DIPLOMARBEIT

Titel

NMR Studies of the Leader Protease of the Foot-and-Mouth Disease Virus

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Verfasserin / Verfasser:	Sabine Schultes
Matrikel-Nummer:	0408332
Studienrichtung (lt. Studienblatt):	A490
Betreuer:	Dr. Georg Kontaxis

Wien, am 7.11.2008

Summary

The leader protease Lb^{pro} of the Foot-and-Mouth Disease Virus (FMDV) belongs to the family of papaine-like cysteine proteinases and is an important virulence determinant of the virus. It is the first encoded protein on the viral genome. During expression in the host cell it is autocatalytically cleaved from the growing polypeptide. Subsequently, it shuts down eukaryotic translation by cleavage of eukaryotic initation factor 4GI and II thus preventing among others IFN synthesis. The protein synthesis from the viral genome is not affected as translation is initiated from an internal ribosomal entry site (IRES).

Crystal structures of Lb^{pro} (1QOL.pdb) as well as of a shortened mutant sLb^{pro} (1QMY.pdb), lacking six C-terminal residues, already existed. They revealed dimerisation of Lb^{pro} by binding of the C-terminal residues to the active site of an adjacent Lb^{pro} molecule and vica versa. This dimerisation was not observed for the shortened mutant. Furthermore the C-terminus was not observable indicating that it is fexible and disordered.
NMR studies of Lb^{pro} (2JQF.pdb) and sLb^{pro} (2JQG.pdb) confirmed the dimerisation of Lb^{pro} but they revealed that Lb^{pro} forms a completely symetric dimer in solution. Additionally the C-terminus of sLb^{pro} was shown to be indeed unstructured and flexible

The aim was to refine the NMR structures of Lb^{pro} as well as of sLb^{pro}. More complete assignment could be achieved with the use of triple resonance TROSY experiments. TROSY experiments have a more favourable ^{15}N relaxation behaviour resulting in spectra whose peaks have smaller widths allowing more accurate peak assignment. For the refinement of the structures, residual dipolar couplings (RDCs) were measured. RDCs provided long range orientational restraints. In the case of Lb^{pro} they are not

only useful to refine the structure but also to determine the precise orientation of the two halves of the dimer to each other.

For both proteins, Lb^{pro} and sLb^{pro}, various dipolar couplings were measured. The proteins were oriented via the phage *PF1* in the magnetic field. The alignment tensor elements were determined by least square fitting of the measured dipolar couplings using the coordinates from the crystal structures. The dipolar couplings were then used as additional restraints for structures calculation.

The structures obtained show a better convergence compared with the existing NMR structures. Their precision is a least as good as the precision of the crystal structures. Additionally the frequency of amino acid residues with energetically favorable dihedral angles, ϕ and ψ, is higher especially for sLb^{pro}. The quarternary structure of the obtained Lb^{pro} compared to the NMR structure (2JQF.pdb) has the same twist angle but differs in the bending angle of about 20-25°.

Zusammenfassung

Die Leader Protease Lb^{pro} des Maul-und-Klauenseuche Viruses gehört zu der Familie der Papain-Cysteinproteasen und ist ein wichtiger Virulenzfaktor des Viruses. Es ist das erste Protein das am Genom des Viruses codiert ist. Während der Expression in der Wirtszelle spaltet es sich autocatalytisch vom wachsenden Polypeptid ab. Danach stoppt es die eukaryotische Translation durch die proteolytische Spaltung des eukaryotischen Translations-initiationfaktors 4GI and II, wodurch unter anderem die Synthese von IFN verhindert wird. Die Proteinsynthese des viralen Genoms ist nicht beeinträchtigt, da die Translation von einer 'Internal Ribosome Entry Site' (IRES) initiiert wird.

Kristallstukturen von Lb^{pro} (1QOL.pdb), als auch von einer verkürzten Mutante, sLb^{pro} (1QMY.pdb), der sechs C-terminale Aminosäuren fehlen, existierten bereits. Diese zeigten, dass Lb^{pro} durch die Bindung der C-terminalen Aminsäuren im aktiven Zentrum eines benachbarten Lb^{pro} Moleküles und umgekehrt dimerisiert. Diese Dimerisierung konnte bei der verkürzten Mutante nicht beobachtet werden. Zusätzlich war bei dieser der C-Terminus nicht sichtbar, was daraufhin weiste, dass dieser unstrukturiert und flexible ist
Untersuchungen mittels NMR von Lb^{pro} (2JQF.pdb) und sLb^{pro} (2JQG.pdb) bestätigten die Dimerisierung von Lb^{pro}, zeigten aber, dass Lb^{pro} in Lösung ein symmetrisches Dimer bildet. Zusätzlich konnte gezeigt werden, dass der C-Terminus von sLb^{pro} wirklich unstrukturiert und flexibel ist.

Das Ziel dieser Arbeit war die NMR Strukturen von Lb^{pro} und sLb^{pro} zu verfeinern. Vollständigere Zuordnung konnte durch die Verwendung von dreifach Resonanz TROSY Experimenten erzielt werden. TROSY Experimente haben ein verbessertes ^{15}N Relaxationsverhalten. Dadurch

haben die Peaks dieser Spektren eine geringere Peakbreite, wodurch diese Peaks einfacher und genauer zugeordnet werden können. Für die Verfeinerung der Strukturen wurden dipolare Kopplungen gemessen. Dipolare Kopplungen erstellten Orientierungsbeschränkungen über große Distanzen. Im Fall für Lb^{pro} sind sie daher nicht nur nützlich um die Struktur zu verfeinern, sondern auch um die genaue Orientierung der beiden Hälften des Dimers relativ zueinander zu bestimmen.

Für beide Proteine, Lb^{pro} and sLb^{pro} wurden unterschiedliche dipolare Kopplungen gemessen. Die Proteine wurden durch den Phagen *PF1* im Magnetfeld orientiert. Die Elemente des Orientierungstensors wurden durch least square fitting der gemessenen dipolaren Kopplungen unter der Zuhilfenahme der Kristallstrukturen berechnet. Die dipolaren Kopplungen wurden dann als zusätzliche Orientierungsbeschränkungen für die Berechnung der Strukturen eingesetzt.

Die erhaltenen Strukturen zeigen eine bessere Konvergenz verglichen mit den existierenden NMR Strukturen. Zusätzlich ist die Anzahl von Aminosäuren, welche energetisch günstigere dihedarale Winkel, ϕ und ψ, aufweisen höher. Dies ist vor allem bei sLb^{pro} der Fall. Die quartäre Struktur, die von Lb^{pro} erhalten wurde, zeigt verglichen mit der NMR-Struktur (2JQF.pdb), dass beide Strukturen denselben Drehungswinkel haben, jedoch unterscheidet sich ihr Beugungswinkel um ca. 20-25°.

Table of Contents

1. Introduction

1.1 Foot-and-Mouth Disease Virus (FMDV)

FMDV is a member of the family Picornaviridae. Picornaviruses are nonenveloped viruses with a single stranded RNA genome of positive sense. The RNA is covaletly linked to a VGP protein at its 5'end and is infectious, meaning it can be directly translated in a host cell. Picornaviruses have an icosahedral capsid with a diameter of about 30nm. Because of this small size the virus has its name, Pico.

Genus	Serotypes	Spezies
Rhinovirus	102 3	Human Rhinovirus (1A, 1B – 100) Bovine Rhinovirus (1 – 3)
Aphthovirus	7 1	Foot-and-Mouth Disease Virus (A, C, O, SAT-1, SAT-2, SAT-3, Asia-1) Equine Rhinitis A Virus
Enterovirus	3 23 6 30 4 2 1 2 1	Poliovirus (1 – 3) Coxsackie Virus A (1 - 22, 24) Coxsackie Virus B (1 – 6) Echovirus (1 - 7, 9, 11 - 21, 24 - 27, 29 - 34) Human Enterovirus (68 – 71) Bovine Enterovirus (1, 2) Porcine Enterovirus A (PEV-8) Porcine Enterovirus B (PEV-9, PEV-10) Simian Enterovirus A
Parechovirus	6 1	Human Parechovirus (1, 2, 3, 5, 5, 6) Ljungan virus
Cardiovirus	1 4	Enzephalomyocarditis Virus (EMCV) Theilovirus (Theiler's Murine Encephalomyelitis Virus (TMEV), Vilyuisk Human Encephalomyelitis Virus

		(VHEV), Theiler-like Virus (TLV), Scaffold virus (SAF-V)
Hepatovirus	2 1	Hepatitis A Virus (HAV) Avian encephalomyelitis-like viruses (AEV)
Teschovirus	1	Porcines Teschovirus (PTV-1 bis -11)
Kobuvirus	1	Aichi Virus Bovine Kobuvirus
Erbovirus	1	Equine Rhinovirus B

Table1: Family of the Picornaviradea [1]

There are seven serotypes of FMDV (A, O, C, Asia1, SAT1, SAT2 and SAT3). Together with Equine rhinitis A virus they constitute the genus Aphtho-virus.(Li et al., 1996; Wutz et al., 1996)

1.1.1 Foot and Mouth Disease

FMDVs are highly contagious. They infect cloven-hoofed animals, causing foot-and-mouth disease. This disease was first described by Frasastorius in 1546 (Fracastorius, 1546). In 1897 Loeffler and Frosch demonstrated foot-and-mouth disease (FMD), as the first disease to be caused by a filterable agent (Loeffler and Frosch, 1897). FMD disease is very infectious and can easily be transmitted by aerosols among infected animals. Also abrasions on the skin or mucous membrane and the oral route by eating FMDV contaminated food can lead to FMDV infections. The incubation period can be between 2 and 14 days, depending on the infection dose and route of infection (Grubman and Baxt, 2004) and the infection can last at least about seven to ten days. The disease is characterized by the appearance of vesicles on the feet and in or around the mouth (Alexandersen et al., 2001) causing foamy salivation and lameness. Adult animals are not as threatened

---

[1] http://www.med.uni-jena.de/virologie/zell/lehre/picornaviren/picornaviren.html

by the disease as younger animals where the disease can lead to inflammation of the heart. A big danger emanates from infected animal which retain asymptotic. They can act as carriers of the disease which is a risk for healthy animals (Hughes et al., 2002).

After an outbreak the only way of getting rid of the disease is achieved by slaughtering of the complete animal live stock in the areas where the disease occurred. Together with additional trading controls an outbreak of the disease lead to high financial losses. Therefore much effort was invested in the development of vaccines. Nowadays vaccines and fast diagnostics are available, leading to improved control of the disease in developed countries. Nevertheless due to the great economic impact of the disease there is interest in completely understanding the molecular mechanisms underlying infection and transmission of the virus.

1.1.2 Structure

First studies with electron microscopy of the FMDV revealed a rather smooth round particle. (Bachrach, 1968) Later x-ray crystallographic studies could determine its structure at atomic resolution (Chan and Wool, 1990). The viral capsid is build up of four structural proteins (VP1-4). From each protein 60 copies assemble forming a T3 icosahedral architecture (Jackson et al., 2003).

VP1-3 are located on the outside and fold into wedge shaped β-barrels. VP4 is located at the inner surface and possesses a myristoyl group which may be involved in capsid assembly or in entry of virus into cells (Belsham et al., 1991; Chow et al., 1987).
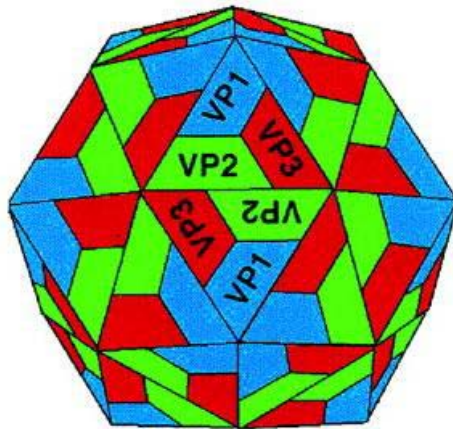
Fig.1: Iscosahedral capsid of the FMDV

VP1-4 build one protomer. 5 protomer assemble of one pentamer. 12 pentamers assemble to a T3 iscosahedral architecture. VP4 is not visible as it is inside the particle. Adapted from (Oliveira et al., 1999)

At a pH below seven the capsid dissociates into twelve pentamers by protonating histidine residues of VP3 leading to weakened polar interaction between the pentamers. (Brown and Cartwright, 1961; van Vlijmen et al., 1998). VP1 consist of a hypervarible sequence (the GH loop) containing a highly conserved tripeptide Arg-Gly-Asp (RGD). This motive is important for attachment on the cellular receptor, integrin. (Berinstein et al., 1995; Jackson et al., 2000; Neff et al., 2000). It is also known that after several passages through cell lines the virus can bind to heparansulfate proteoglycanes (Sa-Carvalho et al., 1997).

1.1.3 Viral infectious Cycle

1.1.3.1 Entry

The first event in infection is the binding of the virus to its cellular receptors. After penetration into the cell the virus enters the acidic endosome. The lowering pH mediates the capsid disassembling and the RNA is released into the cytoplasm. The Vpg protein of the viral RNA is cleaved by cellular

proteases.

1.1.3.2 Translation

The translation of the viral RNA is mediated via an IRES whereas the eukaryotic host cell protein synthesis is shut off by cleavage of both isoforms of eIF4G, I and II, by the viral leader protease L^{pro} (Kuehnel et al., 2004). eIF4G acts as a scaffold protein in eukaryotic translation initiation. Its N-terminal part interacts with eIF4E which binds to the 5´cap of the host cell mRNA and its C terminal part interacts with eIF3 binding to the 40S ribosomal subunit. Thus eIF4G is involved in the connection of the 40S ribosomal subunit with the host cell mRNA. After cleavage, the host cell translation is disabled, which also impairs the host cell immune response. The C-terminal part of eIF4G can bind to the IRES, resulting in an interaction of the 40S ribosomal subunit with the viral RNA (Lopez de Quinto and Martinez-Salas, 2000).
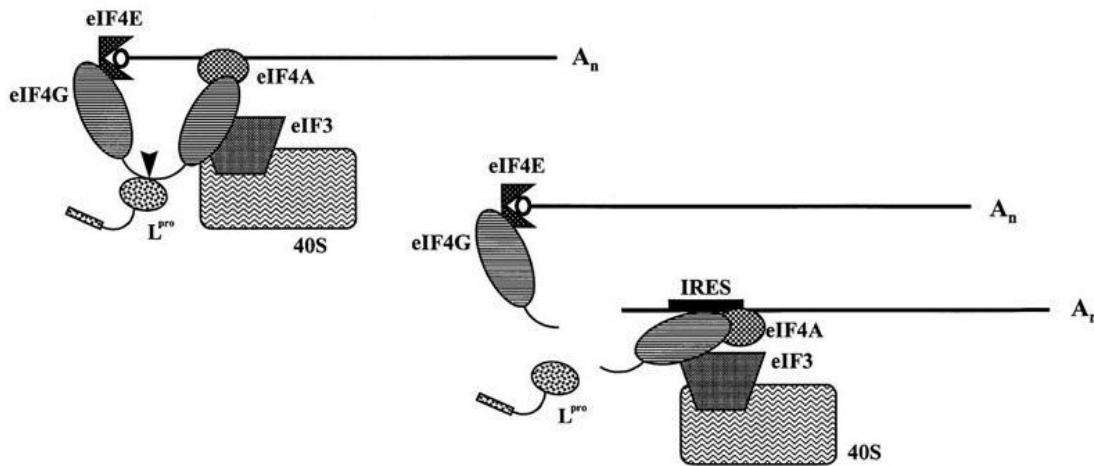
Fig.2: Cleavage of eIF4G by Lb^{pro}

Lb^{pro} cleaves both isoforms of eIF4G thus disabling host cell translation. The C-terminal part of eIF4G is still able to forms an initiation complex with IRES-containing mRNAs. Adopted from (Guarne et al., 1998)

The viral RNA is translated into a single polyprotein which is subsequently cleaved by the viral proteases, L^{pro}, $2A^{pro}$ and $3C^{pro}$creating the individual viral proteins. Then the replication cycle starts.

1.1.3.3 Replication

In this procedure the viral proteins 2B, 2C (putative helicase) 3A, 3B (primer) $3C^{pro}$ and 3D (viral RNA polymerase) are involved (Lyle et al., 2002; Nayak et al., 2006; Suhy et al., 2000). First the minus RNA strand synthesis takes place leading to the formation of a double stranded RNA molecule. On the basis of this replicative form further plus RNA strands are produced, which are used for further translation and for the assembly of new virus particles.

1.1.3.4 Assembly

The synthesized positive RNA strands are packed within a capsid that is formed by the structural proteins (VP0-3). A maturation cleavage of VP0 into VP2 and VP4 yields fully infectious viral particles which are subsequently released from the cell by cell lysis.

1.1.4 Genome of FMDV

The genome of the foot and mouth disease virus is single stranded RNA of positive sense. It consists of about 8500bp.
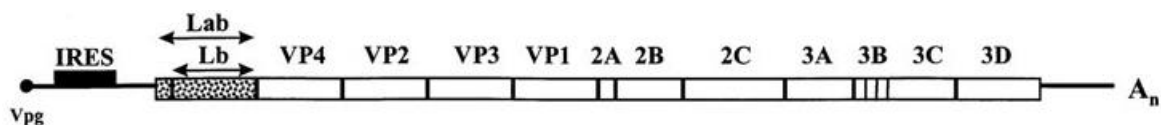
Fig.3: Genome organization of the foot and mouth disease virus (Guarne et al., 1998)

Every genomic RNA strand is covalently linked to a protein (VPG) at its

5′terminus. This is the primer (encoded from 3B) which remains on the RNA after replication (Wimmer, 1982). After the VPG protein the 5'-UTR follows. Downstream of the 5'-UTR the polyprotein is encoded followed by the 3'-UTR. At the 3' end the RNA consists of a polyA tail.

1.1.4.1 5'-UTR

The 5-'UTR is over 1300bp in length (Forss et al., 1984). The first fragment is a long stem loop with about 360bp in length. This is the S-fragment which is assumed to prevent the RNA genome from digestion by exonucleases which are incident in the infected cell (Grubman and Bachrach, 1979). Investigation on the poliovirus genome suggest that a structured 5' end plays a role in genome stability and is presumably implicated in replication (Barton et al., 2001). Downstream the S-fragment is a polyC tract which consists of 90% Cytosines and is over 100bp long. Studies on the poliovirus genome revealed that the polyC tract might be involved in regulating the switch from translation to replication (Walter et al., 2002). After the polyC tract are RNA psuedoknot structures whose function are still undiscovered. Then a hairpin loop structure follows which is known as the the cis-acting replicative element (cre). The cre element consists of a conserved AAACA sequence which is required for genome replication (Mason et al., 2002). The last part on the 5'-UTR is the internal ribosome entry site (IRES). For picornaviridae three different IRES have been identified (Pilipenko et al., 1989) and the foot and mouth disease virus consists an IRES of type two. An IRES binds with eIFs, who interact with the cellular RNA polymerase enabling transcription of the uncapped RNA. In the case of FMDV the C-terminal cleavage product of eIF4G generated by Lb[pro] is responsible for this task. Via eIF3 it connects the cellular RNA polymerase to the viral RNA. Downstream of the IRES there are two AUG startcodons, where the translation of the viral polyprotein can start.

1.1.4.2 Encoded Polyprotein Region

The first encoded protein is the leader protease L^{pro}. Due to the two AUG startcodons which are separated by 84nt the leader protease can exist in two forms, the longer form Lab^{pro} and the 28aa shorter form Lb^{pro} (Kuhn et al., 1990; Sangar et al., 1987). Investigations by Cao et al suggest that in vivo the shorter form, Lb^{pro} is the functionally important form (Cao et al., 1991).

Downstream of L^{pro} the P1, the P2 and the P3 regions follows. The P1 region encodes the four structural virus proteins - VP1, VP2, VP3 and VP4. The P2 and the P3 region code for the nonstructural proteins which participate in RNA replication, structural protein folding and assembly.

The viral proteins are translated as a single polyprotein which is subsequently cleaved by encoded viral proteases - Lb^{pro}, 2A^{pro} and 3C^{pro}. The leader protease cleaves itself off the polyprotein during translation. The second cleavage is performed by 2A^{pro} between 2A and 2B but investigations from Donnelly et al. suggests that this could also be a premature release of the polyprotein during translation (Donnelly et al., 2001a; Donnelly et al., 2001b) The other cleavages (except the maturation cleavage, which is an autocatalytic cleavage (Harber et al., 1991; Knipe et al., 1997; Lee et al., 1993) are performed by the protease 3C^{pro} to obtain the viral proteins.

1.1.4.3 3'-UTR

The 3'-UTR is folded into a stem loop structure and ends with a polyA tail. Both features are required for replication. The stem loop structure is able to bind viral proteins which are involved in replication (Melchers et al., 1997; Pilipenko et al., 1996; Rohll et al., 1995) and the poly A tail is thought to be involved in the initiation of replication (Barton et al., 2001; Herold and Andino, 2001). Furthermore it has been suggested that the polyA tail is also involved in translation (Lopez de Quinto et al., 2002). Investigtigation by (Saiz et al., 2001) indicated that the 3'UTR is specific for each picornavirus.

1.1.5 The Leader Protease

The leader protease is an important determinant of virulence. It is responsible for the shut off of host cell protein translation including INF-γ synthesis which is important for antiviral host defense. As mentioned, the leader protease can exist in two forms, Lab[pro] and Lb[pro] whereby (Cao et al., 1991) revealed that Lb[pro] is the preferentially synthesized form in vivo. For this reason Lb[pro] is the more interesting form to investigate.

The crystal structures of the inactive mutant Lb[pro] C51A as well as a shortened form lacking the six C-terminal amino acids, sLb[pro] C51A, were studied by x-ray crystallography (Guarne et al., 2000; Guarne et al., 1998). These structures confirmed that Lb[pro] is related to the family of papain-like cysteine proteinases. Although sequence identity of Lb[pro] shared with papain is no higher than 15% (Gorbalenya et al., 1991; Skern et al., 1998) characteristic structural features are conserved. The active site of papaine like cysteine proteinases is characterized by a catalytic triad of cysteine, histidine and asparagine. In Lb[pro] this is Cys51, His148 and in place of an asparagine is Asp163, which fulfills the same task as the asparagine in papain.

Lb[pro] has, like other papain-like proteases, a globular region from which a unique 18aa C-terminal extension (CTE) extrudes. In the crystal structure the last six amino acids of this extension reach into the active site of an adjacent molecule, forming a dimer, which is slightly asymmetric. In the six amino acid shortened form, dimer formation was not observed. Further the CTE was not observable in the crystal structure, indicating that it is disordered.

To prove that the dimer formation of Lb[pro] is not an artefact of crystal packaging and to address the issue of the characteristics of the C-terminus of sLb[pro], the solution structures of Lb[pro] C51A as well as sLb[pro] C51A were studied by NMR (Cencic et al., 2007). They revealed that dimer formation of

Lb^{pro} also occurs in solution. But in contrast to the crystal structure, the dimer was found to be symmetric. In the NMR structures of sLb^{pro} C51A the CTE was observable indicating that it is indeed flexible. This was additionally proved by NMR relaxation studies.
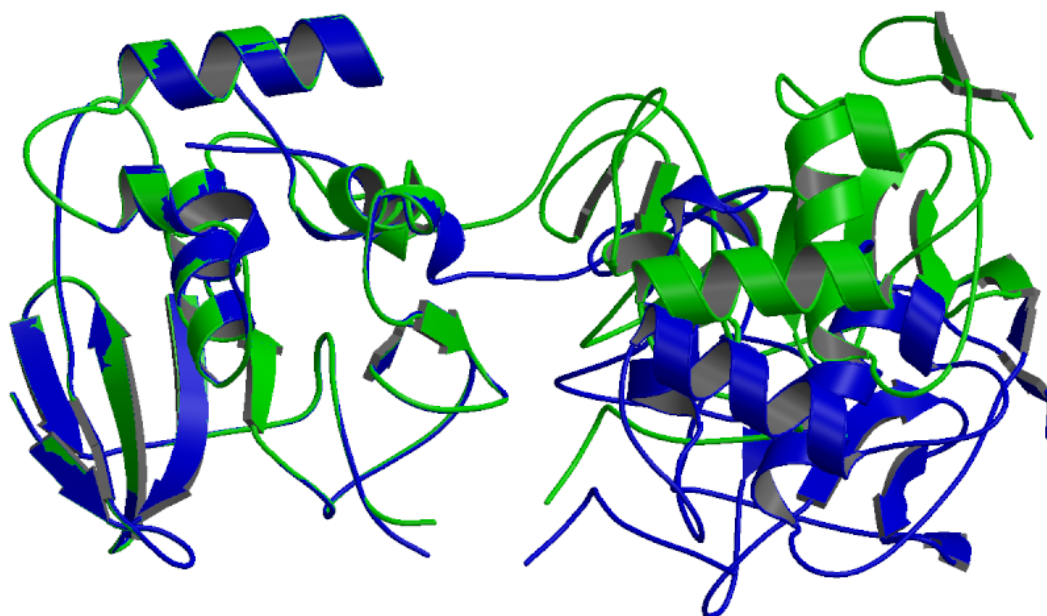
Fig.4: Illustration of the difference between the crystal structure and the NMR structure of the dimeric Lb^{pro} C51A

The crystal structure 1QOL.pdb is coloured blue and the NMR structure 2JQF.pdb is coloured in green. One half of each dimer was overlayed to show the difference in orientation of the monomers to each other. Enforcement of the observed 2-fold symmetry results in a change of the bending angle by about 25-30°.

Fig.5: NMR structure (2JQF.pdb) of the dimeric Lb^{pro} C51A

In one half of the dimeric Lb^{pro} C51A α-helices are colored in green and β-sheets in violet. The active side residues are presented as a ball-and–stick model. The C-terminus of each monomer reaches into the active site of an adjacent molecule, resulting in dimer formation.

The globular region of the leader protease ranges from Met29 to Tyr183. It can be divided into two subdomains, a α-helical and a β-sheet subdomain. The α-helical domain starts at the N-terminus. At the N-terminus there are two short β-sheets, β_1-β_2, followed by four α-helices, α_1-α_4, whereby the longest α-helix, α_1, contains the catalytic Cys51. The β-sheet subdomain contains seven β-strands, six parallel (β_4-β_9) and one antiparallel (β_3 and β_4). In the turn, connecting β_5 and β_6 the catalytic His148 is located. Between the two subdomains there is a deep cleft, where Cys51 and His148 are located at the top, facing each other to allow interaction. The orientation of His148 with respect to Cys51 is stabilized by a hydrogen bond to the side chain oxygen of Asp163. As mentioned before, in papain this task is usually fulfilled by an asparagine residue. Furthermore the fact that this hydrogen bond is exposed to solvent is unusual for papain-like proteases is. Usually this hydrogen bond is protected by two tryptophan residues. During proteolysis, the negative charge development on the carbonyl oxygen is stabilized by the

backbone amide of Cys51 and also (unlike compared to papain) by the side chain amide of Asn46. In papain, a glutamine residue fulfills this task. Despite differences between papain like proteases and Lb^{pro}; the spatial arrangement of the catalytic residues in the active site of Lb^{pro} is well preserved. The differences that have evolved result in different specificity of the protease allowing for three different proteolytic reactions: the self-processing and the cleavage of the eIF4GI and II

Self-processing

The fact that the CTE of Lb^{pro} is found in the active site of an adjacent molecule leads to the alternative assumption that this reaction is intermolecular, but modeling studies showed that the CTE is flexible enough to fold back into the active site of the same protein. This suggests that the reaction can also occur intramolecularly. Actually both reactions are possible, however (Glaser et al., 2001) indicated that self-processing is much more efficient if Lb^{pro} is part of the same chain as the cleavage site, suggesting that this reaction is an intramolecular one in vivo.

Cleavage of eIF4GI and II

The eIF4GI and II are functionally homologues proteins which share an amino acid identity of about 46% (Gradi et al., 1998). The cleavage of both proteins occurs simultaneously and rapidly within virus infected cells (Gradi et al., 2004). eIF4GI is cleaved between Gly674 and Arg675 (Kirchweger et al., 1994) and eIF4GII is cleaved between Gly700 and Ser701 (Gradi et al., 2004). Although the two proteins are homologous, the cleavage site of each protein does not correspond to the homologous site of the other protein. The cleavage of both eIF4G proteins can also take place, if Lb^{pro} is still bound to the polyprotein chain before self-processing (Glaser et al., 2001).

Cleavage sites of Lb[pro]

	↓
eIF4GI	SFANLG RTTLST
eIF4GII	PLLNVG SRRSQP
Lb[pro]/VP4	VQRKLK GAGQSS
	↑
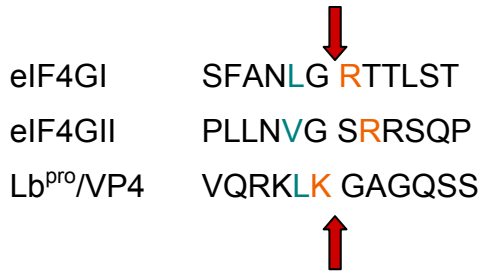
By comparing the sequences surrounding the three cleavage sites it is obvious that Lb[pro] requires (like all other papain like enzymes) a hydrophobic residue at P2 (coloured in blue) for substrate binding. Before the discovery of eIF4GII it was assumed that an additional basic residue providing a positive charge (coloured in orange) is necessary at either the P1 or the P1' position (Glaser et al., 2001). However, it seems possible that the arginine (R) at the P2' site of eIF4GII fulfills this task.

1.1.5.1 CTE –Unique Feature of Lb[pro]

In the prototype protease (papain), no comparable feature like the CTE was found. The development of the CTE prepares Lb[pro] for self-processing. Furthermore it is involved in efficient cleavage of both eIF4G homologues. This is a multiple step process, involving initial recognition and cleavage. For the recognition Lb[pro] interacts via the C-terminal residues 183-195 and the Cys133 with the eIF4G homologues. Cys133 is located in the loop connecting the β-sheets four and five, which is close in space to residues at the beginning of the CTE (Foeger et al., 2002). Out of all CTE residues, that are involved in binding of eIF4G homologues, Asp184 and Glu186 are especially highly conserved in all seven FMDV serotypes, (George et al., 2001; van Rensburg et al., 2002). The binding site on the eIF4GI is already identified and includes the residues 645-657 (Foeger et al., 2005), which is located 17 amino acid residues apart from the cleavage site. For the following cleavage step none of the initial recognition residues, of neither Lb[pro] nor eIF4GI, take part.

Lb^pro forms a dimer by linking the two molecules via their CTEs. Thereby the six C-terminal amino acids are found in the active site of an adjacent molecule. In the shortened form (lacking the six C-terminal amino acids) Lb^pro remains monomeric and the CTE is flexible and disordered.

1.2 NMR Spectroscopy

NMR spectroscopy is a powerful method for studying the structure and also dynamic properties of biomolecules. This is important for the understanding of their functions. NMR measurements are usually performed in solution. The advantage of measuring in solution is that biomolecules experience nearly physiological conditions. Recently NMR experiments in the solid state have become an important method too, although the experiments are more complicated and difficult to interpret.

1.2.1. Principles

1.2.1.1 Nuclear Spin

Subatomic particles, such as electrons neutrons and protons possess a quantum mechanic phenomena, called spin. A nucleus only has a spin if either the number of neutrons or the number of protons or both are odd. Nuclei with spins ≠ 0 have an angular momentum p. These nuclei produce due to their charge a magnetic moment μ.

$$\mu = \gamma * p \qquad (1)$$

For a given spin I there are 2I+1 possible spin states (m=-1,…0…+1). For NMR studies of biomolecules only nuclei with a spin quantum number I=½ are important. Theses are mainly ^{1}H, ^{15}N, ^{13}C and ^{31}P. For those spins there are 2 possible spin states, m=+½ referred as α state and m=-½, referred as β state.

If nuclei with a magnetic moment experience a magnetic field, their magnetic moments orient and the originally degenerate energy levels of the spin states split. Nuclei with spin state m=+½ orient parallel to the magnetic field and

nuclei with m=-½ orient antiparallel. The energy difference of the split spin states linearly depends on the strength of the applied magnetic field B_0.
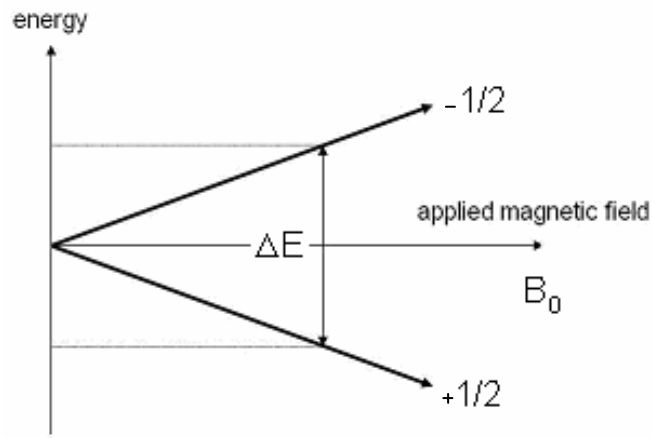
Fig.6: Energy levels for a nucleus with spin quantum number I=1/2

It can be described by the following equation:

$$\Delta E = \gamma h B_0 / 2\pi = h \nu_0 \qquad (2)$$

where B_0 is the applied magnetic field strength and γ is the gyromagnetic ratio.

Due to the angular momentum the spins precess around the magnetic field. The frequency of this precession is given by

$$\nu_0 = \gamma B_0 / 2\pi \qquad [rad/s] \qquad (3)$$

or

$$\omega_0 = -\gamma B_0 \qquad [Hz] \qquad (4)$$

This frequency is called lamor frequency.

The equilibrium populations of the energy levels are described by the Boltzmann equation:

$$\frac{N^{low}}{N^{up}} = e^{-\frac{\Delta E}{kT}} \qquad (5)$$

At equilibrium temperature the lower energy state is more populated but there is only a very small difference. For example for protons in a 18.8 Tesla magnetic field (ν_0 = 800 MHz) at room temperature, the population ratio will be 0.999872. Because every spin produces a magnetic field this population difference gives rise to a net magnetization which is aligned along the direction of the static magnetic field (also described as z-direction). Such a small population difference presents a significant sensitivity problem for NMR because only the net magnetization can be detected.

If electromagnetic energy is applied which equals the difference between the energy levels (ΔE) of the spin states, a transition between these energy levels is induced. After a short time, populations of the spin states return to equilibrium. This process is called relaxation.

1.2.1.2 NMR Spectroscopy

In NMR spectroscopy, high magnetic field strengths are used to enhance the sensitivity of the method. The energy difference between the α and the β spin states lies in the radio frequency (RF) range. If an RF-pulse with resonance frequency corresponding to the Lamor frequency of a distinct nuclei type is applied to spins in equilibrium, the magnetization vector is rotated, depending on the duration of the pulse. Usually 90° and 180° pulses are used. The magnetization can be transferred between nuclei to obtain information about the molecule. Finally the magnetization is rotated back to the xy-plane for

detection. As the magnetization precesses it induces a small current in a coil (this coil is acting as transmitter and also as receiver), which is dedected. The decaying magnetization is measured as a time dependent signal. This is recorded as so called free induction decay (FID). Fourier transformation converts the recorded time domain signal into a frequency domain signal – a NMR spectrum. In NMR-spectroscopy many different types of spectra, which are described below, can be recorded to obtain important structural information.

1.2.1.3 Chemical Shift

In an NMR spectrum every nucleus has its specific resonance frequency according to its chemical environment. The chemical environment of a nucleus is presented in its electron cloud. As electrons are moving charges, a magnetic filed will induce currents in this electron cloud. These currents produce a small local magnetic field which is proportional to B_0. Every nucleus experiences the sum of both, the external magnetic field and its local field. As a result every spin in a molecule has a different resonance frequency

$$\omega_{0,j} = -\gamma\left(1 - \sigma_j\right)B_0 \qquad (6)$$

σ is the shielding constant, which reflects the chemical environment.
To compare spectra from different sources the dependence on the magnetic field has to be removed. Therefore the chemical shift δ is calculated relative to the frequency of a reference compound and it is typically quoted in ppm.

$$\delta = \left(\omega_0 - \omega_{ref}\right)/\omega_{ref} * 10^6 \qquad (7)$$

1.2.1.4 Scalar Coupling

Scalar coupling is an interaction of two nuclei via their chemical bonds mediated by electrons. The magnetic moment of a spin interacts with the electron spin inducing a weak polarization of the electron cloud of the binding electrons. This influences the energy of the second spin involved in the chemical bonding. Depending on the spin state of one nucleus there are two resonance frequencies for the second nucleus resulting in two peaks in the spectra. The peaks are separated by the coupling constant J of the two nuclei. The frequencies of the two peaks are the lamor frequency of the nuclei plus ½ J and the lamor frequency minus ½ J splitting the resonance line into a doublet. If only one peak for a certain nuclei should be observed in a NMR spectrum then the coupling partner has to be decoupled (as described lower).

Via scalar coupling, magnetization from one nucleus can be transferred to a second nucleus through their chemical bond. Therefore correlation of the frequencies of different nuclei can be achieved which is at the basis of multidimensional NMR-spectroscopy.

1.2.1.5 Dipolar Interaction

Two spins can interact directly via their magnetic moments, through space. This interaction is dependent on the distance between the two nuclei, r_{AB}, on their gyromagnetic ratios, $\gamma_{A,B}$ and on the orientation of their internuclear vector with respect to the static magnetic field.

The energy of this interaction is

$$E_{\mu_A\mu_B} = \frac{\mu_0}{4\pi}\frac{1}{r_{AB}^3}\mu_A\mu_B\left[1 - 3\cos^2\theta\right] \qquad (8)$$

$\mu_{A,B}$ is the magnitude of the magnetic dipole moments and is proportional to the respective gyromagnetic ratios. μ_0 is the magnetic permittivity of the vacuum, γ_A and γ_B are the gyromagnetic rations of the spins A and B and r_{AB} is the distance between the spins.

An important application of dipole-dipole interaction for structure determination is the measurements of residual dipolar couplings (RDCs) and the Nuclear Overhauser Effect-effect (NOE).

1.2.1.5.1 NOE (Nuclear Overhauser Effect)

The dipolar interaction mechanism gives rise to magnetization transfer during longitudinal relaxation between spins interacting via their magnetic moments. This is the so called Nuclear Overhauser Effect. The magnitude of the magnetization transfer depends on the distance between the nuclei, making the NOE a powerful tool for distance determination. The magnetization transfer rate follows $1/r^6$ making NOEs observable only for nuclei that are in close spatial proximity. In practice, for two protons this means a distance of max. 6Å. As a proton in a folded protein is surrounded by numerous other protons within this distance many NOE correlations may occur. This provides local distance information which is essential for structure determination. NOEs for protons in a protein are measured in multidimensional NOESY (NOE-SpectroscopY) experiments.

1.2.1.5.2 Dipolar Coupling

1.2.1.5.2.1 Introduction

If two spins interact through space each spin is influenced by the magnetic field produced from the magnetic moment of the other spin. The energy level of one spin is dependant only on the z-component (either α or β) of this field. Due to this fact dipolar coupling like scalar coupling can only be observed in

spectra which are not decoupled.

The Hamiltonian for heteronuclear dipol dipol coupling is given by:

$$H_{dd} = D_{max}^{AB} \left\langle 2I_Z^A I_Z^B \left(3\cos^2\theta - 1\right)/2 \right\rangle \qquad (9)$$

whereby θ is the angle between the internuclear vector and the static magnetic field, the anguler brackets denote time average over motions and

$$D_{max}^{AB} = -\mu_0 \left(h/2\pi\right) S\gamma_A\gamma_B \big/ 4\pi^2 r_{AB}^3 \qquad (10)$$

is the magnitude of the static dipole dipole coupling constant in SI units. h is the Planck's constant and γ_A and γ_B are the gyromagnetic rations of the spins Dipolar couplings are influenced by internal dynamics. They are scaled by the order parameter S which can be derived by relaxation studies. For structured proteins S is nearly one and can be neglected. (Tjandra et al., 1997)

1.2.1.5.2.2 Alignment and RDC (Bax et al., 2001)

As mentioned, dipolar interaction between 2 spins depends on the orientation of their internuclear vector with respect to the static magnetic field.

In isotropic solution, as the molecules tumble quickly around, the dipolar coupling between two spins averages to zero. If the proteins are in a medium where preferential orientation in a magnetic field exists, the dipolar couplings no longer fully average to zero and residual dipolar couplings can be measured. These contain information about how particular internuclear vectors of two distinct nuclei are oriented with respect to the static magnetic field. Therefore residual dipolar couplings provide structure and importantly long range orientational information
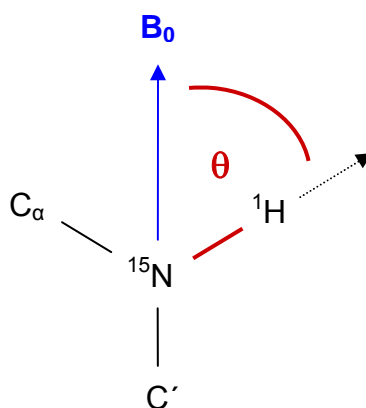
Fig.7: Orientation of the internuclear vector r_{NH} with respect to the static magnetic filed

1.2.1.5.2.3 Order matrix

In anisotropic solution the molecules orientate in the magnetic field due to the magnetic susceptibility of the alignment media. The preferential orientation of the molecule is described by the order matrix or also called Saupe matrix (Saupe and Englert, 1963)

$$S_{ij} = \frac{3}{2}\left\langle \cos\beta_i \cos\beta_j \right\rangle - \frac{1}{2}\delta_{ij} \qquad (11)$$

The order matrix is traceless ($S_{xx} + S_{yy} + S_{zz} = 0$) and symmetric ($S_{ij} = S_{ji}$), and therefore contains only five independent elements. β_ι defines the orientation of the three axes of an arbitrary molecular coordinate system relative to the static magnetic field and δ_{ij} is the Kronecker delta function.
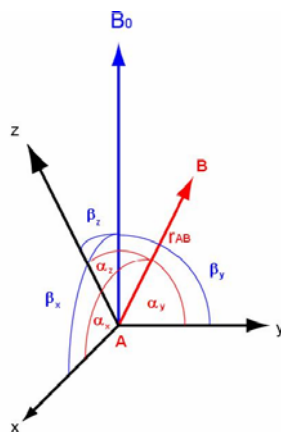
Fig.8: orientation of a internuclear vector r$_{AB}$ with respect to the static magnetic field
The angles α_x α_y and α_z describes the orientation of an internuclear vector with respect to the molecular coordinates x, y, z. β_x, β_y, and β_z are the time-dependent angles between the molecular coordinates and the static magnetic field.

The residual dipolar splitting between two spins can be described in the following way:

$$D^{AB}\left(\alpha_x, \alpha_y, \alpha_z\right) = D_{max}^{AB} \sum_{i,j=\{x,y,z\}} S_{ij} \cos\alpha_i \cos\alpha_j \qquad (12)$$

where α_i are the angles between the internuclear vector and the molecular coordinate frame.

1.2.1.5.2.4 Alignment Tensor

As the order matrix is real and symmetric it can be diagonalized. In this principal axis frame of the alignment tensor equation 12 simplifies to

$$D^{AB}\left(\alpha_x, \alpha_y, \alpha_z\right) = D_{max}^{AB} \sum_{i=\{x,y,z\}} S_{ii} \cos^2\alpha_{ii} \qquad (13)$$

The five independent parameters are an axial and a rhombic component, that

32

correspond to the eigenvalues of the order matrix and three Euler angles that accomplishes the diagonalization of the order matrix. As only the anisotropic part of molecular orientation contributes to residual dipolar couplings the directions of the alignment tensor A_{xx}, A_{yy} and A_{zz} is defined by

$$A_{ii} = \left\langle \cos \beta_{ii}^2 \right\rangle - \frac{1}{3} \qquad (14)$$

and the following definitions are valid:

$$\left| A_{zz} \right| > \left| A_{yy} \right| > \left| A_{xx} \right| \qquad (15)$$

$$A_{yy} + A_{xx} = -A_{zz} \qquad (16)$$

The axial and the rhombic component are defined as the following:

$$A_a = \frac{3}{2} A_{ZZ} \qquad (17)$$

$$A_R = A_{xx} - A_{yy} \qquad (18)$$

The direction of an internuclear vector $\cos \alpha_{ii}$ with respect to the principal axis frame of the alignment tensor an also be expressed in polar coordinates using the angles ϕ and θ.
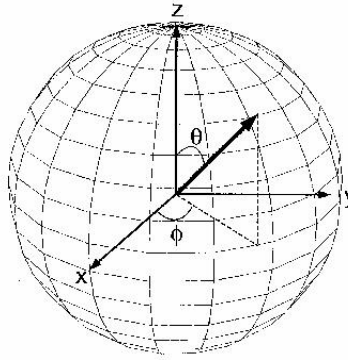
Fig.9: Internuclear vector with respect to principal axis of the alignment tensor. Adopted from (Ramirez and Bax, 1998)

Using this expression dipolar splitting can be described as follows (Clore et al., 1998b):

$$D^{AB}(\theta,\phi) = D^{AB}\left[\left(3\cos^2-1\right) + \frac{3}{2}R\sin^2\theta\cos^2\phi\right] \qquad (19)$$

D_a is the magnitude of the alignment tensor describing the degree of alignment (which correlates with the amount of the alignment media).

$$D_a = \frac{1}{2}D_{max}^{AB}A_a \qquad (20)$$

R is referred as the rhombicity.

$$R = A_r/A_a \qquad (21)$$

1.2.1.5.2.4.1 Determination of the Alignment Tensor Components:

If the structure is known the five independent elements of the order matrix can be calculated if at least five dipolar couplings are available. Usually the number of measured dipolar couplings is much higher. Therefore the system is overdetermined and the elements of the order matrix are then calculated by singular value decomposition.

Another possibility is the prediction of the orientation tensor of the molecule directly from its three-dimensional structure and its charge distribution (Zweckstetter and Bax, 2000; Zweckstetter et al., 2004). Knowledge of the elements of the order matrix allows the calculation of D_a and R. Both parameters are needed for refinement of the existing structure using observed residual dipolar couplings.

For unknown structures the components of the alignment tensor can be obtained from a histogram showing the frequencies of measured RDCs (Clore et al., 1998a). In the case of uniform and isotropic distribution of the internuclear vectors the histogram has the shape of a powder pattern of chemical shift anisotropy observed in solid state NMR (Fig.10). As individual types of internuclear vectors are oriented slightly different the probability of isotropic uniform distribution of the internuclear vectors is higher for an ensemble of different internuclear vector types (Lee et al., 1997). To be comparable the different types of dipolar couplings are normalized relative to a distinct dipolar coupling, usually $^1D_{NH}^N$:

$$D^{NH} = D^{AB}\left(\gamma_N\gamma_H\left\langle r_{NH}^{-3}\right\rangle \bigg/ \gamma_A\gamma_B\left\langle r_{AB}^{-3}\right\rangle\right) \qquad (22)$$

The extreme values of this histogram (the highest, the lowest and the most frequent dipolar coupling) correspond to the x-, y- and z-component of the alignment tensor. The largest absolute value describes the z-component, the

opposite extreme value describes the y-component and the most frequent value describes the x-component.
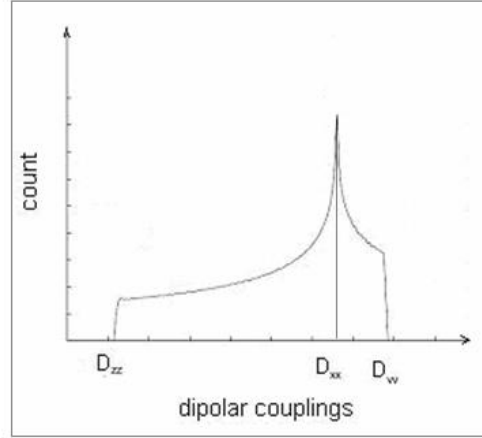
Fig.10: Theoretical distribution of residual dipolar couplings
The components of the alignment tensor can be obtained out of the powder pattern of the histogram showing the frequencies of RDCs

As it can be seen in Fig.9 the z-component of the alignment tensor is derived if $\theta=0°$, the y-component is obtained if both $\theta=90°$ and $\phi=90°$ and the x-component is derived if $\theta=90°$ and $\phi=0°$. By insertion of the respective angles in equation 19 it follows that

$$D_{zz}^{NH} = 2D_a^{NH} \qquad (23)$$

$$D_{yy}^{NH} = -D_a^{NH}\left(1+1.5R\right) \qquad (24)$$

$$D_{xx}^{NH} = -D_a^{NH}\left(1-1.5R\right) \qquad (25)$$

and so the rombicity R, and the magnitude D_a^{AB} can be determined. Knowing these parameters restricts the orientation of the internuclear vector to lie on a double cone. Measurements of RDCs in different alignment media restricts this orientation to at least two directions (one and its inverse) which cannot

be distinguished.

1.2.1.6 Relaxation

For measuring NMR-spectra the magnetization is perturbed. Relaxation is the process by which the magnetization returns to equilibrium. For biomolecular NMR spectroscopy relaxation is the most important limiting factor. Long relaxation times are required for triple resonance experiments. Too fast relaxation leads to too large line widths of the signals in the final NMR spectra and as a consequence to small signal to noise ratios.

1.2.1.6.1 Relaxation Processes

Relaxation is described by two processes, the longitudinal relaxation, also called spin lattice relaxation and the transversal relaxation, also called the spin spin relaxation

1.2.1.6.1.1 Longitudinal Relaxation

After an RF pulse the equilibrium magnetization of particular spins is disturbed. Given time, the original equilibrium magnetization will be reached again. In equilibrium, only z-magnetization is present. The process that describes the return of the bulk magnetization to z-magnetization is called longitudinal relaxation. The longitudinal relaxation rate is $1/T_1$, whereby T_1 describes the average lifetime of a spin staying in the excited spin state before it returns to the lower energy spin state. The energy set free by this process is transferred to the surroundings (called lattice). Therefore longitudinal is an enthalpic process. During this relaxation process a second form of relaxation also occurs, the transverse relaxation.

1.2.1.6.1.2 Transverse Relaxation

Transverse magnetization describes the net magnetization in the xy-plane which is zero in the equilibrium.. Transverse relaxation is an entropic process leading to dephasing, or loss of the phase coherence, of the magnetization in the xy-plane until there is no net magnetization left. This process is described with the time constant T_2 whereby $1/T_2$ is the transverse relaxation rate.

1.2.1.6.2 Origin of Relaxation

The source of both relaxation processes are fluctuating local magnetic fields which influence individual magnetic moments resulting in induced transitions and finally relaxation. In small molecules these fluctuating fields average faster than in larger molecules because small molecules tumble very fast. Therefore their relaxation properties are very different from large molecules e.g. proteins. The major sources of local fields are the dipolar mechanism and the chemical shift anisotropy mechanism.

1.2.1.6.2.1 Dipolar Relaxation Mechanism

As described, two spins can interact via their magnetic dipole moments directly through space (equ. 8). The magnetic field of one spin interacts with the magnetic field of a second spin changing its orientation. The rotational tumbling of a molecule causes field fluctuation leading to relaxation. A further important point is that the magnetic field experienced by one nucleus depends on the spin state of the second nucleus. If the spin state of the second nucleus changes then the first spin experience a change in magnetic field, also leading to relaxation.

1.2.1.6.2.2 Chemical Shift Anisotropy (CSA)

As described before, the static magnetic field induces a small local field in the electron cloud of a nucleus. The size of this local field depends on the

orientation of the molecule. As the molecule tumbles in solution each nucleus experiences an average local field. So the nuclei of one molecule have average chemical shifts giving rise to their resonance frequencies. As every molecule tumbles in a different way there is a fluctuating spread of resonance frequencies which is a source of relaxation. The magnitude of the local field change with respect to orientation depends on the type of nucleus. Therefore the impact of this relaxation mechanism is different for different types of nuclei.

1.2.1.6.3 Cross Correlated Relaxation

Cross correlation is the interference of dipolar relaxation and relaxation due to chemical shift anisotropy if one spin experiences both relaxation mechanisms. Both relaxation mechanisms are dependent on the orientation of the molecule. Therefore they have the same time dependence. As both relaxation mechanisms are correlated they can influence each other. This fact is called cross correlation. If one spin is considered it experience the local field from its CSA and a local field from a neighboring spin. The fluctuations of these fields due to the tumbling of a protein are correlated. As described the local dipolar field experienced by a spin depends on the spin state of a second spin. According to this spin state the local fields can reinforce or cancel each other, leading to a higher relaxation rate in the first case and to a lower in the second case, resulting in different relaxation rates within a multiplet. The effect that cross correlation can result in lower relaxation rates when exploited property is used in TROSY (Transverse Relaxation Optimized SpectroscopY) experiments where only multiplet components with slow effective relaxation rates are selectively detected.

1.2.2 Multidimensional NMR Experiments:

1.2.2.1 Introduction

For bimolecular NMR, multidimensional experiments are used in order to avoid overcrowded spectra.

A general scheme for 1D-, 2D- and 3D-NMR experiments is presented in Fig.11
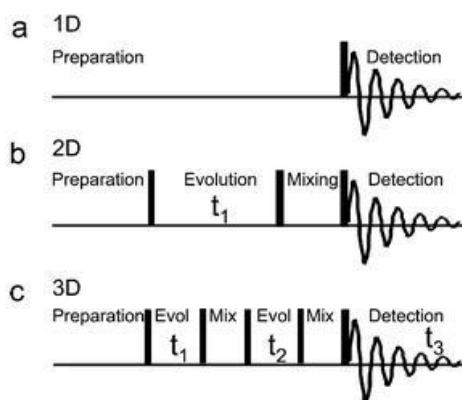
Fig.11 Representative pulse schemes for 1D 2D and 3D spectra [2]

All NMR-experiments start with a preparation period, where the equilibrium of the nuclear spin system is disturbed. In 1D-experiments this non-equilibrium is detected afterwards. In the preparation time of multidimensional experiments magnetization is transferred between nuclei (either by scalar coupling or via dipolar interaction) who then evolve during the evolution period t_1. Next is the mixing period during which the magnetization is transferred to other nuclei. Depending on the dimension of the spectra there can be more evolution and mixing periods. The last mixing time transfers the magnetization to the nuclei, whose FIDs are recorded during the acquisition period. These are usually the most sensitive nuclei (H).

---

[2] http://pubs.rsc.org/ej/AN/2004/b403435j/

1.2.2.2 Applications for multidimensional NMR Experiments

1.2.2.2 1 INEPT - Insensitive Nuclei Enhanced by Polarization Transfer (Morris and Freeman, 1979)

For the magnetization transfer from one type of nuclei to another type of nuclei usually an INEPT pulse sequence is used. The magnetization is transferred via scalar coupling while the evolution of the chemical shift of both nuclei is refocused by a pair of 180° pulses in the middle of the transfer period τ. The time τ that is needed for complete magnetization transfer is equal to $1/2J_{AB}$.
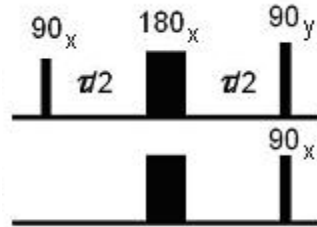
Fig.12: INEPT (Insensitive Nuclei Enhanced by Polarization Transfer)
The INEPT sequence is used for magnetization transfer in multidimensional NMR experiments

In multidimensional experiments sensitivity enhancement can be achieved if the magnetization is transferred from sensitive nuclei with a high gyro magnetic ratio to less sensitive nuclei during the preparation time.

1.2.2.2.2 Decoupling

As described, coupling (scalar and dipolar) between two spins changes the energy of a spin dependent on the spin state of the coupling partner resulting in peak splitting. This leads to overcrowded spectra with reduced intensity which presents an obstacle for peak assignment. In order to obtain only one

peak per recorded spin the coupling partners have to be decoupled during the evolution and acquisition periods. This decoupling can be achieved with a 180° pulse in the middle of the evolution time or with a series of alternating 180° pulses leading to rotation of the respective magnetization during acquisition of a FID. Now the recorded spins experience the overall averaged magnetization of their coupling partners, resulting in only one peak. This peak has the average frequency and the sum of the intensities of the split peaks.

1.2.2.2.3 TROSY Experiments (Pervushin et al., 1997)

TROSY experiments take advantage of cross correlation to reduce relaxation rates by selectively detecting the slowly relaxing multipltt component. For biomolecules that are large in size this is a big advantage because relaxation here is an extensive limiting factor.

As described, fluctuation of local fields due to CSA and dipolar relaxation mechanisms can reinforce or cancel one another. The magnitude of the reinforcement or cancellation depends on the field generated by CSA and the field of the dipolar interaction partner. When they are parallel then the cross correlation is at its maximum.

This is the case of the ^{15}N-^1H bond in proteins whatever the orientation of the molecule is. If a ^{15}N doublet (where ^{15}N is coupled to the ^1H) is considered the different relaxation rates become apparent in the different widths of the two peaks (see Fig.13).
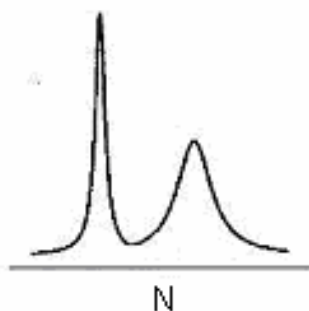
Fig.13: ^{15}N doublet (^{15}N coupled to ^{1}H)
The different peak widths represent the different relaxation rates due to cross correlation.

One peak corresponds to the ^{1}H spin being in the α state and one peak corresponds to spin ^{1}H being in the β state. Because the relaxation rate of ^{15}N is dependant on the spin state of ^{1}H, one peak is sharp and high and one peak is broadened. The same effect also occurs for the proton.

If a ^{15}N-^{1}H spectrum, where the frequencies of ^{1}H and ^{15}N are correlated, is recorded, allowing coupling of ^{15}N and ^{1}H with each other, 4 peaks occur per ^{15}N-^{1}H-group (two in each dimension).
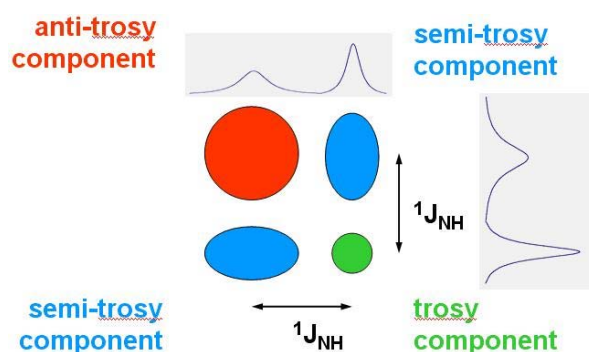
Fig.14: Signals of a ^{15}N-^{1}H-group in a ^{15}N-^{1}H spectrum with ^{15}N-^{1}H-coupling
Due to cross correlation one component of a doublet has a smaller width in each dimension. The trosy component is the peak that has the smallest width in each dimension, which makes the peak the most intense of the multiplet.

One peak of a doublet in each dimension has a smaller width, resulting in one peak having a small width in each dimension due to the decreased effective relaxation rates. In TROSY experiments only this peak will be recorded, resulting in spectra with much higher signal to noise ratios.

1.2.2.3 Assignment

Assignment is the process where the resonances of each signal in the NMR spectra are associated with the corresponding nucleus of the molecule which is investigated. The resonance frequency of a nucleus is the same in each spectrum. By comparing different spectra where the resonance frequencies of nuclei are correlated to each other in various complementary ways assignment can be achieved and different types of nuclei are correlated to each other.

1.2.2.3.1 Backbone Assignment

Different types of triple resonance experiments have to be performed, recording the signals of the backbone H^N, N and CO and $C^{\alpha/\beta}$. Via scalar coupling these nuclei are correlated. For backbone assignment, spectra which have signals from both, inter- and intra-residual atoms and spectra which detect only inter-residual atoms are recorded. By comparison of these pairs of spectra spin systems can be linked. Using residue specific chemical shifts they can be mapped onto the primary sequence thereby assignment can be achieved.

1.2.2.3.2 Side Chain Assignment

Intraresidual signals of aliphatic amino acid side chains can be correlated via scalar coupling using COSY/TOCSY methods and combined with intra and interresidual NOE signals of the respective groups.

1.2.2.4 Measurement of Residual Dipolar Couplings.

As already described, residual dipolar couplings can only be measured in anisotropic solution because in isotropic solution they average to zero. The residual dipolar coupling between two nuclei leads to additional peak splitting in the recorded spectrum. The observed splitting results from scalar coupling (J) plus dipolar coupling (D). In order not to impair the magnetization transfer via the INEPT only a small degree of alignment is desired. This also makes sure that only nearby nuclei give rise to dipolar couplings. In isotropic solution the peak splitting results only from the scalar coupling between the two spins. Therefore each experiment has to be performed once in isotropic solution and once after addition of alignment media. In order to obtain distinct coupling constants the experiments have to be recorded without decoupling of the respective atoms.

2 Materials and Methods

2.1 Protein Expression and Purification

For protein expression the pET System of Novagen was used. A pET11d vector carrying the protein sequence under control of a T7 promotor was introduced in E.coli BL21(DE3)pLysS cells. These cells have a coding sequence for the T7 RNA polymerase under control of a lac operator. The overexpression of the protein was initiated by addition of IPTG, as expression of the T7 polymerase was induced. The T7 promoter is a viral promoter, leading to a rapid expression of the target protein.

As the protein should be analyzed by NMR triple resonance measurements it has to be labelled with ^{15}N and ^{13}C. To achieve this, the bacteria which were used to express the target protein were grown in M9 minimal medium with ^{15}NH$_4$Cl as nitrogen source and ^{13}C$_6$-D-glucose as carbon source.

2.1.1 Expression of Lb^{pro} and sLb^{pro}

2.1.1.1 Solutions and Buffers

LB medium:

10g tryptone

5g yeast extract

10g NaCl

1L H$_2$O,

$\Rightarrow$ autoclaved at 20°C for 20min

LB-ampicillin-agar-plates:

1L LB medium

1ml 100µg/ml ampicillin

1ml 100µg/ml canamycin

15g agar-agar

M9-medium

6g $Na_2HPO_4 \cdot 2H_2O$

3g KH_2PO_4

0.5g NaCl

1g $^{15}NH_4Cl$ (sole

nitrogen source)

$\Rightarrow$ filled up to 1l with millipore H_2O and autoclaved for 20 minutes at 120°C.

Afterwards the following chemicals were added:

2ml 1M $MgSO_4$

0.3ml $CaCl_2$

10ml trace elements	trace elements:
20ml 20% $^{12}C_6$-D- glucose solution	5g EDTA
or 4g $^{13}C_6$-D_glucose	0.83g $FeCl_3 \cdot 6H_2O$
1ml 100µg/ml of	84mg $ZnCl_2$
appropriate	13mg $CuCl_2 \cdot 2H_2O$
antibiotics (ampicillin, canamycin)	10mg $CoCl_2 \cdot 6H_2O$
1ml biotin (1mg/ml)	10mg H_3BO_3
1ml thiamine (1mg/ml)	1.6mg $MnCl_2 \cdot 6H_2O$
	$\Rightarrow$ filled up to 1l with
	millipore H_2O

IPTG-solution

1M isopropyl-β-D- thiogalactoside

2.1.2.2 Transformation

100µl of competent E. coli BL21(DE3)pLysS cells were thawed on ice for 10 minutes. Then 1µl of the pET11d plasmid containing the target protein was added and mixed. After 30minutes incubation the bacteria were heat shocked for 90 seconds at 42°C. Immediately 300µl Lb medium was added and the cells were incubated in the air shaker at 37°C at 225rpm for 45

minutes. 100µl of the cells suspension was plated on LB agar-plates containing ampicillin and canamycin which were incubated over night at 37°C.

2.1.2.3 Starter Culture

One well separated colony was selected and suspended in 20ml LB-medium containing 20µl ampicillin and 20µl canamycin (100µg/ml). The bacteria solution was incubated in the air shaker at 37°C until a OD_{600} of approximately 1.2 was reached.

2.1.2.4 Expression Culture

10ml from the starter culture was added to 1l M9 media and incubated in the air shaker at 37°C. After reaching an OD_{600} of between 0.4 and 0.5 IPTG was added to induce protein expression. The cells were grown overnight and harvested on the next day by centrifugation for 15 minutes at 4000rpm at 4°C. The bacterial cell pellet was resuspended in 30ml lysis buffer A. For storage the bacterial suspension was frozen in liquid nitrogen and stored at -20°C.

2.1.2.5 Expression Control

1ml of the 1l bacterial culture (before harvesting the cells) was taken and centrifuged for 5 minutes at 4000rpm. The cell pellet was resuspended in 200µl lysis buffer A. From this suspension 10µl was separated. The remaining suspension was centrifuged at 18000rpm for 5 minutes at 4°C. Then 10µl from the supernatant were separated. Both samples were analysed by SDS-PAGE

2.1.3 Purification of Lb^{pro} and sLb^{pro}

2.1.3.1 Solutions and Buffers

Lysis buffer A
50mM TRIS-HCl, pH 8.0
50mM NaCl
1mM EDTA
5% glycerol
5mM DTT

Lysis buffer B
50mM TRIS-HCl, pH 8.0
1M NaCl
1mM EDTA
5% glycerol
5mM DTT

NMR buffer
50mM NaCl
5mM DTT
20mM NaH$_2$PO$_4$/Na$_2$HPO$_4$ pH 7.0

2.1.3.2 Cell Lysis

The harvested cells were cooled on ice and lysed by ultrasound using a Bandelin Sonopuls sonicator, (5 cyles at 60% power) for at least 15 minutes, until there were no lumps in the suspension. This suspension was centrifuged at 18000rpm for 20 minutes at 4°C. The supernatant was used for the following purification of Lb^{pro}, whereby the pellet was discarded.

2.1.3.3 Protein Precipitation

The first purification step was the precipitation of unwanted proteins in the supernatant with 30% ammoniumsulfate. After 40 minutes cooling on ice the protein suspension was centrifuged at 18000rpm for 20 minutes at 4°C. The second precipitation with 60% ammoniumsulfate was performed for 40 minutes on ice. This time Lb^{pro} was precipitated. Again the protein suspension was centrifuged at 18000rpm for 20 min at 4°C. The pellet was

resuspended in 5-8ml lysis buffer A. Insoluble proteins were pelleted by centrifugation at 18000rpm for 20 minutes at 4°C in order to discard them.

2.1.3.4 Protein Purification

The Lb^{pro} and sLb^{pro} was purified from the protein solution by fast protein liquid chromatography (FPLC). An ÄKTA-explorer system with UV-detector was used. The salt of the sample was removed using an size exclusion column (HiPrep 26/10 desalting) separating low molecular weight substances form high molecular weight substances. The column was equilibrated with lysis buffer A. The sample was loaded on the column and eluted with lysis buffer A. From the protein containing fractions, Lb^{pro} was purified by anion exchange using the resource Q column. Equilibration was done with lysis buffer A and B according to the instructions of the manufacturers. Fractions containing proteins were pooled and loaded on the column. The proteins were eluted with an gradient starting from 100% lysis buffer A going to 100% lysis buffer B. The obtained fractions were analyzed with SDS-PAGE to find out which of them contained Lb^{pro}. These fractions were pooled. It the volume was over 10ml and it was concentrated with Amicon Ultra–15 Centrifugal Filter Unit 10kDa from Millipore

The next purification step was achieved by size exclusion using the superdex 26/60 gel filtration column. Equilibration and eluation were both performed with the NMR buffer. The fractions containing protein were again analysed by SDS-PAGE in order to find out if Lb^{pro} was obtained in its pure form, or if it was contaminated with other proteins. These fractions were pooled, and again concentrated with the 10kDa centrifugal filter to a final protein concentration of at least 1mM, required to use the protein for NMR triple resonance measurements.

2.1.4 SDS-Polyacrylamid Gel Electrophorese (PAGE)

SDS-PAGE (containing 16% polyacrylamide) was used to detect Lb^{pro} and sLb^{pro} in the collected fractions from the FPLC and in the expression control. The samples used for SDS-PAGE were mixed with sample buffer. The mercaptoethanol reduces disulfide bonds if present and SDS unfolds the proteins and charges them negatively. Therefore the proteins migrate to the positive cathode. Due to size, smaller proteins migrates faster in the polyacrylamide gel, larger proteins are more retained. This causes separation of the proteins according to sizes. The proteins were stained with Commassie blue. In order to determine the size of the proteins, a SDS-7 protein marker for comparison was used.

Stacking Gel	Separation Gel
5% polyacrylamide	16% polyacrylamide
125mM Tris-HCl, pH 6.6	0.15M Tris-HCl, pH 9.0
0.05% SDS	0.01% SDS
0.125% TEMED	0.05% TEMED
0.05% APS	0.07% APS

2.1.5 Determination of the Protein Concentration

The concentration of the protein in the solution was analyzed by absorbance measurements at 280nm using an Eppendorf Biophotometer. At 280nm the π-electrons of the double bonds are excited, therefore the absorption mainly depends on the content of aromatic amino acids tyrosine and tryptophan and phenylalanine. Knowing the extinction coefficient of the protein (Lb^{pro}, sLb^{pro}: $\varepsilon_{280}=43555M/1cm^{-1}$) and the diameter of the cuvette the concentration can easily be calculated from the measured absorption using the Lambert-Beer law.

2.1.6 NMR Sample Preparation

The NMR sample contains about 400μl of concentrated protein solution. If a Shigemi tube is used 250μl are sufficient for measurement. The buffer that is used should not give signals during NMR measurements. So phosphate buffers are preferred for the last purification step of the protein. The protein solution is mixed with about 5% D_2O to provide the frequency lock signal. The verification of the stability of this signal is necessary to ensure stabilization of the NMR instruments during measurement. Additionally, a very low amount (about 2.5%) of NaN_3 is added to prevent microbial growth in the sample.

For measurement of dipolar couplings the proteins have to be aligned. As alignment media the filamentous Phage *Pf1* of the strain LP11-92 solved in potassiumphosphate buffer was used.

2.2. NMR Experiments and Structure Determination

2.2.1 Experiments for Backbone Assignment

2.2.1.1 ^{15}N-HSQC (Heteronuclear Single Quantum Correlation)

The ^{15}N-HSQC correlates the frequencies of the amide nitrogen atom with its attached amide proton. Therefore every backbone peptide group (except proline) of a protein gives rise to a signal. Additionally also the side chain aminde groups of asparagine, glutamine and arginine side chains as well as tryptophan indole ^{15}N-^{1}H are observable. The ^{15}N-HSQC serves as a starting point in spin system classification in the process of sequential amino acid assignment.

The magnetization is transferred in the following way:

$$^{1}H^{N}_{i} \xrightarrow{J_{NH^{N}}} {}^{15}N_{i}\,(t_1) \xrightarrow{J_{NH^{N}}} {}^{1}H^{N}_{i}(t_2)$$
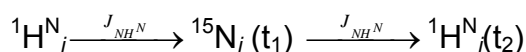
2.2.1.2 Triple Resonance Experiments

2.2.1.2.1 HN(CO)CA (Bax and Ikura, 1991)

The HNCOCA correlates the frequencies of an amide proton and an amide nitrogen with the ^{13}C$^{\alpha}$ of the preceding amino acid.

The magnetization is transferred in the following way:

$$^{1}H^{N}_{i} \xrightarrow{J_{NH^{N}}} {}^{15}N_{i} \xrightarrow{J_{NC'}} {}^{13}C'_{i-1} \xrightarrow{J_{C'C^{\alpha}}} {}^{13}C^{\alpha}_{i-1}(t_1) \xrightarrow{J_{C'C^{\alpha}}} {}^{13}C'_{i-1} \xrightarrow{J_{NC^{\alpha}}}$$

$$^{15}N_{i}(t_2) \xrightarrow{J_{NH^{N}}} {}^{1}H^{N}_{i}(t_3)$$
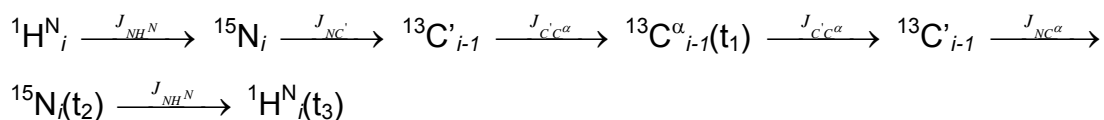
2.2.1.2.2 HNCA (Kay et al., 1990)

The HNCA correlates the frequencies of a amide proton and the nitrogen

with and the $^{13}C^\alpha$ of the same amino acid and the $^{13}C^\alpha$ of the preceding amino acid because the J coupling constant of N to $^{13}C^\alpha$ is similar for intra ($^1J_{NiC}{}^\alpha{}_i$) and inter ($^2J_{NiC}{}^\alpha{}_{i-1}$) residues.

The magnetization is transferred in the following way:

$$^1H_i \xrightarrow{J_{NH^N}} {}^{15}N_i \xrightarrow{J_{NC^\alpha}} {}^{13}C^\alpha{}_{i,i-1}(t_1) \xrightarrow{J_{NC^\alpha}} {}^{15}N_i(t_2) \xrightarrow{J_{NH^N}} {}^1H_i(t_3)$$
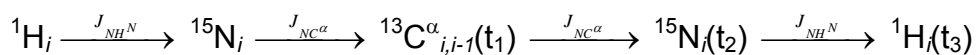
Together with the HNCOCA the HNCA provides a method for sequential assignment, as it can be used to sequentially link residues. But usually $^{13}C^\alpha$ shift degeneracy precludes complete sequential assignment.

2.2.1.2.3 CBCA(CO)NH (Grzesiek and Bax, 1992)

The CBCA(CO)NH correlates the frequencies of a $^{13}C^\alpha$ and a $^{13}C^\beta$ of an amino acid with the amide proton and nitrogen of the next amino acid.

The magnetization is transferred in the following way:

$$^1H^\alpha{}_i/{}^1H^\beta{}_i \xrightarrow{J_{C^{\alpha/\beta}H}} {}^{13}C^\alpha{}_i/C^\beta{}_i(t1) \xrightarrow{J_{C^\alpha C^\beta}} {}^{13}C^\alpha{}_i \xrightarrow{J_{C'C^\alpha}} {}^{13}C'{}_{i+1} \xrightarrow{J_{NC^\alpha}}$$

$$^{15}N_{i+1}(t_2) \xrightarrow{J_{NH^N}} {}^1H^N{}_{i+1}(t_3)$$
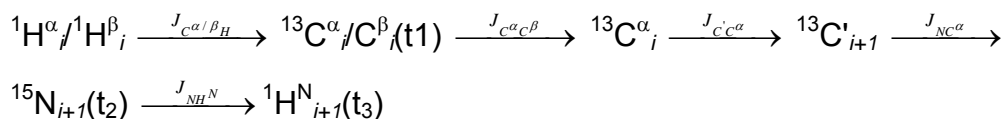
2.2.1.2.4 HNCACB (Wittekind and Mueller, 1993)

The HNCACB correlates the amide proton and nitrogen with the C^α and C^β of the same amino acid and the $^{13}C^\alpha$ and $^{13}C^\beta$ of the preceding amino acid

The magnetization is transferred in the following way:

$$^1H^N{}_i \xrightarrow{J_{NH^N}} {}^{15}N_i \xrightarrow{J_{NC^\alpha}} {}^{13}C^\alpha{}_{i,i-1} \xrightarrow{J_{C^\alpha C^\beta}} {}^{13}C^\alpha{}_{i,i-1}/{}^{13}C^\beta{}_{i,i-1}(t_1) \xrightarrow{J_{C^\alpha C^\beta}}$$

$$^{13}C^\alpha{}_{i,i-1} \xrightarrow{J_{NC^\alpha}} {}^{15}N_i(t_2) \xrightarrow{J_{NH^N}} {}^1H_i(t_3)$$
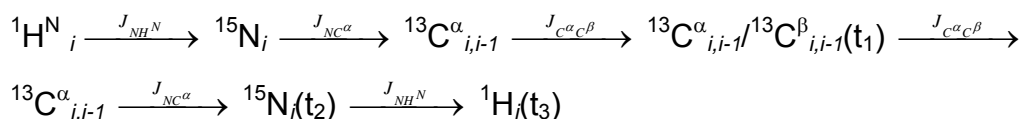
Used together with the CBCA(CO)NH the HNCACB provides a method for

sequential assignment due to larger $^{13}C^\beta$ shift ranges. In combination with $^{13}C^\alpha$ shifts almost complete assignment can be obtained. The disadvantage of both experiments is that there is a long time during which the magnetization is transverse on a ^{13}C atom. During this time period much magnetization is lost by transverse relaxation. This can lead to sensitivity problems, especially for large molecules.

2.2.1.2.5 HNCO (Kay et al., 1990)

The HNCO correlates the frequencies of the amide proton and the nitrogen of an amino acid with the $^{13}C'$ of the preceding amino acid.

The magnetization is transferred in the following way:

$$^1H_i \xrightarrow{J_{NH^N}} {}^{15}N_i \xrightarrow{J_{NC\alpha}} {}^{13}C'_{i\text{-}1}(t_1) \xrightarrow{J_{NC\alpha}} {}^{15}N_i(t_2) \xrightarrow{J_{NH^N}} {}^1H_i(t_3)$$
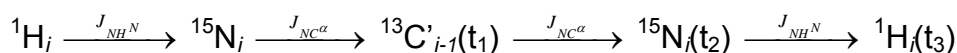
2.2.1.2.6 HN(CA)CO (Clubb et al., 1992)

The HN(CA)CO correlates the frequencies of the amide proton and the nitrogen of a amino acid with the $^{13}C'$ of the same and the preceding amino acid

The magnetization is transferred in the following way:

$$^1H^N_i \xrightarrow{J_{NH^N}} {}^{15}N_i \xrightarrow{J_{NC\alpha}} {}^{13}C^\alpha_{i,i\text{-}1} \xrightarrow{J_{C'C\alpha}} {}^{13}C'_{i,i\text{-}1}(t_1) \xrightarrow{J_{NC\alpha}} {}^{15}N_{i,i\text{-}1}(t_2)$$
$$\xrightarrow{J_{NH^N}} {}^1H^N_{i,i\text{-}1}(t_3)$$
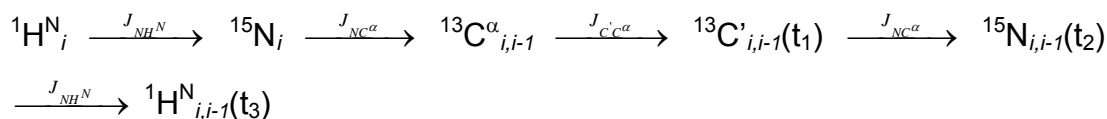
With the HN(CA)CO residues can be linked by their $^{13}C'$ shifts which can be useful for unfolded proteins. Together with the HNCO experiment sequential assigning can be achieved.

HNCA

H^N_i-N_i-C^α_i

H_i-N_i-C^α_{i-1}

HN(CO)CA

H^N_i-N_i-C^α_{i-1}

HN(CA)CO

H^N_i-N_i-C'_i

H^N_i-N_i-C'_{i-1}

HNCO

H^N_i-N_i-C'_{i-1}

HNCACB

H^N_i-N_i-C^α_i-C^β_i

H^N_i-N_i-C^α_{i-1}-C^β_{i-1}

CBCA(CO)NH

C^β_{i-1}- C^β_{i-1}-N_i-H^N_i
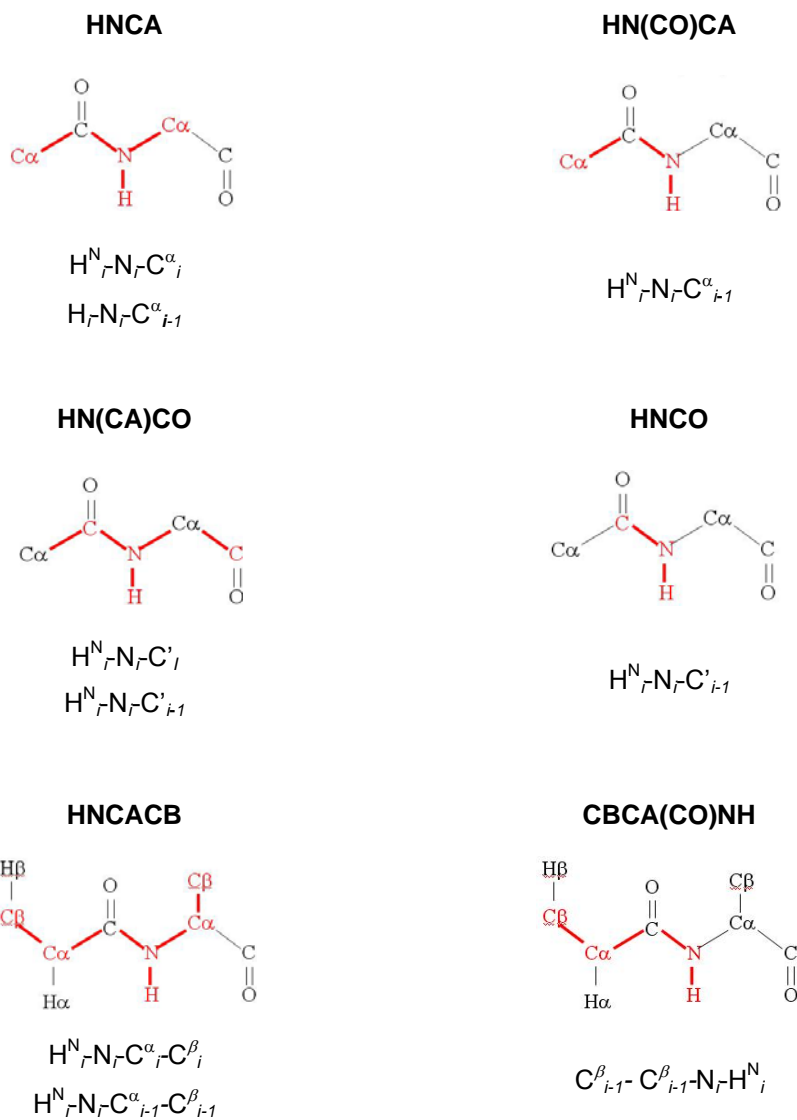
Table 2: Triple resonance experiments for backbone assignment
On the left side of the table experiments in which record intra and inter residual signals are listed. On the right side are the corresponding experiments which only record intra-residual signals.

2.2.1.8 NOESY-^{15}N-HSQC

A NOESY-^{15}N-HSQC is a NOESY followed by a ^{15}N-HSQC experiment. In the NOESY experiment magnetization is transferred between protons that are close in space. The magnetization that is transferred to a proton, attached to a nitrogen, is recorded by the following ^{1}H-^{15}N-HSQC. Therefore protons that are near backbone amide protons can be detected

The correlation pathway is:

$$^1H_A(t_1) \xrightarrow{NOE} {}^1H_B \xrightarrow{J_{NH^N}} {}^{15}N_B(t_2) \xrightarrow{J_{NH^N}} {}^1H_B(t_3)$$
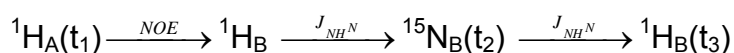
This experiment provides distance restraints, and is also a good complementary experiment for backbone assignment because the amide protons of two neighboring amino acids also give rise to an NOE signal.

2.2.2. Experiments for Side Chain Assignment

2.2.2.1 ^{13}C-HMQC (Heteronuclear Multiple Quantum Correlation) or ^{13}C-HSQC (Heteronuclear Single Quantum Correlation)

The ^{13}C-HMQC or ^{13}C-HSQC correlates the frequencies of a carbon atom with its attached proton. Therefore every CH, CH_2 or CH_3 group of a side chain of a protein gives rise to a signal. For this experiment either multiple quantum or single quantum correlation is chosen for the ^{13}C evolution time depending which one has the slower effective relaxation rate.

The magnetization is transferred in the following way:

$$^1H_i \xrightarrow{J_{HC}} {}^{13}C_i(t_1) \xrightarrow{J_{HC}} {}^1H_j(t_2)$$
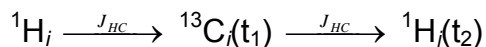
2.2.2.2 HCCH-TOCSY (TOtal Correlation SpectroscopY)

(Bax et al., 1990)

A TOCSY experiment uses a spin lock element or isotropic mixing sequence to transfer the magnetization through scalar couplings within a whole spin system (e.g. one amino acid residue). This mixing sequence consists a series of 180° pulses, separated by very small delays, keeping the magnetization aligned in the transverse plane. This prevents evolution of the chemical shifts with all couplings interactions being active, leading to magnetization transfer. The spin lock is applied to ^{13}C rather than for ^{1}H for reason of better transfer efficiency ($^{3}J_{HH}$~7-8Hz, $^{1}J_{CC}$~35-40Hz). The HCCH-TOCSY experiment correlates all aliphatic ^{1}H atoms within a side chain with each other. The signals are spread according to their ^{13}C frequencies.
The correlation pathway is:

$$^{1}H_i(t_1) \xrightarrow{J_{HC}} {}^{13}C_i(t_2) \xrightarrow{J_{CC(spinlock)}} {}^{13}C_j \xrightarrow{J_{HC}} {}^{1}H_j(t_3)$$
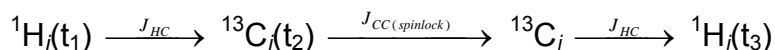
where $^{13}C_i$ and $^{13}C_j$ are part of the same side chain spin network. The J_{HC} coupling is a one bond coupling whereby due to the spin lock sequence the J_{CC} coupling evolve to every ^{13}C atom in the side chain.

2.2.2.3 NOESY-^{13}C-HMQC or ^{13}C-HSQC

A NOESY-^{13}C-HMQC or NOESY-^{13}C-HSQC is a NOESY followed by a ^{13}C-HMQC or ^{13}C-HSQC. In the NOESY experiment magnetization is transferred between protons that are close together. The magnetization that is transferred from protons to protons, attached to an aliphatic side chain carbon atom, is detected by the following ^{13}C-HMQC or HSQC.
The correlation pathway is:

$$1H_i(t_1) \xrightarrow{NOE} {}^{1}H_j \xrightarrow{J_{CH}} {}^{15}C_j(t_2) \xrightarrow{J_{CH}} {}^{1}H_j(t_3)$$
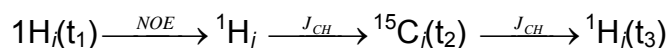
In combination with a HCCH-TOCSY it can be examined which NOE signals arise from intraresidual protons and which arise from interresidual protons. Additionally, the NOE signals provide the important distance restrains for structure calculation.

2.2.3. Experiments for Measuring Scalar (J)

and Dipolar Couplings (D)

Coupling interactions occur in NMR-experiments which are recorded without decoupling of the respective atoms during the evolution and, or the acquisition time of the respective coupling partners. In the resulting spectra the peaks are split according to the respective coupling. Under isotropic conditions only the J-couplings contribute to the peak splitting whereby under anisotropic conditions peak splitting accords to (J+D)-couplings. Each spectrum is measured once in isotropic and once in anisotropic solution. By taking the difference between anisotropic (J+D) and isotropic couplings (J) the dipolar couplings (D) are obtained.

2.2.3.1 ^{15}N-^{1}H^{N}-Couplings - IPAP (In Phase Anti Phase) ^{15}N-HSQC

To obtain ^{15}N-^{1}H^{N}-couplings a ^{15}N-HSQC could be recorded without decoupling the proton during the nitrogen evolution time. The obtained spectrum would contain in the nitrogen dimension for every nitrogen 2 peaks, split by ^{1}(J or J+D)$_{NH^{N}}$. Because ^{1}(J or J+D)$_{NH^{N}}$ is quite large it would be difficult to assign the spectrum. To overcome this problem an IPAP ^{15}N-HSQC (Fig.15) is recorded where two different spectra are obtained. It starts like a conventional ^{15}N-HSQC with an INEPT where magnetization is transferred from the proton to the nitrogen. After the INEPT, magnetization is transverse on ^{15}N and in anti-phase with respect to ^{1}H. The red marked element allows generation of in phase ^{15}N magnetization. One experiment is recorded with this element resulting in a spectrum with ^{15}N anti-phase

doublets consisting one positive and one negative peak and one experiment is recorded without this element resulting in a spectrum with ^{15}N in in-phase doublets where both peaks are positive.
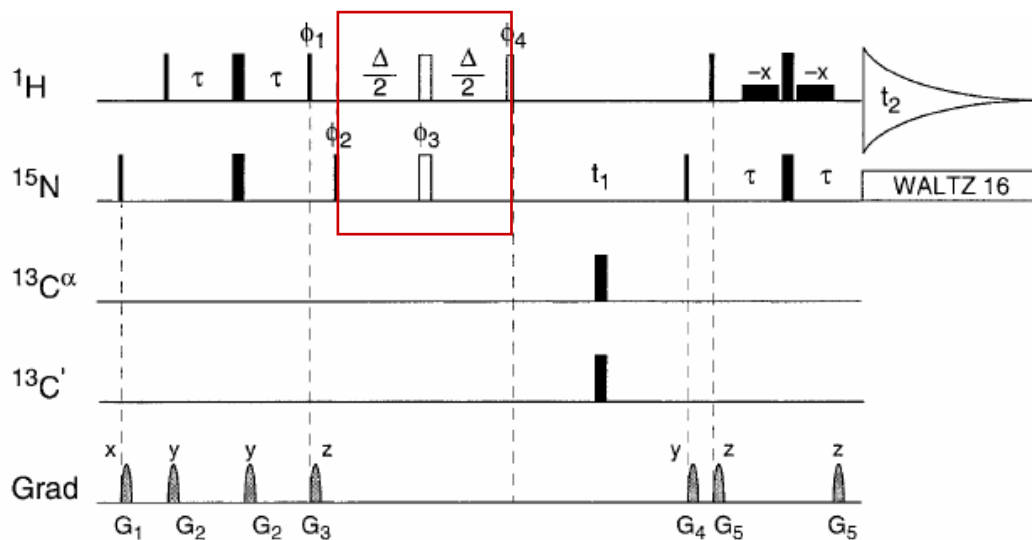
Fig.15: IPAP ^{15}N-HSQC

The red marked element allows generation of ^{15}N in-phase with respect to the ^{1}H. The spectrum is recorded twice, with and without this element, obtaining once anti-phase ^{15}N doublets and once in-phase ^{15}N doublets. Adapted from (Ottiger et al., 1998)

By taking the sum and difference of the two subspectra, two spectra containing only one peak for each nitrogen are obtained, whereby the difference of the resonance frequencies of the peaks in both spectra corresponds to ^{1}(J or J+D)$_{NH^{N}}$.
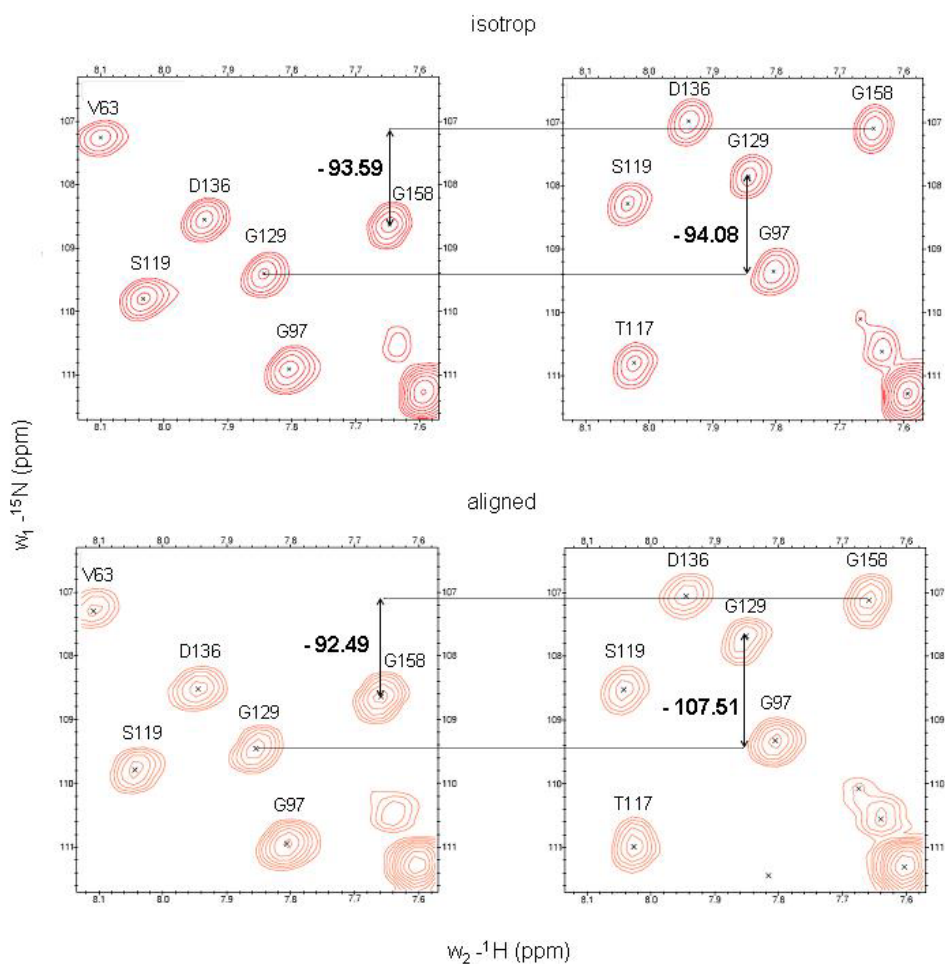
Fig.16: IPAP-^{15}N-HSQC

The figure shows two spectra obtained under isotropic conditions (peaks coloured in red) and two spectra obtained after alignment of sLb^{pro} (peaks coloured in orange). The spectra on the left side correspond to difference of the spectra obtained with an IPAP-^{15}N-HSQC and the spectra on the right side correspond to the sum. The differences of the resonance frequencies of the spectra obtained under isotropic conditions correspond to $^1(J)_{NH}^{N}$ and the differences of the resonance frequencies of the spectra obtained after alignment correspond to $^1(J+D)_{NH}^{N}$. Data are shown for selected residues of sLb^{pro}.

2.2.3.2 ^{1}H^N-^{13}C'- and ^{15}N-^{13}C'-Couplings -

^{15}N-HSQC-TROSY-^{13}C'-coupled (Wang et al., 1998)

The ^{15}N-HSQC-^{13}C'-coupled is a ^{15}N-HSQC without decoupling the carbonyl

atom during the evolution time of the nitrogen and during the acquisition time of the proton. In the obtained spectrum cross peaks have an E.cosy type pattern (Griesinger et al., 1986). Peak splittings in the proton dimension correspond to $^2(J$ or $J+D)_{H^NC'}$ whereby peak splittings in the nitrogen dimension correspond to $^1(J$ or $J+D)_{NC'}$.
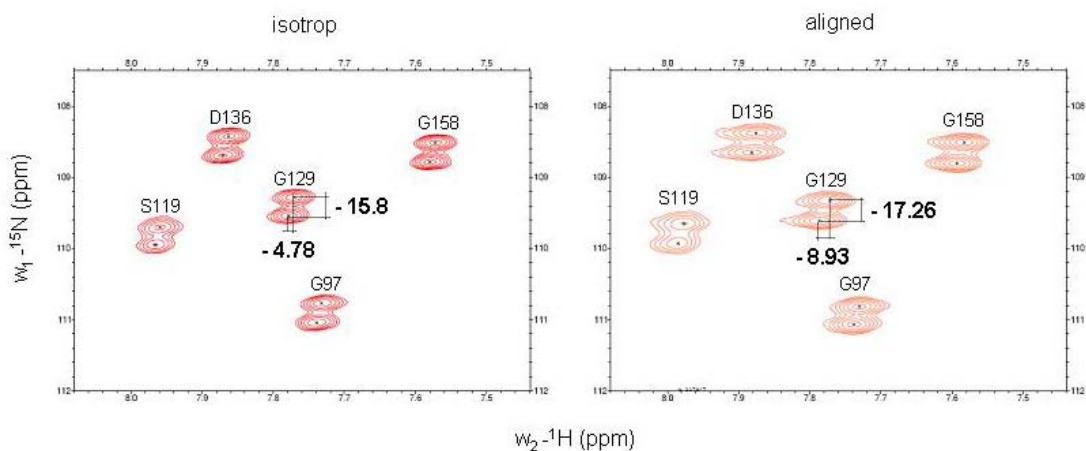
Fig.17: ^{15}N-HSQC-TROSY-^{13}C'-coupled

Peak splittings in the ^{1}H dimension corresponds to $^1H^N$-^{13}C'-couplings and peak splittings in the ^{15}N dimension correspond to ^{15}N-^{13}C'-couplings. Under isotropic conditions the left spectrum (peaks are coloured in red) is obtained from which $^2J_{H^NC'}$ and $^1J_{NC'}$ can be extracted and after alignment the right spectrum (peaks are coloured in orange) is obtained from which $^2(J+D)_{H^NC'}$ and $^1(J+D)_{NC'}$ can be extracted. Data are shown for selected residues of sLb^{pro}.

2.2.3.3 ^{13}C$^\alpha$-^{13}C'- Couplings- HNCO-^{13}C$^\alpha$-coupled

The HNCO-^{13}C$^\alpha$-coupled is a HNCO without decoupling of the ^{13}C$^\alpha$ atom during the evolution time of the carbonyl atom. In the 3D spectrum obtained the ^{13}C' peaks are split according to. $^1(J$ or $J+D)_{C^\alpha C'}$.
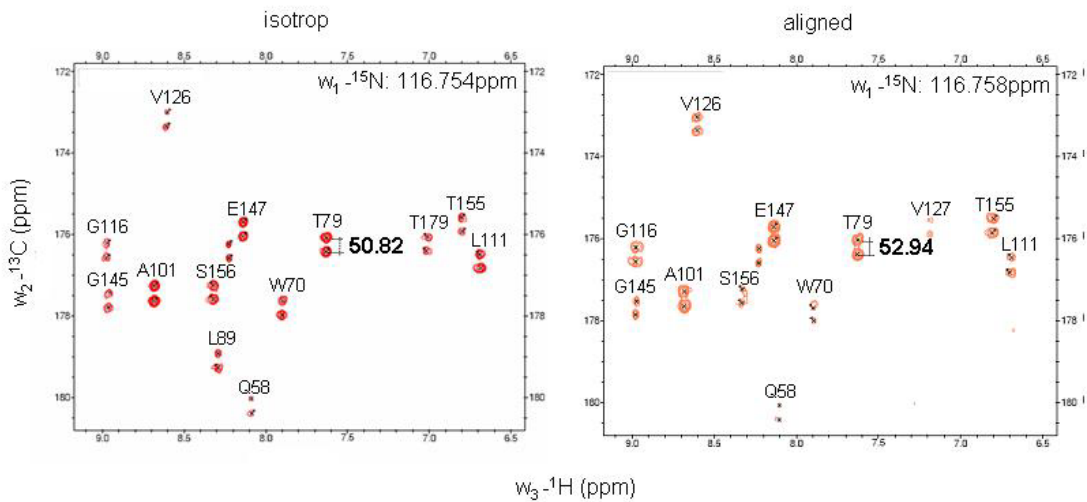
Fig.18: HNCO-^{13}C$^{\alpha}$-coupled

The left spectrum shows a HNCO-^{13}C$^{\alpha}$-coupled obtained under isotropic conditions (peaks are coloured in red) and the right spectrum shows a HNCO-^{13}C$^{\alpha}$-coupled after alignment. Peak splittings in the left spectrum corresponds to $^{1}J_{C}{}^{\alpha}{}_{C'}$ and peak splittings in the right spectrum corresponds to $^{1}(J+D)_{C}{}^{\alpha}{}_{C'}$. Data are shown for selected residues of sLb^{pro}.

2.2.4 Structural Restraints obtained by NMR Experiments

2.2.4.1 Backbone Angle Restraints

For backbone angle prediction from the chemical shift the program TALOS (Cornilescu et al., 1999) is used. TALOS is based on the principle that chemical shifts are dependent on the chemical environment of a nucleus and therefore on the secondary structure. As the chemical shifts of atoms of the protein backbone are not only depended on the residue type but also on the backbone angles, information about the dihedral angles of a protein backbone can be provided. TALOS combines the residue type information and the chemical shift for predicting backbone torsion angles. As input, the chemical shifts of H$^{\alpha}$, C$^{\alpha}$, C$^{\beta}$, C' and N^{H} for each residue are used. TALOS uses the secondary shifts (difference between measured chemical shifts and the corresponding random coil values) of 3 neighboring protein backbone residues and compares them to secondary shifts of proteins whose high resolution structures have been determined independently. From the 10 best

matches that have consistent values for ϕ and ψ average values for the dihedral angles and the standard deviation are calculated which are used for structure calculation.

2.2.4.2 Distance Restraints - NOE

NOE measurements supply information about the distances between protons. There are a lot of multidimensional NEOSY experiments. Depending on the NOESY experiments, NOEs between distinct protons can be measured. For biomolecular NMR NOESY experiments often used are a NOESY-^{15}N-HSQC and a NOESY-^{13}C-HSQC. With a NOESY-^{15}N-HSQC NOEs between amide protons and protons, that are close in space, can be measured. A NOESY-^{13}C-HSQC measures NOEs between protons, attached to a carbon atom. The distance information is derived from the intensities of the NOE crosspeaks which are proportional to $1/r^6$. This rapid decay allows observation of NOE crosspeaks only within a short range (<5-6Å). These local distance restraints provide information about the 3-dimensional structure of the molecule.

2.2.4.3 Orientational Restrains - Residual Dipolar Couplings

Residual dipolar couplings are obtained from measurements of dipolar splittings in anisotropic media (where the molecules can align in a preferred direction). This preferred direction is described by an alignment tensor which can be calculated from the residual dipolar coupling values. The degree of alignment is described by the magnitude of the alignment tensor. Residual dipolar couplings give information about the directions of distinct bonds with respect to the alignment tensor and therefore provide orientational constraints.

2.2.5 Structure Calculation

For structure calculation the Xplor NIH package was used (Schwieters et al., 2006; Schwieters et al., 2003). This program uses known geometric data and a force field and together with NMR data it calculates final structures with minimum energy. Inputs derived from NMR measurements can be e.g. backbone angle restraints obtained by TALOS or ^{3}J-couplings, distance restraints from NOE measurements, and orientational restrains obtained by measurements of the residual dipolar couplings.

The starting coordinates can be the protein in an arbitrary extended and randomized conformation or, if available, the pdb file of the crystal structure which should be refined. The energy minimization is performed by simulated annealing. Because there are often more local minima for a structure the procedure starts with a heating step in order to overcome the energy barriers between these minima. Due to high thermal energy the protein structure is almost completely randomized under these conditions. The temperature is then slowly reduced, resulting in a higher weight of the force field and the NMR restraints resulting that the molecule becomes more and more structured. The output is a bundle of low energy structures, satisfying the restraints as well as possible.

2.2.6 Structure Validation

The accuracy of a structure is assessed by a number of parameters. One quality parameter is the convergence of the structures. A small coordinate r.m.s.d (root mean square deviation) between rigid regions of the protein structures denotes a high convergence. This is usually a good sign but does not provide the information, if incorrect restraints led to false minima during structure determination. Another parameter reporting on the quality of the determined structures is the quality of covalent geometry ant the occurrence of the energetically favorable binding and torsion angles. There are distinct

angle combinations which are, as a result of steric conformation, more preferred than others, resulting in regions of the Ramachandran plot being more preferred than others.

If RDCs are available the accuracy of a structure can also be assessed by evaluating the agreement of the structure with the observed RDCs. The RDCs of the structure are obtained after calculating its Saupe matrix by finding the best fit between the corresponding and the experimental RDCs. The correlation of the observed and the calculated dipolar couplings is assessed by their r.m.s.d.

$$rms\left(D^{obs} - D^{calc}\right) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (D_i^{obs} - D_i^{calc})^2} \qquad (26)$$

In the best case the calculated dipolar couplings are equal to the observed values, then the r.m.s.d. is 0.

Additionally, the quality factor Q is defined which indicates the goodness of the correlation (Clore and Garrett, 1999; Cornilescu et al., 1998) and is a more sensitive parameter than the linear correlation coefficient R.

$$Q = \frac{rms\left(D^{obs} - D^{calc}\right)}{rmsD^{obs}} \qquad (27)$$

It can be shown that the $rmsD^{obs}$ for an assemble of randomized bond vectors, is:

$$rmsD^{obs} = \sqrt{D_a^2\left(4 + 3R^2\right)/5} \qquad (28)$$

In the ideal case this value is also 0.

3. Results:

3.1 Sample preparation of sLb^{pro} and Lb^{pro}

for NMR measurements

For structure determination of Lb^{pro} and the 6 amino acid shortened from sLb^{pro} both proteins were expressed as inactive mutants, C51A to provide better long term stability. The expression media contained $^{15}NH_4Cl$ and $^{13}C_{6-}$ D-glucose for double labelling. Both proteins were purified and concentrated to a minimum concentration of 1mM, typical required for NMR triple resonance experiments.

3.2 NMR - Measurements

For the NMR-experiments Varian Unity Inova 800MHz and 500MHz spectrometers and a 600MHz spectrometer with a cryo cold probe was used. All measurements were performed at 25°C. Experiments, which involved the ^{15}N and $^1H^N$ frequencies, were recorded as TROSY experiments because of more favourable relaxation properties.

3.3 Assignment

Backbone assignment of sLb^{pro} and Lb^{pro} as well as side chains of sLb^{pro} were already available (Cencic, 2005); (Mayer, 2007). The assignments of both structures are available in the BMRB database (Lb^{pro} entry number: 15277, sLb^{pro} entry number: 152778)

In order validate the assignment of Lb^{pro} and to assign the remaining signals of the backbone, triple resonance TROSY spectra for backbone assignment were recorded as the spectra were very crowded. The TROSY versions were used to obtain peaks with narrower line widths leading to a higher accuracy in peak picking.

3.4 Determination of Residual Dipolar Couplings (RDCs)

3.4.1 Sample preparation for measurements of the RDCs

In order to measure the dipolar couplings of Lb^{pro} and sLb^{pro} alignment media was added to both samples. As sLb^{pro} is smaller than Lb^{pro} slightly more alignment media was added to achieve approximately the same degree of alignment.

- Lb^{pro} contained 16,9mg *Pf1*/ml
- sLb^{pro} contained 15,4mg *Pf1*/ml

3.4.2 Measuring of isotropic (J) and anisotropic splittings (J+D)

The following couplings were measured for sLb^{pro} and Lb^{pro}:

- ^{15}N-^{1}H^{N}: using an IPAP -^{15}N-HSQC
- ^{1}H-^{13}C' and
- ^{15}N^{H}-^{13}C': using a ^{15}N-HSQC-TROSY-C'-coupled
- ^{13}C$^{\alpha}$-^{13}C' coupling constant using a HNCO-C$^{\alpha}$-coupled

Each spectrum was measured twice, once in the absence and once in the presents of alignment medium. From the measurements without alignment medium J and from the measurements with the alignment medium J+D was

obtained. The dipolar couplings D were obtained by taking the difference of the isotropic couplings values and the anisotropic couplings values.

3.4.3 Determination of the alignment tensor components

The alignment tensor was determined by least square fitting of the observed dipolar couplings to the calculated using the coordinates of the crystal structure.

As calculated dipolar couplings tend to be smaller than observed values the alignment tensor parameters tend to be underestimated.

Lb^{pro}: D_a^{NH}: 10.2

R: 0.28

The corresponding fitted D_{xx}, D_{yy} and D_{zz} values for $^1D_{NH}^N$ are:

D_{xx}:   -5.9

D_{yy}:   -14.5

D_{zz}:   20.3

From the powder pattern of the histogram showing the frequencies of all observed RDCs of Lb^{pro} normalized relative to $^1D_{NH}^N$ the following D_{xx}, D_{yy} and D_{zz} values observed:

D_{xx}:   -10

D_{yy}:   -25

D_{ZZ}:   35
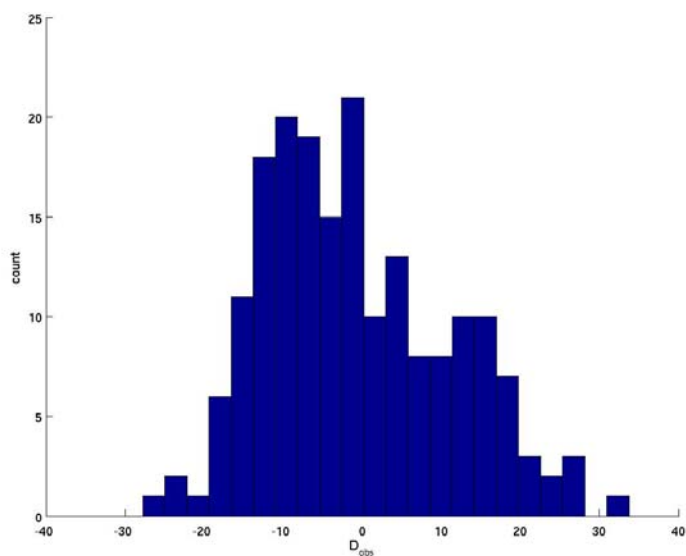
Fig.19: Histogram showing the frequencies of all observed RDCs of Lb^{pro}

The values of the observed RDCs of Lb^{pro} ($^1D_{NH}{}^N$, $^1D_{C^\alpha C'}$, $^2D_H{}^N{}_{C'}$ and $^1D_{NC'}$) were normalized relative to $^1D_{NH}{}^N$ and their frequencies of occurrence were plotted.

sLb^{pro}:   D$_a{}^{NH}$:   8.4

R: 0.55

The corresponding fitted D$_{xx}$, D$_{yy}$ and D$_{zz}$ value are:

D$_{xx}$:   -1.5

D$_{yy}$:   -15.3

D$_{zz}$:   16.8

From the powder pattern of the histogram showing the frequencies of all observed RDCs of sLb^{pro} normalized relative to $^1D_{NH}{}^N$ the following D$_{xx}$, D$_{yy}$ and D$_{zz}$ values observed:

D$_{xx}$:   -5

D$_{yy}$:   -25

D$_{zz:}$:   30
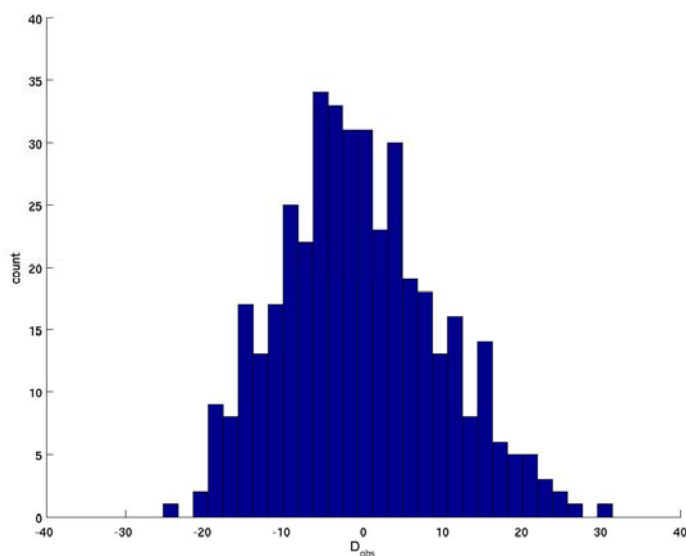
Fig.20: Histogram showing the frequencies of all observed RDCs of sLb^{pro}

The values of the observed RDCs of sLb^{pro} ($^{1}D_{NH}^{N}$, $^{1}D_{C^{\alpha}C'}$, $^{2}D_{H}^{N}{}_{C'}$ and $^{1}D_{NC'}$) were normalized relative to $^{1}D_{NH}^{N}$ and their frequencies of occurrence were plotted.

3.5 Structure calculation

Structure calculations were performed with the Xplor NIH package version 2.17.0 using the alignment tensor parameters obtained by the fit. Additional a conformational database potential was used to bias structure towards energetically favourable dihedral angles (Kuszewski et al., 1996). NOE restraints (Mayer, 2007), as well as measured dipolar couplings were used as input. As starting structures the pdb files of the crystal structures (1QMY.pdb for sLb^{pro} or 1QOL.pdb for Lb^{pro}) were used.

To prove, if the qualities of the structures change upon variation of the alignment tensor parameters, additional structure calculation were performed using the alignment tensor parameters obtained from the pattern of the histograms but there was no significant or systematic difference observed.

The RDCs contribute to the overall energy of the molecule during structure calculation with the following energy term:

$$E_{diploar} = k_{dipolar}(D_{calc} - D_{obs})^2 k_{AB}$$

For different sets of RDCs different force constants (k_{AB}) were used to consider that inaccuracies or measurements errors for small coupling constants are higher than for large coupling constants.

Force constants k_{AB} (empirically optimized):

- k_{NH^N}:   1
- $K_{C^{\alpha}C}$:   3
- $K_{H^N C}$:   2
- K_{NC}:   3

The simulated annealing process was performed with the following parameters:

	initial value	final value
temperature	800K (5000 steps)	800→0K (30000 steps) (20 cycles)
covalent forces		
bond	0.1	1
angle	0.1	1
improper	0.1	1
nonbonded		
radius	1	0.8
van d. Waals	0.01	1
ramachandran		
short range	0.002	0.1
long range	0.0002	0.01
noe	3	30
dihedral	300	30
dipolar	0.01	1
harmonic (C_α)	10	0.1

Table 3: Initial and final values of the force constants used for the simulated annealing procedure

3.6 Solution Structure of Lb^{pro}

10 lowest energy structures calculated for Lb^{pro} were selected and an average structure was calculated. Regions with regular secondary structures of one half of the dimer were overlayd. The backbone r.m.s.d. of the average structure and the 10 lowest energy structure of Lb^{pro} is 0.30+/-0.05Å.

Ramachandran plots of each amino acid residue show that the dihedral angles, ψ and ϕ are very similar for the crystal structure 1QOL.pdb (Guarne et al., 1998), the NMR structure 2JQF.pdb (Cencic et al., 2007) and the refined solution structures of Lb^{pro}.

The most noticeable deviations of the dihedral angles of Lb^{pro} compared with the crystal structure 1QOL.pdb are listed:

residue	$\Delta\phi/\Delta\psi$	residue	$\Delta\phi/\Delta\psi$
96	41/-24	71	41/-33
144	-9/60	138	21/42
145	56/-61	145	48/-99
146	44/12	146	72/-4

Table 4: Differences of the dihedral angles ϕ and ψ when compared the lowest energy structure of Lb^{pro} with the crystal structure 1QOL.pdb

Fig.21 and 22 show an overlay of one half of the dimer of the crystal structure of Lb^{pro} 1QOL.pdb shown in blue, the NMR structure 2JQF.pdb shown in green and the lowest energy NMR structure shown in red. The secondary structure elements are in agreement with the previous published structures. As seen in Fig.21 the twist angle of one half of the dimer compared with the crystal structure is in agreement with the former published NMR structure as the red and the green α-helices (α_3) are parallel. There is only a slightly sideward shift. But these two helices are moved parallel indicating that the structures differ in their bending angle. This is shown more clearly in Fig.22. The bending angle of about 20-25° is a result of the additional RDCs restraints used in structure calculation.
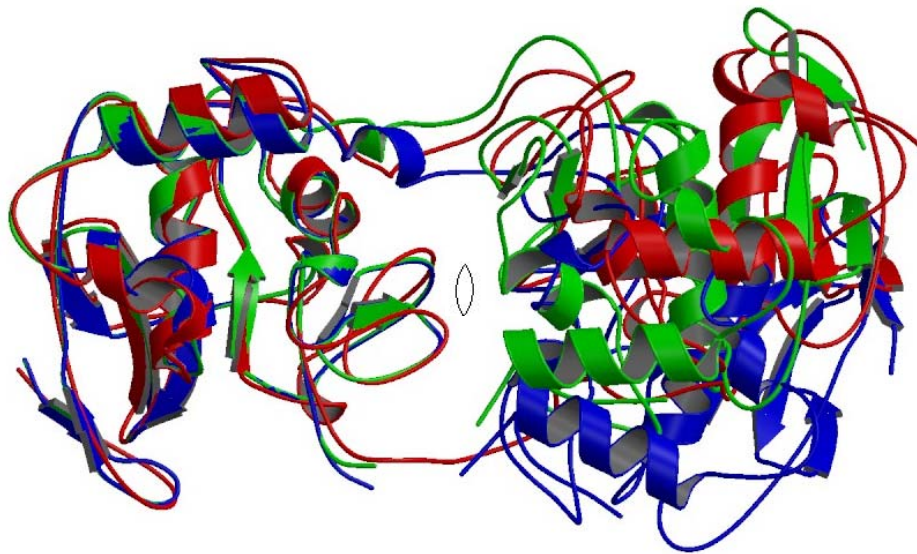
Fig.21: Overlay of 1QOL.pdb 2JQF.pdb and the lowest energy structure of Lbpro

The crystal structure of Lb^{pro} 1QOL.pdb (Guarne et al., 1998) is shown in blue, the NMR structure 2JQF.pdb published by (Cencic et al., 2007) is shown in green and the lowest energy NMR structure is shown in red. One half of each dimer was overlaid. The C_2-symmetry axis is marked with a cross. In this view the twist between the 2 halves of the dimer can be observed
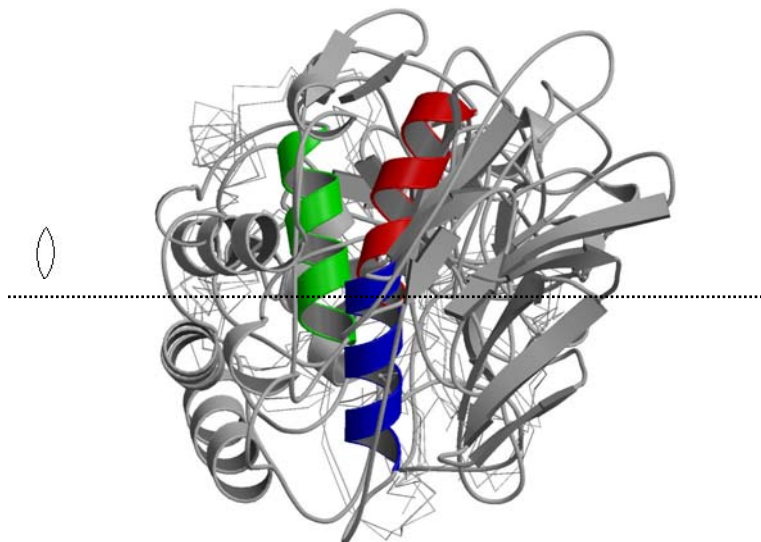
Fig.22: Side view of the overlay from above

Only one half of the dimer shows the secondary structure elements, the other halve is just a trace of the backbone. The dashed line shows the C_2-symetry axis. The α_1-helix from each structure is coloured as described above to accentuate the difference in the bending angle between the structures

3.7 Solution Structure of sLb^{pro}

Regions of regular secondary elements of the 10 lowest energy structures obtained for sLb^{pro} overlay with their average structure with a backbone r.m.s.d. of 0.27+/-0.07Å.

For comparison, the bundle of equivalent structures of sLb^{pro} calculated without RDCs has a backbone r.m.s.d. of 1.60+/-0.53Å (Cencic et al., 2007). This indicates that the additional restraints provided by the RDCs were competent for obtaining more precise NMR structures.

(a)   (b)
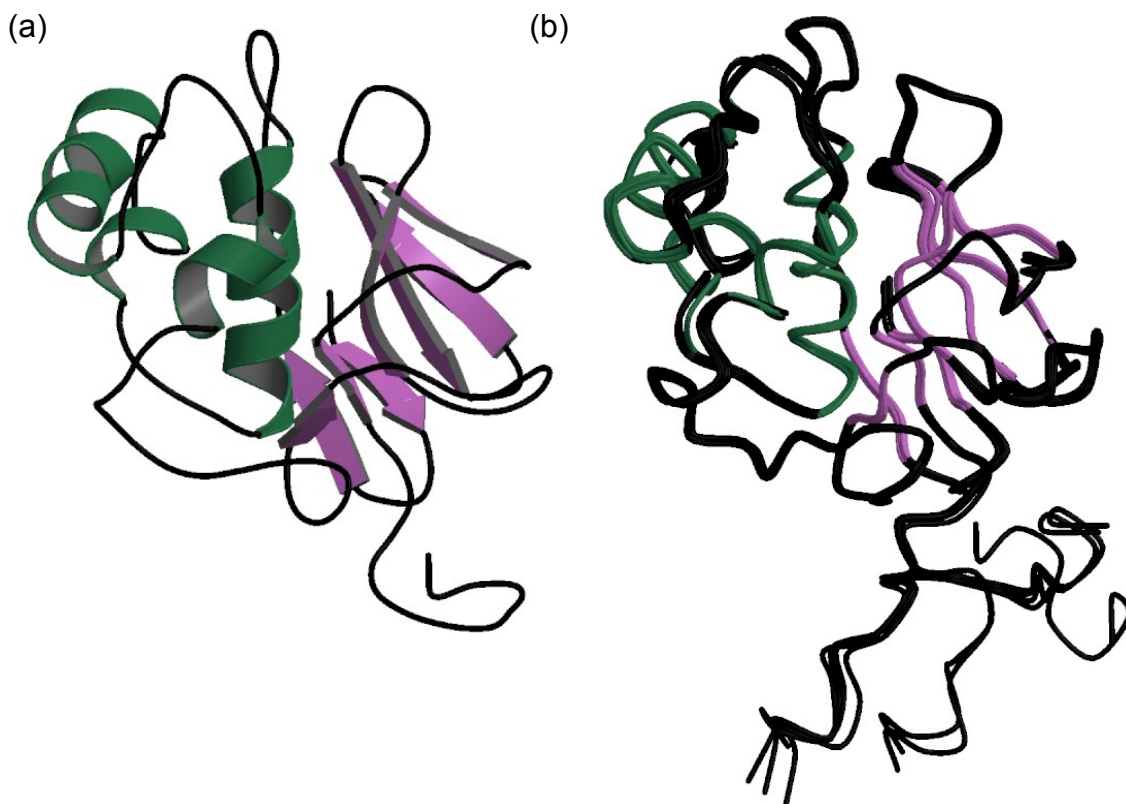
Fig.23: Final converged structures of sLb^{pro}

The α-helices are coloured in green and the β-sheets in magenta.

(a) Lowest energy structure of sLb^{pro}. (b) Bundle of the 10 lowest energy structures of sLb^{pro} where regulary secondary elements are overlaid. As previously published (Cencic et al., 2007) the C-terminus of sLb^{pro} is unstructured and flexible.

The secondary structure elements are also in agreement with the former published structures. An overlay of the 10 lowest energy structures obtained for sLb^{pro} shows that the C-terminus is, as expected, unstructured and flexible.

Ramachandran plots of each amino acid residue show that the occurence of energetically favourable dihedral angles is higher in refined structures of sLb^{pro} than in the NMR structure 2JQG.pdb (Cencic et al., 2007). Overally, the dihedral angles of the obtained structures of sLb^{pro} are in closer agreement with the crystal structure 1QMY.pdb (Guarne et al., 1998) than with the NMR structure 2JQG.pdb.

The most prominent deviations of the dihedral angels ϕ and ψ of lowest energy structure of sLb^{pro} from the crystal structure 1QMY.pdb are listed:

residue	$\Delta\phi/\Delta\psi$	residue	$\Delta\phi/\Delta\psi$
33	-68/-33	127	-104/28
34	52/-18	128	-53/36
45	1/-44	130	12/50
49	25/60	132	14/-44
50	-44/-31	145	-24/173
67	10/43	146	-155/135
71	36/-43	147	-143/500
76	29/-44	158	36/42
93	-40/-24	170	26/104
108	45/-20	171	-105/-31
126	23/86	182	2/18

Table 5: Prominent differences of the dihedral angles ϕ and ψ when compared the lowest energy structures of sLb^{pro} with the crystal structure 1QMY.pdb
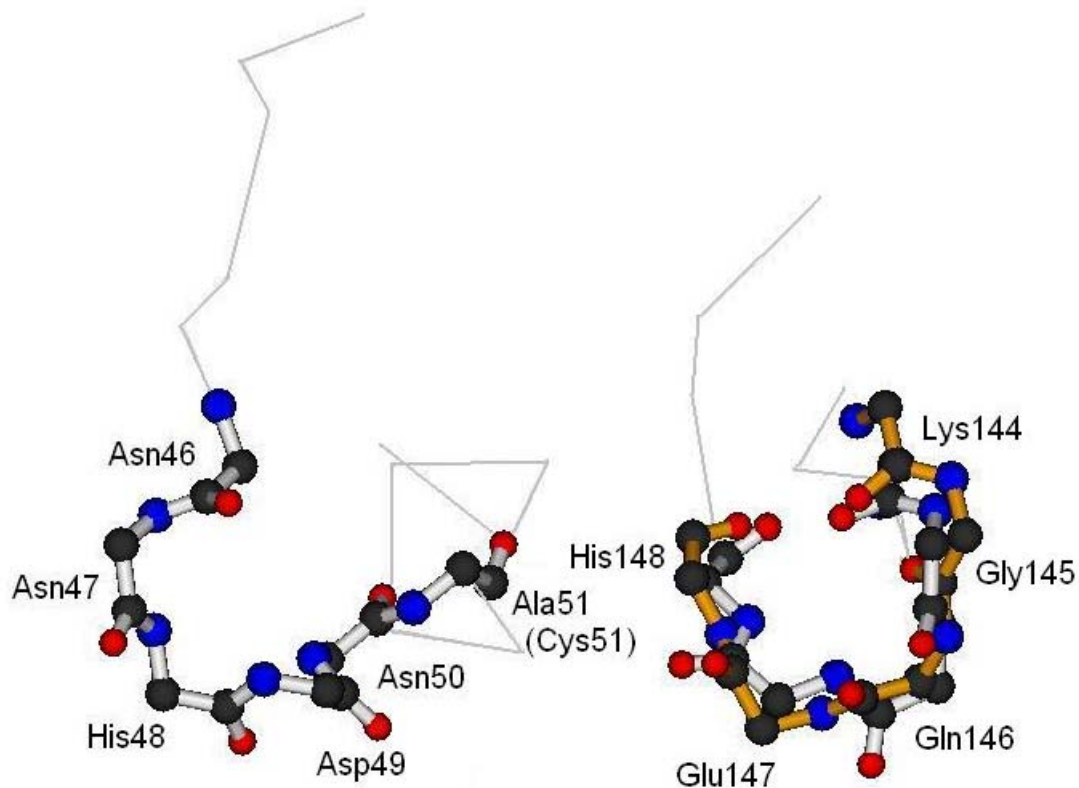
Fig.24: Most prominent differences of the dihedral angles ϕ and ψ of the lowest energy structure of Lb^{pro} with respect to the crystal structure 1QOL.pdb

The figure shows an overlay of the crystal structure 1QMY.pdb (grey) and lowest energy structure of sLb^{pro} (orange). The dihedral angles of the residues Gly145, Gln146 and Glu147 have the largest differences between these two structures. They are located at the top of one active site loop which contains the catalytic residue His148 (loop shown as backbone trace from Gly140 to Ala152 and the backbone of the residues Lys144 to His148 are presented in ball-and-stick). At the other active site loop (containing the catalytic residue Cys51 which is mutated to Ala51) the dihedral angles of the two structures match each other. Therefore only the crystal structure 1QMY.pdb is shown as representative (backbone trace from Phe41 to Leu57 and ball-and-stick model from Asn46 to Ala51).

## 3.8 Cross Validation of RDCs

Validation of the RDCs was performed by calculating how well observed RDCs agree with RDCs predicted for the calculated low energy structures. This agreement is estimated by calculating either a linear correlation

78

coefficient R or better a quality factor Q (equ. 26). For RDCs that were used as restraints in structure calculation the dipolar energy term enforces this agreement. Therefore it can be always expected that the RDCs calculated for the obtained structures agree closely with the observed ones used for structure calculations.

For structure validation a way has to be found to evaluate how good observed values not used as restraints agree with the obtained structures. In crystallography the quality of the structures is assessed by the R_{free} factor. This factor is obtained be refinement of the structure without some reflections. The agreement of these reflections with those predicted from the crystallographic model is reflected in the R_{free} factor.

In NMR studies the accuracy of RCDs can be assessed in the same way. Structure calculations are performed by leaving out one set of RDCs. Then it is evaluated how well the resulting (low-energy) structures agree with those dipolar couplings that have not been used as restraints in the structure calculation. This agreement is again estimated by calculating a linear correlation coefficient R and a quality factor Q. The values are compared with the R and the Q values that are obtained when determining the agreement of the observed RDCs with the RDCs calculated from the crystal structures (Guarne et al., 2000; Guarne et al., 1998) and the previous published NMR structures (Cencic et al., 2007). If an improvement can be observed this is an unbiased way of assessing structural quality.

This can be rationalized by considering four types of RDCs ($^1D_{NH}^{N}$, $^1D_C^{\alpha}{}_{C'}$, $^2D_H^{N}{}_{C'}$ and $^1D_{NC'}$) in one peptide plane. Three types of RDCs already uniquely define the orientation of a peptide plane. The fourth type is redundant and can be used for structure validation.
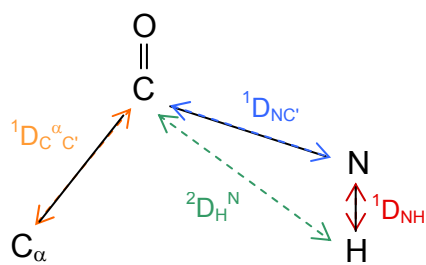
Fig.25: RDCs measured for one peptide plane

The RDCs ($^1D_{NH}{}^N$, $^1D_C{}^\alpha{}_{C'}$, $^2D_H{}^N{}_{C'}$ and $^1D_{NC'}$) are indicated with red dashed double arrow. They determine the precise orientation of each peptide plane.

In practice this situation is even better due to the limitations for the orientation of the peptide planes imposed by covalent geometry.

3.8.1 Cross Validation of the RDCs of Lb^{pro}

3.8.1.1 Cross Validation of $1D_{NH}{}^N$

	R	Q
1QOL.pdb	0.897	0.466
2JQF.pdb	0,880	0,519
Lb^{pro} w. all D$_{obs}$	0.999	0.055
Lb^{pro} w.o. $^1D_{NH}{}^N$	0.824	0,701
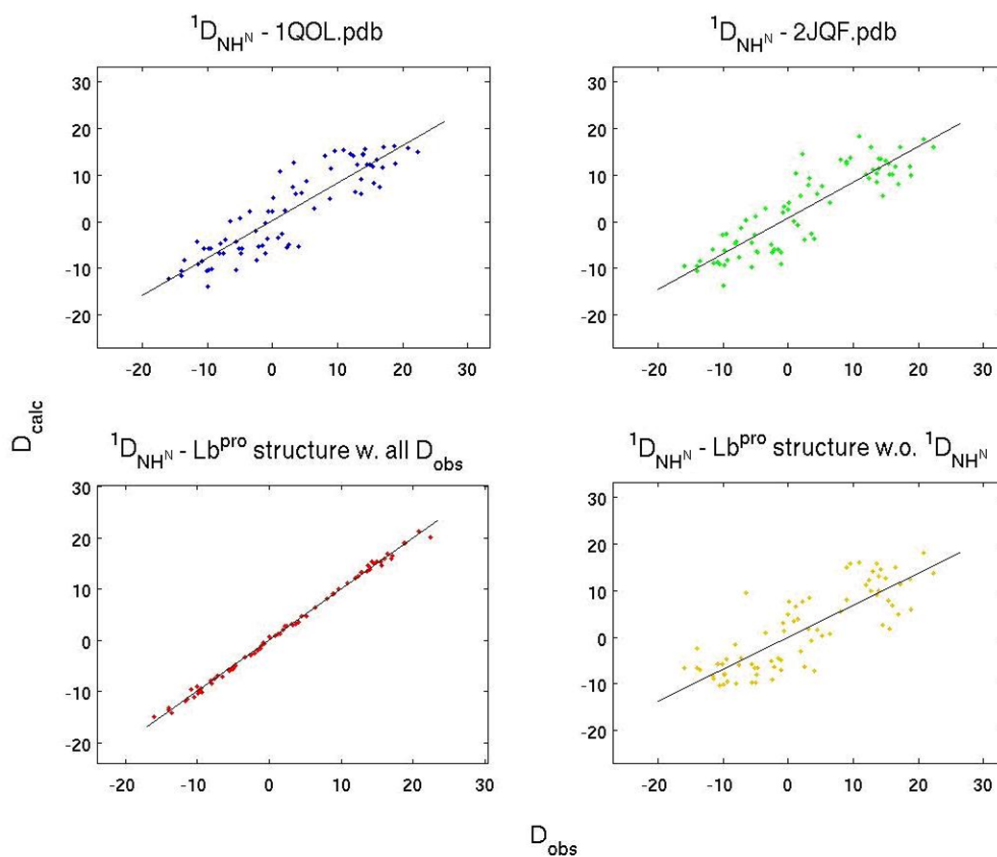
Fig.26: Correlation between observed $^1D_{NH}{}^N$ of Lb^{pro} and calculated $^1D_{NH}{}^N$ of Lb^{pro}
$^1D_{NH}{}^N$ were calculated from the crystal structure of Lb^{pro} 1QOL.pdb (Guarne et al., 1998) (blue), the NMR structure 2JQF.pdb (Cencic et al., 2007) (green), the lowest energy structure of Lb^{pro}, calculated with all sets of observed RDCs, (red) and the lowest energy structure of Lb^{pro} calculated without $^1D_{NH}{}^N$ (yellow)

3.8.1.2 Cross Validation of $^1D_{C^\alpha C'}$

	R	Q
1QOL.pdb	0.793	0.488
2JQF.pdb	0,817	0,450
Lb^{pro} w. all D$_{obs}$	0.957	0.193
Lb^{pro} w.o. $^1D_{C^\alpha C'}$	0.805	0.493
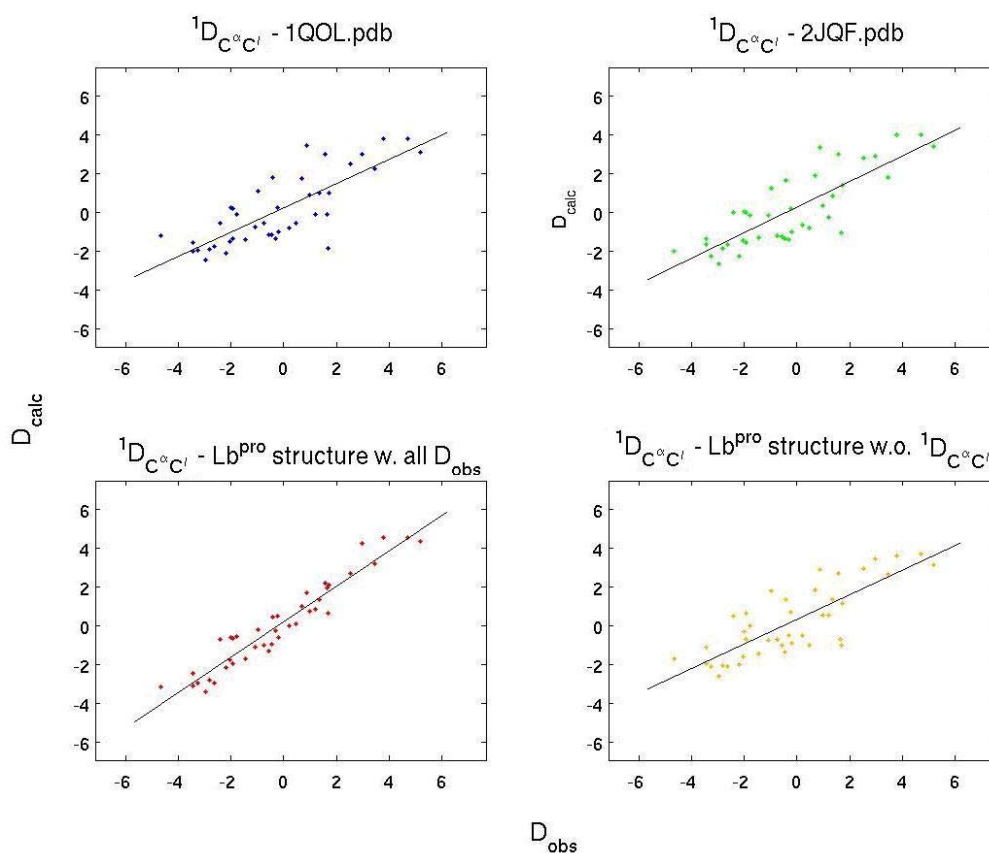
Fig.27: Correlation between observed $^1D_{C^\alpha C'}$ and calculated $^1D_{C^\alpha C'}$ of Lb^{pro}

$^1D_{C^\alpha C'}$ were calculated from the crystal structure of Lb^{pro}, 1QOL.pdb (Guarne et al., 1998) (blue), the NMR structure of Lb^{pro}, 2JQF.pdb (Cencic et al., 2007) (green), the lowest energy structure of Lb^{pro}, calculated with all sets of observed RDCs (red) and the lowest energy structure of Lb^{pro} calculated without $^1D_{C^\alpha C'}$ (yellow)

3.8.1.3 Cross Validation of $^2D_H{}^N{}_{C'}$

	R	Q
1QOL.pdb	0.854	0.382
2JQF.pdb	0,844	0,417
Lb^{pro} w. all D$_{obs}$	0.957	0.194
Lb^{pro} w.o. $^2D_H{}^N{}_{C'}$	0.865	0.362
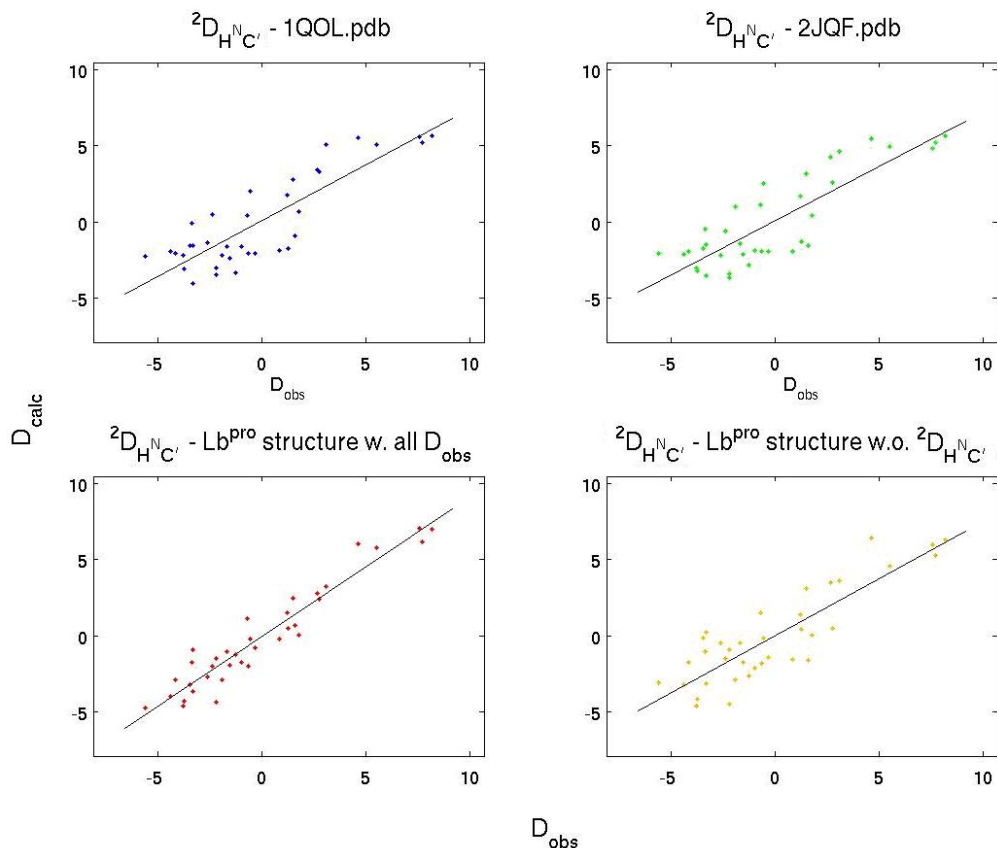
Fig.28: Correlation between observed $^2D_H{}^N{}_{C'}$ and calculated $^2D_H{}^N{}_{C'}$ of Lb^{pro}

$^2D_H{}^N{}_{C'}$ were calculated from the crystal structure of Lb^{pro} 1QOL.pdb (Guarne et al., 1998) (blue), the NMR structure of Lb^{pro}, 2JQF.pdb (Cencic et al., 2007) (green), the lowest energy structure of Lb^{pro}, calculated with all sets of observed RDCs (red) and the lowest energy structure of Lb^{pro} calculated without $^2D_H{}^N{}_{C'}$ (yellow)

3.8.1.4 Cross Validation of the $^1D_{NC'}$

|	R	Q
1QOL.pdb	0.749	0.236
2JQF.pdb	0,773	0,214
Lb^{pro} w. all D$_{obs}$	0.848	0.177
Lb^{pro} w.o. $^1D_{NC'}$	0.754	0.234
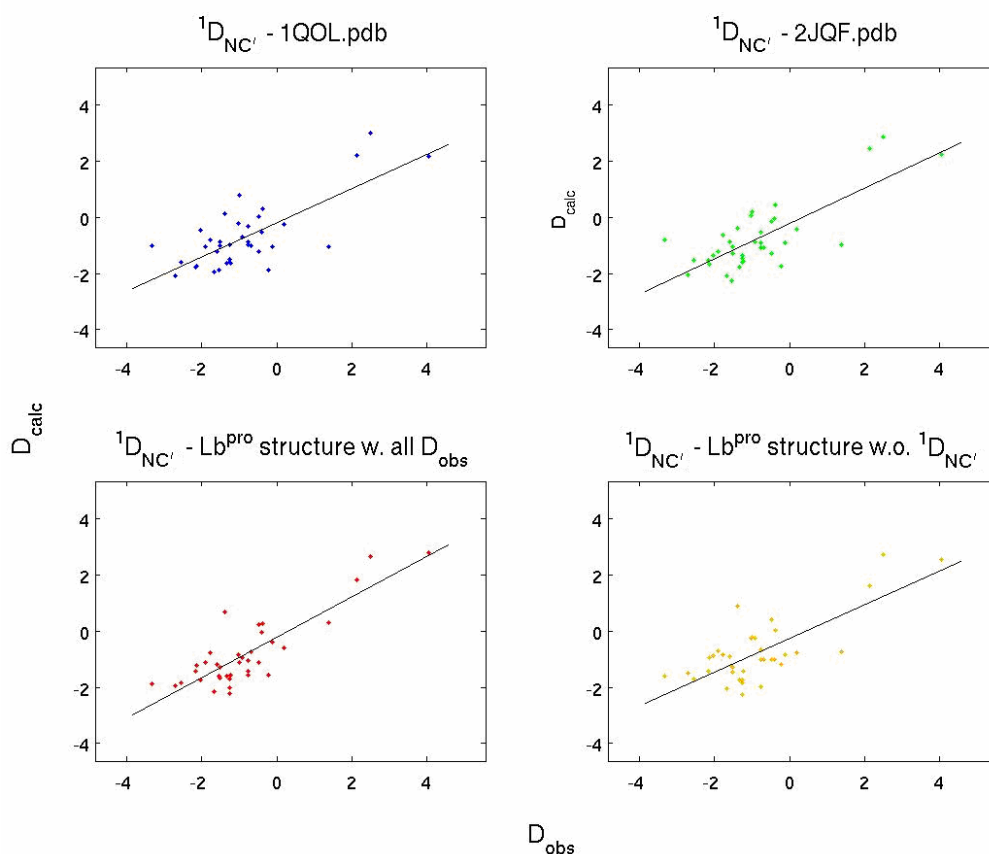
Fig.29: Correlation between observed $^1D_{NC'}$ and calculated $^1D_{NC'}$ of Lb^{pro}

$^1D_{NC'}$ were calculated from the crystal structure of Lb^{pro} 1QOL.pdb (Guarne et al., 1998) (blue), the NMR structure of Lb^{pro}, 2JQF.pdb (Cencic et al., 2007) (green), the lowest energy structure of Lb^{pro}, calculated with all sets of observed RDCs (red) and the lowest energy structure of Lb^{pro} calculated without $^1D_{NC'}$ (yellow)

As expected, the NMR structures calculated with all sets of measured RDCs used as restraints agree with each set of RDCs better than the crystal structure (1QOL.pdb) and the NMR structure (2JQF.pdb).

Independent cross validation of the RDCs shows that the remaining dipolar couplings could refine the structures to match each missing set of observed RDCs comparably well to the crystal structure (1QOL.pdb) or the NMR structure (2JQF.pdb). The structures calculated without $^2D_H{}^N{}_{C'}$ agree with these RDS even better than the other structures.

The only exceptions in this trend of good cross validation is the omission of $^1D_{NH}{}^N$, because they are the most precise RDCs and such have the biggest impact on the structure calculation. Without them the remaining RDCs, which are very small and potentially affected by large relative errors, are not sufficient for convergence.

The fact that the NMR structure, 2JQF.pdb agrees better with the observed $^1D_{NH}{}^N$ than the structure calculated without $^1D_{NH}{}^N$ but with all the other RDCs is a trivial and slightly misleading result, because $^1D_{NH}{}^N$ have been used as restraints in structures calculations of 2JQF.pdb, where the dipolar energy term has enforced the agreement.

3.8.2 Cross Validation of the RDCs of sLb^{pro}

3.8.2.1 Cross Validation of $^1D_{NH}{}^N$

	R	Q
1QMY.pdb	0.863	0.603
2JQG.pdb	0.666	1.150
sLb^{pro} w. all D$_{obs}$	0.992	0.147
sLb^{pro} w.o. $^1D_{NH}{}^N$	0.851	0.640
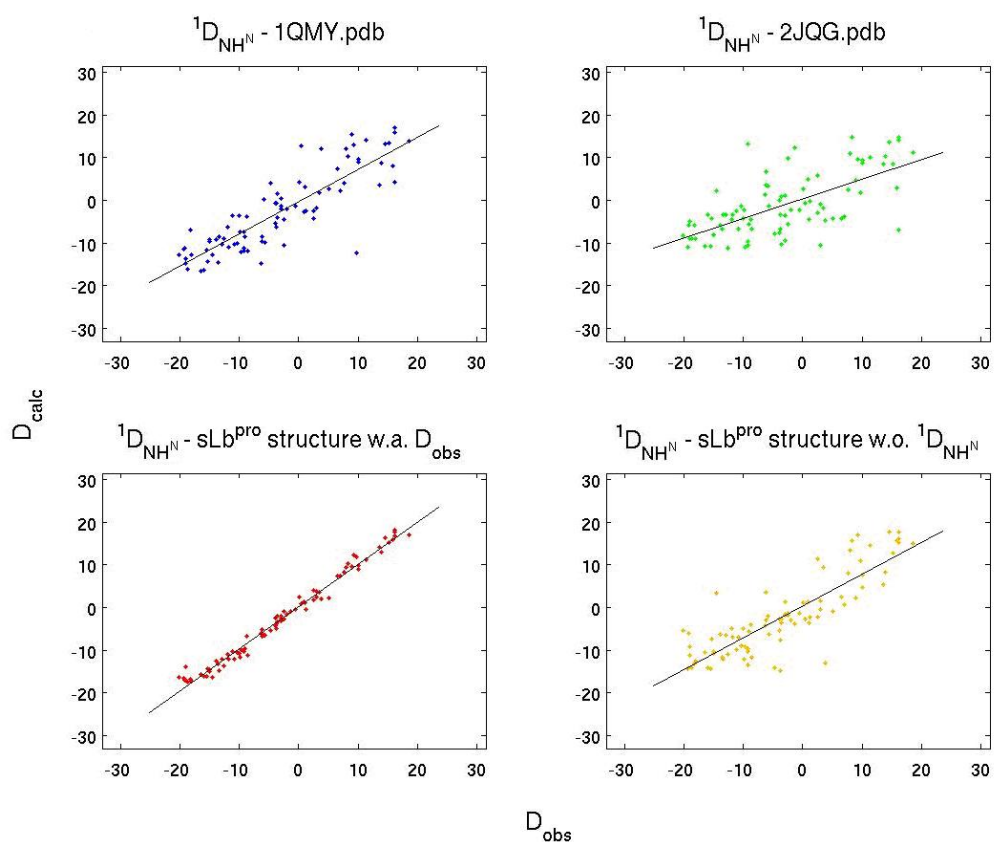
Fig.30: Correlation between observed $^1D_{NH}{}^N$ and calculated $^1D_{NH}{}^N$ of sLb^{pro}

$^1D_{NH}{}^N$ were calculated from the crystal structure of sLb^{pro} 1QMY.pdb (Guarne et al., 2000) (blue), the NMR structure of sLb^{pro}, 2JQG.pdb (Cencic et al., 2007) (green), the lowest energy structure of sLb^{pro}, calculated with all sets of observed RDCs (red) and the lowest energy structure of sLb^{pro} calculated without $^1D_{NH}{}^N$ (yellow)

3.8.2.2 Cross Validation of $^1D_C{}^\alpha{}_{C'}$

	R	Q
1QMY.pdb	0.778	0.489
2JQG.pdb	0.596	0.857
sLb^{pro} w. all D$_{obs}$	0.848	0.347
sLb^{pro} w.o. $^1D_C{}^\alpha{}_{C'}$	0.754	0.511
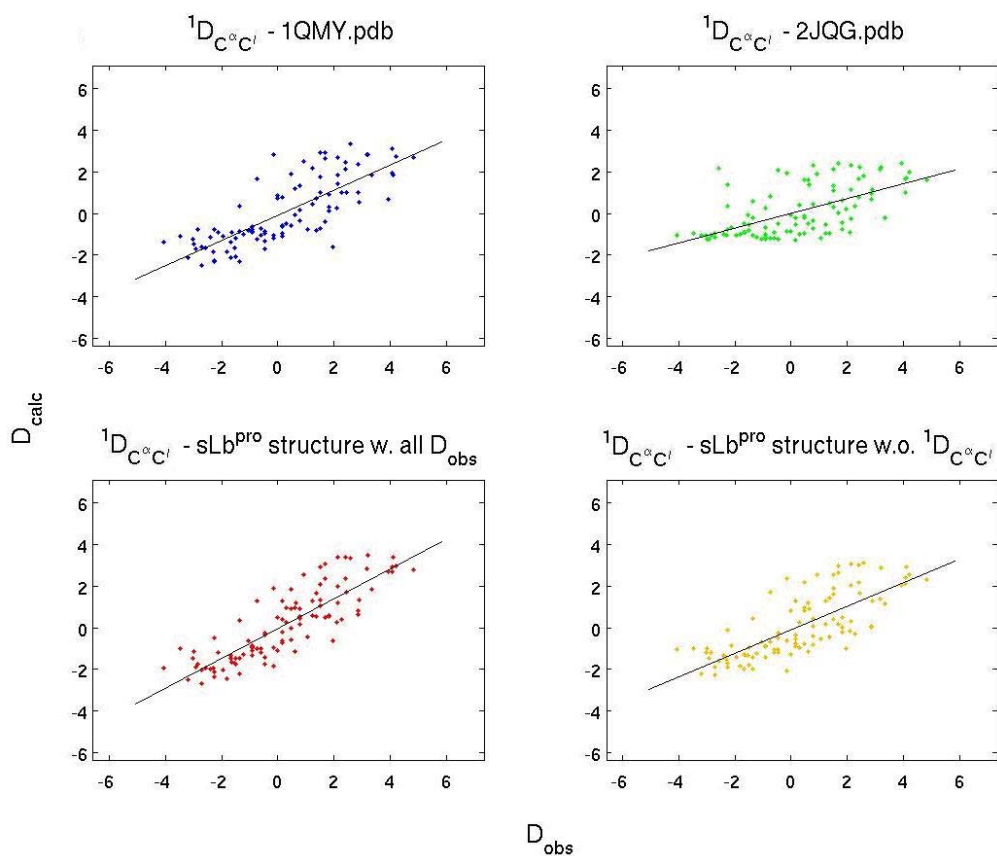
Fig.31: Correlation between observed $^1D_C{}^\alpha{}_{C'}$ and calculated $^1D_C{}^\alpha{}_{C'}$ of sLb^{pro}

$^1D_C{}^\alpha{}_{C'}$ were calculated from the crystal structure of sLb^{pro} 1QMY.pdb (Guarne et al., 2000) (blue), the NMR structure of sLb^{pro}, 2JQG.pdb (Cencic et al., 2007) (green), the lowest energy structure of sLb^{pro}, calculated with all sets of observed RDCs (red) and the lowest energy structure of sLb^{pro} calculated without $^1D_C{}^\alpha{}_{C'}$ (yellow)

3.8.2.3 Cross Validation of $^2D_H{}^N{}_{C'}$

	R	Q
1QMY.pdb	0.731	0.569
2JQG.pdb	0.623	0.853
sLb^{pro} w. all D$_{obs}$	0.879	0.304
sLb^{pro} w.o. ^{2}D$_H{}^N{}_{C'}$	0.776	0.499
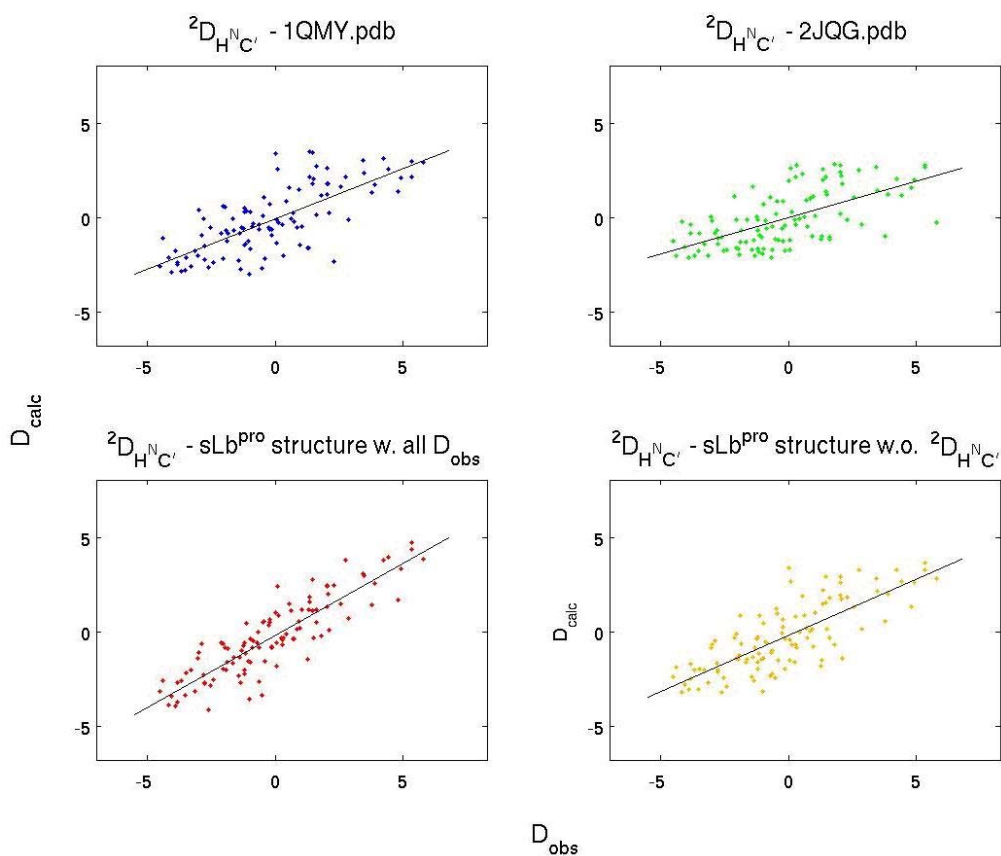
Fig.32: Correlation between observed ^{2}D$_H{}^N{}_{C'}$ and calculated ^{2}D$_H{}^N{}_{C'}$ of sLb^{pro}

^{2}D$_H{}^N{}_{C'}$ were calculated from the crystal structure of sLb^{pro} 1QMY.pdb (Guarne et al., 2000) (blue), the NMR structure of sLb^{pro}, 2JQG.pdb (Cencic et al., 2007) (green), the lowest energy structure of sLb^{pro}, calculated with all sets of observed RDCs (red) and the lowest energy structure of sLb^{pro} calculated without ^{2}D$_H{}^N{}_{C'}$ (yellow)

3.8.2.4 Cross Validation of $^1D_{NC'}$

|	R	Q
1QMY.pdb	0.841	0.222
2JQG.pdb	0.658	0.429
sLb^{pro} w. all D$_{obs}$	0.851	0.219
sLb^{pro} w.o. $^1D_{NC'}$	0.809	0.258
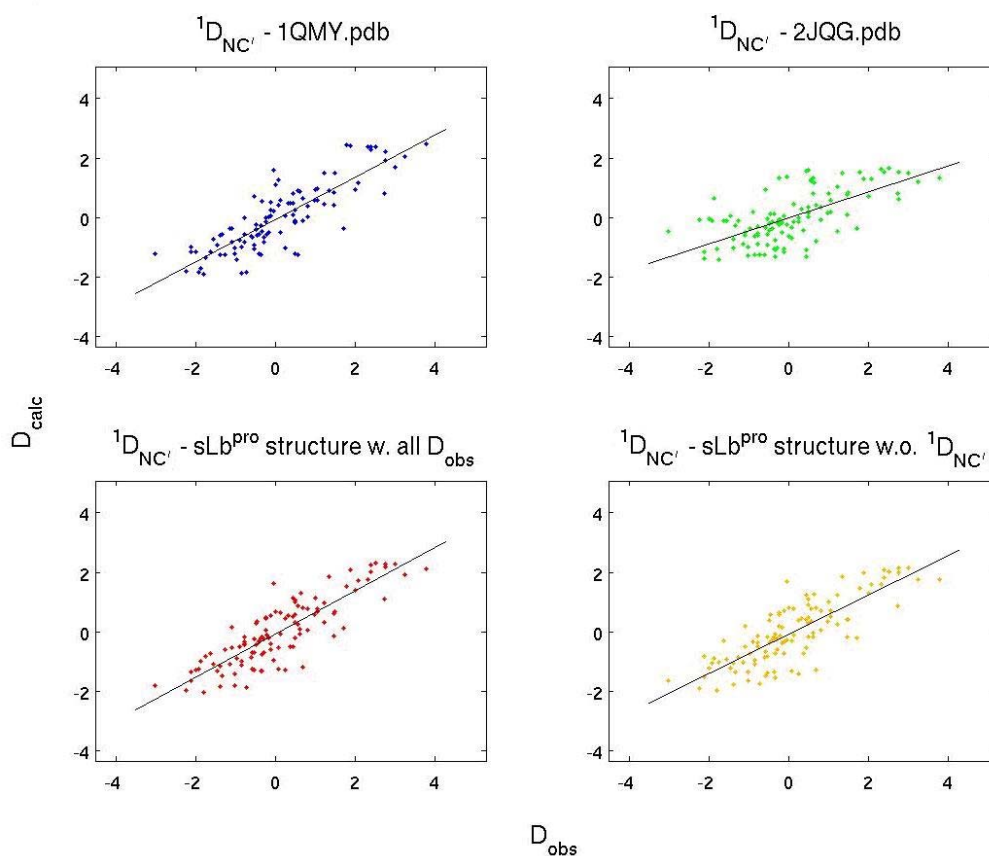
Fig.33: Corellation between observed $^1D_{NC'}$ and calculated $^1D_{NC'}$ of sLb^{pro}
$^1D_{NC'}$ were calculated from the crystal structure of sLb^{pro} 1QMY.pdb (Guarne et al., 2000) (blue), the NMR structure of sLb^{pro}, 2JQG.pdb (Cencic et al., 2007) (green), the lowest energy structure of sLb^{pro}, calculated with all sets of observed RDCs (red) and the lowest energy structure of sLb^{pro} calculated without $^1D_{NC'}$ (yellow)

As expected again the NMR structures calculated with all sets of measured RDCs agree with each set of RDCs better than the crystal structure (1QMY.pdb) and the NMR structure (2JQG.pdb).

Independent cross validation of the RDCs shows that the remaining dipolar couplings could refine the structure to improve the agreement with each missing set of observed RDCs in a way comparable to that how the crystal structure (1QMY.pdb) matches these sets of RDCs. Furthermore in the case of $^2D_H{}^N{}_{C'}$ the structures calculated omitting these RDCs agree better than the crystal structure.

As the crystal structure is very precise with a resolution of 1.9Å these results indicate that also the NMR structures obtained for sLb^{pro} are very precise structures and partially even better than the x-ray structure.
Compared with the NMR structure (2JQG.pdb) the impact of the RDCs for structure refinement is obvious.

4 Discussion:

Lb^{pro} is an important viral protein in the life cycle of the foot and mouth disease virus.

Crystallographic studies first revealed the structure of Lb^{pro} as well as the shortened form sLb^{pro}. Lb^{pro} was observed in a dimeric form and the shortened form sLb^{pro} remained monomeric. The C-terminus of sLb^{pro} could not be observed suggesting that it is flexible. NMR studies (Cencic et al., 2007) showed that the C-terminus of sLb^{pro} is indeed flexible and that the dimer formation of Lb^{pro} also occurs in solution. Additionally the dimer was observed to be completely symmetric, as opposed to an asymmetric dimer in the crystal.

The fact that Lb^{pro} forms a dimer in solution too is quite surprisingly since this was considered as an artefact of the crystal structure. Dimerisation occurs through the C-termini which lie in the active site of an adjacent molecule and vice versa. The relevance of this dimer formation is not completely clear althought it is suggested to be an inhibition mechanism of Lb^{pro}, leading to dimer formation at a distinct concentration. It is also possible however, that the dimer could also be the biologically competent form of substrate recognition. The interaction with the substrate then dissociates the dimmer releasing the active form.

In this study the structures of sLb^{pro} and Lb^{pro} were refined using residual dipolar couplings (RDCs) $^{1}D_{NH}{}^{N}$, $^{1}D_{C}{}^{\alpha}{}_{C'}$, $^{2}D_{H}{}^{N}{}_{C'}$ and $^{1}D_{NC'}$. When compared to the structures calculated by (Cencic et al., 2007), the final structures demonstrate that the additional orientational restraints provided by RDCs are very powerful in obtaining more precise NMR structures, whose precisions are comparable to, or even exceeding that of x-ray structures.

This makes RDC a useful tool to study the relative orientation of the two

halves of the Lb^{pro} dimer relative to each other in a more detailed way. Due to the lack of NOE restraints between the two halves of the dimer their orientation is not very well defined by conventional NMR experiments. As RDCs define the direction of individual bond vectors within the whole Lb^{pro} dimer orientational restraints for the whole molecule are provided. This led to a more precise determination of the relative orientation of the two halves of the dimer.

Using TROSY triple resonance experiments more signals from residues in the active site of Lb^{pro} could be assigned, leading also to a refinement with RDCs in this region.

For cross validation of the observed RDCs of Lb^{pro}, structures were calculated with one set of RDCs left out each time. The resulting structures agree with the missing sets of RDCs comparably well as the crystal structure 1QOL.pdb and the NMR structure 2JQF.pdb agree with these RDCs. The $^{2}D_{H}{}^{N}{}_{C}$ are even matched better by structures calculated without these RDCs.

The only exceptions are structures calculated without $^{1}D_{NH}{}^{N}$. The remaining RDCs are very small and as only a smaller number of them could be measured they were not sufficient for convergence. However, these results demonstrate that the quality of the calculated structure of Lb^{pro} using all RDCs is at least as good as the crystal structure 1QOL.pdb which has a resolution of 3Å.

In the case of sLb^{pro} cross validation of the RDCs show that all sets of RDCs not involved in structure calculation can be matched by structures calculated with the remaining RDCs comparably to the agreement oft these RDCs with the crystal structure 1QMY.pdb. This is also true for structures calculated without $^{1}D_{NH}{}^{N}$ as enough remaining RDCs were availabel. Also for sLb^{pro}

$^2D_H^N{}_{C'}$ are matched better by structures calculated without $^2D_H^N{}_{C'}$ than by the crystal structure. This means that the quality of the structure of sLb^{prp} calculated with all RDCs is at least as good and partially even better as the crystal structure of sLb^{pro} which has a resolution of 1.9 Å. This is a very high resolution for a NMR structures and its precision far exceeds the precision usually achieved by routine NMR studies.

Further investigations are concerned with creating a full length monomeric form of Lb^{pro}. The longer relaxation time of a monomeric Lb^{pro} should make it possible to find the remaining signals of the C-terminus. Another possibility to achieve a longer relaxation time is to deuterate Lb^{pro} in the sidechaines. Knowing the signals of the C-terminus would allow to study the dimerisation of Lb^{pro} and the behaviour of the C-terminus during substrate binding.

5 References

Alexandersen, S., Oleksiewicz, M.B. and Donaldson, A.I. (2001) The early pathogenesis of foot-and-mouth disease in pigs infected by contact: a quantitative time-course study using TaqMan RT-PCR. *J Gen Virol*, **82**, 747-755.

Bachrach, H.L. (1968) Foot-and-mouth disease. *Annu Rev Microbiol*, **22**, 201-244.

Barton, D.J., O'Donnell, B.J. and Flanegan, J.B. (2001) 5' cloverleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis. *Embo J*, **20**, 1439-1448.

Bax, A., Clore, M. and Gronenborn, A.M. (1990) 1H-1H Correlation via isotropic mixing of 13C magnetization, a new three-dimensional approach for assigning 1H and 13C spectra of 13C-enriched proteins. *J. Magn. Reson.*, **87**, 425-429.

Bax, A. and Ikura, M. (1991) An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins. *J Biomol NMR*, **1**, 99-104.

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

Belsham, G.J., Abrams, C.C., King, A.M., Roosien, J. and Vlak, J.M. (1991) Myristoylation of foot-and-mouth disease virus capsid protein precursors is independent of other viral proteins and occurs in both mammalian and insect cells. *J Gen Virol*, **72 (Pt 3)**, 747-751.

Berinstein, A., Roivainen, M., Hovi, T., Mason, P.W. and Baxt, B. (1995) Antibodies to the vitronectin receptor (integrin alpha V beta 3) inhibit binding and infection of foot-and-mouth disease virus to cultured cells. *J Virol*, **69**, 2664-2666.

Brown, F. and Cartwright, B. (1961) Dissociation of Foot-and -Mouth Disease Virus into its Nucleic Acid and Protein Components. *Nature*, **192**, 1163-1164.

Cao, X.M., Bergmann, I.E. and Beck, E. (1991) Comparison of the 5' and 3' untranslated genomic regions of virulent and attenuated foot-and-mouth disease viruses (strains O1 Campos and C3 Resende). *J Gen Virol*, **72 (Pt 11)**, 2821-2825.

Cencic, R. (2005) Structure-Function Relationships in Picornaviral Proteases.

Cencic, R., Mayer, C., Juliano, M.A., Juliano, L., Konrat, R., Kontaxis, G. and Skern, T. (2007) Investigating the substrate specificity and oligomerisation of the leader

protease of foot and mouth disease virus using NMR. *J Mol Biol*, **373**, 1071-1087.

Chan, Y.L. and Wool, I.G. (1990) The primary structure of rat ribosomal protein S20. *Biochim Biophys Acta*, **1049**, 93-95.

Chow, M., Newman, J.F., Filman, D., Hogle, J.M., Rowlands, D.J. and Brown, F. (1987) Myristylation of picornavirus capsid protein VP4 and its structural significance. *Nature*, **327**, 482-486.

Clore, G.M. and Garrett, D.S. (1999) R-factor, free R, and complete cross-validation for dipolar coupling refinement of NMR structures. *Journal of the American Chemical Society*, **121**, 9008-9012.

Clore, G.M., Gronenborn, A.M. and Bax, A. (1998a) A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J Magn Reson*, **133**, 216-221.

Clore, G.M., Gronenborn, A.M. and Tjandra, N. (1998b) Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J Magn Reson*, **131**, 159-162.

Clubb, R.T., Thanabal, V. and Wagner, G. (1992) A Constant time three-dimensional triple resonance pulse scheme to correlate intraresidue HN, 15N and 13C' chemical shifts in 15N-13C labeled proteins. *J. Magn. Reson.*, **97**, 213-217.

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

Cornilescu, G., Marquardt, J.L., Ottiger, M. and Bax, A. (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *Journal of the American Chemical Society*, **120**, 6836-6837.

Donnelly, M.L., Hughes, L.E., Luke, G., Mendoza, H., ten Dam, E., Gani, D. and Ryan, M.D. (2001a) The 'cleavage' activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring '2A-like' sequences. *J Gen Virol*, **82**, 1027-1041.

Donnelly, M.L., Luke, G., Mehrotra, A., Li, X., Hughes, L.E., Gani, D. and Ryan, M.D. (2001b) Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *J Gen Virol*, **82**, 1013-1025.

Foeger, N., Glaser, W. and Skern, T. (2002) Recognition of eukaryotic initiation factor 4G isoforms by picornaviral proteinases. *J Biol Chem*, **277**, 44300-44309.

Foeger, N., Kuehnel, E., Cencic, R. and Skern, T. (2005) The binding of foot-and-

mouth disease virus leader proteinase to eIF4GI involves conserved ionic interactions. *Febs J*, **272**, 2602-2611.

Forss, S., Strebel, K., Beck, E. and Schaller, H. (1984) Nucleotide sequence and genome organization of foot-and-mouth disease virus. *Nucleic Acids Res*, **12**, 6587-6601.

Fracastorius, H. (1546) De sympathia et antipathia rerum liber unus. De contagione et contagiosis morbis et eorum curatione liber I.

George, M., Venkataramanan, R., Gurumurthy, C.B. and Hemadri, D. (2001) The non-structural leader protein gene of foot-and-mouth disease virus is highly variable between serotypes. *Virus Genes*, **22**, 271-278.

Glaser, W., Cencic, R. and Skern, T. (2001) Foot-and-mouth disease virus leader proteinase: involvement of C-terminal residues in self-processing and cleavage of eIF4GI. *J Biol Chem*, **276**, 35473-35481.

Gorbalenya, A.E., Koonin, E.V. and Lai, M.M. (1991) Putative papain-related thiol proteases of positive-strand RNA viruses. Identification of rubi- and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubi-, alpha- and coronaviruses. *FEBS Lett*, **288**, 201-205.

Gradi, A., Foeger, N., Strong, R., Svitkin, Y.V., Sonenberg, N., Skern, T. and Belsham, G.J. (2004) Cleavage of eukaryotic translation initiation factor 4GII within foot-and-mouth disease virus-infected cells: identification of the L-protease cleavage site in vitro. *J Virol*, **78**, 3271-3278.

Gradi, A., Imataka, H., Svitkin, Y.V., Rom, E., Raught, B., Morino, S. and Sonenberg, N. (1998) A novel functional human eukaryotic translation initiation factor 4G. *Mol Cell Biol*, **18**, 334-342.

Griesinger, C., Sorensen, O.W. and Ernst, R.R. (1986) Correlation of Connected Transitions by Two-Dimensional Nmr-Spectroscopy. *Journal of Chemical Physics*, **85**, 6837-6852.

Grubman, M.J. and Bachrach, H.L. (1979) Isolation of foot-and-mouth disease virus messenger RNA from membrane-bound polyribosomes and characterization of its 5' and 3' termini. *Virology*, **98**, 466-470.

Grubman, M.J. and Baxt, B. (2004) Foot-and-mouth disease. *Clin Microbiol Rev*, **17**, 465-493.

Grzesiek, S. and Bax, A. (1992) Correlating Backbone Amide and Side-Chain Resonances in Larger Proteins By Multiple Relayed Triple Resonance NMR. *Journal of the American Chemical Society*, **114**, 6291-6293.

Guarne, A., Hampoelz, B., Glaser, W., Carpena, X., Tormo, J., Fita, I. and Skern, T.

(2000) Structural and biochemical features distinguish the foot-and-mouth disease virus leader proteinase from other papain-like enzymes. *J Mol Biol*, **302**, 1227-1240.

Guarne, A., Tormo, J., Kirchweger, R., Pfistermueller, D., Fita, I. and Skern, T. (1998) Structure of the foot-and-mouth disease virus leader protease: a papain-like fold adapted for self-processing and eIF4G recognition. *Embo J*, **17**, 7469-7479.

Harber, J.J., Bradley, J., Anderson, C.W. and Wimmer, E. (1991) Catalysis of poliovirus VP0 maturation cleavage is not mediated by serine 10 of VP2. *J Virol*, **65**, 326-334.

Herold, J. and Andino, R. (2001) Poliovirus RNA replication requires genome circularization through a protein-protein bridge. *Mol Cell*, **7**, 581-591.

Hughes, G.J., Mioulet, V., Kitching, R.P., Woolhouse, M.E., Alexandersen, S. and Donaldson, A.I. (2002) Foot-and-mouth disease virus infection of sheep: implications for diagnosis and control. *Vet Rec*, **150**, 724-727.

Jackson, T., King, A.M., Stuart, D.I. and Fry, E. (2003) Structure and receptor binding. *Virus Res*, **91**, 33-46.

Jackson, T., Sheppard, D., Denyer, M., Blakemore, W. and King, A.M. (2000) The epithelial integrin alphavbeta6 is a receptor for foot-and-mouth disease virus. *J Virol*, **74**, 4949-4956.

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

Kirchweger, R., Ziegler, E., Lamphear, B.J., Waters, D., Liebig, H.D., Sommergruber, W., Sobrino, F., Hohenadl, C., Blaas, D., Rhoads, R.E. and et al. (1994) Foot-and-mouth disease virus leader proteinase: purification of the Lb form and determination of its cleavage site on eIF-4 gamma. *J Virol*, **68**, 5677-5684.

Knipe, T., Rieder, E., Baxt, B., Ward, G. and Mason, P.W. (1997) Characterization of synthetic foot-and-mouth disease virus provirions separates acid-mediated disassembly from infectivity. *J Virol*, **71**, 2851-2856.

Kuehnel, E., Cencic, R., Foeger, N. and Skern, T. (2004) Foot-and-mouth disease virus leader proteinase: specificity at the P2 and P3 positions and comparison with other papain-like enzymes. *Biochemistry*, **43**, 11482-11490.

Kuhn, R., Luz, N. and Beck, E. (1990) Functional analysis of the internal translation initiation site of foot-and-mouth disease virus. *J Virol*, **64**, 4625-4631.

Kuszewski, J., Gronenborn, A.M. and Clore, G.M. (1996) Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases. *Protein Sci*, **5**, 1067-1080.

Lee, L.K., Rance, M., Chazin, W.J. and Palmer, A.G., 3rd. (1997) Rotational diffusion anisotropy of proteins from simultaneous analysis of 15N and 13C alpha nuclear spin relaxation. *J Biomol NMR*, **9**, 287-298.

Lee, W.M., Monroe, S.S. and Rueckert, R.R. (1993) Role of maturation cleavage in infectivity of picornaviruses: activation of an infectosome. *J Virol*, **67**, 2110-2122.

Li, F., Browning, G.F., Studdert, M.J. and Crabb, B.S. (1996) Equine rhinovirus 1 is more closely related to foot-and-mouth disease virus than to other picornaviruses. *Proc Natl Acad Sci U S A*, **93**, 990-995.

Loeffler, F. and Frosch, P. (1897) Summarischer Bericht uber die Ergebnisse der Untersuchungen zur Erforschung der Maul- und Klauenseuche. *Zentbl. Bakteriol. Parasitenkd Abt. I 22:257-259*.

Lopez de Quinto, S. and Martinez-Salas, E. (2000) Interaction of the eIF4G initiation factor with the aphthovirus IRES is essential for internal translation initiation in vivo. *Rna*, **6**, 1380-1392.

Lopez de Quinto, S., Saiz, M., de la Morena, D., Sobrino, F. and Martinez-Salas, E. (2002) IRES-driven translation is stimulated separately by the FMDV 3'-NCR and poly(A) sequences. *Nucleic Acids Res*, **30**, 4398-4405.

Lyle, J.M., Clewell, A., Richmond, K., Richards, O.C., Hope, D.A., Schultz, S.C. and Kirkegaard, K. (2002) Similar structural basis for membrane localization and protein priming by an RNA-dependent RNA polymerase. *J Biol Chem*, **277**, 16324-16331.

Mason, P.W., Bezborodova, S.V. and Henry, T.M. (2002) Identification and characterization of a cis-acting replication element (cre) adjacent to the internal ribosome entry site of foot-and-mouth disease virus. *J Virol*, **76**, 9686-9694.

Mayer, C. (2007) Investigation of the Structure and Enzyme - Substrate Interactions of the Foot - and -Mouth Disease Virus.

Melchers, W.J., Hoenderop, J.G., Bruins Slot, H.J., Pleij, C.W., Pilipenko, E.V., Agol, V.I. and Galama, J.M. (1997) Kissing of the two predominant hairpin loops in the coxsackie B virus 3' untranslated region is the essential structural feature of the origin of replication required for negative-strand RNA synthesis. *J Virol*, **71**, 686-696.

Morris, G.A. and Freeman, R. (1979) Enhancement of Nuclear Magnetic-Resonance Signals by Polarization Transfer. *Journal of the American Chemical Society*, **101**, 760-762.

Nayak, A., Goodfellow, I.G., Woolaway, K.E., Birtley, J., Curry, S. and Belsham, G.J. (2006) Role of RNA structure and RNA binding activity of foot-and-mouth disease virus 3C protein in VPg uridylylation and virus replication. *J Virol*, **80**, 9865-9875.

Neff, S., Mason, P.W. and Baxt, B. (2000) High-efficiency utilization of the bovine integrin alpha(v)beta(3) as a receptor for foot-and-mouth disease virus is dependent on the bovine beta(3) subunit. *J Virol*, **74**, 7298-7306.

Oliveira, A.C., Ishimaru, D., Goncalves, R.B., Smith, T.J., Mason, P., Sa-Carvalho, D. and Silva, J.L. (1999) Low temperature and pressure stability of picornaviruses: implications for virus uncoating. *Biophys J*, **76**, 1270-1279.

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

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

Pilipenko, E.V., Blinov, V.M., Romanova, L.I., Sinyakov, A.N., Maslova, S.V. and Agol, V.I. (1989) Conserved structural domains in the 5'-untranslated region of picornaviral genomes: an analysis of the segment controlling translation and neurovirulence. *Virology*, **168**, 201-209.

Pilipenko, E.V., Poperechny, K.V., Maslova, S.V., Melchers, W.J., Slot, H.J. and Agol, V.I. (1996) Cis-element, oriR, involved in the initiation of (-) strand poliovirus RNA: a quasi-globular multi-domain RNA structure maintained by tertiary ('kissing') interactions. *Embo J*, **15**, 5428-5436.

Ramirez, B.E. and Bax, A. (1998) Modulation of the alignment tensor of macromolecules dissolved in a dilute liquid crystalline medium. *Journal of the American Chemical Society*, **120**, 9106-9107.

Rohll, J.B., Moon, D.H., Evans, D.J. and Almond, J.W. (1995) The 3' untranslated region of picornavirus RNA: features required for efficient genome replication. *J Virol*, **69**, 7835-7844.

Sa-Carvalho, D., Rieder, E., Baxt, B., Rodarte, R., Tanuri, A. and Mason, P.W. (1997) Tissue culture adaptation of foot-and-mouth disease virus selects viruses that bind to heparin and are attenuated in cattle. *J Virol*, **71**, 5115-5123.

Saiz, M., Gomez, S., Martinez-Salas, E. and Sobrino, F. (2001) Deletion or substitution of the aphthovirus 3' NCR abrogates infectivity and virus replication. *J Gen Virol*, **82**, 93-101.

Sangar, D.V., Newton, S.E., Rowlands, D.J. and Clarke, B.E. (1987) All foot and

mouth disease virus serotypes initiate protein synthesis at two separate AUGs. *Nucleic Acids Res*, **15**, 3305-3315.

Saupe, A. and Englert, G. (1963) High-Resolution Nuclear Magnetic Resonance Spectra of Orientated Molecules. *Physical Review Letters*, **11**, 462-&.

Schwieters, C.D., Kuszewski, J.J. and Clore, G.M. (2006) Using Xplor-NIH for NMR molecular structure determination. *Progress in Nuclear Magnetic Resonance Spectroscopy*, **48**, 47-62.

Schwieters, C.D., Kuszewski, J.J., Tjandra, N. and Clore, G.M. (2003) The Xplor-NIH NMR molecular structure determination package. *J Magn Reson*, **160**, 65-73.

Skern, T., Fita, I. and Guarne, A. (1998) A structural model of picornavirus leader proteinases based on papain and bleomycin hydrolase. *J Gen Virol*, **79 (Pt 2)**, 301-307.

Suhy, D.A., Giddings, T.H., Jr. and Kirkegaard, K. (2000) Remodeling the endoplasmic reticulum by poliovirus infection and by individual viral proteins: an autophagy-like origin for virus-induced vesicles. *J Virol*, **74**, 8953-8965.

Tjandra, N., Omichinski, J.G., Gronenborn, A.M., Clore, G.M. and Bax, A. (1997) Use of dipolar 1H-15N and 1H-13C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat Struct Biol*, **4**, 732-738.

van Rensburg, H., Haydon, D., Joubert, F., Bastos, A., Heath, L. and Nel, L. (2002) Genetic heterogeneity in the foot-and-mouth disease virus Leader and 3C proteinases. *Gene*, **289**, 19-29.

van Vlijmen, H.W., Curry, S., Schaefer, M. and Karplus, M. (1998) Titration calculations of foot-and-mouth disease virus capsids and their stabilities as a function of pH. *J Mol Biol*, **275**, 295-308.

Walter, B.L., Parsley, T.B., Ehrenfeld, E. and Semler, B.L. (2002) Distinct poly(rC) binding protein KH domain determinants for poliovirus translation initiation and viral RNA replication. *J Virol*, **76**, 12008-12022.

Wang, Y.X., Marquardt, J.L., Wingfield, P., Stahl, S.J., Lee-Huang, S., Torchia, D. and Bax, A. (1998) Simultaneous measurement of H-1-N-15, H-1-C-13 ', and N-15-C-13 ' dipolar couplings in a perdeuterated 30 kDa protein dissolved in a dilute liquid crystailine phase. *Journal of the American Chemical Society*, **120**, 7385-7386.

Wimmer, E. (1982) Genome-linked proteins of viruses. *Cell*, **28**, 199-201.

Wittekind, M. and Mueller, L. (1993) HNCACB, a High-Sensitivity 3D NMR Experiment to Correlate Amide-Proton and Nitrogen Resonances with the Alpha-

and Beta-Carbon Resonances in Proteins. *Journal of Magnetic Resonance*, **101**, 201-205.

Wutz, G., Auer, H., Nowotny, N., Grosse, B., Skern, T. and Kuechler, E. (1996) Equine rhinovirus serotypes 1 and 2: relationship to each other and to aphthoviruses and cardioviruses. *J Gen Virol*, **77 (Pt 8)**, 1719-1730.

Zweckstetter, M. and Bax, A. (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. *Chem. Soc.*, **122**, 3791–3792.

Zweckstetter, M., Hummer, G. and Bax, A. (2004) Prediction of charge-induced molecular alignment of biomolecules dissolved in dilute liquid-crystalline phases. *Biophys. J.*, **86**, 3444–3460.

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mir diese Diplomarbeit ermöglicht haben und die mir währenddessen hilfreich zur Seite gestanden sind.

Bei Georg möchte ich mich bedanken für seine unermüdliche Betreuung dieser Arbeit, für die geduldige Beantwortung all meiner NMR spezifischen Fragen, sowie auch immer wieder für seine computertechnischen Ratschläge.

Ebenfalls bedanke ich mich bei allen weiteren Mitgliedern der NMR-Gruppe, für die Bereitstellung aller nötigen Mitteln, für die freundliche und hilfsbereite Unterstützung im Labor und für die netten Unterhaltungen zwischendurch.

Natürlich möchte ich mich bei meinen Eltern und meiner Familie bedanken, ohne die es nicht möglich gewesen wäre so viel Zeit und Energie für diese Arbeit zu investieren.

Ich bedanke mich bei all meinen Freunden, die mich während dieser Zeit, sowie auch sonst, so gut möglich unterstützten.

Curriculum Vitae

Sabine Schultes

Date of birth 11.08.1985
Place of birth: Vienna

Education:

- 1999-2004: Secondary College for Chemical Technology
  Rosensteingasse 79, 1170 Vienna
  Educational Focus: Biochemistry, Biotechnology and Genetic Engineering
  Diploma Project: Synthesis of N-(4-(2-fluoro)-pyridinyl-5-phenyl-thiazol-2-yl)-2,2-dimethylpropanamid, a Potential Fungizid for Usage in Agriculture
  Graduation: with highest honors
- Since October 2004: Studies of Molecular Biology at the University of Vienna,
  Emphases: Structural Biology, Microbiology/Immunology, Neuroscience
- Since July, 2007:Diploma Thesis at the Department of Biomolecular
  Structural Chemistry, University of Vienna,
  Title: **NMR Studies of the Leader Protease of the Foot and Mouth Disease Virus**

Training:

- EMBO Practical Course 2008: Multidimensional NMR in Structural Biology

working Experience:

- *Internship* DonauChem, Vienna (1.7.-31.7.2001)
- *Internship* Austrian Agency for Health and Food Safety, Vienna (5.8.-30.8.2002)
- *Praktical Work of the Diploma Project* TU-Vienna, Institute for Applied Syntetic Chemistry (24.6-17.7.2003)
- *Internship* NUA-environmental Analysis Agency of Lower Austria, Südstadt (2.8.-10.9.2004)
- *Internship* Gerot Pharmazeutika (1.7.-31.7.2005)
- *project work* Ce-M-M Research Center for Molecular Medicine, Vienna (11.8.-6.10.2006)
- *Diploma Thesis* Departement of Biomolecular Structural Chemistry of the University of Vienna (since 7.9.2007)

Talks:

- *VBC Student Symposium* (Vienna, October 2007)
  NMR Analysis of the Leader Protease of the Foot and Mouth Disease Virus – Structure, Oligomerisation State and Functional Analysis

Scholarships:

- *Leistunsstipendium*
  (a competitive scholarship of the University of Vienna) for the academic year 2005/06
- *NÖ Top-Stipendium*
  of the government of the federal State of Lower Austria

further Activities

- Since April 2005 I am tutoring mathematics and chemistry in different coaching institutions ("Lernoase", Mödling; "Lernen mit Erfolg", Mödling).