. Bothnerby, J. S. Lindsey, *J. Phys. Chem.* **1988**, *92*, 3400-3403.
- [86] A. A. Bothner-By, in *Encyc. Nucl. Magn. Reson.* (Eds.: D. M. Grant, R. K. Harris), Wiley, Chichester, **1995**, pp. 2932-2938.
- [87] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Proc. Natl. Acad. Sci.* **1995**, *92*, 9279-9283.
- [88] P. Ram, J. H. Prestegard, *Biochim. Biophys. Acta* **1988**, *940*, 289-294.
- [89] C. R. Sanders, J. P. Schwonek, *Biochemistry* **1992**, *31*, 8898-8905.
- [90] A. Bax, G. Kontaxis, N. Tjandra, *Methods Enzymol.* **2001**, *339*, 127-173.
- [91] J. A. Losonczi, M. Andrec, M. W. Fischer, J. H. Prestegard, *J. Magn. Reson.* **1999**, *138*, 334-342.
- [92] M. Zweckstetter, A. Bax, *J. Am. Chem. Soc.* **2000**, *122*, 3791-3792.
- [93] J. W. Emsley, in *Encyc. Nucl. Magn. Reson.* (Eds.: D. M. Grant, R. K. Harris), Wiley, London, **1996**, p. 2788-2799.
- [94] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Nat. Struct. Biol.* **1997**, *4*, 292-297.
- [95] J. R. Tolman, H. M. Al-Hashimi, L. E. Kay, J. H. Prestegard, *J. Am. Chem. Soc.* **2001**, *123*, 1416-1424.
- [96] J. Sass, F. Cordier, A. Hoffmann, M. Rogowski, A. Cousin, J. G. Omichinski, H. Löwen, S. Grzesiek, *J. Am. Chem. Soc.* **1999**, *121*, 2047-2055.
- [97] M. Ottiger, A. Bax, *J. Biomol. NMR* **1998**, *12*, 361-372.
- [98] J. A. Losonczi, J. H. Prestegard, *J. Biomol. NMR* **1998**, *12*, 447-451.
- [99] C. R. Sanders, J. H. Prestegard, *Biophys. J.* **1990**, *78*, 281-289.
- [100] R. S. Prosser, V. B. Volkov, I. V. Shiyonovskaya, *BioChem. Cell. Biol.* **1998**, *76*, 443-451.
- [101] S. Cavagnero, H. J. Dyson, P. E. Wright, *J. Biomol. NMR* **1998**, *13*, 387-391.
- [102] M. R. Hansen, L. Mueller, A. Pardi, *Nat. Struct. Biol.* **1998**, *5*, 1065-1074.
- [103] M. R. Hansen, P. Hanson, A. Pardi, *Methods Enzymol.* **2000**, *317*, 220-240.

- [104] H. Wang, M. Eberstadt, E. T. Olejniczak, R. P. Meadows, S. W. Fesik, *J. Biomol. NMR* **1998**, *12*, 443-446.
- [105] R. S. Prosser, I. V. Shiyankovskaya, *Conc. Magn. Reson.* **2001**, *13*, 19-31.
- [106] G. Cho, B. M. Fung, V. B. Reddy, *J. Am. Chem. Soc.* **2001**, *123*, 1537-1538.
- [107] G. M. Clore, M. R. Starich, A. M. Gronenborn, *J. Am. Chem. Soc.* **1998**, *120*, 10571-10572.
- [108] H. J. Sass, G. Musco, S. J. Stahl, P. T. Wingfield, S. Grzesiek, *J. Biomol. NMR* **2000**, *18*, 303-309.
- [109] M. Rückert, G. Otting, *J. Am. Chem. Soc.* **2000**, *122*, 7793-7797.
- [110] S. Meier, D. Häussinger, S. Grzesiek, *J. Biomol. NMR* **2002**, *24*, 351-356.
- [111] L. G. Barrientos, C. Dolan, A. M. Gronenborn, *J. Biomol. NMR* **2000**, *16*, 329-337.
- [112] K. Fleming, D. Gray, S. Prasanna, S. Matthews, *J. Am. Chem. Soc.* **2000**, *122*, 5224-5225.
- [113] L. G. Barrientos, J. M. Louis, A. M. Gronenborn, *J. Magn. Reson.* **2001**, *149*, 154-158.
- [114] J. Trempe, F. G. Morin, Z. Xia, R. H. Marchessault, K. Gehring, *J. Biomol. NMR* **2002**, *22*, 83-87.
- [115] H. Desvaux, J. C. P. Gabriel, P. Berthault, F. Camerel, *Angew. Chem. Int. Ed.* **2001**, *40*, 373-376.
- [116] B. Luy, K. Kobzar, H. Kessler, *Angew. Chem. Int. Ed.* **2004**, *43*, 1092-1094.
- [117] K. Schmidt-Rohr, H. Spiess, *Multi-dimensional Solid State NMR and Polymers*, Academic Press, London, **1994**.
- [118] E. Brunner, J. Ogle, M. Wenzler, H. R. Kalbitzer, *Biochem. Biophys. Res. Co.* **2000**, *272*, 694-698.
- [119] S. Gaemers, A. Bax, *J. Am. Chem. Soc.* **2001**, *123*, 12343-12352.
- [120] D. D. Ojennus, R. M. Mitton-Fry, D. S. Wuttke, *J. Biomol. NMR* **1999**, *14*, 175-179.
- [121] R. Tycko, F. J. Blanco, Y. Ishii, *J. Am. Chem. Soc.* **2000**, *122*, 9340-9341.
- [122] J. J. Chou, S. Gaemers, B. Howder, J. M. Louis, A. Bax, *J. Biomol. NMR* **2001**, *21*, 377-382.
- [123] G. T. Montelione, G. Wagner, *J. Am. Chem. Soc.* **1989**, *111*, 5474-5475.
- [124] S. J. Archer, M. Ikura, D. Torchia, A. Bax, *J. Magn. Reson.* **1991**, *95*, 636-641.
- [125] M. Billeter, D. Neri, G. Otting, Y. Q. Qian, K. Wüthrich, *J. Biomol. NMR* **1993**, *2*, 257-274.
- [126] A. Bax, G. W. Vuister, S. Grzesiek, F. Delaglio, A. C. Wang, R. Tschudin, G. Zhu, *Methods Enzymol.* **1994**, *239*, 19-105.

- [127] B. Brutscher, J. Boisbouvier, A. Pardi, D. Marion, J. P. Simorre, *J. Am. Chem. Soc.* **1998**, *120*, 11845-11851.
- [128] M. Czisch, R. Boelens, *J. Magn. Reson.* **1998**, *134*, 158-160.
- [129] P. Permi, T. Sorsa, I. Kilpelainen, A. Annala, *J. Magn. Reson.* **1999**, *141*, 44-51.
- [130] T. Facke, S. Berger, *J. Magn. Reson.* **1996**, *119A*, 260-263.
- [131] A. Ross, M. Czisch, T. A. Holak, *J. Magn. Reson.* **1996**, *118A*, 221-226.
- [132] A. Meissner, J. O. Duus, O. W. Sørensen, *J. Magn. Reson.* **1997**, *128*, 92-97.
- [133] A. Meissner, J. O. Duus, O. W. Sørensen, *J. Biomol. NMR* **1997**, *10*, 89-94.
- [134] M. Ottiger, F. Delaglio, A. Bax, *J. Magn. Reson.* **1998**, *131*, 373-378.
- [135] M. H. Lerche, A. Meissner, F. M. Poulsen, O. W. Sørensen, *J. Magn. Reson.* **1999**, *140*, 259-263.
- [136] P. Dosset, J. C. Hus, D. Marion, M. Blackledge, *J. Biomol. NMR* **2001**, *20*, 223-231.
- [137] G. M. Clore, A. M. Gronenborn, A. Bax, *J. Magn. Reson.* **1998**, *133*, 216-221.
- [138] G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **1998**, *120*, 6836-6837.
- [139] J. Meiler, J. J. Prompers, W. Peti, C. Griesinger, R. Bruschweiler, *J. Am. Chem. Soc.* **2001**, *123*, 6098-6107.
- [140] A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, G. L. Warren, *Acta Cryst.* **1998**, *54D*, 905-921.
- [141] C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, *J. Magn. Reson.* **2003**, *160*, 66-74.
- [142] N. Tjandra, J. G. Omichinski, A. M. Gronenborn, G. M. Clore, A. Bax, *Nat. Struct. Biol.* **1997**, *4*, 732-738.
- [143] M. Nilges, A. M. Gronenborn, A. T. Brünger, G. M. Clore, *Prot. Eng.* **1988**, *2*, 27-38.
- [144] C. A. Bewley, K. R. Gustafson, M. R. Boyd, D. G. Covell, A. Bax, G. M. Clore, A. M. Gronenborn, *Nat. Struct. Biol.* **1998**, *5*, 571-578.
- [145] G. M. Clore, M. R. Starich, C. A. Bewley, M. L. Cai, J. Kuszewski, *J. Am. Chem. Soc.* **1999**, *121*, 6513-6514.
- [146] Y. X. Wang, N. Nemati, J. Jacob, I. Palmer, S. J. Stahl, J. D. Kaufman, P. L. Huang, H. E. Winslow, Y. Pommier, P. T. Wingfield, S. Lee-Huang, A. Bax, D. Torchia, *Cell* **1999**, *99*, 433-442.
- [147] S. Patel, M. Latterich, *Trends Cell Biol.* **1998**, *8*, 65-71.
- [148] K. J. Koller, M. J. Brownstein, *Nature* **1987**, *325*, 542-545.
- [149] D. Moir, S. E. Stewart, B. C. Osmond, D. Botstein, *Genetics* **1982**, *100*, 547-563.

- [150] K. U. Frohlich, H. W. Fries, M. Rudiger, R. Erdmann, D. Botstein, D. Mecke, *J. Cell Biol.* **1991**, *114*, 443-453.
- [151] J. M. Peters, M. J. Walsh, W. W. Franke, *EMBO J.* **1990**, *9*, 1757-1767.
- [152] V. Pamnani, T. Tamura, A. Lupas, J. Peters, Z. Cejka, W. Ashraf, W. Baumeister, *FEBS Lett.* **1997**, *404*, 263-268.
- [153] D. M. Epstein, S. J. Benkovic, P. E. Wright, *Biochemistry* **1995**, *34*, 11037-11048.
- [154] A. G. Palmer, J. Williams, P. E. Wright, A. McDermott, *J. Chem. Phys.* **1996**, *100*, 13293-13310.
- [155] L. E. Kay, *Nat. Struct. Biol.* **1998**, *5 Suppl*, 513-517.
- [156] J. J. Prompers, A. Groenewegen, C. W. Hilbers, H. A. Pepermans, *Biochemistry* **1999**, *38*, 5315-5327.
- [157] J. Ye, K. L. Mayer, M. J. Stone, *J. Biomol. NMR* **1999**, *15*, 115-124.
- [158] A. Allerhand, D. Doddrell, R. Komoroski, *J. Chem. Phys.* **1971**, *55*, 189-198.
- [159] A. Abragam, *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford, **1961**.
- [160] Y. Hiyama, C. H. Niu, J. V. Silverton, A. Bavoso, D. A. Torchia, *J. Am. Chem. Soc.* **1988**, *110*, 2378-2383.
- [161] N. Tjandra, P. Wingfield, S. Stahl, A. Bax, *J. Biomol. NMR* **1996**, *8*, 273-284.
- [162] M. R. Boyd, C. Redfield, *J. Am. Chem. Soc.* **1999**, *121*, 7441-7442.
- [163] C. D. Kroenke, M. Rance, A. G. Palmer, *J. Am. Chem. Soc.* **1999**, *121*, 10119-10125.
- [164] A. L. Lee, A. J. Wand, *J. Biomol. NMR* **1999**, *13*, 101-112.
- [165] M. Bloom, L. W. Reeves, E. J. Wells, *J. Chem. Phys.* **1965**, *42*, 1615-1624.
- [166] G. M. Clore, A. Szabo, A. Bax, L. E. Kay, P. C. Driscoll, A. M. Gronenborn, *J. Am. Chem. Soc.* **1990**, *112*, 4989-4991.
- [167] G. M. Clore, P. C. Driscoll, P. T. Wingfield, A. M. Gronenborn, *Biochemistry* **1990**, *29*, 7387-7401.
- [168] J. M. Schurr, H. P. Babcock, B. S. Fujimoto, *J. Magn. Reson.* **1994**, *105B*, 211-224.
- [169] N. Tjandra, S. E. Feller, R. W. Pastor, A. Bax, *J. Am. Chem. Soc.* **1995**, *117*, 12562-12566.
- [170] Z. Zheng, J. Czaplicki, O. Jardetzky, *Biochemistry* **1995**, *34*, 5212-5223.
- [171] D. Fushman, R. Weisemann, H. Thüring, H. Rüterjans, *J. Biomol. NMR* **1994**, *4*, 61-78.
- [172] M. J. Stone, W. J. Fairbrother, A. G. Palmer, J. Reizer, M. H. Saier, P. E. Wright, *Biochemistry* **1992**, *31*, 4394-4406.
- [173] V. M. Bloomfield, W. O. Dalton, K. E. Van Holde, *Biophys. J.* **1967**, *76*, 3044-3057.
- [174] V. M. Bloomfield, *Science* **1968**, *161*, 1212-1219.

- [175] J. d. I. T. Garcia, H. y. B. M. L. Carrasco, *J. Magn. Reson.* **2000**, *147*, 138-146.
- [176] L. E. Kay, D. Torchia, A. Bax, *Biochemistry* **1989**, *28*, 8972-8979.
- [177] R. Brüschweiler, X. Liao, P. E. Wright, *Science* **1995**, *268*, 886-889.
- [178] L. K. Lee, M. Rance, W. J. Chazin, A. G. Palmer, *J. Biomol. NMR* **1997**, *9*, 287-298.
- [179] A. M. Mandel, M. Akke, A. G. Palmer, *J. Mol. Biol.* **1995**, *246*, 144-163.
- [180] R. Cole, J. P. Loria, *J. Biomol. NMR* **2003**, *26*, 203-213.
- [181] J. W. Peng, G. Wagner, *Biochemistry* **1992**, *31*, 8571-8586.
- [182] J. P. Carver, R. E. Richards, *J. Magn. Reson.* **1972**, *6*, 89-105.
- [183] T. D. Goddard, D. G. Kneller, University of California, San Francisco, **2002**.
- [184] V. Y. Orekhov, D. E. Nolde, A. P. Golovanov, D. M. Korzhnev, A. S. Arseniev, *Appl. Magn. Reson.* **1995**, *9*, 581-588.
- [185] A. G. Palmer,  
<http://cpmcnet.columbia.edu/dept/gsas/biochem/labs/palmer/software/curvefit.html>,  
Columbia University, **1998**.
- [186] P. I. Hanson, R. Roth, H. Morisaki, R. Jahn, J. E. Heuser, *Cell* **1997**, *90*, 523-535.
- [187] M. Shtilerman, G. H. Lorimer, S. W. Englander, *Science* **1999**, *284*, 822-825.
- [188] W. Peti, J. Meiler, R. Brüschweiler, C. Griesinger, *J. Am. Chem. Soc.* **2002**, *124*, 5822-5833.
- [189] V. Truffault, Ph.D. Thesis, Pg. No. 47, Technische Universität München (Munich), **2002**.
- [190] J. L. Baber, A. Szabo, N. Tjandra, *J. Am. Chem. Soc.* **2001**, *123*, 3953-3959.
- [191] M. Mayer, B. Meyer, *Angew. Chem. Int. Ed.* **1999**, *38*, 1784-1788.
- [192] S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **1996**, *274*, 1531-1534.
- [193] M. Ormo, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, S. J. Remington, *Science* **1996**, *273*, 1392-1395.
- [194] H. H. Meyer, J. G. Shorter, J. Seemann, D. Pappin, G. Warren, *EMBO J.* **2000**, *19*, 2181-2192.
- [195] X. Zhang, A. Shaw, P. A. Bates, R. H. Newman, B. Gowen, E. Orlova, M. A. Gorman, H. Kondo, P. Dokurno, J. Lally, G. Leonard, H. Meyer, M. van Heel, P. S. Freemont, *Mol. Cell.* **2000**, *6*, 1473-1484.
- [196] C. Strub, C. Schlieker, B. Bukau, A. Mogk, *FEBS Lett.* **2003**, *553*, 125-130.
- [197] P. Brazeau, W. Vale, R. Burgus, N. Ling, M. Butcher, J. Rivier, R. Guillemin, *Science* **1973**, *179*, 77-79.
- [198] S. Froidevaux, A. N. Eberle, *Biopolymers* **2002**, *66*, 161-183.

- [199] G. Weckbecker, I. Lewis, R. Albert, H. A. Schmid, D. Hoyer, C. Bruns, *Nat. Rev. Drug Discov.* **2003**, *2*, 999-1017.
- [200] S. Reichlin, *N. Engl. J. Med.* **1983**, *309*, 1495-1501.
- [201] S. Reichlin, *N. Engl. J. Med.* **1983**, *309*, 1516-1563.
- [202] Y. C. Patel, T. Wheatley, *Endocr.* **1983**, *112*, 220-225.
- [203] J. C. Reubi, *Endocr. Rev.* **2003**, *24*, 389-427.
- [204] C. Bruns, G. Weckbecker, F. Raulf, K. Kaupmann, P. Schoeffter, D. Hoyer, H. Lubbert, *Ann. NY Acad. Sci.* **1994**, *733*, 138-146.
- [205] Y. C. Patel, *Front. Neuroendocrinol.* **1999**, *20*, 157-198.
- [206] E. Pohl, A. Heine, G. M. Sheldrick, Z. Dauter, K. S. Wilson, J. Kallen, W. Huber, P. J. Pfäffli, *Acta Cryst.* **1995**, *51D*, 48-59.
- [207] C. Wynants, G. van Binst, H. R. Loosli, *Int. J. Pept. Prot. Res.* **1985**, *25*, 608-614.
- [208] C. Wynants, G. van Binst, H. R. Loosli, *Int. J. Pept. Prot. Res.* **1985**, *25*, 615-621.
- [209] C. Wynants, D. Tourwe, W. Kazmierski, V. J. Hruby, G. van Binst, *Eur. J. Biochem.* **1989**, *185*, 371-381.
- [210] H. Widmer, A. Widmer, W. Braun, *J. Biomol. NMR* **1993**, *3*, 307-324.
- [211] G. Melacini, Q. Zhu, M. Goodman, *Biochemistry* **1997**, *36*, 1233-1241.
- [212] A. Otte, J. Mueller-Brand, S. Dellas, E. U. Nitzsche, R. Herrmann, H. R. Mäcke, *Lancet* **1998**, *351*, 417-418.
- [213] W. A. Breeman, M. de Jong, D. J. Kwekkeboom, R. Valkema, W. H. Bakker, P. P. Kooij, T. J. Visser, E. P. Krenning, *Eur. J. Nucl. Med.* **2001**, *28*, 1421-1429.
- [214] M. de Jong, W. H. Bakker, E. P. Krenning, W. A. Breeman, M. E. van der Pluijm, B. F. Bernard, T. J. Visser, E. Jermann, M. Behe, P. Powell, H. R. Mäcke, *Eur. J. Nucl. Med.* **1997**, *24*, 368-371.
- [215] C. Waldherr, M. Pless, H. R. Mäcke, T. Schumacher, A. Crazzolara, E. U. Nitzsche, A. Haldemann, J. Mueller-Brand, *J. Nucl. Med.* **2002**, *43*, 610-616.
- [216] A. Heppeler, S. Froidevaux, H. R. Mäcke, E. Jermann, M. Béhé, P. Powell, M. Hennig, *Chem. Eur. J.* **1999**, *5*, 1974-1981.
- [217] M. Henze, J. Schuhmacher, P. Hipp, J. Kowalski, D. W. Becker, J. Doll, H. R. Mäcke, M. Hofmann, J. Debus, U. Haberkorn, *J. Nucl. Med.* **2001**, *42*, 1053-1056.
- [218] M. Hofmann, H. Mäcke, R. Borner, E. Weckesser, P. Schoffski, L. Oei, J. Schumacher, M. Henze, A. Heppeler, J. Meyer, H. Knapp, *Eur. J. Nucl. Med.* **2001**, *28*, 1751-1757.
- [219] M. Piotto, V. Saudek, V. Sklenar, *J. Biomol. NMR* **1992**, *2*, 661-665.
- [220] V. Sklenar, M. Piotto, R. Leppik, V. Saudek, *J. Magn. Reson.* **1993**, *102*, 241-245.

- [221] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, H. J. Dyson, E. Oldfield, J. L. Markley, B. D. Sykes, *J. Biomol. NMR* **1995**, *6*, 135-140.
- [222] D. H. Live, D. G. Davis, W. C. Agosta, D. Cowburn, *J. Am. Chem. Soc.* **1984**, *106*, 1939-1941.
- [223] A. Bax, S. Subramanian, *J. Magn. Reson.* **1986**, *67*, 565-569.
- [224] L. Braunschweiler, R. R. Ernst, *J. Magn. Reson.* **1983**, *53*, 521-528.
- [225] U. Piantini, O. W. Sørensen, R. R. Ernst, *J. Am. Chem. Soc.* **1982**, *104*, 6800-6801.
- [226] C. Griesinger, O. W. Sørensen, R. R. Ernst, *J. Am. Chem. Soc.* **1985**, *107*, 6394-6396.
- [227] G. Bodenhausen, R. Freeman, *J. Magn. Reson.* **1977**, *28*, 471-476.
- [228] A. Bax, M. F. Summers, *J. Am. Chem. Soc.* **1986**, *108*, 2093-2094.
- [229] C. Griesinger, R. R. Ernst, *J. Magn. Reson.* **1987**, *75*, 261-271.
- [230] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **1979**, *71*, 4546-4553.
- [231] I. Bertini, C. Luchinat, M. Piccioli, *Proc. Natl. Acad. Sci.* **1994**, 91-141.
- [232] H. Kessler, M. Gehrke, C. Griesinger, *Angew. Chem. Int. Ed.* **1988**, *27*, 490-536.
- [233] S. Liu, J. Pietryka, C. E. Ellars, D. S. Edwards, *Bioconjug. Chem.* **2002**, *13*, 902-913.
- [234] S. Aime, M. Botta, M. Fasano, M. P. Marques, C. F. Geraldès, D. Pubanz, A. E. Merbach, *Inorg. Chem.* **1997**, *36*, 2059-2068.
- [235] S. Aime, A. Barge, M. Botta, M. Fasano, J. D. Ayala, G. Bombieri, *Inorg. Chim. Acta* **1996**, *246*, 423-429.
- [236] H. Günther, *NMR Spectroscopy*, Wiley, New York, **1980**.
- [237] J. D. Heise, D. Raftery, B. K. Breedlove, J. Washigton, C. F. Kubiak, *Organometallics* **1998**, *17*, 4461-4468.
- [238] T. F. Havel, in *DISGEO: Quantum chemistry exchange program* (Ed.: Exchange No. 507), Indiana University, **1986**.
- [239] T. F. Havel, *Prog. Biophys. Mol. Biol.* **1991**, *56*, 43-78.
- [240] G. M. Crippen, T. F. Havel, *Distance geometry and molecular conformation*, John Wiley & Sons, New York, **1998**.
- [241] T. Havel, K. Wüthrich, *Bull. Math. Biol.* **1984**, *46*, 673-698.
- [242] Discover, 2.9.7/95.0/3.0.0 ed., Biosym/MSI, San Diego, **1995**.
- [243] A. E. Torda, R. M. Scheek, W. F. van Gunsteren, *Chem. Phys. Lett.* **1989**, *157*, 289-294.
- [244] A. E. Torda, R. M. Scheek, W. F. van Gunsteren, *J. Mol. Biol.* **1990**, *214*, 223-235.
- [245] D. A. Pearlman, P. A. Kollman, *J. Mol. Biol.* **1991**, *220*, 457-479.
- [246] A. P. Nanzer, W. F. van Gunsteren, A. E. Torda, *J. Biomol. NMR* **1995**, *6*, 313-320.

- 
- [247] G. Hessler, Ph.D. Thesis, Technische Universität München (Munich), **1997**.
- [248] C. Roelz, Ph.D. Thesis, Technische Universität München (Munich), **2000**.
- [249] L. A. Kelley, S. P. Gardner, M. J. Sutcliffe, *Prot. Eng.* **1996**, 9, 1063-1065.



## 9. Appendix

### 9.1. RDC Studies on VAT-N

#### 9.1.1. Pulse Program Implemented: $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC

```

;mvdNHipap.f2      ;mvdesh  08/01/02
;IPAP-[15N,1H]-HSQC (1H-coupled in F1(N) )
;
;Reference:
;Ottiger, M., Delaglio, F. & Bax, A.: J. Magn. Reson. 131, 373-378 (1998)

;##### NO CARBON DECOUPLING HERE #####

#include <Avance.incl>
#include <Grad.incl>

;####Pulses to be set#####
;p1      proton 90 at p11
"p2=p1*2"
;p3      nitrogen 90 at p12
"p4=p3*2"
;pcpd2      : 90 degree soft pulse on X (f2) at p112

;#####Gradient pulses (may have to be set manually):#####
;"p20=2.0m"
;"p11=1.0m"
;"p17=0.4m"

;#####delays to be set#####
"d0=in0*0.5 - p3*0.63 -p1"
;d1      :relaxation delay (>= 1 sec)
;d4      :1/4JNH*0.7 (about 2.2ms)
"d5=d4-p17-2u"
"d11=10m"      ;I/O delay
"d12=10u"      ;increment delay
"d13=25m"
;d19      :Watergate Delay (~80 u)
;"d22=p2"
;"d23=p3"
"d27=p3*1.26 -p1"
"d25=p1*2"
;d16      :>= 150u (gradient recovery)

;#####Acquisition info#####
;Use pseudo 3D experimental setup with F3=1H, F2=15N and F1= dummy with
;TD=2 (any nucleus would do!)The first FID recorded here will generate IP
;component and the second one ;will generate AP component of the splitting.
;ND0= 2, ND10= 2, TD0=1, CPDPRG2=garp (you will see the coupling in the
;indirect dimension. Replace in0 by in10 as the increments depends upon
;in0. used l2 = 1
;GPNAM2= sine.100
;GPNAM= sine.100

```

```

;Gradient strengths which we have used during test runs were: gpz0=40% and
;gpz2=10%
;with 1 ms gradient pulses (p16 and p17)

#####Processing info#####
;create 2D from the 3D data set.
;Enter slice No. "1" for IP component and "2" for AP component.
;mc2=states,
;Make REVERSE false in both the dimensions
;Addition of IP and AP part:
;copy the processed datasets into new process nos., define multipliers
;"alpha" and "gamma" respectively.
;( used 1 and 1.1)

#####calculated parameters#####
define delay wg
define delay cen18
define delay cen24
"cen24=(p4-p2)/2"
"wg=p1*4.77+d19*10"
"cen18=(wg-p4*2-6us)/2"
"d24=d4-p16-d16-600u" ;600u compensate of J-evolution during 3919 sequence
"l3=(td2/2)"

#define WG (p1*0.231 ph14 d19*2 p1*0.692 ph14 d19*2 p1*1.462 ph14 d19*2
p1*1.462 ph15 d19*2 p1*0.692 ph15 d19*2 p1*0.231 ph15):f1

1      ze
      10u ru2
2      d11 do:f2
      d12
3      d12*3
4      d12*3
5      10u do:f4
      10u p12:f2
      d1
      1m UNBLKGRAD
      10u p11:f1
      (p3 ph0):f2          ;eliminate Boltzmann
      p16:gp0*2
      3m
      (p1 ph0)
      2u
      p17:gp2
      d5
      (cen24 p2 ph0):f1 (p4 ph6):f2
      2u
      p17:gp2
      d5
      (p1 ph6)
      6u
      p16:gp0*-0.8
      1m p11:f1
      (p3 ph3):f2
      if "l2==1" goto 20
d4
      (cen24 p2 ph1):f1 (p4 ph5):f2
d4
      d27
      (p1 ph4):f1
20     d0
      d25 ; (p1*2 ph0) ;no 180 for F1-coupled spectrum

```

```

    d0
    if "l2==2"goto 40
    (p3 ph7):f2
    goto 41
40    (p3 ph17):f2
41    2u
    p16:gp0*0.8
    2m p11:f1
    (p1 ph0)
    5u
    d24
p16:gp0
d16
WG (cen18 p3 ph1 3u p4 ph0 3u p3 ph1):f2
p16:gp0
d16 p112:f2
d24
    go=2 ph31 cpd2:f2
    d11 wr #0 if #0 zd
    1m BLKGRAD
    10u do:f2
    d12 iu2
lo to 3 times 2
    d12 ru2
    d12 dp7
    d12 dp17
lo to 4 times 2
    d12 id0
    d12 ip31
    d12 ip31
    lo to 5 times 13
    d12 do:f4
    d12 do:f2
    1m BLKGRAD
    100u

exit

ph0=0
ph1=1
ph3=0 0 2 2
ph4=1 1 1 1 1 1 1 1 3 3 3 3 3 3 3 3
ph5=0 0 0 0 1 1 1 1 2 2 2 2 3 3 3 3
ph6=1 3
ph7=0
ph17=1 1 1 1 3 3 3 3
ph10=0
ph14=0
ph15=2
ph31=0 2 2 0

```

### 9.1.2. Comparison of RDCs Obtained by $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC and Tr-SmTr Approach

Highlighted residues showed differences in the two measurements, *i.e.*, TROSY-Semi-TROSY and  $^{15}\text{N}$ - $^1\text{H}$ -IPAP-HSQC. To see the cause of the differences these residues were carefully looked up in the corresponding spectra. The reason of the differences is written in the comments column.

Residues	Tr-SmTr [Hz]	IPAP-HSQC [Hz]	Errors [Hz]	Comments
6 GLY	-3.630	1.020	4.000	
8 ILE	-21.500	-15.680	2.000	Overlap in IPAP-HSQC
9 LEU	-18.160	-15.680	2.000	
12 ALA	14.530	12.790	2.000	
24 VAL	18.160	18.620	2.000	
25 ARG	3.630	15.680	2.000	Overlap in IPAP-HSQC
28 GLU	-3.630	-2.940	2.000	
29 SER	3.630	4.900	2.000	
30 SER	-3.640	1.960	2.000	Overlap in IPAP-HSQC
31 ARG	-3.640	-2.940	2.000	
41 VAL	-0.010	10.780	2.000	Overlap in IPAP-HSQC
42 VAL	0.010	-1.960	2.000	
43 GLU	-3.630	-8.820	4.000	Overlap in IPAP-HSQC
44 ILE	-7.27	----	2.000	
45 GLU	-14.530	-9.800	2.000	Overlap in IPAP-HSQC
49 LYS	-7.270	-12.740	2.000	Overlap in IPAP-HSQC
50 THR	-18.170	-16.660	2.000	
51 VAL	-10.890	-11.900	2.000	
52 GLY	-10.900	-7.430	2.000	Line shape IPAP-HSQC
54 VAL	0.000	----	2.000	
55 TYR	7.270	14.700	2.000	No reason
66 ILE	0.010	2.940	2.000	
67 VAL	7.270	13.080	2.000	Very Strong Overlap in IPAP-HSQC

Residues	Tr-SmTr [Hz]	IPAP-HSQC [Hz]	Errors [Hz]	Comments
68 ARG	3.640	13.720	2.000	Overlap in IPAP-HSQC
71 SER	-3.640	-1.350	2.000	
72 VAL	0.010	0.980	2.000	
73 MET	-10.900	----	2.000	
74 ARG	-3.630	-1.060	2.000	
75 ASN	-3.640	----	2.000	
77 CYS	-2.590	-4.980	2.000	
85 VAL	-7.260	-9.800	2.000	
86 LYS	-18.160	----	2.000	
88 ARG	-10.910	-10.780	2.000	
90 VAL	-3.640	0.980	2.000	Overlap in IPAP-HSQC
97 LYS	-3.630	-2.940	2.000	
98 VAL	-3.630	-1.960	2.000	
99 THR	-3.640	-0.980	2.000	
100 LEU	3.640	0.980	4.000	
101 ALA	-8.970	-13.720	2.000	Overlap in IPAP-HSQC
103 ILE	-0.010	-3.920	2.000	
116 ILE	-10.890	-11.760	2.000	
117 GLU	-18.170	-16.450	2.000	
122 ARG	-14.540	-13.720	2.000	
124 LEU	-14.540	-14.500	2.000	
129 MET	3.640	----	2.000	
134 ASN	-10.900	-8.820	2.000	
135 ILE	-21.810	-18.610	4.000	
149 LEU	-0.010	6.410	2.000	Overlap in IPAP-HSQC
150 PHE	-14.540	-----	2.000	
151 LYS	-18.160	-15.680	4.000	
152 VAL	-7.270	-7.840	2.000	
153 VAL	-3.640	1.960	2.000	No reason
154 LYS	3.630	0.980	3.000	
155 THR	-3.630	----	2.000	

---

Residues	Tr-SmTr [Hz]	IPAP-HSQC [Hz]	Errors [Hz]	Comments
162 VAL	3.640	14.450	2.000	
163 GLU	7.260	6.860	2.000	
169 LYS	-5.910	-2.940	3.000	
171 GLU	-10.900	-7.760	2.000	
172 ILE	0.010	-0.980	2.000	

---

### 9.1.3. Example of RDC Input File for CNS Calculation

Following file was used for refining structure of VAT-N (only the first 8 and the last residue are shown).

```

assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 6 and name HN)
  (resid 6 and name N) 1.02 2
assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 8 and name HN)
  (resid 8 and name N) -21.50 2
assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 9 and name HN)
  (resid 9 and name N) -15.68 2
assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 12 and name HN)
  (resid 12 and name N) 12.79 2
assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 24 and name HN)
  (resid 24 and name N) 18.16 2
assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 25 and name HN)
  (resid 25 and name N) 3.63 2
assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 28 and name HN)
  (resid 28 and name N) -2.94 2
assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 29 and name HN)
  (resid 29 and name N) 4.90 2
.....
assign(resid 500 and name OO)
  (resid 500 and name Z)
  (resid 500 and name Y)
  (resid 500 and name X)
  (resid 172 and name HN)
  (resid 172 and name N) 0.01 2

```

## 9.2. $^{15}\text{N}$ Backbone Relaxation Rates ( $R_1$ and $R_2$ ) and $^{15}\text{N}$ - $^1\text{H}$ Heteronuclear-NOE Used for the Model-free Analysis of VAT-N

Symbol  $\Delta$  stands for the error estimated for the respective rate. The error estimated on NOE dataset were 10%.

Residues	$R_1(\text{s}^{-1})$	$\Delta R_1$	$R_2(\text{s}^{-1})$	$\Delta R_2$	NOE	$\Delta\text{NOE}$
I7	1.33	0.053	12.224	0.194	0.806	0.08
L9	1.248	0.036	12.836	0.24	0.785	0.078
V11	1.355	0.034	13.44	0.284	0.799	0.079
A12	1.514	0.039	13.517	0.272	0.849	0.084
E13	1.372	0.07	13.572	0.183	0.811	0.081
G20	1.481	0.069	12.487	0.355	0.779	0.077
M21	1.466	0.037	13.267	0.482	0.777	0.077
V24	1.377	0.028	12.357	0.271	0.793	0.079
R25	1.438	0.039	12.81	0.26	0.845	0.084
D27	1.426	0.044	12.863	0.23	0.824	0.082
E28	1.409	0.044	12.515	0.244	0.834	0.083
S29	1.454	0.04	12.57	0.328	0.834	0.083
S30	1.465	0.05	13.67	0.358	0.898	0.089
R31	1.452	0.033	13.273	0.255	0.825	0.082
L34	1.369	0.05	12.088	0.177	0.806	0.08
E37	1.208	0.041	11.453	0.182	0.73	0.073
V41	1.355	0.048	12.01	0.229	0.756	0.075
V42	1.466	0.054	13.44	0.105	0.799	0.079
E43	1.303	0.044	13.583	0.195	0.813	0.081
K46	1.292	0.03	13.063	0.147	0.851	0.085
T50	1.227	0.029	13.046	0.179	0.849	0.084
V51	1.299	0.052	13.289	0.18	0.875	0.087
G52	1.341	0.028	13.236	0.286	0.793	0.079
V54	1.456	0.046	12.436	0.195	0.877	0.087
Y55	1.472	0.058	12.61	0.524	0.867	0.086
A57	1.43	0.036	11.983	0.404	0.796	0.079
R58	1.426	0.045	12.178	0.201	0.76	0.076



---

Residues	$R_1(\text{s}^{-1})$	$\Delta R_1$	$R_2(\text{s}^{-1})$	$\Delta R_2$	NOE	$\Delta\text{NOE}$
E60	1.374	0.043	13.173	0.256	0.798	0.079
E62	1.432	0.029	12.561	0.221	0.804	0.08
N63	1.311	0.04	11.948	0.316	0.823	0.082
K64	1.359	0.04	12.275	0.246	0.846	0.084
G65	1.386	0.036	12.247	0.275	0.803	0.08
I66	1.398	0.036	13.935	0.187	0.823	0.082
V67	1.448	0.033	13.668	0.19	0.844	0.084
R68	1.384	0.039	12.55	0.206	0.825	0.082
I69	1.439	0.042	13.457	0.251	0.822	0.082
S71	1.461	0.07	13.379	0.361	0.845	0.084
V72	1.416	0.04	12.722	0.194	0.873	0.087
R74	1.452	0.049	13.259	0.195	0.886	0.088
C77	1.406	0.026	12.448	0.257	0.816	0.081
G78	1.471	0.035	13.304	0.172	0.861	0.086
S80	1.249	0.041	12.181	0.184	0.845	0.084
V87	1.254	0.041	12.732	0.408	0.815	0.081
K89	1.317	0.057	12.866	0.212	0.752	0.075
T92	1.2	0.048	10.886	0.053	0.739	0.073
E93	1.283	0.039	12.013	0.207	0.821	0.082
I94	1.29	0.048	11.158	0.192	0.778	0.077
A95	1.395	0.053	12.38	0.346	0.796	0.079
K96	1.403	0.035	11.957	0.252	0.783	0.078
V98	1.34	0.045	12.25	0.259	0.804	0.08
T99	1.324	0.036	12.759	0.189	0.835	0.083
L100	1.312	0.04	12.112	0.179	0.798	0.079
A101	1.246	0.038	13.285	0.277	0.835	0.083
I104	1.357	0.057	11.496	0.335	0.73	0.073
R105	1.444	0.063	11.953	0.283	0.632	0.063
D107	1.508	0.036	11.203	0.256	0.663	0.066
F112	1.307	0.049	9.852	0.345	0.473	0.047
G113	1.305	0.049	8.136	0.378	0.282	0.028
G115	1.266	0.038	9.95	0.451	0.643	0.064

---

Residues	$R_1(s^{-1})$	$\Delta R_1$	$R_2(s^{-1})$	$\Delta R_2$	NOE	$\Delta$ NOE
I116	1.326	0.054	13.446	0.173	0.7	0.07
E117	1.303	0.063	13.394	0.189	0.781	0.078
R122	1.295	0.048	13.087	0.204	0.775	0.077
L124	1.263	0.04	12.439	0.301	0.82	0.082
I125	1.34	0.037	12.725	0.171	0.81	0.081
R126	1.318	0.053	12.663	0.263	0.851	0.085
R127	1.371	0.048	12.338	0.188	0.8	0.08
M129	1.427	0.025	12.108	0.229	0.827	0.082
N134	1.238	0.032	13.564	0.26	0.862	0.086
I135	1.276	0.033	13.531	0.278	0.799	0.079
S136	1.307	0.056	14.355	0.192	0.745	0.074
V137	1.317	0.044	11.771	0.302	0.736	0.073
G139	1.357	0.063	12.647	0.898	0.872	0.087
L140	1.34	0.037	12.433	0.203	0.743	0.074
L142	1.415	0.035	11.508	0.27	0.57	0.057
G144	1.565	0.07	9.416	0.597	0.653	0.065
T146	1.414	0.046	9.074	0.706	0.506	0.05
G147	1.559	0.075	8.163	0.709	0.548	0.054
L149	1.344	0.043	12.253	0.238	0.659	0.065
F150	1.23	0.036	13.852	0.196	0.829	0.082
K151	1.231	0.044	13.779	0.244	0.833	0.083
V152	1.367	0.049	12.774	0.306	0.813	0.081
V153	1.372	0.042	12.319	0.21	0.791	0.079
K154	1.464	0.035	12.275	0.16	0.772	0.077
T155	1.419	0.052	12.44	0.182	0.818	0.081
L156	1.414	0.031	11.984	0.119	0.851	0.085
S158	1.451	0.04	12.528	0.293	0.826	0.082
V160	1.109	0.042	9.699	0.206	0.558	0.055
V162	1.484	0.032	13.497	0.181	0.817	0.081
E163	1.412	0.034	11.811	0.191	0.854	0.085
I164	1.393	0.05	12.561	0.095	0.859	0.085
G165	1.35	0.033	13.13	0.382	0.807	0.08
E166	1.245	0.05	13.764	0.186	0.842	0.084

---

Residues	$R_1(\text{s}^{-1})$	$\Delta R_1$	$R_2(\text{s}^{-1})$	$\Delta R_2$	NOE	$\Delta\text{NOE}$
T168	1.282	0.057	12.958	0.135	0.775	0.077
K169	1.38	0.023	11.786	0.187	0.795	0.079
I170	1.356	0.038	11.334	0.162	0.76	0.076
E171	1.331	0.03	12.704	0.269	0.737	0.073
R173	1.273	0.043	13.185	0.269	0.768	0.076
S178	1.447	0.052	9.523	0.573	0.539	0.053
L181	1.52	0.038	6.997	0.338	0.213	0.021

---

### 9.3. Chemical Shift and NOE Tables of Ga<sup>III</sup> and Y<sup>III</sup>-DOTATOC

Following conventions are used for the chemical shift tables:

- <sup>u</sup> up-field shift on the frequency scale,
- <sup>d</sup> for down-field shift on the frequency scale, and
- \* COOH is modified to CH<sub>2</sub>OH.

Following conventions are used for the NOE tables:

- \* in the structure calculations, a pseudo atom used and the upper limits adjusted by the appropriate pseudo atom correction (Ref: Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1983**, *169*, 949-961.)

#### 9.3.1. Proton Chemical Shifts for Ga<sup>III</sup>-DOTATOC (290 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
NH	8.78	8.45	8.28	8.74	8.48	8.13	8.10	7.81*
H $\alpha$	4.94	4.89	4.80	4.44	4.03	4.48	4.91	4.03
H $\beta^u$	3.11	2.98	3.03	3.08	1.45	4.54	3.11	4.18
H $\beta^d$	3.36	3.15	2.92	3.22	1.76		3.39	
H $\gamma^u$	--	--	--	he3: 7.73	0.50	1.41	--	1.33
H $\gamma^d$	--	--	--	he1: 10.31	0.70		--	
H $\delta^u$	7.50	--	7.29	hd1: 7.31	1.49	--	--	--
H $\delta^d$		--		hita2: 7.43		--	--	
H $\epsilon^u$	7.54	--	7.04	hz2: 7.68	2.87	--	--	--
H $\epsilon^d$		--		hz3: 7.37		--	--	
Aromatic								3.74 and 3.84
(H $\xi$ ) or	7.56	--	--	--	--	5.77	--	(H $\beta$ )
other						(OH)		5.69 and 5.84 (OH)

9.3.2. Carbon Chemical Shifts for Ga<sup>III</sup>-DOTATOC (290 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
CO	175.80	173.01	173.68	178.45	178.03	175.51	175.40	64.50*
C $\alpha$	55.21	55.82	57.99	58.97	59.66	63.02	57.72	59.58
C $\beta$ /Me	40.45	44.36	40.55	29.07	32.41	69.73	42.04	69.18
C $\gamma$	132.21	--	132.21	111.94	24.43	22.05	--	21.94
C $\delta$	130.28	--	133.70	C $\delta$ <sup>1</sup> 127.50 C $\delta$ <sup>2</sup> 130.05	29.09	--	--	--
C $\epsilon$	131.80	--	118.50	C $\epsilon$ <sup>1</sup> 139.82 C $\epsilon$ <sup>2</sup> 121.20	42.03	--	--	--
C (mis.)	132.40 (C $\omega$ )	--	157.97 (C $\omega$ )	C $\zeta$ <sup>3</sup> 122.30 C $\eta$ <sup>2</sup> 125.00 C $\zeta$ <sup>2</sup> 114.90	--	--	--	--

9.3.3. Proton Chemical Shifts for Y<sup>III</sup>-DOTATOC (Major-290 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
NH	9.60	8.81	8.30	8.73	8.48	8.10	8.37	7.83*
H $\alpha$	4.90	4.86	4.85	4.38	4.04	4.50	4.88	4.04
H $\beta$ <sup>u</sup>	3.42	3.32	3.05	3.19	1.76	4.59	3.43	4.20
H $\beta$ <sup>d</sup>	3.05	3.11	3.23	3.03	1.45	--	3.12	--
H $\gamma$ <sup>u</sup>	--	--	--	he3: 7.72	0.50	1.43	--	1.33
H $\gamma$ <sup>d</sup>	--	--	--	he1: 10.31	0.70	--	--	--
H $\delta$ <sup>u</sup>	7.54	--	7.34	hd1: 7.30	1.49	--	--	--
H $\delta$ <sup>d</sup>	--	--	--	hita2: 7.44	--	--	--	--
H $\epsilon$ <sup>u</sup>	7.62	--	7.09	hz2: 7.68	2.88	--	--	--
H $\epsilon$ <sup>d</sup>	--	--	--	hz3: 7.39	--	--	--	--
Aromatic or other	7.54	--	--	--	--	5.68 (OH)	--	3.69 (H $\beta$ ) 5.40, 5.51 (OH)

9.3.4. Proton Chemical Shifts for  $Y^{III}$ -DOTATOC (Minor-290 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
NH	9.36	8.39	8.06	8.66	--	--	8.14	7.73*
H $\alpha$	4.84	4.68	4.75	4.23	--	--	4.65	--
H $\beta^u$	3.32	3.03	3.35	3.15	--	--	3.23	--
H $\beta^d$	2.98	2.88	3.23	2.97	--	--	2.89	--
H $\gamma^u$	--	--	--	he3: --	--	--	--	--
H $\gamma^d$	--	--	--	he1: --	--	--	--	--
H $\delta^u$	7.53	--	--	hd1: --	--	--	--	--
H $\delta^d$	--	--	--	hita2: --	--	--	--	--
H $\epsilon^u$	7.58	--	--	hz2: --	--	--	--	--
H $\epsilon^d$	--	--	--	hz3: --	--	--	--	--
Aromatic or other	7.48	--	--	--	--	-- (OH)	--	-- (H $\beta$ ) -- (OH)

9.3.5. Carbon Chemical Shifts for  $Y^{III}$ -DOTATOC (Major-290 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
CO	172.49	171.05	172.99	175.83	176.85	173.75	172.82	64.12*
C $\alpha$	59.52	55.13	58.19	59.02	59.54	63.04	56.23	59.54
C $\beta$ /Me	40.96	44.66	40.66	29.06	32.27	69.51	42.04	69.09
C $\gamma$	137.59	--	131.82	108.79	24.44	22.11	--	21.97
C $\delta$	130.70	--	133.70	C $\delta^1$ 127.50 C $\delta^2$ 131.57	29.04	--	--	--
C $\epsilon$	132.30	--	118.20	C $\epsilon^2$ 137.03 C $\epsilon^3$ 121.20	42.07	--	--	--
C (mis.)	132.50 (C $\omega$ )	--	155.21 (C $\omega$ )	C $\zeta^3$ 122.40 C $\eta^2$ 125.00 C $\zeta^2$ 114.90	--	--	--	75.01*

9.3.6. Carbon Chemical Shifts for  $Y^{III}$ -DOTATOC (Minor-290 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
CO	172.57	171.02	172.98	175.87	--	--	173.05	64.14*
C $\alpha$	59.58	55.98	58.46	59.01	--	--	56.32	--
C $\beta$ /Me	40.91	44.99	40.04	29.03	--	--	41.52	--
C $\gamma$	138.05	--	--	--	--	--	--	--
C $\delta$	130.30	--	--	C $\delta^1$ --	--	--	--	--
				C $\delta^2$ --				
C $\epsilon$	132.00	--	--	C $\epsilon^2$ --	--	--	--	--
				C $\epsilon^3$ --				
				C $\zeta^3$ --				
C (mis.)	132.10	--	--	C $\eta^2$ --	--	--	--	--
	(C $\omega$ )			C $\zeta^2$ --				

9.3.7. Proton Chemical Shifts for  $Y^{III}$ -DOTATOC (Major-275 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
NH	9.61	8.83	8.32	8.75	8.49	8.12	7.94	7.82
H $\alpha$	4.87	4.66	4.63	4.16	3.81	4.29	4.77	3.85
H $\beta^u$	3.22	3.14	3.06	2.99	1.56	4.41	3.21	3.99
H $\beta^d$	2.87	2.90	2.83	2.80	1.24		2.91	
H $\gamma^u$	--	--	--	he3: 7.52	0.48	1.23	--	1.12
H $\gamma^d$				he1: 10.30	0.24			
H $\delta^u$	7.34	--	7.15	hd1: 7.10	1.28	--	--	--
H $\delta^d$				hita2: 7.23				
H $\epsilon^u$	7.41	--	6.89	hz2: 7.47	2.66	--	--	--
H $\epsilon^d$				hz3: 7.19				
Aromatic or other	7.34	--	--	--	--	5.48 (OH)	--	3.72 & 3.62 (H $\beta^*$ ), 5.20 & 5.31 (OH)

9.3.8. Proton Chemical Shifts for  $Y^{III}$ -DOTATOC (Minor-275 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
NH	9.38	8.41	8.07	8.67	8.49	8.10	7.77	7.74*
H $\alpha$	4.79	4.46	4.53	4.08	--	--	4.66	--
H $\beta^u$	3.15	2.81	3.15	2.94	--	--	3.19	--
H $\beta^d$	3.04	2.65	2.77	2.74	--	--	2.90	--
H $\gamma^u$	--	--	--	he3: --	--	--	--	--
H $\gamma^d$	--	--	--	he1: --	--	--	--	--
H $\delta^u$	7.32	--	--	hd1: --	--	--	--	--
H $\delta^d$	7.37	--	--	hita2: --	--	--	--	--
H $\epsilon^u$	7.37	--	--	hz2: --	--	--	--	--
H $\epsilon^d$	--	--	--	hz3: --	--	--	--	--
Aromatic or other	7.27	--	--	--	--	-- (OH)	--	-- (H $\beta$ ) -- (OH)

9.3.9. Carbon Chemical Shifts for  $Y^{III}$ -DOTATOC (Major-275 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
CO	170.99	168.56	170.02	172.68	173.62	170.46	169.79	61.20*
C $\alpha$	56.32	52.09	55.40	56.12	56.63	60.11	52.30	56.65
C $\beta$ /Me	38.10	41.58	37.81	26.03	29.31	66.53	38.93	66.13
C $\gamma$	134.67	--	128.58	106.05	21.61	19.24	--	19.05
C $\delta$	127.80	--	130.80	C $\delta^1$ 124.50 C $\delta^2$ 128.57	26.20	--	--	--
C $\epsilon$	129.40	--	115.50	C $\epsilon^2$ 133.97 C $\epsilon^3$ 118.20	39.10	--	--	--
C (mis.)	129.60 (C $\omega$ )	--	151.30 (C $\omega$ )	C $\zeta^3$ 119.50 C $\eta^2$ 122.10 C $\zeta^2$ 111.90	--	--	--	72.33*



9.3.10. Carbon Chemical Shifts for  $Y^{III}$ -DOTATOC (Minor-275 K)

AA	D-Phe <sup>1</sup>	Cys <sup>2</sup>	Tyr <sup>3</sup>	D-Trp <sup>4</sup>	Lys <sup>5</sup>	Thr <sup>6</sup>	Cys <sup>7</sup>	Thr <sup>8</sup>
CO	171.10	168.44	170.20	172.45	--	--	169.94	--
C $\alpha$	56.45	53.07	55.74	56.12	--	--	53.72	--
C $\beta$ /Me	37.00	42.27	37.09	26.08	--	--	38.59	--
C $\gamma$	134.46	--	--	--	--	--	--	--
C $\delta$	127.60	--	--	C $\delta^1$ --	--	--	--	--
				C $\delta^2$ --				
C $\epsilon$	129.00	--	--	C $\epsilon^2$ --	--	--	--	--
				C $\epsilon^3$ --				
				C $\zeta^3$ --				
C (mis.)	129.30	--	--	C $\eta^2$ --	--	--	--	--
	(C $\omega$ )			C $\zeta^2$ --				

9.3.11. NOEs Used for the Peptide Structure Calculation of Ga<sup>III</sup>-DOTATOC

From	To	Upper Limit	Lower Limit
Phe <sup>2</sup> H <sup>N</sup>	Phe <sup>2</sup> H <sup>β*</sup>	2.86	4.40
Phe <sup>2</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>N</sup>	3.50	6.50
Cys <sup>3</sup> H <sup>N</sup>	Phe <sup>2</sup> H <sup>α</sup>	2.30	2.80
Cys <sup>3</sup> H <sup>N</sup>	Phe <sup>2</sup> H <sup>β*</sup>	3.08	4.66
Cys <sup>3</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>β*</sup>	2.80	4.32
Cys <sup>3</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>N</sup>	4.00	6.50
Tyr <sup>4</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>α</sup>	2.00	2.80
Tyr <sup>4</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>α</sup>	2.72	3.32
Tyr <sup>4</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>N</sup>	4.00	6.50
Tyr <sup>4</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>α</sup>	5.00	6.40
Tyr <sup>4</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>β</sup>	3.53	4.32
Tyr <sup>4</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>γ2*</sup>	4.50	7.00
Trp <sup>5</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>α</sup>	2.00	2.60
Trp <sup>5</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>β*</sup>	4.50	7.00
Trp <sup>5</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>α</sup>	2.71	3.31
Trp <sup>5</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>β*</sup>	2.24	4.10
Trp <sup>5</sup> H <sup>α</sup>	Trp <sup>5</sup> H <sup>δ1</sup>	3.00	3.80
Trp <sup>5</sup> H <sup>α</sup>	Trp <sup>5</sup> H <sup>ε3</sup>	2.56	3.13
Lys <sup>6</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>α</sup>	2.24	2.74
Lys <sup>6</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>β1</sup>	4.50	5.50
Lys <sup>6</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>β2</sup>	4.50	5.50
Lys <sup>6</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>α</sup>	2.75	3.36
Lys <sup>6</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>β*</sup>	2.35	5.18
Lys <sup>6</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>γ*</sup>	2.66	4.74
Lys <sup>6</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>N</sup>	3.01	3.68
Lys <sup>6</sup> H <sup>α</sup>	Trp <sup>5</sup> H <sup>δ1</sup>	3.32	4.05
Lys <sup>6</sup> H <sup>α</sup>	Trp <sup>5</sup> H <sup>ε3</sup>	3.45	4.22
Lys <sup>6</sup> H <sup>α</sup>	Trp <sup>5</sup> H <sup>ζ3</sup>	4.24	5.18

From	To	Upper Limit	Lower Limit
Lys <sup>6</sup> H <sup>α</sup>	Lys <sup>6</sup> H <sup>γ*</sup>	2.43	4.48
Lys <sup>6</sup> H <sup>β1</sup>	Trp <sup>5</sup> H <sup>ε3</sup>	5.00	6.40
Lys <sup>6</sup> H <sup>β2</sup>	Trp <sup>5</sup> H <sup>ε3</sup>	5.00	6.40
Lys <sup>6</sup> H <sup>γ1</sup>	Trp <sup>5</sup> H <sup>δ1</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>γ2</sup>	Trp <sup>5</sup> H <sup>δ1</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>γ1</sup>	Trp <sup>5</sup> H <sup>ε1</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>γ2</sup>	Trp <sup>5</sup> H <sup>ε1</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>γ1</sup>	Trp <sup>5</sup> H <sup>ζ2</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>γ2</sup>	Trp <sup>5</sup> H <sup>ζ2</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>δ*</sup>	Trp <sup>5</sup> H <sup>ε1</sup>	2.95	7.00
Lys <sup>6</sup> H <sup>δ1</sup>	Trp <sup>5</sup> H <sup>ζ2</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>δ2</sup>	Trp <sup>5</sup> H <sup>ζ2</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>ε1</sup>	Trp <sup>5</sup> H <sup>ε1</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>ε2</sup>	Trp <sup>5</sup> H <sup>ε1</sup>	4.50	6.00
Lys <sup>6</sup> H <sup>*ε</sup>	Trp <sup>5</sup> H <sup>ζ2</sup>	3.43	5.09
Thr <sup>7</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>α</sup>	2.76	3.74
Thr <sup>7</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>β*</sup>	3.57	5.26
Thr <sup>7</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>α</sup>	2.59	3.16
Thr <sup>7</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>β</sup>	2.60	3.18
Thr <sup>7</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>N</sup>	3.00	6.50
Cys <sup>8</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>α</sup>	5.00	6.50
Cys <sup>8</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>β*</sup>	5.00	6.40
Cys <sup>8</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>α</sup>	2.76	3.37
Cys <sup>8</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>β</sup>	3.12	3.81
Cys <sup>8</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>β1</sup>	3.60	4.77
Cys <sup>8</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>β2</sup>	2.48	3.03
Cys <sup>8</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>N</sup>	3.00	4.20
Thr <sup>9</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>α</sup>	2.23	2.72
Thr <sup>9</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>β1</sup>	2.98	3.64
Thr <sup>9</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>β2</sup>	3.51	4.29
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>α</sup>	2.77	3.39

---

From	To	Upper Limit	Lower Limit
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>β</sup>	2.83	3.46
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>γ2*</sup>	3.00	5.50
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>C*</sup>	2.72	3.57
Thr <sup>9</sup> H <sup>α</sup>	Thr <sup>9</sup> H <sup>γ2*</sup>	2.86	4.49
Thr <sup>9</sup> H <sup>γ2*</sup>	Thr <sup>9</sup> H <sup>C*</sup>	3.32	5.25

---

9.3.12. NOEs Used for the Peptide Structure Calculation of  $Y^{III}$ -DOTATOC (Major)

From	To	Lower Bound	Upper Bound
Cys <sup>3</sup> H <sup>N</sup>	Phe <sup>2</sup> H <sup>β*</sup>	2.49	4.77
Cys <sup>3</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>α</sup>	2.73	3.33
Cys <sup>3</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>β*</sup>	2.18	3.56
Tyr <sup>4</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>β*</sup>	1.79	3.65
Tyr <sup>4</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>β*</sup>	1.78	3.17
Tyr <sup>4</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>β</sup>	3.47	4.25
Tyr <sup>4</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>α</sup>	3.70	4.52
Trp <sup>5</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>α</sup>	1.90	2.32
Trp <sup>5</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>α</sup>	2.53	3.10
Trp <sup>5</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>β*</sup>	1.78	2.82
Lys <sup>6</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>α</sup>	2.16	2.63
Lys <sup>6</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>β*</sup>	2.87	5.21
Lys <sup>6</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>α</sup>	2.81	3.44
Lys <sup>6</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>β*</sup>	1.78	3.54
Thr <sup>7</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>α</sup>	3.08	3.76
Thr <sup>7</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>β*</sup>	2.53	4.82
Thr <sup>7</sup> H <sup>α</sup>	Thr <sup>7</sup> H <sup>γ2*</sup>	2.17	4.48
Thr <sup>7</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>β</sup>	2.41	2.94
Cys <sup>8</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>γ2*</sup>	2.70	5.12
Cys <sup>8</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>β*</sup>	1.78	2.93
Thr <sup>9</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>α</sup>	2.11	2.58
Thr <sup>9</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>*β</sup>	2.17	3.88
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>α</sup>	2.76	3.38
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>β</sup>	2.95	3.61
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>C*</sup>	2.59	3.42
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>γ2*</sup>	2.45	4.82
Thr <sup>9</sup> H <sup>C*</sup>	Thr <sup>9</sup> H <sup>γ2*</sup>	1.78	4.67

9.3.13. NOEs Used for the Peptide Structure Calculation of Y<sup>III</sup>-DOTATOC (Minor)

From	To	Upper Limit	Lower Limit
Cys <sup>3</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>α</sup>	2.72	3.32
Cys <sup>3</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>β*</sup>	2.19	3.50
Tyr <sup>4</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>α</sup>	2.00	2.45
Tyr <sup>4</sup> H <sup>N</sup>	Cys <sup>3</sup> H <sup>β*</sup>	1.90	4.04
Tyr <sup>4</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>α</sup>	2.64	3.23
Tyr <sup>4</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>β*</sup>	2.00	4.16
Tyr <sup>4</sup> H <sup>α</sup>	Tyr <sup>4</sup> H <sup>β*</sup>	1.78	3.62
Trp <sup>5</sup> H <sup>N</sup>	Tyr <sup>4</sup> H <sup>α</sup>	2.01	2.46
Trp <sup>5</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>α</sup>	2.73	3.33
Trp <sup>5</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>β*</sup>	1.78	2.96
Trp <sup>5</sup> H <sup>α</sup>	Trp <sup>5</sup> H <sup>β*</sup>	1.93	4.08
Lys <sup>6</sup> H <sup>N</sup>	Trp <sup>5</sup> H <sup>α</sup>	2.00	2.44
Lys <sup>6</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>α</sup>	2.71	3.31
Lys <sup>6</sup> H <sup>α</sup>	Lys <sup>6</sup> H <sup>β*</sup>	1.78	3.51
Thr <sup>7</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>α</sup>	2.85	3.48
Thr <sup>7</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>γ2*</sup>	2.40	4.75
Cys <sup>8</sup> H <sup>N</sup>	Lys <sup>6</sup> H <sup>α</sup>	3.77	4.61
Cys <sup>8</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>α</sup>	2.55	3.11
Cys <sup>8</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>β</sup>	3.26	3.98
Cys <sup>8</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>α</sup>	2.61	3.19
Cys <sup>8</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>α</sup>	4.40	5.38
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>α</sup>	4.14	5.06
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>7</sup> H <sup>β</sup>	4.94	6.03
Thr <sup>9</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>α</sup>	2.21	2.70
Thr <sup>9</sup> H <sup>N</sup>	Cys <sup>8</sup> H <sup>β*</sup>	2.35	4.08
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>α</sup>	2.63	3.21
Thr <sup>9</sup> H <sup>N</sup>	Thr <sup>9</sup> H <sup>β</sup>	3.14	3.84
Thr <sup>9</sup> H <sup>β</sup>	Thr <sup>9</sup> H <sup>C*</sup>	2.73	3.50

## 9.4. List of Publications

### Teile dieser Arbeit sind bereits erschienen:

1. F Kramer, **M V Deshmukh**, H Kessler, and S J Glaser “Residual Dipolar Coupling Constant: An elementary derivation of key equations” *Concepts in Magnetic Resonance*, 21A (1), 10–21, 2004.
2. **M V Deshmukh**, Georg Voll, Angelika Kühlewein, Helmut Mäcke, Jörg Schmitt, Horst Kessler and Gerd Gemmecker “NMR Studies of the Gallium and Yttrium Complexes of DOTA-D-Phe<sup>1</sup>-Tyr<sup>3</sup>-Octreotide: Structural Differences Correlate with Differences in Potency” Submitted to *Journal of Medicinal Chemistry* am 19.05.2004

### Following part of the results belonging to this thesis has been already published:

1. F Kramer, **M V Deshmukh**, H Kessler, and S J Glaser “Residual Dipolar Coupling Constant: An elementary derivation of key equations” *Concepts in Magnetic Resonance*, 21A (1), 10–21, 2004.

### Following part of the results belonging to this thesis has been communicated for publication:

1. **M V Deshmukh**, Georg Voll, Angelika Kühlewein, Helmut Mäcke, Jörg Schmitt, Horst Kessler and Gerd Gemmecker “NMR Studies of the Gallium and Yttrium Complexes of DOTA-D-Phe<sup>1</sup>-Tyr<sup>3</sup>-Octreotide: Structural Differences Correlate with Differences in Potency” Submitted to *Journal of Medicinal Chemistry* on 19.05.2004

## 9.5. Zusammenfassung für das Jahrbuch der TUM

### *auf Deutsch*

In letzter Zeit wurden dipolare Restkopplungen (*residual dipolar couplings*, RDCs) als Mittel zur Strukturbestimmung von Biomolekülen eingeführt. Das physikalische Konzept und die mathematische Ableitung der RDCs werden hier vorgestellt. RDCs wurden ebenso zur Strukturverfeinerung und Berechnung der Orientierung der zwei Teildomänen von VAT-N (20,5 kDa) verwendet. VAT-N ist die substraterkennende Domäne des hexameren VAT-Komplexes, der eine dreigeteilte Domänenstruktur hat: N-D1-D2. Die Interdomänendynamik von VAT-N wurde mit Hilfe der  $^{15}\text{N}$  Rückgratrelaxationsparameter untersucht. Nach diesen Untersuchungen hat VAT-N eine parallele Ausrichtung und geringe relative Bewegung der Teildomänen. Es werden ebenfalls Ergebnisse von Titrations von VAT-N mit SsrA, Ubiquitin, Kasein und Barstar präsentiert. Mit Hilfe NMR-spektroskopischer Untersuchungen wurden schließlich die unterschiedlichen Wirksamkeiten der Ga-III und Y-III komplexe von Somatostatin-Analagon, DOTATOC, untersucht.

### *in English*

Residual Dipolar Couplings (RDCs) have been recently established as a structural restraint for biomolecular structure calculation. Here the physical concept of the RDC constant is demonstrated using an easy mathematical approach. RDCs were used for structure refinement and to calculate the subdomain orientation of VAT-N (20.5 kDa). VAT-N is the substrate recognition domain of the hexameric VAT complex formed by a tripartite domain structure, N-D1-D2. The inter-domain dynamics of VAT-N has been studied by  $^{15}\text{N}$  backbone relaxation parameters in the model-free framework. In conclusion, the relative orientation of VAT-N subdomains is parallel (kidney shape) and no large amplitude motion exists between them. Results of titration of VAT-N with SsrA, Ubiquitin, casein and Barstar are presented. The differences in the potency of Ga-III and Y-III chelated somatostatin analogue (DOTATOC) are studied by NMR.



## 9.6. Lebenslauf

**Name:** Deshmukh  
**Vorname:** Mandar Vinayakrao  
**Geboren:** 07.08.1975 in Nanded, Indien  
**Familienstand:** verheiratet mit Mrunal Pendke  
**Dienstanschrift:** Organische Chemie  
Technische Universität München  
Lichtenbergstr. 4, D-85747 Garching  
Tel.: (089) 289-13760  
Fax.: (089) 289-13210  
**E-Mail:** mvdeshmukh@ch.tum.de  
**Internet:** [www.org.chemie.tu-muenchen.de/people/mvdesh](http://www.org.chemie.tu-muenchen.de/people/mvdesh)  
**Akademischer Grad:** Master of Science (M. Sc.)



## SCHULAUSSBILDUNG

06/80 - 06/90      Grundschule (Secondary School Certificate) Nanded, Indien  
07/90 - 06/92      Higher Secondary School Certificate, Latur, Indien.

## WISSENSCHAFTLICHER WERDEGANG

07/92 - 07/95      „Bachelor of Science“ (B. Sc.), in Physik, Chemie und Mathematik  
bei Dr. B. A. Marathwada University, Aurangabad, Indien  
07/95 - 07/97      „Master of Science“ (M. Sc.), in Organischer Chemie bei Dr. B. A.  
Marathwada University, Aurangabad, Indien  
08/97 - 05/98      Arbeit als R & D Chemiker bei Canpex Chemicals, Pune, Indien  
05/98 - 09/00      Wissenschaftlicher Assistent im „National Chemical Laboratory“,  
Pune, Indien  
seit 10/00          Promotion im Fachgebiet NMR Spectroscopie bei Prof. Dr. H.  
Kessler, Lehrstuhl II, Technische Universität München

## Stipendien und Auszeichnungen:

1. „Open Merit Scholarship“ bei der „middle school scholarship examination“ (Rank 15<sup>th</sup>).
2. Reisestipendium zur 43 ENC (Experimental Nuclear Magnetic Conference) in Asilomar, Pacific Grove, CA, USA, in 2002.

## Publikationen

1. LCST in Poly(*N*-isopropylacrylamide) copolymers: High resolution proton NMR investigations  
**M V Deshmukh**, A A Vaidya, M G Kulkarni, P R Rajamohanan, and S Ganapathy  
Polymer, 41, 7951-7960, 2000.
2. Design and evaluation of new ligands for lysozyme recovery by affinity thermoprecipitation.  
A A Vaidya, B S Lele, **M V Deshmukh**, and M G Kulkarni  
Chemical Engineering Science, 56, 5681-5692, 2001.
3. Residual Dipolar Coupling Constant: An elementary derivation of key equations.  
F Kramer, **M V Deshmukh**, H Kessler, and S J Glaser  
Concepts in Magnetic Resonance, 21A (1), 10–21, 2004.
4. NMR Studies of the Gallium and Yttrium Complexes of DOTA-D-Phe<sup>1</sup>-Tyr<sup>3</sup>- Octreotide: Structural Differences Correlate with Differences in Potency  
**M V Deshmukh**, Georg Voll, Angelika Kühlewein, Helmut Mäcke, Jörg Schmitt, Horst Kessler and Gerd Gemmecker  
Submitted to Journal of Medicinal Chemistry am 19.05.2004

## Hobbies

Cricket und Schach  
Amateurastronomie (Sterngucker)  
Buecher lesen

München, den 29.Juni 2004

Mandar Vinayakrao Deshmukh