Elucidation of the Dynamics of the Autophagosomal Membrane-Associated Protein GABARAP and Structure Calculation of Lipoprotein (CD1348) by Solution NMR Spectroscopy

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Irina Apanasenko
from Kurgan

Jülich, August 2021

from the institute for Physikalische Biologie
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Dieter Willbold
Co-supervisor: Prof. Dr. Birgit Strodel

Date of the oral examination: 30/09/2021

Abstract

Membrane proteins account roughly for a third of all proteins expressed by the genomes of most organisms and carry out some of the most important cellular functions. Unfortunately, due to experimental difficulties that arise during work with these proteins, only 10% of the solved structures in the protein data-based (PDB) are membrane proteins. Nuclear magnetic resonance spectroscopy (NMR) is a unique technique for determining structures, elucidating the dynamics, and studying the binding of lipids, ligands and drugs to membrane proteins. The current work shows two different applications of solution NMR for studying membrane proteins.

The first application is an elucidation of the dynamics of the autophagosomal membrane-associated protein GABARAP. The 117-residue $GABA_A$ receptor-associated protein (GABARAP) interacts with various molecules to perform critical regulatory functions in vesicle transport and fusion events in autophagy [1]. Enzymatic lipidation of the C-terminus of GABARAP allows anchoring to autophagic membranes, which was reported to facilitate membrane hemifusion upon oligomerization. Structure determination of GABARAP using NMR and X-ray crystallography suggests significant conformational heterogeneity and dynamics. Some previous research shows that at high salt concentrations in crystal structure, the N-terminal domain can interact with the hydrophobic binding pockets of a neighbouring protein molecule [2]. Oligomerization of GABARAP *in vivo* might be induced and stabilized via interactions with the other proteins, such as tubulin, or with membranes. In this work, GABARAP was studied together with phospholipids nanodiscs in order to mimic a membrane-associated environment.

Understanding the molecular mechanisms of a multifunctional protein, such as GABARAP, requires knowledge not only its tertiary structure but also of its conformational dynamics. NMR spectroscopy is a powerful tool to study structure and dynamics on virtually all time scales at atomic resolution. In particular, the fast protein dynamics describes the movement of loops, side-chain rotations, and local atomic vibrations in the time range from a picosecond to a few nanoseconds. To determine the dynamics of GABARAP anchored to a nanodisc, we have measured ^{15}N relaxation rates R_1, R_2 and the $\{^1H\}$-^{15}N heteronuclear Nuclear Overhauser Effects (NOEs), which quantify the dynamics of the protein and reveal conformational dynamics of various regions of the tertiary structure. Residues in the termini and loop regions are highly mobile on the nanosecond time-scale as indicated by low order parameters for free GABARP. GABARAP anchored to nanodisc shows decreased internal mobility. The conformation changes of the N-terminal in the presence of a nanodisc were quantified by calculating differences of chemical shifts in three dimensions.

Studying the dynamics of the membrane protein GABARAP anchored to nanodiscs is essential for the characterization of conformational changes on different time scales to achieve a complete picture of GABARAP interacting with a membrane and finally determine the role of this protein in autophagy.

Another membrane protein, the 178-residue lipoprotein CD 1348, was found in the resistance machinery of gram-positive anaerobe bacterium *Clostridium difficile*. This bacterium has a strong line of defence against the innate immune system of the human host, such as cationic antimicrobial peptides (CAMPs), and against different antibiotics and lantibiotics. An operon encoding a three-

Abstract

component ABC-transporter in *C.difficile* is characterized as a resistance system. Also, a two-component system, which includes histidine kinase (HK) and a response regulator (RR), located on the same gene cluster might play an essential role in sensing to lantibiotics. Interestingly, lipoprotein CD 1348 is encoded directly in front of the CprABC transporter and can play an essential role in the resistance. Lipoprotein structure and its function are unknown.

The second application is the determination of 3D structure of lipoprotein by solution NMR. Also, Small-angle X-ray scattering (SAXS) and theoretical calculations using the computation method "TopModel" were used independently to generate the lipoprotein structure. The assignment of the protein backbone side chains was made by the sequential backbone assignment using 3D ^{15}N, ^{13}C Transverse relaxation optimized spectroscopy (TROSY) NMR experiments.

The calculation of the tertiary structure of proteins is one of the most important applications of NMR in structural biology. For this reason, the 3D structure calculation methods by solution NMR were performed here. The experimental 3D structure of the lipoprotein CD1348 was obtained from calculating interproton distances of NOE data from 3D ^{15}N NOESY-TROSY experiments and one 4D ^{13}C, ^{13}C Nuclear Overhauser Effect Spectroscopy (NOESY) HSQC experiment.

Protein structure determines function, given that the specificity of active sites and binding sites depends on the precise three-dimensional conformation. Thus, resolving the protein structure is a considerable step forward for future analysis for each protein. In the current work, the possibility of binding lantibiotics with lipoprotein was analysed via titration experiments by measuring the chemical shifts of lipoprotein with different concentrations of lantibiotic gallidermin. The chemical shifts of some amino acids were observed in the presence of lantibiotic. The calculated structure of the lipoprotein is important for a better understanding of its role in the resistance machinery of the bacterium *C. difficile* against CAMP, antibiotics and lantibiotics.

Zusammenfassung

Membranproteine machen ungefähr ein Drittel aller endogenen Proteine aus, die von dem Großteil aller Organismen exprimiert werden und führen einige der wichtigsten zellulären Funktionen aus. Aufgrund experimenteller Hürden, die während der Arbeit mit diesen Proteinen auftreten können, wird sind lediglich ungefähr 10 % der in Membranproteine. Die Kernspinresonanzspektroskopie (NMR-Spektroskopie) von Membranproteinen ist eine einzigartige Messtechnik, um Strukturen zu lösen, Proteindynamik zu beobachten und die Bindung von Lipiden, Liganden und Wirkstoffen an den Proteinen zu validieren. Die vorliegende Arbeit beschreibt zwei Anwendungen Methoden der Lösungs-NMR Spektroskopie, auf Membran-assoziierte Proteine.

Die erste Anwendung dient der Charakterisierung der Dynamik des autophagosomal Membran-assoziierten Proteins GABARAP. Das aus 117 Aminosäure bestehende $GABA_A$-Rezeptor-assoziierte Protein (GABARAP) interagiert mit verschiedenen Molekülen, um entscheidende regulatorische Funktionen im Vesikeltransport sowie bei Fusionsereignissen in der Autophagie auszuführen [1]. Enzymatische Lipidierung des C-Terminus von GABARAP erlaubt die Verankerung an Autophagie-assoziierten Membranen, was zur Stabilisierung der Membran-Hemifusion nach Oligomerisierung beiträgt. Strukturauflösungen von GABARAP durch NMR und Röntgenkristallographie weisen auf eine signifikante strukturelle Heterogenität und Dynamik hin. Zuvor veröffentlichte Forschungsergebnisse zeigen, dass eine in hohen Salzkonzentrationen vorliegende Alternativstruktur des N-Terminus mit der hydrophoben Bindungsstasche eines benachbarten Moleküls interagieren kann [2]. *In vivo* Oligomerisierung von GAPARAP wird wahrscheinlich durch Interaktionen mit anderen Proteinen wie Tubulin oder mit Membranen induziert und stabilisiert. In der vorliegenden Arbeit wurde GABARAP an Phospholipid-Nanodisks untersucht, um eine Membran-assoziierte Umgebung zu simulieren.

Um die molekularen Mechanismen von solchen multifunktionalen Proteinen zu verstehen, wird nicht nur ein Verständnis der Tertärstruktur benötigt, sondern auch der konformationellen Dynamik. Die NMR-Spektroskope ist ein leistungsstarkes Werkzeug um sowohl die Struktur als auch die Dynamik auf praktisch jeder Zeitskalierung auf atomarer Ebene zu untersuchen. Im Speziellen beschreibt schnelle Proteindynamik die Bewegung von Loop-Regionen, Seitenketten-Rotationen und lokalen atomaren Schwingungen in Auflösungsbereichen von Pico- bis Nanosekunden. Um die Dynamik von GABARAP zu bestimmen haben wir die ^{15}N-Relaxationsraten R_1, R_2 und die $\{^1H\}$-^{15}N heteronuklearer Kern Overhauser Effekt (NOEs) gemessen, welche die Dynamik quantifizieren und die Konformations dynamik verschiedener Regionen der Tertiärstruktur offenbaren. Die Aminosäurereste, die in den Termini- und Loop-Bereichen zu finden sind, sind höchst mobil auf der Nanosekunden Skala. An Nanodisks verankertes GABARAP eine geringere interne Dynamik. Die Konformationsänderungen des N-Terminus in Gegenwart der Nanodisks wurde durch die Berechnung der chemischen Verschiebungen in drei Dimensionen charakterisiert.

Die Untersuchung der Dynamik des Membranproteins GABARAP zusammen mit Nanodisks ist essenziell für die Charakterisierung konformationeller Änderungen auf verschiedenen

Zeitskalierungen, um ein vollständiges Bild von GABARAP und seiner Rolle bei der Autophagie zu erhalten.

Ein weiteres Membranprotein, das aus 178 Aminosäuren bestehende Lipoprotein CD 1348, wurde in der Resistenzmachinerie des Gram-positiven anaeroben Bakteriums *Clostridium difficile* entdeckt. Dieser Bakterienstamm hat eine starke Abwehr gegenüber der angeborenen Immunität des menschlichen Wirtes, wie kationische antimikrobielle Peptide (CAMPs) und verschiedene Antibiotika. Ein Operon welches für einen drei Komponenten ABC Transporter in *C. difficile* codiert, ist durch ein Resistenzsystem charakterisiert. Zudem ist ein Zwei-Komponenten-System, welches eine Histidin Kinase (HK) und einen Response Regulator (RR) beinhaltet, auf demselben Gencluster lokalisiert und spielt wahrscheinlich eine entscheidende Rolle in der Sensorik. Interessanterweise ist das Lipoprotein CD 1384 direkt vor dem CprABC Transporter codiert und könnte daher eine entscheidende Rolle in der Resistenz spielen. Weder die Struktur des Lipoproteins noch seine Funktion sind bisher bekannt.

Die zweite Anwendung ist die Berechnung der 3D-Struktur des Lipoproteins. Andere Methoden wie Kleinwinkel-Röntgenstreuung (SAXS) sowie die theoretische Berechnung mit Hilfe der Computer-gestützten Methode “TopModel” wurden unabhängig benutzt, um die Lipoproteinstruktur zu generieren. Die Zuweisung der Seitenketten des Proteinrückrates wurde durch die sequenzielle Zuweisung über 3D 15N, 13C TROSY-NMR-Experimente erreicht

Die Berechnung der Tertiärstruktur von Proteinen ist eine der wichtigsten Anwendungen von NMR innerhalb der Strukturbiologie. Die experimentelle 3D Struktur des Lipoproteins CD1348 wurde durch die Berechnung von Interproton-Distanzen aus NOE Daten von 3D 15N NOESY-TROSY Experimenten sowie mittels 4D 13C, 13C, NOESY HSQC Experimente erstellt.

Die Proteinstruktur bestimmt die Funktion, da die die Spezifität aktiver Bereiche und Bindungsstellen von der präzisen dreidimensionalen Struktur abhängig ist. Folglich ist die Auflösung unbekannter Proteinstrukturen ein bedeutender Schritt für zukünftige Analysen von Proteinen. In der vorliegenden Arbeit wurde die Möglichkeit der Bindung von Lantibiotika mit Lipoproteinen über Titrationsexperimente mittels der Messung der chemischen Verschiebung bei verschiedenen Konzentrationen des Lantibiotikums Gallidermin analysiert. Hier wurde die chemische Verschiebung einiger Aminosäuren in der Anwesenheit des Lantibiotikums beobachtet. Die berechnete Struktur des Lipoproteins ist wichtig für ein besseres Verständnis seiner Rolle in der Resistenzmachinerie des Bakteriums *C. difficile* gegenüber CAMP, Antibiotika und Lantibiotika.

Content

List of Figures

List of Figures

List of Figures

List of Tables

Chapter I Introduction

In this chapter, the basic information about two membrane-associated proteins GABARAP and lipoprotein, are presented. The GABARAP protein plays a crucial role in Autophagy and especially in autophagosome formation. After the enzymatic lipidation, GABARAP can anchor to the membrane. Lipids nanodiscs were used as a model system of biological membranes. Another protein, lipoprotein, is involved in the resistance machinery of bacteria *C. difficile* against antibiotics and antibiotics. Because of the strong antibiotic-resistant strains in *C. difficile,* a new treatment must be found. Some promising lantibiotics like subtilin, nisin and gallidermin might be good candidates. In this chapter, the dynamics of proteins and structure calculation investigated with the help of solution NMR is described, the same as fundamental principles and extending size limitation problems.

1. Membrane Proteins

Membrane proteins play an essential role in living organisms. They are involved in many cell processes such as ion transport, metabolites, sending and receiving signals, anchoring enzymes and other proteins to specific locations in the cell, to regulate intracellular vesicular transport, form the shape of organelles and others [3]. Approximately 20% and 30% of all genes encode α-helical membrane proteins [4], but this number is more significant because of the presence of β-barrels in some membrane proteins [5]. Membrane proteins can be separated into different groups based on how they are associated with the lipid bilayer. Some are thought to extend across the membrane; others are located in the cytosol and associated with the membrane by one or more covalently attached lipid chains [6].

1.1. GABARAP

1.1.1. Autophagy

Autophagy ("self-eating") is a natural cellular degradation and recycling process. This is one of the main catabolic mechanisms that helps cells to survive in different stress situations such as heat, starvation, infection, oxidative stress, and others. The process of Autophagy is involved in preventing some types of human diseases, including cancer, muscular dystrophies, and neurodegenerative disorders such as Huntington, Alzheimer, and Parkinson [7-13]. There are three types of autophagy processes in mammalian cells: microautophagy, macroautophagy, and chaperone-mediated autophagy (CMA). During microautophagy, the cytoplasmic cargo is directly engulfed by the lysosome during a random process of membrane invagination. Another type of Autophagy is the CMA, where proteins are specifically targeted to lysosomes via signal peptides and coordinated by chaperones located on both sides of the targeted membrane [14]. Macroautophagy is based on the sequestration of cytoplasmic contents in a *de novo* formed double-membrane vesicle "autophagosome" followed by the fusion with the lysosomes where the lysosomal enzymes facilitate the degradation of the sequestered products. In 2016 the Nobel Prize for Physiology or Medicine was awarded to Yoshinori Ohsumi for discoveries of the mechanisms of Autophagy in yeast organisms. The essential roles of ubiquitin-like proteins in forming the double-membrane sequestering vesicle autophagosome were shown, which plays a central role in the macroautophagy process. The process of autophagosome formation is shown in Figure 1. In yeast, autophagosome formation starts at a single perivacuolar site called the phagophore assembly site (PAS) [15, 16]. In mammalian cells, an isolated membrane is known as phagophore appears from vesicles from the endoplasmic reticulum (ER) [17, 18] and other cytosolic membrane structures such as trans-Golgi network and late endosomes [18-20]. The structure of vesicles from ER was forming an Ω-like shape "omegasomes" and was very dynamic and colocalized with autophagy-specific proteins [18]. The formation of omegasomes leads to phagophore formation and correspond to the early stage in the autophagic process [17, 18, 21]. Multiple Atg (Autophagy) proteins are involved in autophagosome formation [22]. Especially Atg 8 proteins family plays a crucial role in intracellular transport and fusing vesicles into autophagosome [23]. Human

analogies of Atg8 such as microtubule-associated protein 1 light chain 3 (LC3) and the gamma-aminobutyric acid receptor-associated protein (GABARAP) family proteins have been identified as a required system for elongation and maturation of the autophagosome [24-26].

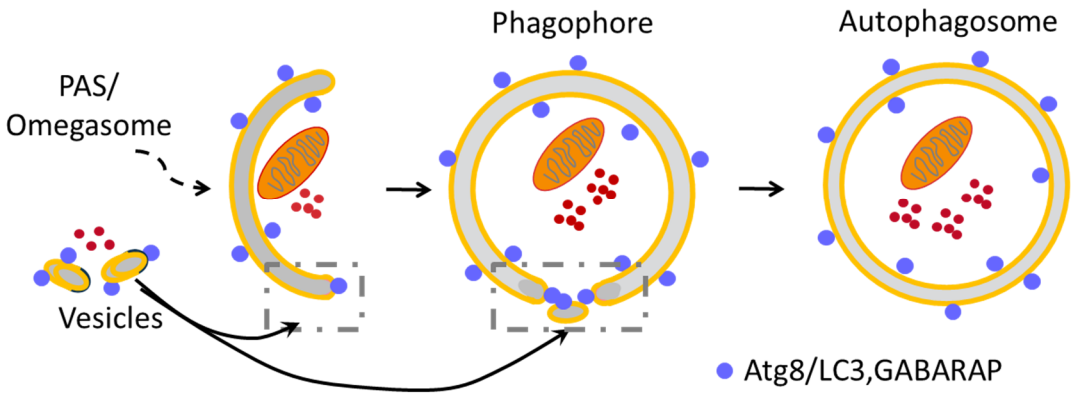


Figure 1. Schematic depiction of autophagosome formation. Phagophore surrounds the content that should be removed from the cells by forming the double membrane. Then vesicles with anchored Autophagy relate proteins on it (blue dots) elongate and close the phagophore. Sequestering vesicle autophagosome is formed.

1.1.2. Autophagy Related Protein GABARAP

The γ-aminobutyric acid type A receptor-associated protein (GABARAP) is a mammalian cells Atg8 homolog, which was found in plants to mammals and can be expressed in a wide variety of organisms and tissues. Mammalian forms of GABARAP have identical amino acid levels, suggesting that the function of GABARAP is essential or beneficial in mammals [1]. GABARAP is a 117 amino acids protein with a molecular weight of around 14 kDa. It belongs to small ubiquitin (Ub)-like proteins (UBLs) covalently anchored to lipid membranes and play an essential role in growing and closing the autophagosome. GABARAP is implicated in various membrane trafficking and fusion of autophagosome formation during the autophagy process [1, 26]. Figure 2 shows the structure of GABARAP. It contains a C-terminal domain (residues 27–117) and a small N-terminal subdomain (residues 1–26). The C-terminal domain shows structural similarity to ubiquitin. N-terminal is not present in ubiquitin and is located on the opposite side of the C-terminal. It has two hydrophobic pockets (hp) where hp1 is formed by Gly 17, Ile 21, Pro 30, Lys 48, Leu 50 and Phe 104, whereas hp2 is formed by Tyr 49, Val 51, Pro 52, Leu 55, Phe 60, Leu 63 and Ile 64.

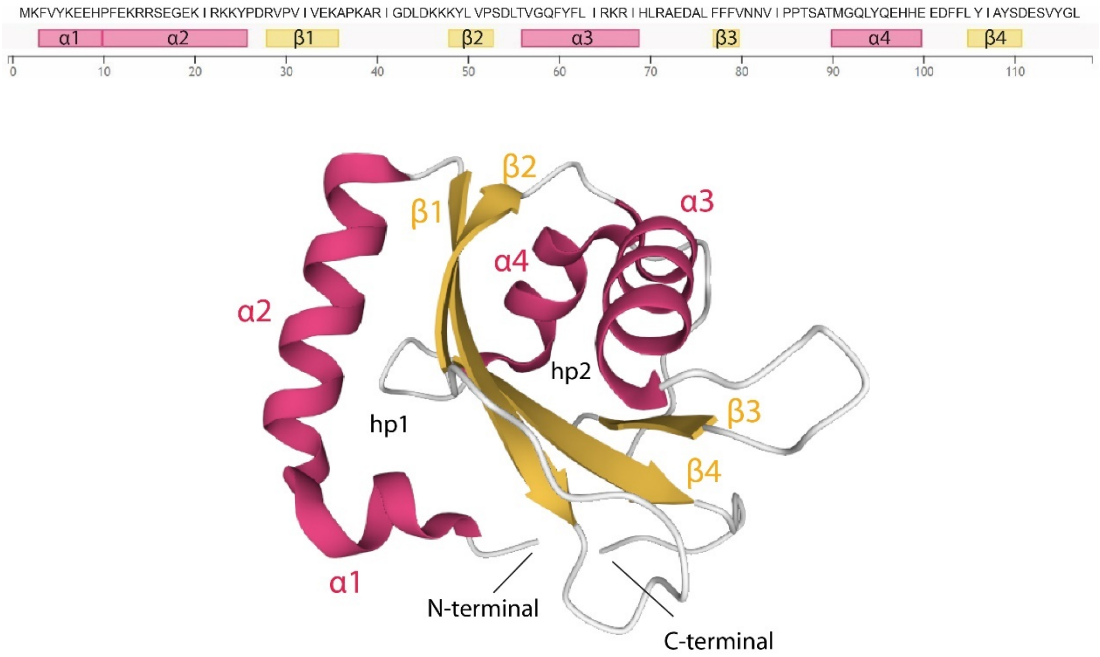


Figure 2. Structure of GABARAP. The picture shows the sequence, secondary and tertiary structure of GABARAP, where pink is the α-helix, yellow is the β-sheet.

A monomeric form mainly represents GABARAP structure at low protein concentration and in solution with salt concentration up to 100 μM [2]. The oligomeric conformation was detected in the crystal structure at the high salt concentration (2.4 M ammonium sulphate). This enhances the hydrophobic interactions dominating the multimerization interface. Oligomerization of GABARAP *in vivo* might be induced and stabilized via interactions with other proteins, such as tubulin, or with membranes [2]. GABARAP protein can exist in two distinct conformations called "open" and "close". In "close" conformation, the first ten amino acids from the N-terminal have some of van der Waal's interactions with residues from the central β sheets (β1 and β2). When the N-terminal is flipped almost 180° relate to the start position, this is "open" conformation. Figure 3 represents the open conformation of GABARAP. The first six amino acids from N-terminal binds a β2 strand of neighbouring molecule in the crystal structure. The N-terminal is flipped almost 180° relate to the start position [2].

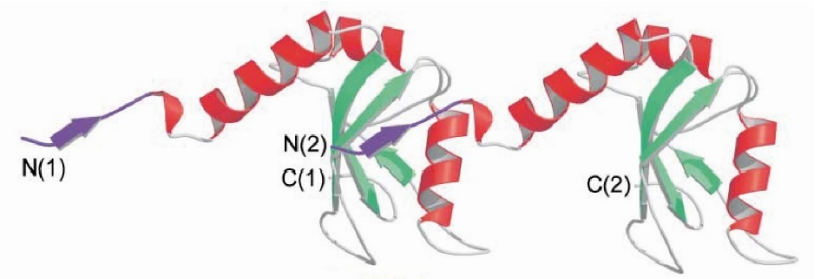


Figure 3. Open conformation of GABARAP. Intermolecular contacts between neighbouring molecules of GABARAP. The N-terminal (N(2)) of one GABARAP molecule binds with the β2-strand of neighbouring GABARAP in the crystal structure. The first six amino acids from N-terminal are coloured violet. β-strands are coloured green, and α-helixes are coloured red. Picture from the ref. [2].

GABARAP protein plays an essential role in vesicle transport and membrane fusion during autophagosome formation [1, 27]. For the association with the membrane, GABARAP is reversibly coupled to membranes in a Ubq-like manner [28]. Enzymatically lipidated C-terminal allows to anchor

to the membrane. GABARAP protein is covalently attached to the membrane via four steps mechanisms (Figure 4). First, Leu 117 in C-terminal was cleaved off by cysteine protease Atg4B, which leads to a truncated form GABARAP-I with Gly 116 in the end. Then it is subsequentially activated by Atg7 and transferred to the E2-like enzyme Atg3. After it, GABARAP-I is enabled to covalently bind to phospholipids (PE, PS) subsequentially transformed into GABARAP-PE. The process of lipidation can be reversed by the cysteine protease Atg 4B back to GABARAP-I.

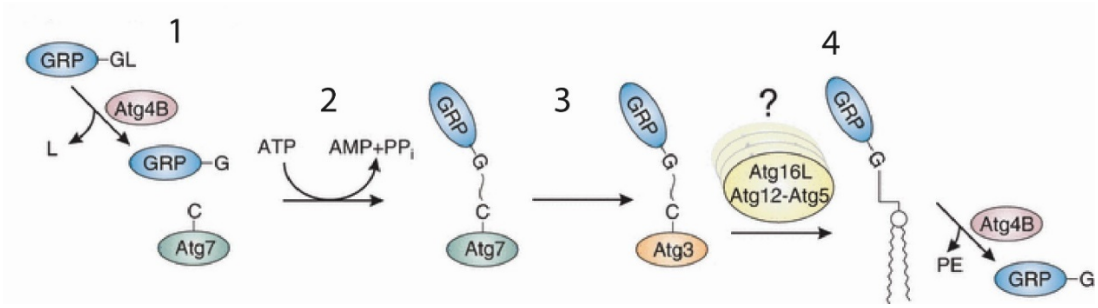


Figure 4. Schematic representation of the liquidation process of GABARAP. (1) The Leu 117 is cleaved off by Atg4B. (2) With the help of ATP, GABARAP-I is activated by Atg7 and transferred (3) to the E2-like enzyme Atg3. (4) The Atg3-GABARAP complex enables the binding of GABARAP with the phospholipids (PS, PE). Atg4B facilitates delipidation to GABARAP-I. The picture from [1].

This modification of GABARAP into GABARAP-PE is proposed to be crucial for its intracellular distribution during the autophagy process [25].

1.1.3. Nanodiscs

Nanodiscs are suitable membrane mimicking systems for membrane proteins, which allow the study of these types of proteins in a native-like environment. Due to low molecular mass, nanodiscs have some advantages for nuclear magnetic resonance (NMR) spectroscopy compared to liposomes. Figure 5 shows the schematic representation of the nanodisc structure [29]. They usually consist of an assembly of phospholipids held together by two copies of amphipathic apolipoproteins, known as membrane scaffold proteins (MSPs), arranged in a disc-shaped lipid bilayer [30, 31]. MSP is an amphipathic helical protein designed to self-assemble into a small lipid bilayer upon detergent removal [32, 33]. A detergent-free environment provides a near native-like lipid membrane system. With MSP's help, the nanodiscs' size might be controlled from 9.5 to 12.8 nm by deleting one or more of the α-helixes [34]. One of the shorted constructs is MSP1D1Δ5 does not have a fifth α-helix and consist of 167 amino acids. It forms nanodiscs with a diameter of 9.2 nm and has a molecular weight of around 95 kDa [35]. Smaller nanodiscs have advantages in solution NMR to reduce the rotational correlation time and thus increase spectral resolution and sensitivity of the spectrum. MSP1D1Δ5 construct shows an efficient application in NMR [31, 35]. The lipid compositions of biological membranes have many variations, and it might be challenging to define the exact lipid composition for a studied membrane protein. Standard phospholipids such as DMPC, DPPC or POPC are most used for nanodiscs preparation [31]. Nanodiscs have some advantages over liposomes due to their smaller size, homogeneity and higher stability [36].

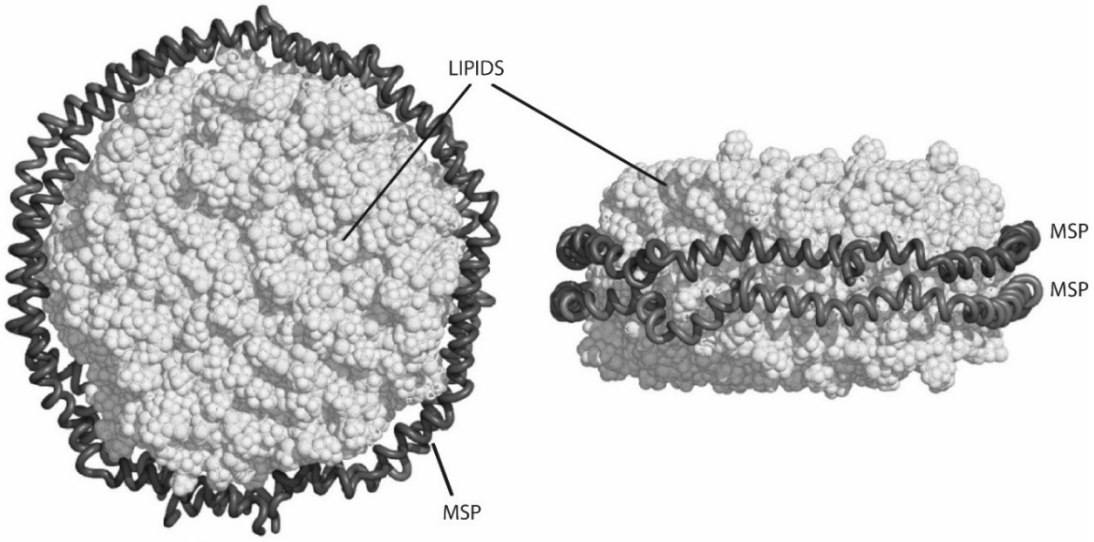


Figure 5. Structure of nanodiscs. Top and side view of the nanodisc. The lipid bilayer is surrounded by two copies of amphipathic helices of belt protein (MSP). The picture was reproduced from [29].

1.2. Lipoprotein CD1348

1.2.1. The Importance of Antibiotics in our Life

Antibiotics ("against life") is a type of medicine against infections caused by bacteria. The first antibiotic, penicillin, was discovered a few decades ago by Sir Alexander Fleming. Nowadays, antibiotics are widely used to treat different bacterial infections and significantly increased the human standard of living. Antibiotics can be separated into antibiotics to induce cell death (bactericidal) or inhibit cell growth (bacteriostatic). Most bactericidal antibiotics are divided by their mode of action, in other words, how they kill the bacteria. Antibiotics inhibit one of four targets in the cell: DNA synthesis, RNA synthesis, cell wall synthesis, or protein synthesis [37]. The cell wall synthesis of bacteria is the most effective target for antibiotic intervention [38]. The bacterial cell wall is critical to the ability of bacteria to survive in an unfavourable environment because successful treatment with cell wall synthesis inhibitor can course cell lysis. The defensive mechanisms of bacteria are developing, leading to antibiotic resistance [39]. The developed mechanisms are various and can defy the activity of many different types of antibiotics. Thus, there is a need for new antimicrobial medicines against human pathogens.

1.2.2. Lantibiotics

Lantibiotics ("*lan*thionine-containing anti*biotic*") are a group of small ribosomally synthesised peptides produced by gram-positive bacteria. Lantibiotics have high antimicrobial activity against gram-positive bacteria. Most lantibiotics bind to Lipid II or other peptidoglycan precursors, causing inhibition of bacterial cell wall synthesis. Lipid II is a membrane-anchored cell-wall peptidoglycan precursor that is important for bacterial cell-wall biosynthesis [40]. Some lantibiotics can form pores in the cell wall, leading to membrane leakage and cell death [41]. The typical lantibiotics contain dehydrated amino acids (didehydroalanine (Dha) and didehydrobutyrine (Dhb)) and other unusual amino acid modifications [42]. They are complex polycyclic molecules formed by the dehydration of selected

residues of Ser and Thr, and the addition of Cys thiols leads to unsaturated amino acids to form methyllanthionine bridges and lanthionine rings [43]. Different names of lantibiotics related to these rings. Based on the pathway by which maturation of the peptide occurs and the presence or absence of antibiotic activity, all lantibiotics can be separated into four classes [43, 44]. Only lantibiotics from two classes were reported as active against microbes [45]. Rings play a crucial role in the antimicrobial activity against a lot of gram-positive bacteria. The most know lantibiotics, nisin, gallidermin, and subtilin, are members of the same class and are shown in Figure 6.

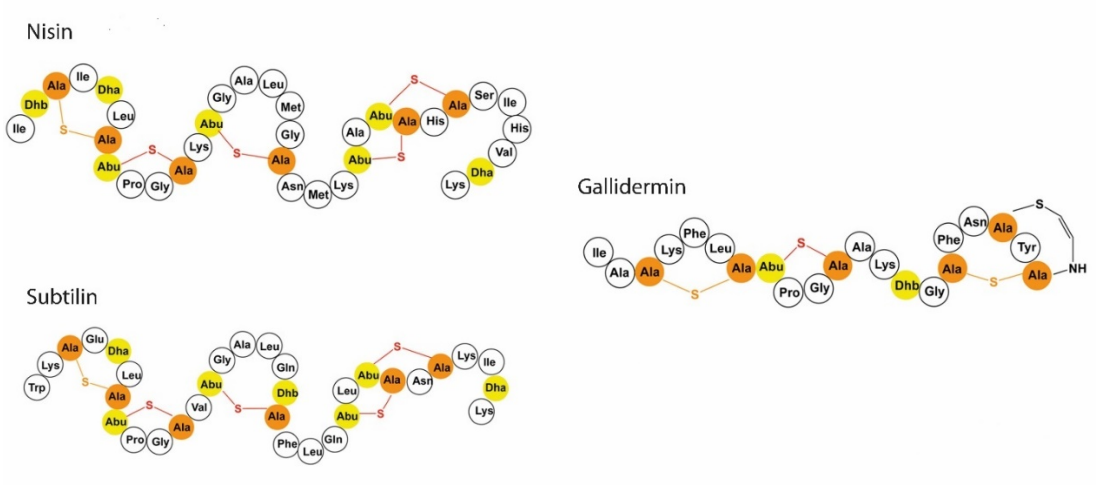


Figure 6. Schematic presentation of lantibiotics nisin, gallidermin and subtilin. Based on [46].

One of the most studied and best-characterized lantibiotics is nisin. Nisin is widely used in the food industry as a food preservative. It is a 34-amino-acid polypeptide with five internal ring structures (Figure 6). Nisin can be divided into three parts: N-terminal is three first lanthionine rings, C-terminal is two intertwined rings and six last amino acids. The flexible hinge region (Asn, Met, Lys) is located between the N-terminal and C-terminal. Nisin has two inhibition mechanisms that bind to peptidoglycan precursor lipid II and initiate pore formation in the cell wall [47]. The mode of action of nisin is shown in Figure 7; the picture was reproduced and simplified from Ref. [40]. Firstly, nisin reaches the bacterial plasma membrane; then, it binds to Lipid II with the help of two rings in the N-terminal; after that, the pore can be formed with an involved stable transmembrane orientation nisin: Lipid II complex. Nisin can binds to all available Lipid II molecules in the membrane to form pores. The pore complexes have a similar structure, including eight nisin molecules and four Lipid II molecules; these complexes are stable. The size of the pore is around 2-2.5 mm [48]. Another two lantibiotics, gallidermin and subtilin, have a similar mode of action [49, 50].

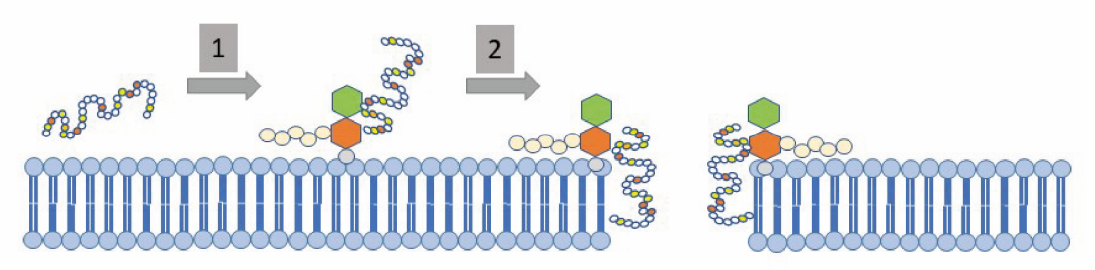


Figure 7. Mode of action of nisin. 1. Two nisin rings bind with lipid II. With the flexible hinge region, nisin can insert its C-terminal inside of the cytoplasmic membrane. 2. During or after assembly, four nisin-lipid II complexes and four additional nisin molecules can form pores.

These types of lantibiotics belong to the large class of cationic antimicrobial peptides (CAMPs). CAMP is an essential natural defence of most living organisms against bacteria. Bacteria that colonize mammalians and cause infections have a huge resistance mechanism to our immune system or, in other words, to CAMP. Much information about antibiotics resistance might be received by studying the bacterial resistance mechanisms against human defensins. One of the examples is the bacteria *Clostridium difficile* that has substantial resistance machinery against CAMP.

1.2.3. Resistance Machinery *Clostridium difficile* to CAMPs

Clostridium difficile (C. difficile) is a gram-positive anaerobe bacterium that enters the host as a dormant spore and germinates in the intestine. It produces different toxins and course diarrheal disease [51]. *C. difficile* infection primarily affects older people, hospitalised patients, and people after antibiotic therapy. People with typical flora and a robust immune system have a high chance to prevent *C. difficile* colonization in the body, but the resistance is lower after antimicrobial therapies [52]. Also, young and previously healthy people had *C. difficile* infection after close contact with infected patients [53]. Inside the body, the bacterium should defend itself against the host's innate immune system, such as cationic antimicrobial peptides (CAMPs) [54, 55]. CAMP production was found almost in all organisms, including bacteria, fungi, plants, and animals. The *C. difficile* has a huge line of defence; it can adapt to the presence of CAMP [56, 57]. The operon of the Cationic antimicrobial peptide resistance (Cpr) system from *C. difficile* consists of an ABC transporter (CprABC) and a two-component system (TCS).

1.2.3.1. ABC Transporter

ATP-binding cassette (ABC) transporters play a unique role in many biological processes. Especially proteins of ABC transporter located across the biological membranes and involved in the transport of substrates. Bacteria *C. difficile* contain an ABC transporter that helps transport or effluxes multiple antibiotics outside of the cell. It consists of four functional domains, two nucleotide-binding domains (NBDs) and two hydrophobic transmembrane domains (TMDs). The NBD is localised in the cytoplasmic face of the cellular membrane and is responsible for ATP binding and hydrolysing to generate the energy for transport [58]. TMD is crossing through the membrane. Figure 8 from Ref. [46] shows CprABC type resistance operon from *C. difficile*. TMD includes two transmembrane proteins (CprB and CprC) with equal stoichiometry, including six α-helices in each. NBD is presented by protein (CprA) that might be responsible for limiting antibiotic susceptibility [59].

TCS includes histidine kinase (HK), and a response regulator (RR) is located on the same gene cluster. The HK of the Cpr system (CprK) contains an extracellular loop that might play an essential role in sensing. As described earlier, the first two rings of lantibiotics such as nisin, gallidermin, and subtilin bind with the lipid II complex. Thus, these two rings might be recognised by CprK, resulting in the signalling of the TCS RR (CprR) [56]. After adding nisin, CprR can initiate the expression of CprABC [56, 57]. TCS and ABC transporter contribute the resistance of C. difficile to many lantibiotics, which is a big problem in treating this infection.

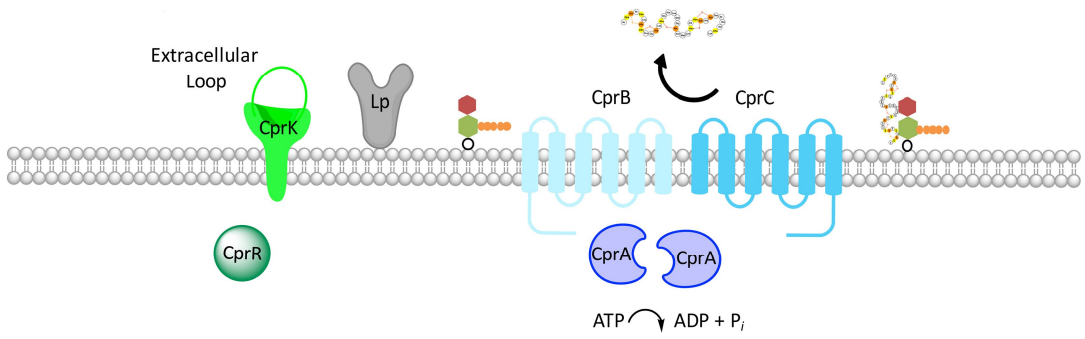


Figure 8. Schematic representation of the CprABC resistance system of *C.difficile*. CprABC type resistance operon is shown here with ABC transporter CprABC (blue colour) and two two-component systems (green colour). The picture shows the ABD transporter, two-component system, extracellular loop, lipoprotein (grey colour), lantibiotic nisin and lipid-protein II.

Also, close to cprABC transporter, a membrane-associated resistance protein (lipoprotein CD1348) is present. It is known that lipoprotein is not upregulated by the presence or absence of lantibiotics or CAMP and has a basal expression level [56].

1.2.3.2. lipoprotein

Lipoprotein is a resistance protein identified directly in front of the CprABC transporter from the bacteria *Clostridium difficile.* It was produced by the gene CD630_13480 [60]. Lipoprotein contains 178 amino acids and has a molecular weight of 20.2 kDa (calculated via EXPASY ProtParam). The original sequence (black) and his-tag version (red) are shown:

```
                      20                  40                  60
  1 MNKIAVSFLIIATTLLSTAC MDYSISAVELVDSKESAVVK KDEDAKEETTSKMINSKKTT
  1       HHHHHHHHHHSSGA MDYSISAVELVDSKESAVVK KDEDAKEETTSKMINSKKTT

                      80                 100                 120
 61 KIPIEIISKDEKIVKYLQID EESSLKDKLRLILDTLSNEY FNGLDMEVEVKEKDNLVKIN
 61 KIPIEIISKDEKIVKYLQID EESSLKDKLRLILDTLSNEY FNGLDMEVEVKEKDNLVKIN

                     140                 160
121 LIEPDKKSRVSWKDDYLNEQ NIIYTINNIIKNVIQEEDNS IWIEEVEIYYNGKLIELR
121 LIEPDKKSRVSWKDDYLNEQ NIIYTINNIIKNVIQEEDNS IWIEEVEIYYNGKLIELR
```

A crystallography test was performed for characterisation of the structure of the CD1348 lipoprotein. The yield and the quality were enough for crystallization, but the optimised plates do not show any crystals. Therefore, the overall shape of the protein was measured by Small-angle X-ray scattering (SAXS). The *ab initio* model is shown in Figure 9A from Ref. [60]. The model has a long flexible N-terminal domain and a compact C-terminal domain [60]. The theoretical model from computation

methods “TopModel” is shown in Figure 9B [61]. The model contains a long flexible N-terminal loop domain, five ß-sheets and two α-helices. The model will be observed in more detail in the discussion part of Chapter III.

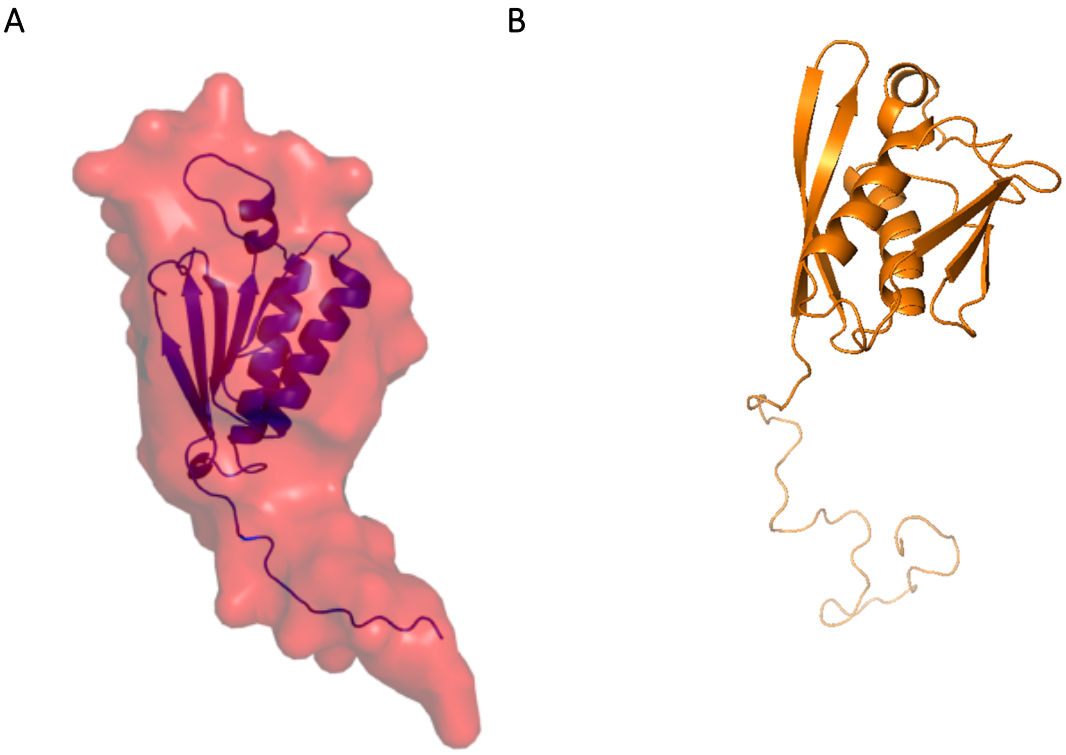


Figure 9. The 3D model of lipoprotein from the theoretical data. A is *ab initio* model from SAXS with fitted the theoretical calculation model inside. B is the theoretical model from TopModel.

Previously in these independent calculations of the lipoprotein CD1348 structure, both models are showing high similarity. The experimental data of the lipoprotein structure were not found in the literature.

2. Nuclear Magnetic Resonance

Nuclear Magnetic Resonance (NMR) spectroscopy is a powerful technique for characterising biomolecules at atomic resolution. Nowadays, NMR is widely used for studying the structure, properties, and dynamics of biological macromolecules. One of the significant advantages of solution NMR spectroscopy is data on molecular dynamics over timescales ranging from picoseconds to seconds.

NMR methods are based on the interaction of an external magnetic field with nuclei that have a magnetic moment. Nuclei are exposed to a certain radio frequency pulse, move to another energy level; after the pulse is turned off, they return to their original state while emitting electromagnetic radiation. As a result, an image of damped resonance oscillations is obtained from all nuclei. This free induction decay (FID) depends on the chemical composition and physical state of the analyte and the number of resonating nuclei.

2.1. Basic Principles of NMR

NMR is based on the fact that nuclei of atoms have magnetic properties that can be used to receive the information about the quantum mechanical property of a nucleus, their moving etc. NMR active nuclei are characterised by nuclear spin quantum number (*I*) is different from zero. Nuclei with nonzero spin angular momentum (^{1}H, ^{13}C, ^{31}P, ^{15}N, ^{19}F etc.) have nuclear magnetic moments. For protein NMR spectroscopy, ^{1}H, ^{13}C, and ^{15}N elements are most interesting because these isotopes have spin-½ (i.e., $I = ½$) and are visible in NMR. The natural abundance of ^{13}C (1.07 %) and ^{15}N (0.4 %) active isotopes is low. Thus, isotope labelling is necessary to maximize the sensitivity of these types of atoms.

The nucleus of spin quantum number *I* is allowed to accept only a fixed number of quantized energy levels. The spin can have 2*I*+1 angular momentum states from -*I* to *I*. Thus, spin ½ have only two possible energy levels -½ and +½. It is known as Zeeman levels. The quantum probability of a given nuclear spin adopting one of these two energy levels creates a nuclear magnetic moment (μ) proportional to the gyromagnetic ratio *γ*. Energies are presented by:

$$E = -\mu * B_0 \quad (1)$$

Where μ is the nuclear magnetic moment, B_0 is a presented magnetic field.

For a single nucleus with $I=½$ and positive *γ*, the magnetic energy is minimal when the spin angular momentum is parallel to B_0, and it has maximum energy when the spin angular momentum is antiparallel to the magnetic field. As spins normally have some initial orientation in respect to B_0, this creates a spin angular momentum that forces the spin precess around the magnetic field at the Larmor frequency ω_0:

$$\omega_0 = \gamma * B_0 \quad (2)$$

The Larmor frequency depends on the type of nucleus (gyromagnetic ratio) and the external static magnetic field B_0 value.

For the most used nuclei in biomolecular NMR the gyromagnetic ratios have the following values:

Gyromagnetic ratio for the most used nuclei in biomolecular NMR.

Nuclei	Gyromagnetic ratio (C* kg^{-1}) or (Hz/T)
^{1}H	$2.68 * 10^{8}$
^{15}N	$-2.71 * 10^{7}$
^{13}C	$6.73 * 10^{7}$

The observed resonance depends not only on the gyromagnetic ratio γ and external static magnetic field B_0 but also on the molecule's surroundings. Each proton is in a unique chemical environment; each will be exposed to a different magnetic field depending on the electron density distribution.

During an NMR experiment, the equilibrium of nuclear spin processes around B_0 can be disturbed by a radio frequency (RF) pulse. Changes of the spin orientation and relaxation back to equilibrium is detected as a signal in an NMR experiment. The signal is called the free induction decay (FID) and has the form of a damped sine wave. The observed FID is the sum of all FID for each nucleus. The FID is a function of time and might be transformed into a function of frequency with the help of Fourier transformation, which has individual peaks for each nucleus at the Larmor frequency. The scale used in NMR is chemical shift δ which can be defined as:

$$\delta = \frac{\nu - \nu_{ref}}{\nu_{ref}} \text{ x } 10^{-6} \qquad (3)$$

Where ν is the resonance frequency of the nucleus in the protein, ν_{ref} is the absolute resonance frequency of a reference substance. The chemical shift δ is defined in parts per million (ppm). The advantage of using δ is the fact that chemical shifts don't depend on B_0.

After applying an RF pulse, the spins return to equilibrium state through two different relaxation processes describing by FID. Longitudinal relaxation time T_1 corresponds to the process when the magnetisation returns to thermal equilibrium, after being perturbed by the RF pulse, along the direction of the static magnetic field B. Transversal relaxation time T_2 corresponds to the loss of coherence of the transverse nuclear spin magnetization. Relaxation times T_1 and T_2 are usually several seconds for small molecules, while for proteins, T_2 is measured in the order of milliseconds. Thus, the FID has a rapid decay, leading to broadening the linewidths in the NMR spectrum because T2 is inversely proportional to the linewidth. Large proteins have broader linewidth because their transverse relaxation rate R_2 ($1/T_2$) is proportional to the rotational correlation time τ_c. For more giant molecules, the correlation time is large and thus broader linewidth. Structures of proteins above 30 kDa are challenging to solve by solution NMR [62]. That is why some additional methods must be applied to increase the quality of spectra (section 2.5).

2.2. Chemical Shift Assignments

The assignment of resonances in the NMR spectra is one of the first steps in obtaining the protein structure and dynamic information. It is possible to do with multidimensional NMR [63]. The signals observed for correlated nuclei in an N-dimensional spectrum looks like cross-peaks. One of the most useful NMR experiments in protein NMR spectroscopy is the Heteronuclear Single Quantum Correlation (HSQC) [64], which provides a 'fingerprint' of the protein structure on a 2D plane. All amino acids, except Pro, give rise to a peak in the spectrum, including two additional peaks for Gln and Asn side chain NH_2 and one additional peak for each Trp indole group. 2D ^{1}H-^{15}N HSQC spectrum can provide a good way to assess the quality of the protein sample, folded or unfolded, to permit optimisation of some sample conditions (pH, temperature, ionic strength) and optimised acquisition parameters. It is important to identify which amino acid is represented by the given HSQC cross-peaks for further analysis. In application to proteins, multiple protons can absorb energy at the same frequency, and the signals are overlapping. For this purpose, multidimensional NMR experiments such as three-dimensional (3D) heteronuclear experiments, 4D [65] and even 5D [66] have been established because the chemical shift ranges of ^{15}N and ^{13}C are much broader. Multiple-resonance experiments have been developed to obtain sequence-specific chemical shift assignments, using interactions between spins of active nuclei. The interactions between spin-spin can be through bonds or space. Spin-spin interactions that occur via bonding electrons are called spin coupling, or J-coupling. It provides information on the chemical connectivity between atoms and dihedral angles according to the Karplus low [67]:

$$J(\phi) = A\cos^2\phi + B\cos\phi + C \tag{4}$$

Where $J(\phi)$ is the coupling constant of the *J*-coupling, and ϕ is the dihedral angle. The parameters *A, B* and *C* are dependent on the atoms and substituents attached to the bond of interest and are semi-empirically derived by correlating observed ^{3}J-coupling constants with corresponding dihedral angles measured in high-resolution structures. Typical coupling constant values in the proteins are shown in Figure 10.

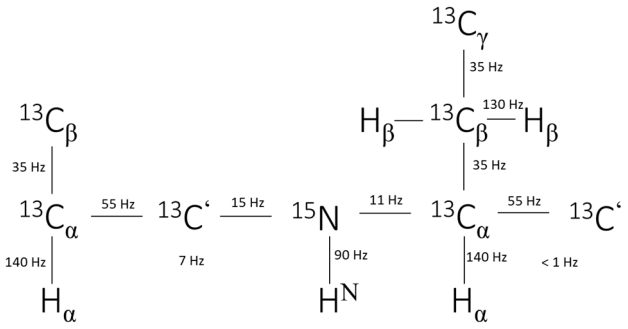


Figure 10. ^{1}J- and ^{2}J-coupling constants in peptides and proteins. Picture from [68].

J-coupling is widely used for correlation NMR signals of atoms that are chemically bound to each other, and this is a base principle for many multidimensional NMR experiments. Manipulation with

magnetisation by exploiting specific coupling constants allows measuring the signal from specific resonance. It is extremely useful for protein NMR assignment. Set of standard 2D or triple-resonance experiments (Figure 11), based on J-couplings over one or two bonds ($^{1}J/^{2}J$) of ^{13}C and ^{15}N isotopically labelled protein samples, are used to obtain backbone chemical shift assignments.

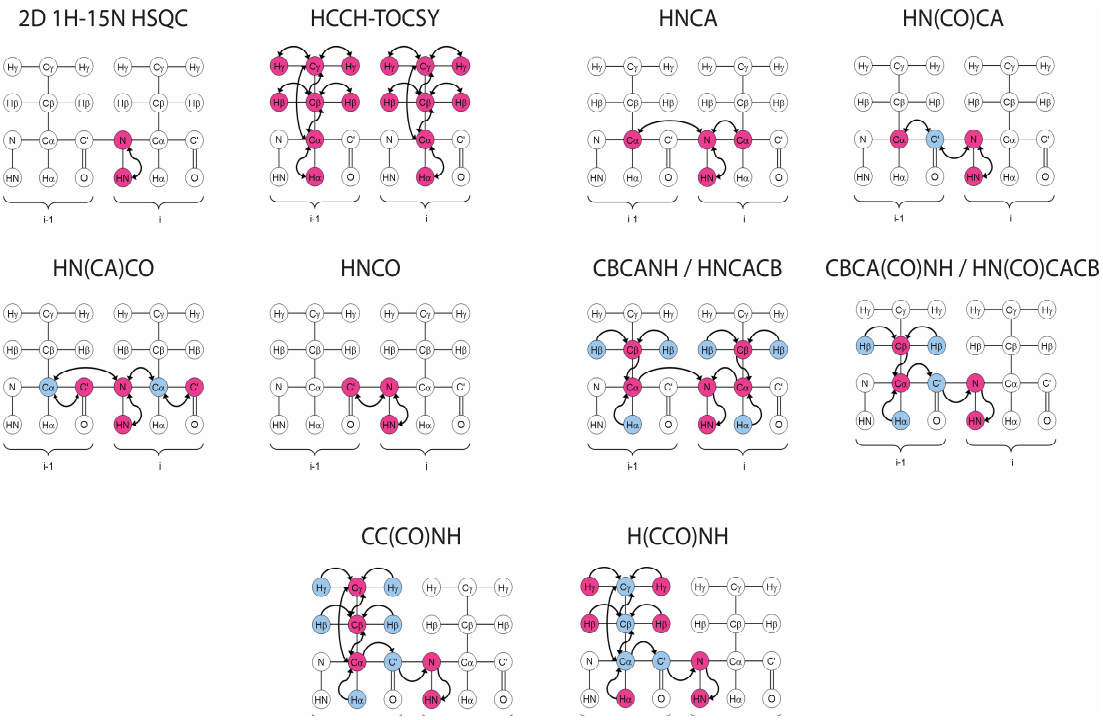


Figure 11. Schematic depiction of standard 2D and 3D NMR experiments for protein assignment. The diagrams show which nuclei are involved in each NMR experiment. Black arrows indicate the magnetisation transfer in each experiment. Nuclei coloured in pink are observed chemical shifts in the resulting spectra, whereas the blue colour of nuclei is only used for magnetisation transfer. Figure adapted from https://www.protein-nmr.org.uk.

Combining the triple-resonance experiment allows a sequential backbone assignment (Figure 12), where amino acids are connected by the exact positions of ^{13}C chemical shift on different ^{15}N planes and different proton chemical shifts.

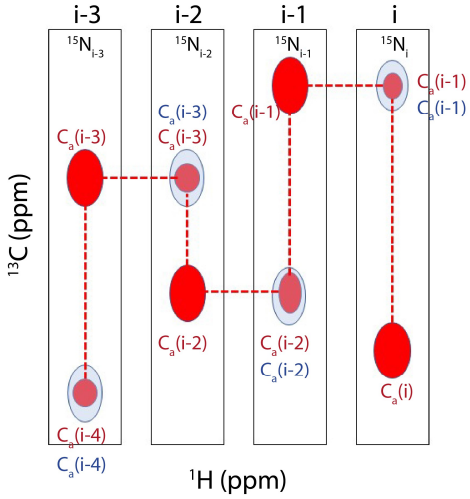


Figure 12. A general example of a sequential backbone assignment. Four 15N-strips from a combination of two triple resonance experiments, HNCA (red) and HN(CO)CA (blue), showing the $^{13}C_{\alpha}$ of residue (i) and receding residue (i-1). The resonance for the HN(CO)CA experiment represents only one peak for a previous amino acid.

Another common spin-spin interaction is through space and is called the nuclear Overhauser effect (NOE). NOE allows estimate distances between not directly covalently attached nuclei (usually at to 5 Å). The obtained information about distance restraints has a wide usage for structure determination of proteins.

2.3. Structure Calculation of Proteins Using NMR

The structure and dynamics of proteins are essential for understanding their function. Knowing the structure is a central aspect of structural biology. One of the most popular methods for determining protein structures is X-ray crystallography, solution and solid-state NMR spectroscopy, and cryo-electron microscopy. With the help of solution NMR spectroscopy has been solved of more than 13000 proteins three-dimensional (3D) structures in the Protein Data Bank (www.rcsb.org). The solution NMR has a limit to the size of the proteins; that is why structures above 30 kDa are challenging to solve by NMR. X-ray crystallography can be used for larger proteins or complexes, but the NMR method is more useful when the protein of interest has resisted attempts at crystallization, which is a common problem for many membrane proteins.

The NMR structure determination of a protein is a lengthy process because it involves some steps of preparation as uniformly $^{13}C/^{15}N$-labeling of the protein, the acquisition of a set of 2D, 3D or 4D NMR experiments, the data processing, all chemical shift assignment (see section 2.2), NOE assignment and collection of conformational restraints, structure calculation, refinement, and validation [69].

After the assignment of all individual nuclei, backbone torsion angles (ϕ,φ) and the χ1 torsion angle (Figure 13) of side chains can be easily predicted based on ^{1}HN, ^{15}N, $^{13}C\alpha$, $^{1}H\alpha$ and $^{13}C'$ chemical shifts, using a program such as TALOS-N [70]. Also, based on the Karplus low (equation 4), the $^{3}J_{HNH\alpha}$, $^{3}J_{NH\beta2/3}$ and the $^{3}J_{H\alpha H\beta2/3}$ couplings can be used to determine the ϕ-angle of the protein backbone and the certain side-chain χ1 torsion angle [71].

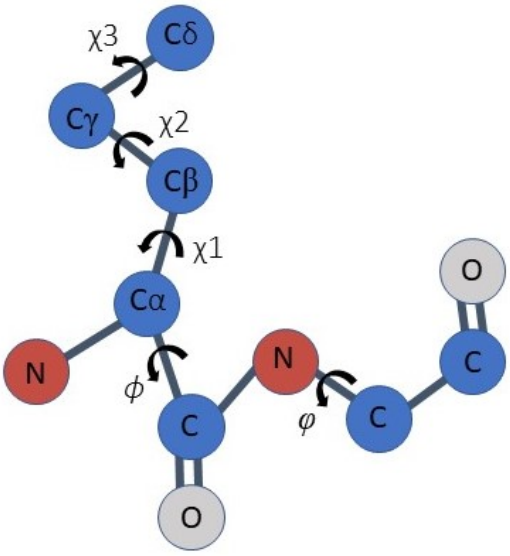


Figure 13. Backbone torsion angles (ϕ,φ) and the $\chi 1,2,3$ torsion angles of the amino acids.

The primary source for collecting conformational restraints from NMR data is derived from the nuclear Overhauser effect (NOE). NOE is a cross-relaxation mechanism between two spins through space interaction, generally less than 5Å. This interaction is completely independent of the nature of the covalent binding between the two interacting spins [72]. NOE data are obtained from Nuclear Overhauser Enhancement SpectroscopY (NOESY) experiments. In multidimensional NOESY experiments, the resulting NOEs of cross-relaxation between two nuclei are observed as cross-peaks correlating the chemical shifts of the two involved nuclei. The intensity of the cross-peaks connecting protons depends mainly on the distance between them, and it is proportional to the inverse sixth power of the internuclear distance r^{-6}. Figure 14 shows the example of using NOE for structure calculation.

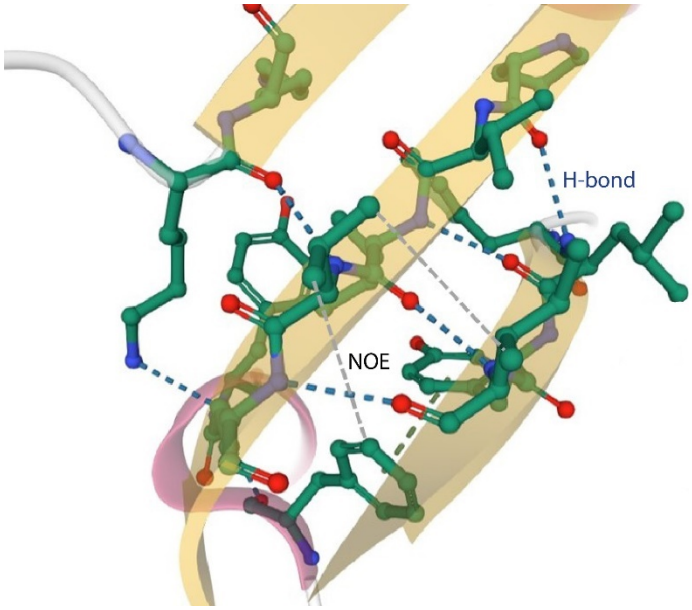


Figure 14. Illustration of structural restraints with hydrogen bond and NOE. The figure shows the H-bond (blue dotted lines) between H-O and NOE (grey dotted lines).

The intensity of cross-peaks depends on the distance between and can be used to measure the internuclear distance between nuclei in a protein. The recorded NOE intensities can be converted to NOE distance restraints, which are the most helpful information for calculating protein structure. NOE derived interproton distance restraints together with hydrogen bond information, and torsion angle restraints are used for studying globular proteins.

2.4. Using NMR to Study Dynamics in Proteins

2.4.1. Integrated NMR for studying dynamics in proteins

If the structure of proteins is well known, the next step is understanding proteins function. The motion of each protein is induvial and can bring much information about its conformational dynamics. Some physical methods such as NMR, polarization-resolved fluorescence spectroscopy, and molecular dynamics (MD) simulations are widely used to study the protein dynamics on multiple time scales. NMR is a perfect technique for studying protein dynamics on the different time scales from pico- to seconds. Figure 15 shows the timescale of different protein dynamical processes and NMR

experiments used for characterization of them. Fast protein dynamic describes the movement of loops, side-chain rotations and local atomic vibration on the picosecond to low nanosecond range. NMR experiments for studying fast protein dynamics are longitudinal spin relaxation R_1, transverse spin relaxation R_2, and the heteronuclear Overhauser enhancement (NOE) experiments. By contrast, the conformational changes such as protein folding, chemical exchange, ligand interaction, enzyme catalysis and allosteric regulation process on the microseconds to seconds range. It is covered by rotating frame relaxation experiments $R_{1\rho}$, Carr-Purcell-Meiboom-Gill (CPMG), Paramagnetic Relaxation Enhancement (PRE), and chemical exchange saturation transfer [73-75].

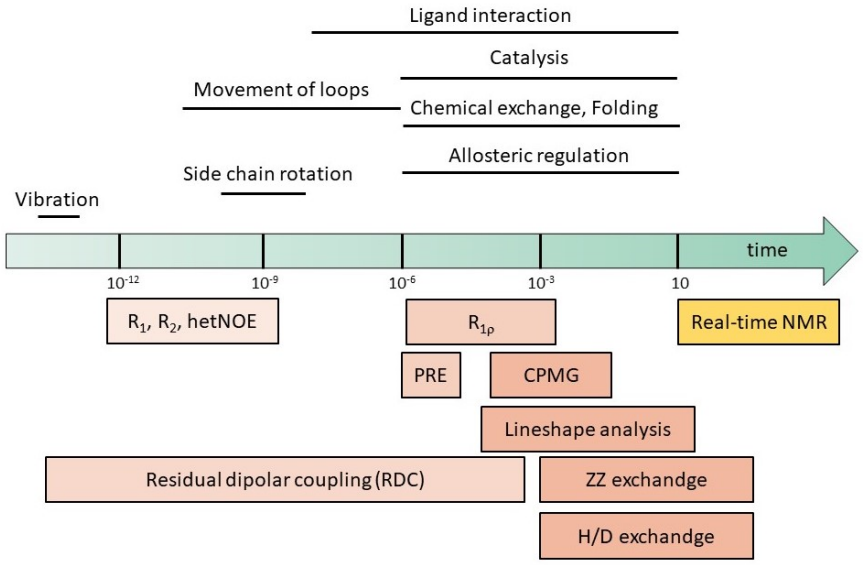


Figure 15. The timescale of different dynamic processes in proteins and NMR experiments corresponding to them. Protein dynamics range from pico- to seconds. The picture is adapted from Ref. [76, 77]

The NNR data helps to represent the parameters that can describe the dynamic with the help of global rotational diffusion and local flexibility (order parameters) that are sensitive to inter-and intramolecular interactions, respectively [78].

2.4.2. Protein dynamic on the pico- to nanosecond time scale.

The backbone dynamics can be described with the help of longitudinal spin relaxation rate R_1, transverse spin relaxation rate R_2, and the hetNOE data measured for each available residue in the protein.

Applying radiofrequency pulses generate perturbs nuclear spins from their equilibrium state. The process by which the spins return to their equilibrium is called relaxation. The relaxation mechanisms are based on the dipolar coupling between two nuclei and the chemical shift anisotropy (CSA), which depends on the orientation of the spin to the external magnetic field. Collisions with solvent molecules lead to translational and rotational diffusion of the protein. Internal and global motion are independent processes; the behaviour of the random fluctuation can be described by the autocorrelation function, which mathematical interpretation function that connects both internal and overall motion is the time autocorrelation function as

$$C_{total}(\tau) = C_i(\tau) \cdot C_g(\tau) \quad (5)$$

Here, $C_i(\tau)$ is an internal autocorrelation function, and $C_g(\tau)$ is a global autocorrelation function.

Two model-independent quantities can describe the fast internal motion: order parameter (S^2), which measures the amplitude of the motion, and an effective correlation time τ_e. The simplest isotropic formalism might be applied for small molecules, simple polymers, and data obtained from one-dimensional NMR measurements of proteins. Since the measurable relaxation parameters are more easily understood regarding the probabilities of motions at specific frequencies rather than times, the autocorrelation function C(τ) can be Fourier transformed. Thus, C(τ) can be described as the spectral density function J(ω):

$$J(\omega) = \frac{2}{5}\left(S^2 \frac{\tau_c}{1+\omega^2 {\tau_c}^2} + (1-S^2)\frac{\tau}{1+\omega^2\tau^2}\right) \quad (6)$$

Which corresponds to an internal correlation function of

$$C_i(\tau) = S^2 + (1-S^2)e^{-\frac{\tau}{\tau_e}} \quad (7)$$

Where S^2 is generalized order parameter, τ_c is the overall isotropic rotational correlation time of the molecule, $\tau=\tau_c\tau_e/(\tau_c+\tau_e)$ where τ_e is a single effective correlation time describing the internal motion.

And correspond the global correlation time:

$$C_g(\tau) = \frac{1}{5}e^{-\frac{\tau}{\tau_e}} \quad (8)$$

For studying backbone protein dynamics, usually, the ^{15}N relaxation rate (R_1), the ^{15}N relaxation rate (R_2), and the heteronuclear nuclear Overhauser effect ($\{^1H\}$-^{15}N NOE) are used. All these parameters are typically measured using 2D ^1H,^{15}N HSQC experiments in which the intensities of peaks are modulated as a function of a time delay placed at a point in the sequence when the relevant relaxation process is active [79].

The relaxation rates of ^{15}N R_1 and R_2 and hetNOE depend on the spectral density function J(w) in the following manner [80]:

$$R_1 = \frac{1}{T_1} = \frac{1}{4}d^2[J(\omega_H-\omega_N) + 3J(\omega_N) + 6J(\omega_H+\omega_N)] + c^2J(\omega_N) \quad (9)$$

$$R_2 = \frac{1}{T_2} = \frac{1}{8}d^2[4J(0) + J(\omega_H-\omega_N) + 3J(\omega_N) + 6J(\omega_H-\omega_N) + 6J(\omega_H)] + \frac{c^2}{6}[4J(0)] + 3J(\omega_N) + R_{ex} \quad (10)$$

$$hetNOE = 1 + \frac{1}{4}d^2\frac{\gamma_H}{\gamma_N}\left[\frac{6J(\omega_H+\omega_N) - J(\omega_H-\omega_N)}{R_1}\right] \quad (11)$$

where ω_H and ω_N are angular Larmor frequencies for ^1H and ^{15}N spins, respectively, R_{ex} is the relaxation rate due to chemical exchange, c is the Chemical shift anisotropy (CSA), and d is dipole-dipole (DD) constants are defined as:

$$c = \frac{1}{3}\omega_N{}^2(\sigma_{\parallel} - \sigma_{\perp})^2 \quad (12)$$

$$d = \frac{1}{4}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_H{}^2\gamma_N{}^2\left(\frac{h}{2\pi}\right)^2 \langle r_{NH}{}^{-3}\rangle^2 \quad (13)$$

Here, h is Planck's constant, γ_H and γ_N are the gyromagnetic ratios of ^{1}H and ^{15}N nuclei, μ_0 is the permeability of free space, and r_{NH} is the internuclear distance between N-H bonds ≈ 1.02 Å. $\sigma_{\parallel}$ and $\sigma_{\perp}$ are the parallel and perpendicular components of the axially symmetric ^{15}N chemical shift tensor CSA = - 172 ppm which is assumed to be coaxial in a first approximation concerning the dipolar interaction.

In the "original" model-free formalism (equation 6), the overall motions are unrestricted with the correlation function that decays exponentially to zero with a single characteristic time scale τ_m. Internal motions are assumed to be restricted because the correlation function for internal motion decays faster. Then the resulting total correlation function (equation 5) in this model has a double exponent with the fast and slow phases representing internal and global motion. The amplitude of the global (slow) phase, that is, the plateau value divided by the initial value C(0), is characterized as the square of the order parameter (S^2) and represents the degree of spatial restriction of the backbone H-N internal motions. The values of S^2 have some limited cases - completely restricted motions described only by the global movement have S^2=1, and completely unrestricted motions, which are described only by the internal movement, have S^2=0. Thus, this parameter helps to explain the dynamic in values.

In the case of only isotropic tumbling, "simplified" model-free formalism from equation 6 might be applied. Here is the internal motion of a bond vector is extremely fast in comparison to the overall tumbling; thus, equation 6 reduces to the form:

$$J(\omega) = \frac{2}{5}\left(S^2\frac{\tau_c}{1+\omega^2\tau_c{}^2}\right) \quad (14)$$

With the help of this model, relaxation data still can be fit to find S^2 and the overall isotropic rotational correlation time of the molecule.

2.5. Extending the Size limitation of NMR

Because of the rotational tumbling of the protein, the application of solution NMR spectroscopy is limited by the size of the studied molecules or complexes (<~30 kDa) [62]. The signal-to-noise ratio is low for high molecular weight specimens due to fast transverse nuclear spin relaxation during the evolution and recording periods. Labelling with deuterium ^{2}H, meaning the replacement of side-chain protons by deuterium, reduces dipolar coupling interactions (because of the lower gyromagnetic ratio of ^{2}H *vs* ^{1}H) and allows to study also larger molecules.

Relaxation T_2 is a dephasing mechanism of the same type of spins over time. The number of spins and couplings in bigger molecules is large, and they evolve in different ways. The bigger system has a broader line width. To increase the resolution and sensitivity of NMR spectra of big molecules, Transverse Relaxation-Optimized SpectroscopY–(TROSY) was used. TROSY takes advantage of the

interference between DD coupling and the CSA in the relaxation of coupled heteronuclear spins to produce narrow resonance lines at high magnetic field strengths. This technique can be applied to two- three-dimensional HSQC, triple resonances, and relaxation experiments. In a non-decoupled $^{1}H,^{15}N$ HSQC spectrum, each correlation for both dimensions appears as four peaks because the average signal is split into both dimensions due to $^{1}J_{HN}$ coupling. The four components of this multiplet are not submitted to the same relaxation rates and have different intensities. When DD and CSA contributions to the relaxation cancel each other out through destructive interferences of their oscillations, only the slowest relaxation can be seeing. This signal has the smallest line width and highest intensity. Figure 16 shows the line with difference between of coupled [$^{1}H,^{15}N$] HSQC spectrum (A), decoupled [$^{1}H,^{15}N$] HSQC spectrum (B) and [$^{1}H,^{15}N$] TROSY-HSQC spectrum (C). The picture was reproduced from Ref [81].

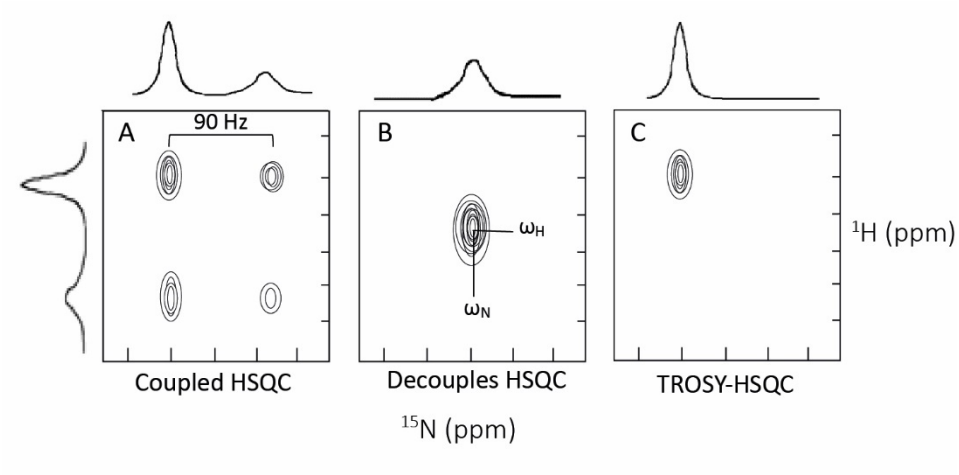


Figure 16. Illustrates the increased resolution in a two-dimensional nitrogen-proton HSQC coupled experiment using the TROSY technique. Panel A shows a region of coupled HSQC spectrum that contains four peaks associated with a single NH group. Every peak is a distinct pair of single-quantum nitrogen and proton transitions. The intensities and line widths vary due to the different relaxation rates of each transition. If decoupling occurs during proton and nitrogen evolution periods, only a single peak will be observed (panel B). The line widths in both directions are the average of the relaxation rates for each transition. Panel C shows the TROSY-HSQC spectrum of the same NH group, where only the most slowly relaxing peak is shown. The remaining three components of the quartet been removed by phase cycling in the pulse sequence. For large proteins or complexes, TROSY spectra have a higher signal-to-noise ratio than in a standard HSQC spectrum because only the most slowly decaying transition is used in the polarization transfer step.

In triple-resonance experiments, the TROSY principle can efficiently reduce relaxation during different coherence transfer steps, e.g. from ^{1}H to ^{15}N, from ^{15}N to $^{13}C\alpha$ or ^{13}CO, which have long transfer periods due to the small $^{1}J(^{15}N, ^{13}C\alpha)$- and $1J(^{15}N, ^{13}CO)$-coupling constants.

3. Aims

The membrane GABARAP protein plays an essential role in the autophagy process. It is involved in autophagosome formation during oligomerization. To this end, it can be anchored to the autophagosomal membrane system. Availability of membrane can help to understand the function of "open" conformation of GABARAP structure. This alternate conformation of the N-terminal region can be important for oligomerization and hemifusion. The aim of this work was a characterization of the dynamics of GABARAP anchored to a lipid membrane by NMR spectroscopy. As a suitable membrane mimic, phospholipids nanodiscs were used. For the better quality of the studied spectra, deuterated GABARAP should be produced. The production of this type of sample is costly; thus, the ratio between protein yield and the price of the sample should be optimized. The dynamics of membrane-anchored GABARAP should be studied by ^{15}N relaxation NMR experiments. These experimental results were supposed to be compared with free GABARAP in solution. The goal was to produce deuterated GABARAP anchored to lipid nanodiscs in a suitable concentration, then study structural details, dynamics, and functional relevance with a membrane by solution NMR methods.

Lipoprotein CD1348, unlike GABARAP, is the new uncharacterized protein founded in the resistance machinery bacteria *C.difficile*. This system has resistance against different lantibiotics. Knowing the structure helps to understand the function and role of the presence of the lipoprotein in the resistance system. The first aim of this work was the assignment of all amino acid resonances and determination of their chemical shifts on the 2D, 3D and 4D NMR spectra. Then the calculation of the tertiary structure of lipoprotein with the help of all chemical shift assignments, NOE assignment and collection of conformational restraints. The second aim was the checking interaction between lipoprotein and lantibiotics by titration NMR experiments. If there is an interaction, which parts of the lipoprotein are involved in it.

Chapter II GABARAP

Elucidation of the Dynamics of the Autophagosomal Membrane-Associated Protein GABARAP by Solution NMR

This chapter is devoted to the GABARAP protein project. GABARAP is a membrane-associated protein that plays an essential role in the elongation and maturation of autophagosomes during their formation. Studying the membrane protein together with the membrane is one of the aims of this work. Laboratory methods like transformation DNA, expression and purification of proteins, isotope labelling, and NMR sample preparation are presented. The extension of the existing methods for full deuteration of protein and then anchoring to nanodiscs are described. Also, analytical methods and NMR techniques used for the sample characterization are discussed. The results part includes expression and purification of TEV protease and belt protein MSP1D1Δ5 for nanodisc production. Also, the expression and purification of fully deuterated GABARAP and GABARAP mutant data. Due to the high price of producing a deuterated sample, two expression methods of fully deuterated GABARAP with high yield are compared. The difference in chemical shifts on NMR spectra of free GABARAP and GABARAP anchored to nanodisc is presented. Also, some preliminary data of the model-free analysis of the ^{15}N relaxation data of GABARAP-nanodisc samples are described and discussed here.

1. Materials

1.1. Chemicals, enzymes, and media components

Chemicals, enzymes, and components for the cell's growth media used in this work correspond to purity grade pro analysis (p.a.). All components were dependent on the requirement experiments and listed in Table 1 and Table 2. Standard chemicals were usually purchased from AppliChem GmbH (Darmstadt, Germany), Carl Roth (Karlsruhe, Germany), Merck KGaA (Darmstadt, Germany), Serva (Heidelberg, Germany) and Sigma-Aldrich (St. Louis, MI, USA).

Table 1: Chemicals used in this work.

Chemical	Manufacturer/Distributor
Agarose	Serva (Heidelberg, Germany)
complete protease inhibitor cocktail	Roche (Basel, Switzerland)
D_2O	Cambridge Isotope (Andover, USA)
Dithiothreitol (DTT)	AppliChem (Darmstadt, Germany)
EDTA-disodium salt	AppliChem (Darmstadt, Germany)
IPTG	UBPbio (Aurora, USA)
TCEP	Sigma-Aldrich (St. Louis, MI, USA)
DMPC	Avanti Polar Lipids (Alabaster, USA)
MPB-PE	Avanti Polar Lipids (Alabaster, USA)
^{15}N-Ammonium chloride	Cambridge Isotope (Andover, USA)
D-Glucose-$^{13}C_6$	Sigma-Aldrich (Steinheim, Germany)
Unstained protein marker	Thermo Fisher Scientific (Waltham, USA)

Table 2: Enzymes used in this work.

Enzyme	Manufacturer/Distributor
DNAse I	AppliChem (Darmstadt, Germany)
DNAse A	AppliChem (Darmstadt, Germany)
Lysozyme	AppliChem (Darmstadt, Germany)

1.2. Bacteria cultures and plasmids

The bacterial strain *E. coli* BL21 (DE3) T1 was used for DNA amplification and recombinant protein expression of GABARAP and MSP1D1Δ5. Another type of strain, *E. coli* BL21 DE3 pRARE2 was used for His-TEV protease expression. Strains and plasmids are listed in Table 3 and Table 4.

Table 3: Bacterial strains used for recombinant protein expression.

Strain	Genotype	Reference/Source
BL21 (DE3)	F^- ompT hsdSB(rB^-mB^-) gal dcm (DE3)	Novagen, Darmstadt

Table 4: Plasmids used for recombinant gene expression.

Plasmids	Resistance	Property	Reference/Source
pET11a_GABARAP	Ampicillin	Expression of GABARAP	Life Technologies, Carlsbad, USA
pET28a_MSP1D1Δ5	Kanamycin	Expression of MSP1D1Δ5	AG Wagner, Harvard Medical School
pRK792_TEV_S219P	Ampicillin/ Chloramphenicol	Expression of His-TEV protease	AG Wagner, Harvard Medical School

1.3. Laboratory equipment

All necessary devices for producing the NMR sample at the biological S2 laboratory are described here.

Table 5: Laboratory equipment used in this work.

Equipment	Manufacturer/Distributor
ÄKTA purifier	GE Healthcare (Chicago, USA)
Branson Sonifier 250	Branson Ultrasonics (Danbury, USA)
Centrifuges 5417R, 5702R, 5804R	Eppendorf (Hamburg, Germany)
Centrifuge Avanti J-20 XP and Rotors	Beckman Coulter (Brea, USA)
Gel Doc XR and gel documentation system	Biorad (Hercules, USA)
Spectrophotometer UV 1800	Shimadzu (Kyoto, Japan)
Lambda 25 UV/Vis spectrophotometer	PerkinElmer (Waltham, USA)
Spectrophotometer (measure OD)	Schott Instruments (Mainz, Germany)

1.4. Column chromatography

There is a list of prepacked columns for high-resolution purification of large biomolecules and protein complexes used with ÄKTA protein purification systems. Also, nickel-charged affinity resin Ni-NTA Agarose was used to purify recombinant proteins containing a polyhistidine (6xHis) sequence.

Table 6: Columns and resins used for column chromatography.

Column/resin	Manufacturer
Ni(2+)-NTA-Agarose	Qiagen (Hilden, Germany)
HiLoad 16/10 SP Sepharose HP	GE Healthcare (Freiburg, Germany)
HiLoad 16/600 Superdex 75 pg	GE Healthcare (Freiburg, Germany)
HiLoad 16/600 Superdex 200 pg	GE Healthcare (Freiburg, Germany)

1.5. NMR spectrometers

Liquid state NMR spectrometers operating at different magnetic fields were used for studying dynamics and chemical shifts changes.

Table 7: NMR spectrometers used for the GABARAP project.

Spectrometers	Manufacturer
Bruker Avance III HD NMR (600 MHz)	Bruker, Billerrica, USA
Bruker Avance III HD NMR (600 MHz)	Bruker, Billerrica, USA
Bruker Avance III HD NMR (700 MHz)	Bruker, Billerrica, USA
Variant (800 MHz)	Agilent (Varian), Santa Clara, USA
Bruker Avance III HD NMR (900 MHz)	Bruker, Billerrica, USA

1.6. Software and databases

The following software and databases were used to evaluate, analyse, and visualise GABARAP protein data.

Table 8: Software and databases used for the GABARAP project.

Software/database	Usage	Reference/distribution
BMRB Databank	Open databank of chemical shifts of biomolecules	http://www.bmrb.wisc.edu [82]
Expasy-ProtParam	Analysis of proteins based on their primary sequence	http://web.expasy.org/protparam [83]
NMRPipe	Processing of NMR spectra	[84]
NMRDraw	Visualisation of processed MNR spectra	[84]
NMRViewJ 8.0.3	Visualisation and analysis of NMR spectra	[85]
MUNIN	Determination of peak intensities from multi-dimensional spectra	[86, 87]
Matlab	Evaluation and analysis of relaxation data	
PDB Database	Structures of biological macromolecules	http://rcsb.org
PyMOL	Visualisation of protein structures	
RasMol 2.7.5	Visualisation of protein structures	[88]
TENSOR2	Fast protein dynamics from relaxation data	[89]
TopSpin	Basic software of Bruker NMR spectrometers	
VnmrJ	Basic software of Bruker NMR spectrometers	
Adobe Illustrator CS5	Generation of figures	
Origin 2017	Chemical shifts calculation and visualisation of relaxation data	

2. Methods

2.1. Microbiological methods

2.1.1. Bacterial growth media

The solid bacterial growth media allows bacteria to grow as colonies on a Petri dish. It is based on the jelly-like substance agar. Required antibiotics for different plasmids were added to the medium-agar mixture when the agar was in a liquid state at a temperature below 60°C.

Liquid growth media were used for protein expression. In this work, there are two different liquid bacterial growth media: LB-medium and M9-mediums (Table 9). M9 medium was used for uniform ^{15}N, ^{13}C-Labeling and deuteration of the protein. The amount of glucose and ammonium chloride for M9+ medium was calculated for growing cells until OD~4. Both mediums were prepared using ultrapure normal or deuterated water. For the M9 medium, in addition, TS2 trace element solution and a vitamin cocktail for proper growth of bacteria were prepared (Table 10). Both media were sterilised in an autoclave at 121°C and then stored at room temperature. To avoid adding non-sterile components to the growth media, vitamin cocktail, antibiotics and TS2 solution were filtered with the help of 0.22 µm pore diameter membrane filters. Before usage, the media were heated to 37°C for the best cell's adaptation.

Table 9: Composition of bacterial growth media for recombinant gene expression.

Regular M9-medium H_2O or D_2O	M9+-medium H_2O or D_2O
NaH_2PO_4 9.1 g/L	K_2HPO_4 19 g/L
KH_2PO_4 3 g/L	KH_2PO_4 5 g/L
NaCl 0.5 g/L	Na_2HPO_4 9 g/L
NH_4Cl 1 g/L	K_2SO_4 2.4 g/L
$CaCl_2$ 100 µM	D-Glucose-$^{13}C_6$ 14 g/L
$MgSO_4$ 2mM	NH_4Cl 4 g/L
Fe(III)-Cl 10 µM	Trace element solution 2 ml/L
D-Glucose-$^{13}C_6$ 2.5 g/L	Vitamines cocktail 1 ml/L
Trace element solution 2 ml/L	$MgCl_2$ 0.95 g/L
Thiaminhydrochlorid 5 mg/ml	Ampicillin 200 µg/ml
Vitamines cocktail 1 ml/L	pH=6.5
$MgCl_2$ 0.95 g/L	
Ampicillin 200 µg/ml	
LB-medium H_2O or D_2O	
Tryptone 10 g/L	
NaCl 10 g/L, Yeast extract 5 g/L	

Table 10: Composition of TS2 solution and vitamin cocktail solution used in M9 medium.

TS2 trace element solution	Vitamin cocktail solution
$MnCl_2 \cdot 4\ H_2O$ 30 mg/L	D-biotin 1 g/L
$ZnSO_4 \cdot 7\ H_2O$ 100 mg/L	Choline chloride 1 g/L
H_3BO_3 300 mg/L	Folic acid 1 g/L
$CoCl_2 \cdot 6\ H_2O$ 200 mg/L	Nicotinamide 1 g/L
$NiCl_2 \cdot 6\ H_2O$ 20 mg/L	Sodium-D-pantothenate 1 g/L
$CuCl_2 \cdot 2\ H_2O$ 10 mg/L	Pyridoxal hydrochloride 1 g/L
$Na_2MoO_4 \cdot 2\ H_2O$ 900 mg/L	Riboflavin 0.1 g/L
Na_2SeO_3 20 mg/L	dissolved in 20 mM phosphate buffer, pH 7
dissolved in H_2O	

2.1.2. Transformation in *E.coli* with plasmid DNA

Bacterium transformation of MSP1D1Δ5

1 µl PET28a_MSP1D1Δ5 plasmid vector with kanamycin resistance was transformed into *E.coli* BL21(DE3)T1 cells. The plasmid was added to competent cells and gently mixed. Then cells with plasmid were incubated in ice for 30 minutes. The cells were heat-shocked for 45 sec at 42°C to create a pressure difference between the outside and the inside parts of the cell, which induces the formation of pores through which plasmid DNA can enter. Then the Eppendorf was placed on ice for 5 min for maximum transformation efficiency after heat shock. 600 µl LB medium was added to cells and incubated for 1 hour at 37°C. Cells were spread on an agar plate with Kanamycin (25 µg/ml) to grow cells that contain the antibiotic resistance vector. The plate was incubated overnight at 37°C.

Bacterium transformation of His-TEV-Protease

1 µl pRK792 (TEVS219P) plasmid were resuspended in 100 µl *E.coli* BL21DE3pRARE2. The cells were incubated on ice for 30 minutes, exposed to a heat shock at 42°C for 45 seconds, and then placed on ice for 2 minutes. 900 µl LB was added, and the cells were incubated for 1 hour at 37°C. The cells were spread on an agar plate with 100 µg/ml Ampicillin and 34 µg/ml Chloramphenicol. The plate was incubated overnight at 37°C.

Bacterium transformation of GABARAP and GABARAP-G116CΔ117

GABARAP or GABARAP-G116CΔ117 mutant pET11a vectors with Ampicillin resistance were transformed into BL21(DE3)T1 *E.coli* cells. 0.2 µl of the plasmid with concentration 0.1 µg/µl was added to 50 µl competent cells and gently mixed. The aliquots with plasmid were incubated in ice for 15-30 minutes. The cells were exposed to a heat shock for 45 seconds at 42°C; then, the Eppendorf was placed immediately on ice for at least 5 min. 600 µl LB medium based on ultrapure water was added, and the cells were incubated for 1 hour at 37°C. Cells were spread on an agar-agar plate with

ampicillin (200 µg/ml) to grow only cells that contain the antibiotic resistance vector. The plate was incubated overnight at 37°C.

2.1.3. Expression of proteins

Expression of MSP1D1Δ5 and His-TEV-Protease

His-TEV-Protease was made for future uses of another protein purification (MSP1D1Δ5 in this work). For the expression of His-TEV-Protease, the protocol [90] was used. For the initial trial expression, the separate colonies of His-TEV-Protease from the agar plate were picked and placed into a 200 ml LB medium with 1% Glucose, 100 µg/ml Ampicillin and 34 µg/ml Chloramphenicol. The picked cells were grown at 37°C with shaking at 130 rpm. After 14 hours, the cell culture was diluted and grown in 1 litre LB medium containing antibiotics (100 µg/ml Ampicillin and 34 µg/ml Chloramphenicol) at 37°C, with shaking at 130 rpm. When the optical density at 600 nm (OD_{600}) reached ~0.6, 1 mM of IPTG was added to induce the protein expression. The cell culture was incubated at 30°C for 4 hours. During induction, the cells use most of their resources to produce the target protein and afterwards do not grow anymore. After induction, cells were harvested by centrifugation at 5,000 g for 20 minutes. The resulting pellets were stored at -20°C.

For the expression of His-MSP1D1Δ5, a well-established protocol was used [91]. First, freshly transformed BL21(DE3)T1 cells with the plasmid were grown overnight in 60 ml LB medium supplemented with 30 mg/ml Kanamycin at 37°C with constant shaking at 130 rpm. The overnight culture was diluted in 3 litres fresh LB medium containing antibiotic Kanamycin (30 mg/ml) and then incubated at 37°C for 2 hours. When OD_{600} reached ~0.7, protein expression was induced by the addition of 1 mM IPTG. After 4 hours at 37°C, *E.coli* were harvested by centrifugation at 5,000 g for 20 minutes. The pellets were stored at -20°C until further use.

High-yield expression of deuterated and labelled GABARAP and GABARAP-G116CΔ117

The big molecules have large linewidths on the NMR spectrum due to slow tumbling and the spectral overlap from a large number of unique signals. As the final system of GABARAP anchored to nanodisc is quite big for solution NMR experiments, fully deuterated GABARAP protein was used in this work. Replacement of non-exchangeable protons by deuterons is a common technique to overcome problems arising from fast ^{1}H and ^{13}C transverse relaxation. Deuteration has significantly improved spectral resolution and sensitivity compared with experiments performed on fully protonated molecules [92-96].

However, protein expression in *E.coli* is usually employed to produce large quantities of protein for structural and functional studies. Because of the high price of deuterium oxide (D_2O) and D-Glucose-$^{13}C_6$, the protocol described in Ref. [97] was used and optimised to reduce the cost of the final NMR sample. The advantage of this method is using essentially less amount of D_2O, but the same amount of glucose and ammonium chloride as for 1 L by optimization of growth conditions where cells can

grow in linear log phase until OD_{600} of up to 10. The main steps of the protocol will be explained below.

Growing *E.coli* cells in a D_2O medium usually have few sub-culturing steps with gradually increasing the deuterium content of the medium to allow cells to adapt to D_2O. Here, two sub-culturing steps force cells to use deuterium water by changing the medium based on H_2O to an identical medium with D_2O. In the first step, few colonies of freshly transformed cells from section 2.1.2 were picked and grown in 1 ml LB/H_2O medium (see Table 9) at 37°C for 3 hours. In the second step, 0.3 ml of cells from LB/H_2O were resuspended in a fresh 5 ml LB/D_2O medium and incubated at 37°C for 5 hours. Growing cells in LB medium based on D_2O requires that *E.coli* cells be adapted to atypical D_2O medium and increase their chance to be adapted in M9/D_2O medium.

For isotope labelling in M9 minimal medium, D-Glucose-$^{13}C_6$ and NH_4Cl were used as the sole carbon and nitrogen sources. In the current work, a modified M9 minimal medium was used to maintain an optimal pH for *E.coli* growth over a wide range of cell densities [97]. It might be reached by extremely high cell density (OD_{600}), where the pH of the medium, oxygen level, and other factors can be regulated and stay optimal [97]. The compounds for this M9+ medium are shown in Table 9.

After growing in LB/D_2O medium, cells were centrifuged at 2,000 g for 10 minutes; the supernatant was discarded. The remaining cells were resuspended in 25 ml of unlabelled M9/D_2O and incubated for 14 hours at 37°C. Grown cells in unlabelled M9 medium were centrifuged and diluted in 250 ml ^{15}N, ^{13}C labelled M9 medium. The start OD_{600} is ~0.7. After 11 hours of incubation at 37°C, the cells reached OD_{600} ~4.0. The maximum cell density depends on the availability of nutrients and, therefore, on the amount of glucose. For the expression of GABARAP or GABARAP-G116CΔ117, 1 mM IPTG was added to the cells. The flask with the cell culture was incubated for 16 hours at 25°C. After overnight incubation, *E.coli* were harvested by centrifugation at 5,000 g for 20 minutes. The pellets were stored at -20°C until further use.

2.2. Preparation of protein samples

2.2.1. Purification of His-TEV-Protease

The buffers and solutions used for purification of His-TEV-Protease were based on ultrapure water and sterilised using filters with 0.2 µm pore diameter (Table 11).

Table 11: Buffers used for the purification of His-TEV-Protease.

Lysis buffer	Wash buffer I
Tris-HCl 20 mM, pH 8 NaCl 500 mM Imidazole 10mM Protease inhibitor 1 tablet/50 mL DNAse I ~300 µl in 50 ml buffer (c_{stock} = 2 mg/ml)	Tris-HCl 20 mM, pH 8 NaCl 500 mM Imidazole 10 mM Gu-HCl 1M Glycerine 10% (v/v)
Wash buffer II	**Wash buffer III**
Tris-HCl 20 mM, pH 8 NaCl 500 mM Imidazole 10 mM Glycerine 10% (v/v)	Tris-HCl 20 mM, pH 8 NaCl 500 mM Imidazole 20 mM Glycerine 10% (v/v)
Elution buffer I	**Elution buffer II**
Tris-HCl 20 mM, pH 8 NaCl 500 mM Imidazole 500 mM Glycerine 10% (v/v)	Gu-HCl 6 M NaPi 100 mM, pH8 Tris 10 mM NaCl 500 mM Imidazole 500mM
Storages buffer	
Tris 50 mM, pH 8 NaCl 25 mM EDTA 0.5 mM Glycerine 10% (v/v)	

Cell pellets that had been obtained from 1 litre of cell culture his-TEV-protease expression (section 2.1.3) were lysed by resuspension in 30 ml lysis buffer. Subsequently, the suspension was sonicated 4 times for 1 minute and 1-minute intervals in ice with a Branson sonicator (50% duty cycle, output control 6). The soluble his-TEV-protease was separated from cell debris by centrifugation with 50,000 g for 30 minutes at 4°C. The supernatant was added to 6 ml of Ni-NTA resin (Qiagen) equilibrated in lysis buffer. The resin was washed by 9 column volumes (CV) with ~54 ml of wash buffer I, 9 CV of wash buffer II, and 2 CV of wash buffer III. After washing, his-TEV-protease was eluted from the resin by 2 CV of elution buffer I, 5 times in different tubes. All tubes with the collected sample were stored in ice during elution. Also, 1 mM EDTA and 1 mM DTT were added to every tube with the sample to avoid disulfate bond formation. The samples were checked for the presence of the his-TEV-protease by SDS-PAGE gel. All tubes that contained protease were pooled and then dialyzed to remove imidazole for 14 hours at 4°C. During dialyse, the buffer was exchanged with the storage buffer (Table

11). The sample was concentrated using the Vivaspin 20 ml concentrator (MWCO 3,500 Da) at 4°C, 3,000 g for 15 minutes. The final concentration of TEV protease was 0.87 mg/ml. The protease was stored in 1 ml aliquots with 20% Glycerine at -80°C for future uses (His tag cleaved MSP1D1Δ5 protein expression in the current work).

2.2.2. Purification of His tag cleaved MSP1D1Δ5 protein

His tag cleaved MSP1D1Δ5 protein purification was based on buffers and solutions prepared using ultrapure water and sterilised using filters with pore diameters of 0.2 µm (Table 12).

Table 12: Buffers used for the purification of MSP1D1Δ5.

Lysis buffer	Wash buffer I
NaPi 20 mM, pH 7.4 Protease inhibitor 1 tablet/50 mL Triton X-100 1% Lysozyme ~ 500 µl in 50 ml buffer (c_{stock} = 25 mg/ml) DNAse I ~250 µl in 50 ml buffer (c_{stock} = 2 mg/ml)	Tris-HCl 40 mM, pH 8.0 NaCl 300 mM Triton X-100 1%
Wash buffer II	**Wash buffer III**
Tris-HCl 40 mM, pH 8.0 NaCl 300 mM Sodium cholate 50 mM	Tris-HCl 40 mM, pH 8.0 NaCl 300 mM
Wash buffer IV	**Elution buffer I**
Tris-HCl 40 mM, pH 8.0 NaCl 300 mM Imidazole 10 mM	Tris-HCl 40 mM, pH 8.0 NaCl 300 mM Imidazole 300 mM
Elution buffer II	**Ni-NTA batch buffer**
Tris-HCl 40 mM, pH 8.0 NaCl 300 mM Imidazole 750 mM	Tris-HCl 20 mM, pH 8.0 NaCl 150 mM Sodium cholate 50 mM
Dialyse buffer	**Assemble buffer**
Tris-HCl 50 mM, pH 8.0 EDTA 0.5 mM DTT 1 mM	Tris-HCl 20 mM, pH 7.4 NaCl 100 mM EDTA 0.5 mM
Ni-NTA wash batch buffer	
Tris-HCl 20 mM, pH 8.0 NaCl 150 mM Sodium cholate 50 mM, Imidazole 750 mM	

A well-established protocol was used with minor changes for the purification of MSP1D1Δ5 protein [98]. The pellet from section 2.1.3. was resuspended in 50 ml lysis buffer and incubated for 30 minutes at room temperature with continuous rolling. The cell suspension was placed on ice, and the cells were lysed by sonication eight times 30 seconds with 30 seconds intervals (50 % duty cycle, output control 5). The lysate was separated by centrifugation at 45,000 g for 30 minutes at 4°C. Then the supernatant was added to 6 ml of the Ni-NTA resin equilibrated by wash buffer I. After adding the protein, the column was washed with 4 CV of the following buffers: wash buffer I, II, III, IV. The protein with His-tag was eluted with 4 CV of elution buffer I and 4 CV of elution buffer II. Afterwards, all fractions were collected and checked for the presence of the protein. Then fractions that contained protein were pooled together with His-TEV-Protease (from section 2.2.2), the sample was dialyzed in dialyse buffer at room temperature. After overnight dialysis, the efficiency of TEV protease was tested using SDS-PAGE. Then again, the sample was dialyzed twice in 2 litres Ni-NTA wash batch buffer without Sodium Cholate for 2 hours. 50 mM Sodium Cholate was added to the MSP1D1Δ5 protein, and the sample was loaded to 6 ml Ni-NTA resin equilibrated by Ni-NTA batch buffer. The flowthrough was collected because it contains the highest concentration of protein without his tag. Afterwards, the column was washed twice with 2 CV Ni-NTA wash batch buffer. After washing, the flowthrough was collected in two tubes and checked for the presence of the protein by SDS gel. Then, tubes that contained protein were pooled and dialyzed overnight in 5 litres assemble buffer at room temperature. The final concentration (2.7 mg/ml) was reached using the concentrator Vivaspin (MWCO 3,000 Da) at 3,000 g for 1 hour. The protein was separated in 1 ml aliquots, frozen in liquid nitrogen, lyophilized, and stored at -20°C.

2.2.3. Purification of wild type GABARAP and mutant GABARAP-G116CΔ117

The mutation of GABARAP with Cys at the 116th position and deleted the last 117th Leu (GABARAP-G116CΔ117) was used in this work as artificially created GABARAP-I (Chapter I, 1.1.2). This mutant will be used further for membrane anchoring. For purification of GABARAP and GABARAP mutant, a lysis buffer, two ion-exchange buffers, an elution buffer and an NMR buffer have been used; see Table 13. All buffers and solutions used for purifying GABARAP and GABARAP-G116CΔ117 were based on ultrapure water, sterilised using 0.2 µm pore diameter filters and degassed in a vacuum for at least 1 hour.

For GABARAP-G116CΔ117, which has Cys amino acid at the end of the sequence, 5 mM β-mercaptoethanol was added to the lysis buffer to avoid the formation of disulphide bonds of cysteine residues. All buffers for ion exchange contained 2 mM TCEP, which was used as a reducing agent to break disulphide bonds within and between proteins. The elution buffer for SEC contained 0.25 mM TCEP. The buffers containing TCEP were stored at -4°C because TCEP is very unstable at room temperature.

Table 13: Buffers used for the purification of GABARAP and GABARAP-G116CΔ117.

Lysis buffer	Ion exchange buffer A
Sodium phosphate buffer 25 mM, pH 6.5 EDTA 100 µM DNaseA ~100 µl in 50 ml buffer (c_{stock} = 5 mg/ml) Protease inhibitor 1 tablet/50 ml KCl 50 mM	Sodium phosphate buffer 25 mM, pH 6.5 KCl 50 mM EDTA 100 µM TCEP 2 mM (for mutant)
Ion exchange buffer B	**Elution buffer**
Sodium phosphate buffer 25 mM, pH 6.5 KCl 1 M EDTA 100 µM TCEP 2 mM (for mutant)	HEPES/KOH 25 mM, pH 7.4 KCl 100 mM TCEP 0.25 mM
NMR buffer	
NaPi 25 mM, pH 6.9 KCl 100 mM NaCl 100 mM EDTA 100 µM	

The human's wild type GABARAP purification has been performed following a well-established protocol with minor changes [2]. First, the pellet obtained from the high-yield expression was resuspended in 50 ml lysis buffer for 10 g pellets. Then the cells were incubated on ice for 30 minutes. To lyse the cells, they were sonicated using the Branson Sonifier eight times for 30 seconds each (50 % duty cycle, output control 5); in-between the sonification cycles, the lysate was placed back on the ice for 30 seconds. The soluble protein in the lysate was separated from cell debris by centrifugation with 50,000 g for 45 minutes at 4°C. The supernatant was added to the Äkta System and purified using cation-exchange chromatography. Purification was made by a HiLoad 16/10 SP Sepharose column with 20 ml column volume. The flow rate of the column was 1.6 ml/min. The salt concentration of KCl was set as 40% (400 mM) so that most of the proteins can be eluted in a reasonable time. After running through the Äkta, the sample was collected in different fractions with 2 ml of the sample in each. The fractions were checked for the presence of the protein and pulled together. For the best purification efficiency, the sample was again loaded to the Äkta System. NMR buffer was used for SEC of GABARAP. Gel filtration was made by HiLoad 16/600 Superdex 75 pg column with 120 ml column volume. The flow rate of the column was 1 ml/min. After gel filtration, the sample was collected in different fractions. All fractions were checked by SDS-PAGE gel for the presence of the protein and then pulled together. The final concentration of GABARAP was reached using concentrator Vivaspin (MWCO 3,000 Da) at 3,500 g for 2 hours at 10°C.

For purification of GABARAP-G116CΔ117, elution buffer was used for SEC instead of NMR buffer. The necessary concentration of GABARAP-G116CΔ117 for future lipidation was reached by concentrator Vivaspin (MWCO 3,000 Da) at 4,500 g for 2 hours.

2.2.4. Lipidation of GABARAP-G116CΔ117

Lipidation of GABARAP-G116CΔ117 has been performed using a lipid resuspension buffer and assemble buffer (*Table 14*). Both buffers were based on ultrapure water and sterilised by 0.2 µm pore diameter filters. The lipid resuspend buffer was stored at -4°C because of the presence of TCEP in it. The assemble buffer was degassed in a vacuum for at least 1 hour.

Table 14: Buffers used for the lipidation of GABARAP.

Lipid's resuspend buffer	Assemble buffer
HEPES/KOH 25 mM, pH 7.6 KCl 100 mM TCEP 0.25 mM Sodium cholate 100 mM	Sodium phosphate buffer 10mM, pH 7.4 NaCl 150 mM

The lipidation of GABARAP-G116CΔ117 was achieved chemically through a maleimide reaction between 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-[4-(p-maleimidophenyl)butyramide] (MPB-PE) and the C-terminus of the protein. MPB-PE reacts with the thiol group of Cys with forming a covalent bond (Figure 17).

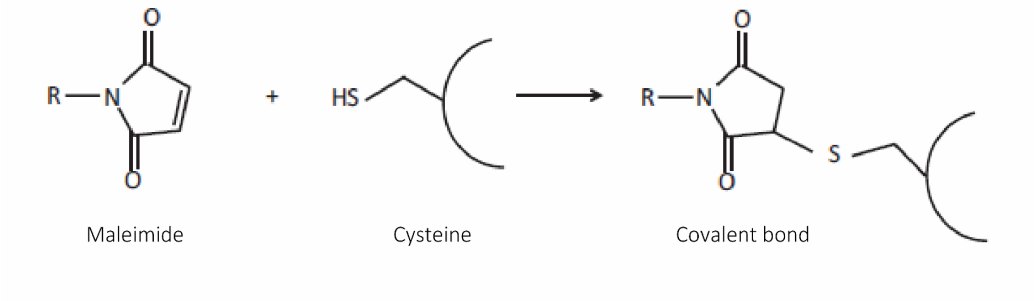


Figure 17. Maleimide reaction of MPB-PE with 116 Cys of GABARAP-G116CΔ117.

For lipidation, six molecules of MPB-PE (1.8 M) per one molecule of GABARAP-G116CΔ117 (300 µM) were used. Before maleimide reaction with the protein, the MPB-PE should be dissolved in lipid resuspend buffer. First, MPB-PE powder was resuspended in chloroform. Subsequently, the chloroform was removed entirely upon incubation of the sample under a constant stream of N_2 for 45 minutes. After the evaporation of chloroform, the film of MPB-PE was formed. The lipid resuspends buffer without TCEP was added to the MPB-PE film and mixed thoroughly. Then the tube with MPB-PE was placed in an ultra-sonic bath for better resuspension for 30 minutes.

For the maleimide reaction: MPB-PE and GABARAP-G116CΔ117 were mixed in equal volumes. The mixture was incubated for 60 minutes at room temperature with constant shaking using Thermomixer

(Eppendorf). The coupling was tested for the presence of lipidated GABARAP-G116CΔ117 (GABARAP-PE) using an SDS-PAGE gel.

2.2.5. Anchoring GABARAP-PE I to nanodiscs

Anchoring GABARAP-PE to NDs was made using the protocol [99] with some changes. Membrane scaffold protein MSP1D1Δ5 and DMPC lipids were used for producing lipid nanodiscs (Chapter I, 1.1.3). First, DMPC lipids suspended in chloroform were placed in a glass tube. The chloroform was evaporated under a constant stream of N_2 for 45 minutes. Then, the tube with a film was placed in the vacuum for 1 hour. The film of DMPC was resuspended in the lipid resuspend buffer (Table 14). The mixture in the glass tube was placed in an ultrasonic bath for 30 minutes. GABARAP-PE I, MSP1D1Δ5 and DMPC have been used at molar ratios 1:1:53, such that every nanodisc contains only one molecule of GABARAP on each side. All necessary compounds were mixed and incubated in a sequence ice-37°C-ice-37°C-ice for 20 minutes each round. The detergent then was dialyzed (MWCO 3,500 Da) by 2.5 litres assemble buffer (Table 14) for 20 hours at 4°C; the buffer was changed twice to keep it fresh. The precipitation of proteins and lipids after dialyzing was removed using a membrane filter (0.22 µm). The clear solubilised protein sample was loaded to the Äkta system and purified using a 16/600 Superdex 200 pg column for the size exclusion chromatography. The assemble buffer was running with a 0.5 ml/min flow rate through the column for the best separation. After gel filtration, the sample was collected in different fractions. All fractions were checked by SDS-PAGE gel for the presents of the protein with nanodisc and then mixed. The changing from the assemble buffer to the NMR buffer was done during concentration. The concentrator Vivaspin (MWCO 10,000 Da) was filled with the sample up to maximal volume. The sample was centrifuged at 600 g for 2 hours at 10°C. The buffer in the filtrate container was checked for the absence of the protein; then, the container was emptied. Afterwards, the concentrator was filled with NMR buffer up to the maximal volume. The final concentration of GABARAP-nanodiscs (390 µM) in the NMR buffer was reached at 600 g for 2 hours at 10°C. This amount of protein was enough for three identical NMR samples using Slot Shigemi tubes [100].

2.3. Analytical methods

2.3.1. SDS-polyacrylamide gel electrophoresis

Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) was developed by Lämmli and is commonly used to separate proteins with different molecular masses [101]. SDS-PAGE was used for controlling the protein size distribution after expression and purification using SEC. Also, the difference between the thickness of lines of GABARAP-PE I and MSP1D1Δ5 may give the approximate calculation of ratio estimation between GABARAP and nanodiscs. The polyacrylamide gels consisted of a 5% (w/v) acrylamide stacking gel and a 15% (w/v) acrylamide separation gel. To separate lipidated and free GABARAP, Lämmli buffer without β-ME was used to prevent the disruption of maleimide coupling between protein and lipid. The compositions used for SDS-PAGE were based on ultrapure water (Table 15).

Table 15: Compositions used for SDS-PAGE gel.

Lämmli buffer (dye 4x)	Coomassie staining solution
Tris-HCl 200 mM, pH 6.8 β-ME 8% (v/v) SDS 8% (w/v) Glycerin 40% (v/v) Bromophenol blue 0.05 (w/v)	Isopropanol 25% (v/v) Acetic acid 10% (v/v) Coomassie Brilliant Blue R250 0.05% (w/v)
SDS-PAGE running buffer	
Tris-HCl 50 mM, pH 8.3 Glycine 385 mM SDS 0.1% (w/v)	

The Lämmli buffer (4 µl) was added to the protein sample (20 µl), and then each sample was boiled for at least 5 minutes at 95°C. First, 5 µl of the unstained protein marker was loaded onto the gel, then 7 µl of each sample were loaded in separate rows on the gel. SDS-PAGE gels were run in SDS-PAGE running buffer with the following conditions: 45 mA, 100 W per gel for 45 min or until the dye front reached the end of the gel.

2.3.2. Quantification and concentration of protein by UV/Vis spectrophotometry

UV/Vis spectrophotometry measures the absorption of UV-visible light of proteins. A typical absorption wavelength for amino acids with aromatic rings is 280 nm, so that the UV/Vis spectrum was measured between 220 nm and 320 nm. The protein concentration was then estimated from the absorbance at the characteristic wavelength λ = 280 nm using the Lambert-Beer-Law:

$$lg\left(\frac{I_0}{I_1}\right) = A_\lambda = \varepsilon_\lambda cl \qquad (15)$$

Here, I_0 is the radiant flux received by the sample, I_1 is the radiant flux transmitted by the sample, A_λ is the absorbance at a given wavelength, l is the path-length in the sample solution used to measure the absorbance, ε_λ is the molar extinction coefficient, and c is the concentration of protein in moles. The extinction coefficient of the used in this work proteins was calculated using the Expasy ProtParam program based on the amino acid sequences. The molar extinction coefficient of GABARAP is 11,920 $M^{-1}cm^{-1}$, TEV protease is 31,970 $M^{-1}cm^{-1}$, and MSP1D1Δ5 is 18,450 $M^{-1}cm^{-1}$. The extinction coefficient $\varepsilon_\lambda=2\varepsilon_{\lambda1}+2\varepsilon_{\lambda2}$ = 60,740 $M^{-1}cm^{-1}$ for the GABARAP-G116CΔ117-nanodisc complex was calculated assuming two belt proteins MSP1D1Δ5 and two GABARAP-G116CΔ117 molecules per nanodisc.

2.4. Nuclear magnetic resonance (NMR) spectroscopy

The NMR measurements were carried out with NMR spectrometers from Varian (Palo Alto, USA) and Bruker (Karlsruhe, Germany), see Table 16.

Table 16: NMR spectrometers used in this work.

Spectrometer	Frequency	NMR CRYO Probe
Bruker (Magnex)	600 MHz	5 mm cryo-QCI H-P/C/N-D
Bruker (Oxford)	600 MHz	5 mm cryo-TCI H-/C/N-D
Bruker	700 MHz	5 mm cryo-TCI H-/C/N-D
Bruker	900 MHz	5 mm cryo-TCI H-/C/N-D
Varian	800 MHz	5 mm H-/C/N-D

2.4.1. NMR samples

All NMR experiments were recorded in the NMR buffer (Table 13) with 10% (v/v) D_2O. The presence of D_2O is necessary for field-locking as well as for shimming purposes. Due to the high salt concentration and a small amount of deuterated sample, special-shaped 5 mm Slot Shigemi tubes [100] are used. The tube has a susceptibility-matched glass cavity with the slot for the sample, which helps keep the pulse lengths shorter due to a smaller sample volume.

NMR experiments were conducted for the following samples:

1. 900 µM [U-^{2}H,^{15}N,^{13}C] - GABARAP

2. 570 µM [U-^{2}H,^{15}N,^{13}C] – GABARAP-G116CΔ117

3. ≈ 390 µM [U-^{2}H,^{15}N,^{13}C] – GABARAP-G116CΔ117 – DMPC nanodiscs

2.4.2. NMR experiments

All NMR experiments discussed in this chapter are listed in Table 17.

Dynamic NMR studies are sensitive to temperature changes; therefore, the temperature calibration of each NMR spectrometer was regularly tested with a perdeuterated methanol sample as ^{1}HN NMR spectrum [102]. The important recording parameters of the measured NMR spectra can be found in Table 17.

Table 17: NMR experiments. The first column is the sample number according to section 2.4.1., the second column is the name of the experiment, then the selected acquisition parameters and references for used pulse programs.

N	Experiment	Spectrometer	SW1 (ppm)	t1 (complex)	SW2 (ppm)	t2 (complex)	^{15}N offset	SW3 (ppm)	t3 (complex)	^{13}C offset	NS	Relaxation delays (ms)	Temp (K)	Pulse
1	^{15}N-TROSY-HSQC	600 Oxford	16	1536	29	256	117.0	-	-	-	48	-	313	[103], [104]
1	3D ^{15}N,^{13}C-TROSY-HNCO	600 Oxford	16	1024	29	44	117.0	12	32	176.0	16	-	313	
1	3D ^{15}N,^{13}C-TROSY-HNCA	600 Oxford	16	1024	29	44	117.0	26	48	56.0	16	-	313	
1	3D ^{15}N,^{13}C-TROSY-HNCACB	600 Oxford	16	1024	29	44	117.0	52	64	43.0	24	-	313	
2	^{15}N-TROSY-HSQC	700 Bruker	16	1024	29	256	116.5	-	-	-	32	-	313	
2	3D ^{15}N,^{13}C-TROSY-HNCO	700 Bruker	16	1024	29	48	116.5	12	32	176.0	16	-	313	
3	^{15}N-TROSY-HSQC	600 Oxford	16	1536	29	256	117.0	-	-	-	48	-	313	
3	^{15}N-TROSY-HSQC	900 Bruker	16	2048	29	256	117.0	-	-	-	32	-	313	
3	3D ^{15}N,^{13}C-TROSY-HNCO	600 Oxford	16	1024	29	44	117.0	12	32	176.0	24	-	313	
3	3D ^{15}N,^{13}C-	600 Oxford	16	1024	29	44	117.0	26	48	56.0	32	-	313	

	TROSY-HNCA													
3	3D ^{15}N,^{13}C-TROSY-HNCACB	600 Oxford	16	1024	29	44	117.0	52	64	43.0	24	-	313	
3	^{15}N-TROSY-T_1	600 Oxford	16	1024	29	96	117.0	-	-	-	16	0, 1280, 80, 1120, 160, 960, 240, 800, 320, 640, 400, 480, 0, 480, 1280	313	[105], [106]
3	^{15}N-TROSY-T_1	900 Bruker	16	1536	29	96	117.0	-	-	-	16	0, 1280, 80, 1120, 160, 960, 240, 800, 320, 640, 400, 480, 0, 480, 1280	313	
3	^{15}N-TROSY-$T_{1\rho}$	600 Oxford	16	1024	29	96	117.0	-	-	-	16	10, 100, 20, 90, 30, 80, 40, 70, 50, 60, 10, 60, 100	313	[105], [106], [107]
3	^{15}N-TROSY-$T_{1\rho}$	900 Bruker	16	1536	29	96	117.0	-	-	-	24	10, 100, 20, 90, 30, 80, 40, 70, 50, 60, 10, 60, 100	313	
3	{^{1}H}-^{15}N-TROSY-NOE	600 Oxford	16	1024	29	96	117.0	-	-	-	24	-	313	[105], [106]
3	{^{1}H}-^{15}N-TROSY-NOE	900 Bruker	16	1536	29	96	117.0	-	-	-	32	-	313	

2.4.3. Processing the NMR spectra

All NMR experiments measure the response of a nuclear spin excitation by applying a radiofrequency magnetic field, usually in the form of a short pulse or pulse sequence. It generates a rotating nuclear magnetic moment, which induces a small oscillating voltage in the probe coil: a 'free induction decay' (FID). The raw experimental data must be converted into a frequency domain to generate the visible spectrum. The data processing methods include manipulating the NMR FID to improve the signal-to-noise ratio and increase the spectral resolution. For this purpose, the NMRPipe software package was used (*Table 8*). It consists of a series of different programs, but in this work, only two of them were used for processing the NMR data: NMRPipe and NMRDraw. Processing NMR spectroscopic data involves the steps shown in Figure 18.

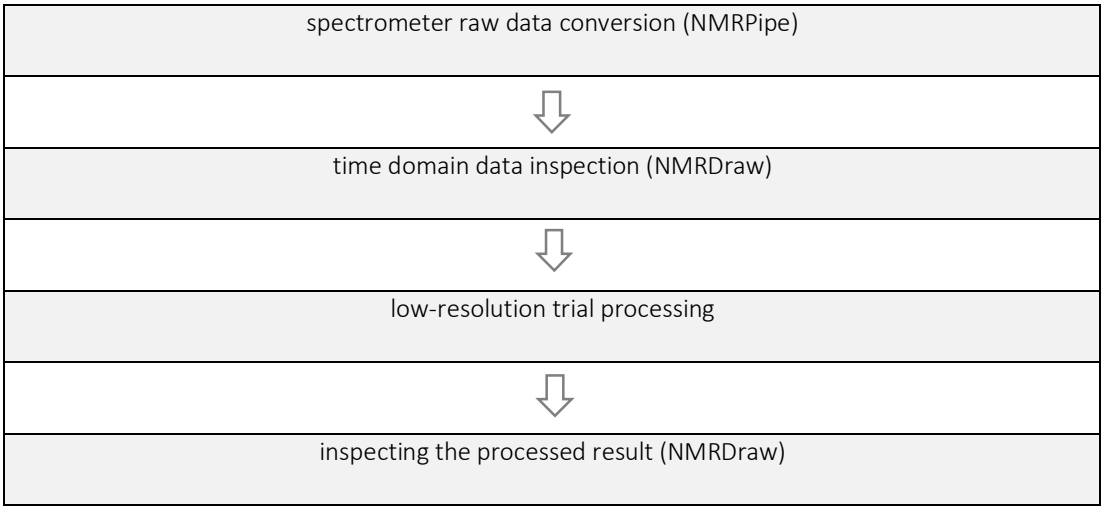


Figure 18. Overview of data processing with the NMRPipe package.

NMRPipe is a UNIX C-shell script pipe command-based program, where a list of commands is presented as a text file. The created conversion script contains one to four dimensions, where all data have a standard format with the same organization of real and imaginary points. For some types of experiments, the script supports special options for complex acquisition schemes, including sensitivity or gradient-enhanced data. Once the time-domain data has been converted, the frequency spectra can be open and inspected using the graphical interface NMRDraw with initial phasing and suppression of the solvent signal. After opening a spectrum in NMRDraw, the first trial processing can be saved and executed as UNIX shell scripts for future use, such that the parameters in the script can be optimised in further processing steps. The optimisation of the spectra in NMRDraw increases the quality of the spectra. The spectral processing parameters include first-point scaling, phase correction, baseline correction, reversed spectra, and left/right swapped spectra. The processed results will be inspected to decide if additional processing is required; if not, the spectrum is saved in a format needed for future analysis. Spectra were saved in an NMRViewJ program format (Table 8).

2.4.4. Sequential resonance assignment

Fully deuterated samples of GABARAP, GABARAP-G116CΔ117 and GABARAP-NDs were recorded using 600, 700 and 900 MHz spectrometers. Assignment of resonances was obtained from J-correlated 3D

triple resonance experiments: 3D TROSY-HNCO, TROSY-HNCA, TROSY-HNCACB. Also, 2D ^{15}N TROSY-HSQC spectra were recorded for H_N, ^{15}N assignment. The assignment of ^{15}N, H_N, $^{13}C_\alpha$, $^{13}C_\beta$, and ^{13}CO was updated from the well-known chemical shifts of the non-deuterated GABARAP sample. In a ^{2}H, ^{15}N, ^{13}C labelled protein, there are shifts of the backbone (^{15}N, H_N, $^{13}C_\alpha$ and ^{13}CO) and $^{13}C_\beta$ side-chain nuclei [108]. Nevertheless, no significant changes between deuterated and non-deuterated samples were observed in chemical shifts.

2.4.5. Chemical shift perturbation analysis

Chemical shifts are one of the most useful indicators of the chemical environment of each molecule. Chemical shift perturbation (CSP) is a common technique for demonstrating changes in chemical shifts. The average weighted chemical shift perturbation between tree different nuclei resonances in the recorded 3D ^{15}N,^{13}C-TROSY-HNCO spectra was calculated using [20]:

$$\Delta\delta_{ave} = \sqrt{\Delta\delta_{HN}{}^2 + \left(\frac{\Delta\delta_N}{10}\right)^2 + \left(\frac{\Delta\delta_C}{4}\right)^2} \qquad (16)$$

Where Δδ is the chemical shift perturbation for the proton, nitrogen, and carbon dimensions.

2.4.6. Fast protein dynamics from pico- to nanoseconds

Applied radiofrequency pulses cause perturbation of nuclear spins from their equilibrium state. The process by which the spins return to their equilibrium is called relaxation and was described in Chapter I, 2.1. NMR relaxation is one of the methods for studying fast protein dynamics on the different time scales from pico- to nanoseconds (Chapter I, 2.4).

Experimental determination of R_1

The relaxation rates R_1 for each possible amino acid were determined with the help of a ^{15}N TROSY-based inversion-recovery experiment. In this experiment, the pulse sequence described in Refs. [105, 106] was used with minor changes. The relaxation decay was sampled for 15 different delays (see *Table 17*) for each relaxation time measurement. The FIDs for these delay times were recorded before increasing the t_1 evolution period. The order of the relaxation delay durations was pseudo-randomized for each value of t_1 [105]. The parameter of relaxation R_1 was received by fitting exponentials to the relaxation curves. The decay of longitudinal relaxation follows an exponential decay

$$I_t = I_0 e^{-R_1 t} \qquad (17)$$

I_t is the amplitude after an inter-pulse delay t, and I_0 is the amplitude at time 0.

The ^{15}N-TROSY-T_1 experiment was recorded at two different magnetic field strengths, 600 and 900 MHz. The parameters of these experiments are summarized in Table 17. In the NMRViewJ program (Table 8), the peak intensities for each available amino acid at different time delays were estimated by MUNIN [86, 87]. Then the data were fitted by equation 17 in MATLAB to determine R_1.

Experimental determination of R_2 with the help of $R_{1\rho}$

T_2 transversal spin-spin relaxation is a parameter characterizing refocusing the chemical shift due to phase coherence in the XY plane after the RF pulse was applied. $T_{1\rho}$ is spin-lattice relaxation time in the rotating frame. In measuring ^{15}N T_2 with $T_{1\rho}$, the transverse magnetization relaxes in the rotating frame while aligned along the applied B_1 or "spin-lock" magnetic field. The relaxation rate R_2 was determined with the help of R_1 and $R_{1\rho}$ according to:

$$R_2 = \frac{1}{\gamma^2 {B_1}^2} \cdot \left(\left(\Omega^2 + \gamma^2 {B_1}^2 \right) R_{1\rho} - \Omega^2 R_1 \right) \tag{18}$$

Here, R_1 is the longitudinal relaxation rate from the ^{15}N-TROSY-T_1 experiment (see above) and $R_{1\rho}$ is the transverse relaxation rate in the rotating frame, γ is the gyromagnetic rate, Ω is resonance offset.

The relaxation rate $R_{1\rho}$ is an analogy to the longitudinal relaxation R_1 and was received by the ^{15}N-TROSY-$T_{1\rho}$ experiment. The experiment is based on the pulse sequence from the following Ref. [105, 107]. In this work, the ^{15}N-TROSY-$T_{1\rho}$ experiment was recorded at two different magnetic fields, 600 MHz and 900 MHz. The parameters of these experiments are summarized in Table 17. The $T_{1\rho}$ is determined by analysing the intensity of the NMR signal for each amino acid in the NMRViewJ program (Table 8) using the MUNIN program [86, 87]. Then the data were fitted by equation 17 in MATLAB to determine $R_{1\rho}$. MUNIN determined the error by the intensity peaks analysis. The resonance offset was calculated from the GABARAP peak list by using the NMRViewJ program.

Experimental determination of heteronuclear NOE

The steady-state hetNOE was determined with the help of $\{^1H\}$-^{15}N NOE experiment, which corresponds to cross-relaxation between two dipolar-coupled spins ^1H and ^{15}N and provides the information about N-H bond dynamics on the fast ps timescale. $\{^1H\}$-^{15}N NOE is unique among the mentioned relaxation parameters R_1, R_2 and $\{^1H\}$-^{15}N NOE because it is regarded as essential for the accurate estimation of the spectral density function (see Chapter I, 2.4.2) at high frequencies (ωH ± ωN), and it is essential for the identification of fast backbone motions [109-111]. For the determination of $\{^1H\}$ -^{15}N NOE values, the steady-state approach was used. It measures the longitudinal polarization at the thermal equilibrium of ^{15}N spin I_0 and then the steady-state longitudinal ^{15}N polarization under ^1H irradiation I_{sat} [112]. The pulse sequence described in Refs. [105, 106] was used in the $\{^1H\}$-^{15}N NOE experiment. The NOE of the N-H bond is calculated from the ratio of cross-peak intensities of each amino acid in two experiments - with and without proton saturation such as:

$$NOE = \frac{I_{sat}}{I_o} \tag{19}$$

Where I_{sat} is the peak intensity with proton saturation and I_0 is the references peak intensity without proton saturation.

{^{1}H}-^{15}N NOE experiments were performed for two different ^{1}H frequencies of 600 and 900 MHz. The parameters are summarized in Table 17. {^{1}H}-^{15}N NOE of a folded protein has a maximum value ~ 1.0, and it only occurs when the internal dynamic (Chapter I, 2.4.2) is absent. NOE values less than 0.65 is considered to be flexible. The volume of the peaks for each amino acid was calculated with the help of NMRviewJ, and the error was determined using the noise level of the spectrum.

Backbone and sidechain dynamics

NMR relaxation experiments are widely used for characterising protein backbone and sidechain dynamics. If the measured ^{15}N relaxation data R_{1i}, R_{2i} and NOE_i are available for each residue *i,* the model-free Lipari-Szabo approach can be used for the determination of the protein backbone dynamics [25, 26]. The spectral density function (Chapter I, equation 6) describes the internal motion and overall motion. In the model-free approach, the internal mobility is characterized by two parameters describing the amplitude order parameter S^2 and a characteristic correlation time of the internal motion τ_i. The overall motion is described by global τ_c. Data analysis is based on statistical analysis by first applying the simplest model sphere, dynamical parameters (S^2, τ_c global, diffusion tensor, τ_c internal, R_{ex}) are then added until the fit is statistically significant. In this work, the program TENSOR2 was used to determine the rotational diffusion tensor from the three-dimensional structure coordinates and ^{15}N relaxation data for one magnetic field. TENSOR2 uses Monte-Carlo sampling methods with appropriate χ2 and F-tests to determine which models are suitable for describing the rotational diffusion tensor. For calculating the global rotation correlation time τ_c, only one magnetic field strength is sufficient because the correlation time is determined by the ratio R_2/R_1 of the relaxation rates. To exclude flexible amino-acid residues, values of the ratio R_2/R_1 that deviate from the average value by more than 20% were removed. The determination of the internal correlation time depends on the field strength. In this work, two different magnetic fields were used (600 MHz and 900 MHz) for the identical 390 µM [U-^{2}H,^{15}N,^{13}C] – GABARAP-G116CΔ117 – DMPC nanodisc sample at 40°C.

2.4.7. Hydrodynamical radius

If GABARAP protein has isotropic symmetry, the hydrodynamic radius can be determined with the help of global rotation correlation time τ_c (Chapter 1, 2.4.2) and Stokes-Einstein equation:

$$\tau_c = 4\pi\eta \frac{{r_H}^3}{3k_B T} \tag{20}$$

Where η is the viscosity of the solution, r_H is the hydrodynamical radius, k_B is the Boltzmann constant, and T is the temperature in Kelvin. The viscosity of the NMR sample, which contains 10% D_2O, was calculated as

$$\eta = 0.9 \cdot \eta(H_2O) + 0.1 \cdot \eta(D_2O) \tag{21}$$

3. Results

3.1. Expression and purification of MSP1D1Δ5 belt protein and TEV protease

The expressions and purifications of MSP1D1Δ5 protein and TEV protease are described below in sections *2.1.3, 2.2.1* and *2.2.2*. Figure 19 (left) shows the SDS-GEL picture of the elution of His-Tev-protease. After cell pellets were lysed and soluble TEV-protease was separated from cell debris by centrifugation, the supernatant was added to Ni-NTA resin and washed by washing buffers (Table 11). After washing, the Tev-protease was eluted five times in different tubes. The elution one (2) contains the biggest concentration of the TEV-protease based on the line intensities, while the last fifth elution (6) contains the smallest concentration. The size of the TEV-protease is 27kDa, while the sample run through the gel has a molecular weight of 30 kDa. The gel picture indicates the successful production of Tev-protease from the described above protocol. Figure 19 (right) shows the SDS-GEL slice picture of the purification of His-tag cleaved MSP1D1Δ5. After cells were lysed and the lysate was separated from cell debris, the supernatant was added to the Ni-NTA resin column. The eluted MSPD1Δ5 (2) was containing His-tag. To remove the His-tag, the TEV-protease were added to the sample, and the sample was dialysed. The efficiency of TEV-protease after dialysing is shown (4,5). Then the sample was loaded again to Ni-NTA resin to remove the rest of TEV-protease. The sample was run through, flow-through was collected (7), also the washed sample was collected (8). The size of the His-tag version of MSP1D1Δ5 is 22.1 kDa, while the protein run through the gel has a molecular weight of ~22 kDa. After His-tag cleavage, the size of MSP1D1Δ5 is 19.5 kDa, and it is shown on the gel around 18.4 kDa. The protein purity of more than 95%.

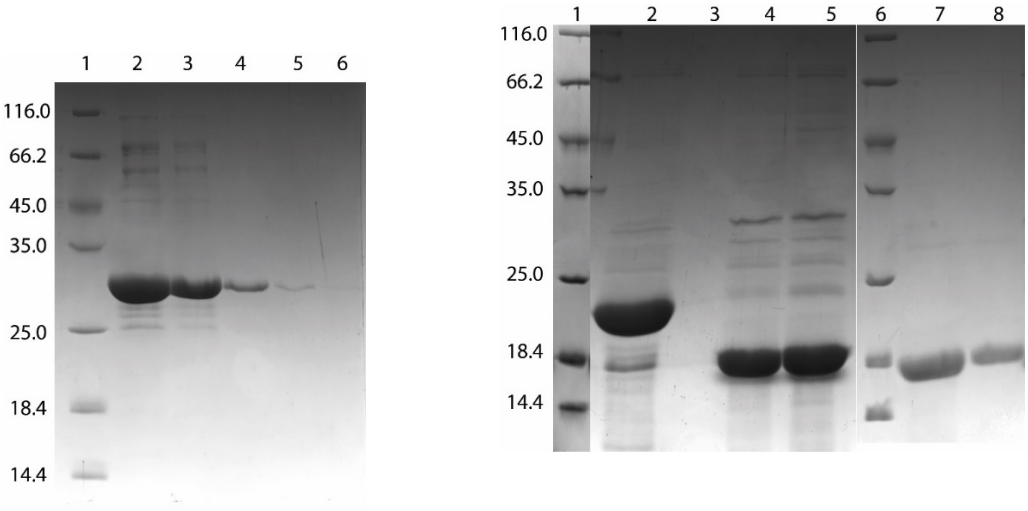


Figure 19. SDS-PAGE (15%) gel slices containing MSP1D1Δ5 protein (right) and TEV protease (left) were taken after purification. The left picture contains the marker line (1) and five different fractions of Tev-protease after elution. The right picture shows the SDS-PAGE of MSP1D1Δ5 before and after His-tag cleavage. (1,6) is a marker, (2) is a His-tag version of MSP1D1Δ5 after elution. (3) is an empty fraction. (4,5) MSP1D1Δ5 after his-tag cleavage. (7) is collected flow through the Ni-NTA, which contains the biggest concentration of MSP1D1Δ5 without His-tag. (8) is the washing of MSP1D1Δ5 without His-tag from the Ni-NTA resin.

3.2. Expression of fully deuterated GABARAP-G116CΔ117 in two different M9 mediums.

The expression of GABARAP-G116CΔ117 was described in sections 2.1.3 using the updated expression protocol from Ref. [97]. The success of using the new protocol for GABARAP protein expression was monitored by SDS-PAGE analysis. Figure 20 shows the two SDS-PAGE gel slices of the GABARAP-G117CΔ117 expression in *E.coli* using the standard recipe (left) and the new recipe of M9+ medium (right).

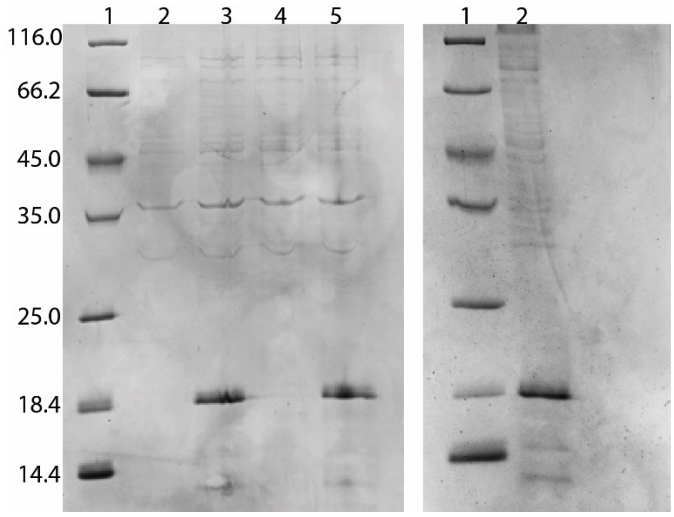


Figure 20. SDS-PAGE (15%) gel slices containing deuterated protein samples before and after expression in two different M9 mediums. There is the expression in the standard M9/D2O medium on the left picture where the first line is a marker, the second and fourth are the sample before adding IPTG in two different flasks, the third and the fifth are the sample shortly before the harvest for both flasks. The right picture shows the expression of GABARAP-G116CΔ117 in the new M9+/D_2O medium where the first line is the marker, and the second line is the sample after 20 hours of expression with 1 mM IPTG.

The SDS-PAGE analysis of both samples showed the strong appearance band in the gel (3 and 5, left) after adding 1 mM IPTG and 20 hours of expression. The same strong band at the same position (2, right) was observed after adding 1 mM IPTG to express the protein in the new M9+ medium. The samples run through the gel have a molecular weight of 18.4 kDa, while the actual molecular weight of GABARAP is 14 kDa. The gel pictures indicate the successful production of GABARAP-G116CΔ117 in standard and modified M9 mediums.

3.3. Purification of fully deuterated GABARAP-G116CΔ117 after expression in new M9 medium.

After cell pellets were lysed and soluble GABARAP protein was separated from cell debris by centrifugation, the supernatant was loaded to the Äkta system using a superloop. For removing many containments from cells, purification of fully deuterated GABARAP-G116CΔ117 was performed in two steps. The first step is ion-exchange chromatography using an HiLoad SP Sepharose column 16/10. Figure 21 shows an anion exchange chromatogram for GABARAP-G116CΔ117, grown in 125 ml of new M9+ mediums. The absorption at 280 nm and % of salt concentration of buffer B (Table 13) are plotted as a function of the volume passed through the column. GABARAP-G116CΔ117 protein was

eluted at a buffer B concentration of approximately 25-26%, corresponding to a salt concentration of 300 mM KCl.

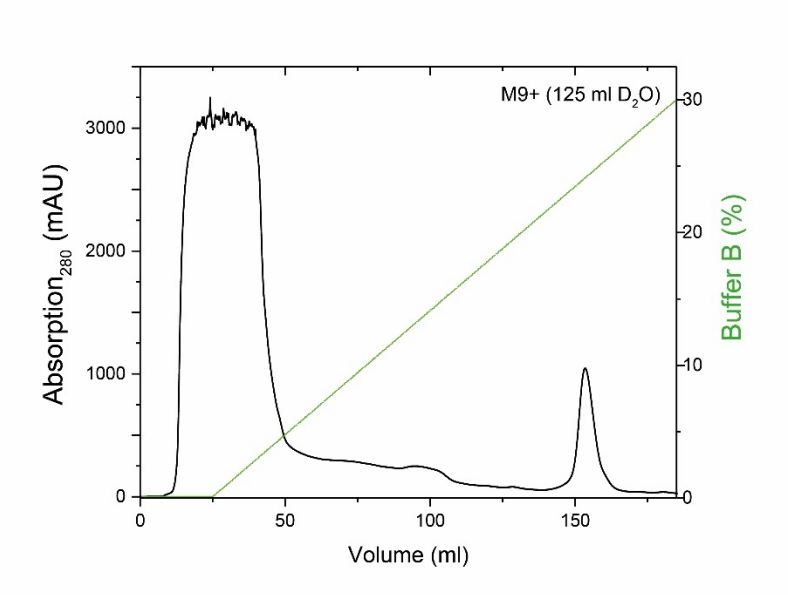


Figure 21. Anion Exchange Chromatography of deuterated GABARAP-G116CΔ117 from M9+ medium. Black absorption line: elution profile of GABARAP-G116CΔ117 protein from 125 ml of deuterated new M9+ medium eluted from HiLoad SP Sepharose 16/10 column with 0-1 M KCl gradient over 9 CV. Green line: Concentration gradient of buffer B in %.

The presence of protein in the eluted fractions was checked by SDS gel. Figure 22 shows SDS-PAGE slices of GABARAP-G116CΔ117 for different fractions after elution. The figure clearly shows that the protein has good purity and a strong band around 18.4 kDa, which corresponds to the size of the protein 14 kDa.

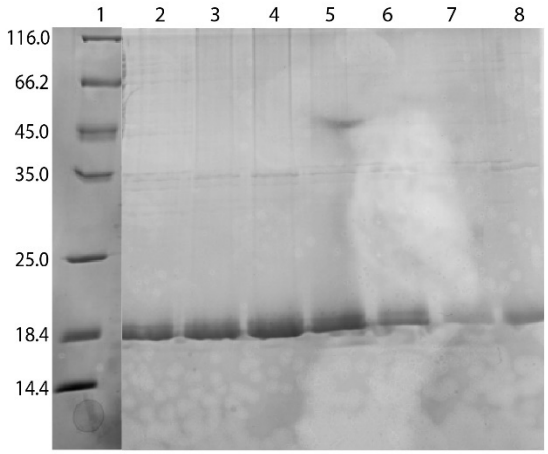


Figure 22. SDS-PAGE (15 %) of fractions after Anion Exchange Chromatography. SDS-PAGE of the protein elution from 125 ml of deuterated new M9+ medium. The first line is the marker; lines 2-8 correspond to the different fractions in the range of 26% of buffer B (around 150 ml, Figure 21).

The second step of protein purification is a Size Exclusion Chromatography made using HiLoad 16/600 Superdex 75 pg column. Figure 23 shows the Size Exclusion chromatogram of GABARAP-G116CΔ117 absorption at 280 nm for M9+ medium. The maximum volume of elution was set as 120 ml. The protein was eluted in the range of 70 ml to 90 ml. The main peak has good separation and high absorption (~1100 mAU), while the impurities are significantly lower.

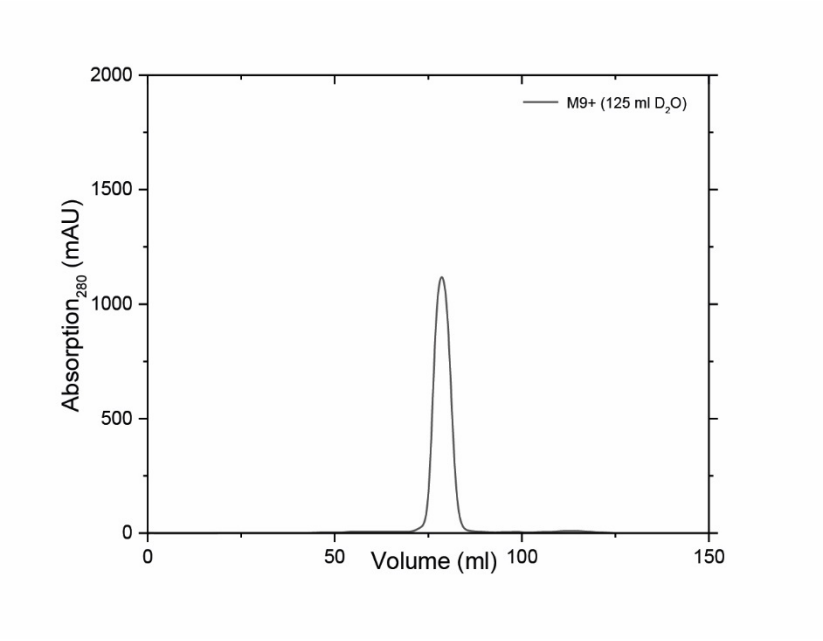


Figure 23. Size Exclusion Chromatography of deuterated GABARAP-G116CΔ117 from M9+ medium. Black absorption line: elution profile of GABARAP-G116CΔ117 protein from 125 ml new M9+ medium using HiLoad 16/600 Superdex 75 pg column.

The purity of the eluted protein was checked by SDS-PAGE. Figure 24 shows the SDS-PAGE slice for the elution of GABARAP-G116CΔ117 protein from 125 ml of deuterated M9+ medium where five fractions (2 ml of the sample in each) from the range 74 ml to 84 ml were checked for presents of the protein. The elution peak of the new M9+ medium has a protein purity of more than 95%. The fractions were collected, and the final concentration of GABARAP-G116CΔ117 in 10 ml buffer was 0.58 mg/ml for 125 ml M9+/D_2O. Fractions were collected, and the volume was reduced using Vivaspin concentrator (MWCO 3,500 Da) until the concentration reached the amount of 0.3 mM. Thus 1.155 mg of deuterated GABARAP-G116CΔ117 was purified from 125 ml expression in a new M9+ medium. This amount of protein was enough for three identical NMR samples.

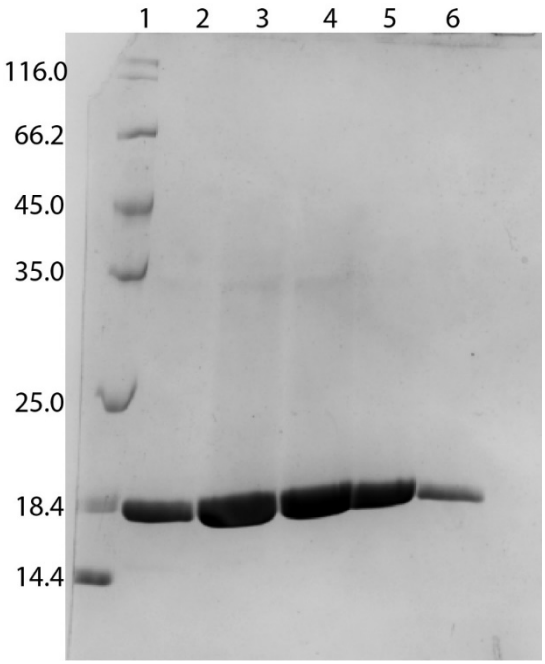


Figure 24. SDS-PAGE (15%) of fractions after Size Exclusion Chromatography. Five elution fractions of GABARAP-G116CΔ117 which was eluted from 125 ml of the M9+ medium.

3.4. Purification of fully deuterated wt GABARAP

Since it was shown that deuteration could change the chemical shift on the spectrum [108]. The new fully deuterated wt protein sample was prepared to compare the chemical shifts of the mutant GABARAP-G116CΔ117 with the wild type of GABARAP.

Expression of fully deuterated GABARAP was made in the M9+ medium using the same protocol as for GABARAP-G116CΔ117 (section 2.1.3). Figure 25A shows the SDS-PAGE of the sample after adding 1 mM IPTG and 19 hours incubation. The protein has a size of around 18.4 kDa based on the marker. Purification of the GABARAP was made using the same methods and devices as before for the mutant (section 2.2.3). Figure 25B shows the slice of SDS-PAGE that corresponds to the main peak (26% buffer B) of GABARAP after ion-exchange chromatography. The eluted peak of GABARAP from the new M9+/D_2O medium has a purity of more than 95%.

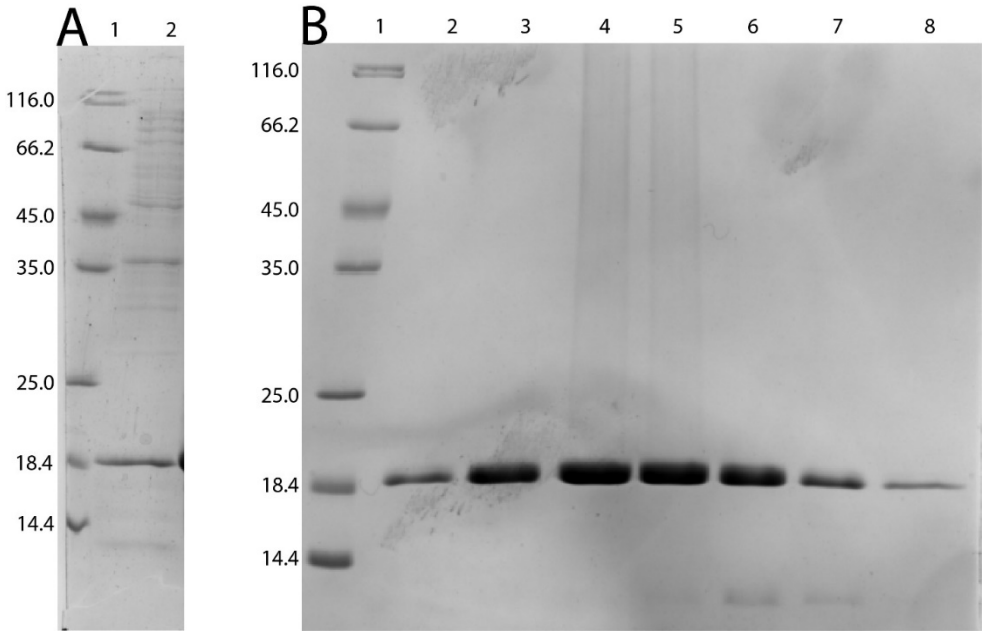


Figure 25. SDS-PAGE (15%) gel slices containing protein samples of GABARAP protein. Figure A shows the expression of GABARAP in the new M9+ medium where the first line is the marker, and the second line is the sample after 19 hours of expression with 1 mM IPTG. Figure B shows the protein purity after anion exchange chromatogram, where line 1 is the marker, lines 2-8 correspond to the different fractions in the range of 26% of buffer B.

These seven fractions were mixed and loaded for SEC. The maximum elution volume was 150 ml, but the protein was eluted in the range of 70 ml to 85 ml. The main peak has good separation from other impurities. The sample volume was reduced using Vivaspin concentrator (MWCO 3,5 kDa) until the concentration is equal to 0.9 mM.

3.5. NMR chemical shifts of GABARAP wt and GABARAP mutant

Making the mutation of the protein can lead to changes in the structure of the protein. In the current work, the artificially created GABARAP I protein was purified. The variant of fully deuterated GABARAP-G116CΔ117 was produced and compared with the fully deuterated wild type of GABARAP to rule out the C-terminal deletion and mutation of the Gly will not change the structure of GABARAP. Figure 26 shows two overlapped NMR spectra of GABARAP and GABARAP-G116CΔ117 samples (see

section 2.4.1). [^{1}H, ^{15}N] -TROSY HSQC spectrum for wt GABARAP and [^{1}H, ^{15}N] HSQC spectrum for GABARAP-G116CΔ117. Due to specific TROSY experiments (see Chapter I, 2.4.3), the TROSY spectrum was shifted in ^{1}H and ^{15}N dimensions to compare with the HSQC spectrum. Some changes in chemical shifts were observed. These resonances are labelled by the respective sequence position. The peak of 117 Leu is absent for the mutant, and 116 Cys has a new chemical shift. The most significant chemical-shift changes can be seen for 115 Tyr 114 Val, 113 Ser, 82 Asn, 40 Arg and 38 Lys, labelled by different colours for both spectra. Nevertheless, no significant changes were observed for the whole protein structure.

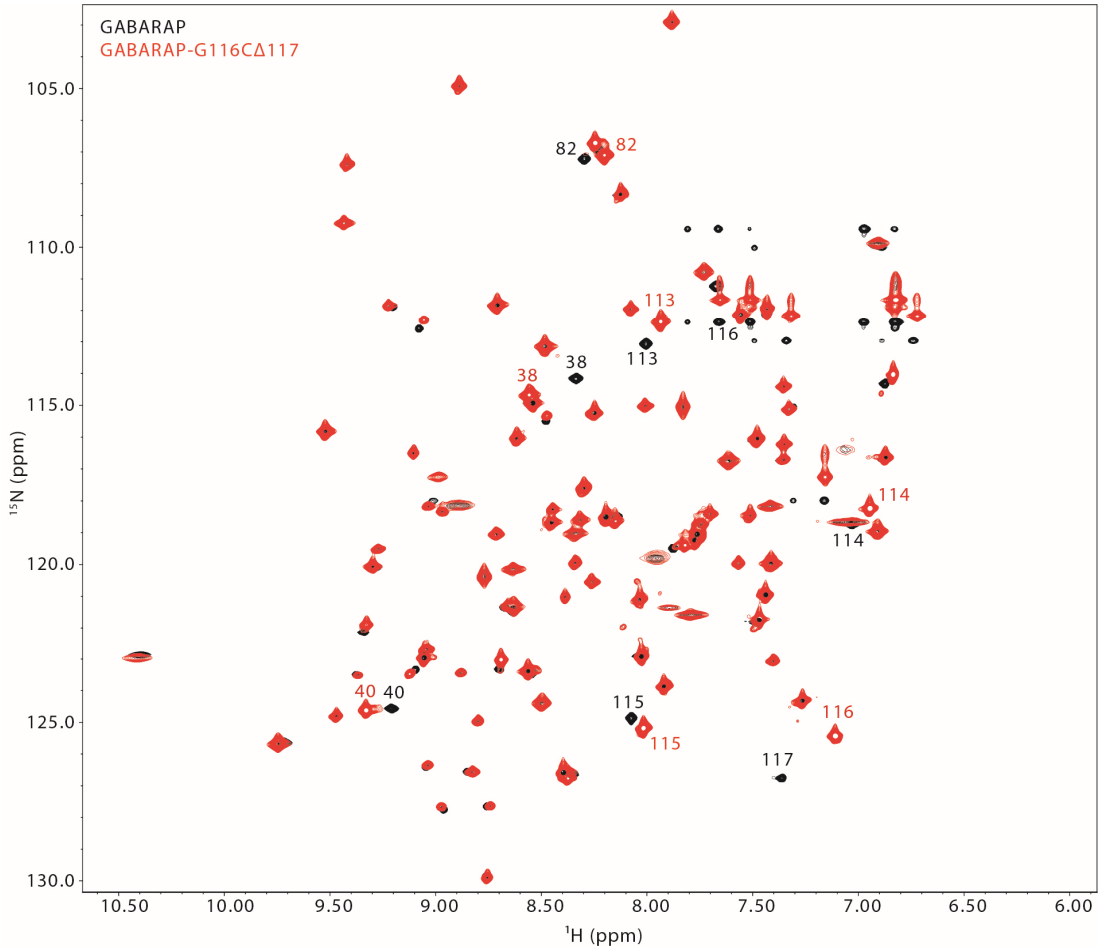


Figure 26. 2D [^{1}H - ^{15}N] TROSY-HSQC spectrum of GABARAP and [^{1}H - ^{15}N] HSQC GABARAP-G116CΔ117. 2D corrected [^{1}H ^{15}N] TROSY HSQC spectra of GABARAP (black) was recorded at 40°C at 600 MHz in NMR buffer (Table 13), which contains 117 amino acids. The overlapped 2D [^{1}H ^{15}N] HSQC spectra of GABARAP-G116CΔ117 (red) was recorded at 40°C at 700 MHz in NMR buffer (Table 13), which contains 116 amino acids and replaced 116 Gly to Cys.

3.6. GABARAP-G116CΔ117 anchored to nanodiscs

In section 2.2.3. the purification of GABARAP-G116CΔ117 was described. First, the 117 Leu was cleaved off, and Gly 116 was replaced with Cys for maleimide coupling. The protocol of lipidation of GABARAP-G116CΔ117 was described in section 2.2.4. MPB-PE molecule was coupled to the mutant to produce GABARAP-PE. The success of the coupling was checked using SDS-PAGE (Figure 27). The lipidated GABARAP-G116CΔ117 (GABARAP-PE) runs faster in the gel than the normal one, and the new line is located slightly lower in the picture. The ration 1:6 was enough for the lipidation of more than 90% of GABARAP- G116CΔ117.

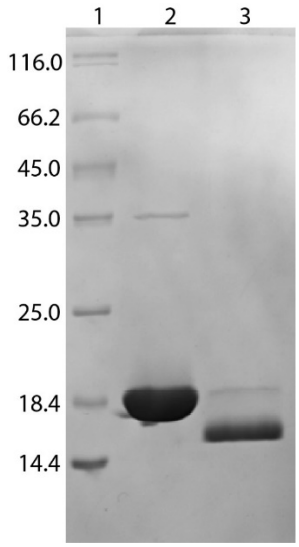


Figure 27. SDS-PAGE (15%) of lipidated GABARAP-PE. Line 1 is the marker, line 2 is GABARAP-G116CΔ117 before MPB-PE coupling and line 3 is the reaction of lipidation of GABARAP-G116CΔ117 in 1:6 ratio after 30 minutes incubation with MPB-PE (GABARAP-PE).

The assembly of GABARAP-PE anchored to ND is described in section 2.2.5. DMPC lipids were mixed with lipidated GABARAP and MSP1D1Δ5 in the ratio 53:1:1 that corresponds to two molecules of GABARAP per ND (one for each side). The size of the nanodisc was determined by the length of belt protein MSP1D1Δ5 and relates about 8.2 - 9.2 nm in diameter [35]. The molecular weight of the MSP1D1Δ5 nanodisc is approximately 95 kDa; thus, nanodisc with two molecules of GABARAP (14 kDa each) has around 123 kDa. To separate free GABARAP and GABARAP anchored to nanodisc, SEC using a HiLoad 16/600 Superdex 200 pg column was made. Figure 28 shows the SEC chromatogram of the absorption at 280 nm from the elution volume. The expected main peak was observed from 50 to 80 ml with a maximum of 70 ml. The approximate elution volume for complex GABARAP-ND was calculated by Christina Möckel [113]. The shape of the peak has a shoulder at 65 ml, around 15% of the main peak. The shoulder was not collected to the final sample after elution.

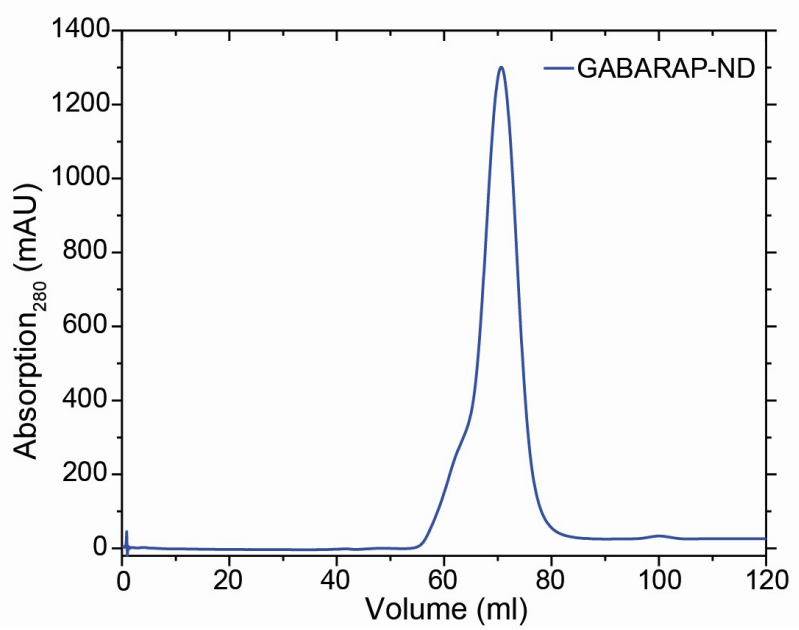


Figure 28. Size Exclusion Chromatography of lipidated GABARAP-G116CΔ117 anchored to DMPC nanodiscs. The absorption line at 280 nm of the elution profile of complex GABARAP-ND using HiLoad 16/600 Superdex 200 pg column.

GABARAP-PE and MSP1D1Δ5 have a good separation on the SDS-PAGE, which allows suggesting the ratio between them based on the line intensities. The upper line corresponds to the belt protein MSP1D1Δ5, while the lower one corresponds to the lipidated GAPARAP. Figure 29 shows the SDS-PAGE of different fractions of GABARAP-PE anchored to NDs after elution. All collected fractions contain both proteins, and the ratio of GABARAP-PE and MSP1D1Δ5 is 1:1, as was expected. Regarding the elution volume and SDS-PAGE, it can be assumed that the DMPC nanodiscs contain two molecules of GABARAP-PE.

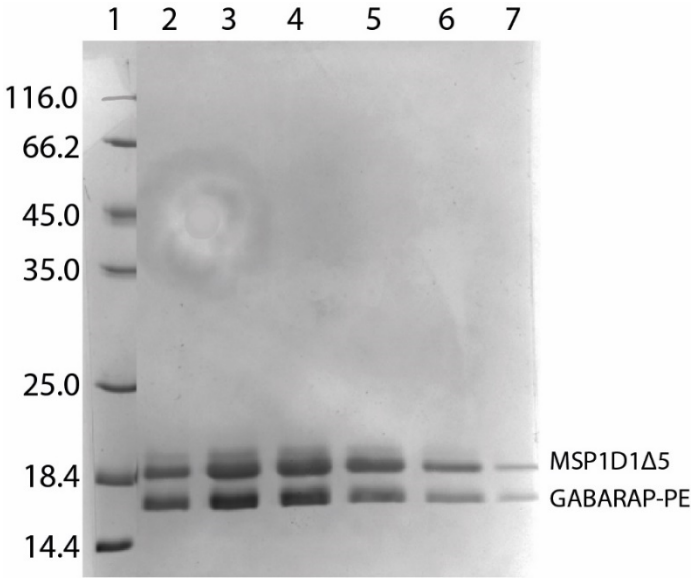


Figure 29. SDS-PAGE (15%) gel slice containing protein sample of GABARAP-PE anchored to nanodiscs. Line 1 is the marker; lines 2 to 7 correspond the six elution fractions from 68 to 80 ml of the size exclusion chromatography of GABARAP-PE anchored to nanodiscs.

3.7. Differences between free GABARAP-G116CΔ117 and GABARAP-G116CΔ117 anchored to ND by NMR spectroscopy

One of the main aims of this work is to see the changes in the structural conformation after anchoring GABARAP to nanodisc. The differences between protein in solution and protein anchored to ND can help understand how GABARAP might interact with the membrane. To compare free protein GABARAP-G116CΔ117 and GABARAP- G116CΔ117 protein anchored to ND, [^{1}H, ^{15}N] – TROSY HSQC spectrum and [^{1}H, ^{15}N] HSQC spectrum were recorded at 40°C. Figure 30 shows the overlay of the corrected TROSY and HSQC spectra of a free mutant of GABARAP and mutant of GABARAP with nanodiscs at 40°C.

Significant changes in chemical shifts can be found in the area of the C-terminus and spatially near amino acid residues. The most prominent shifts are observed for residues: Lys 38, Arg 40, Asn 82, Asp 111, Glu 112, Ser 113, Val 114, Tyr 115 and Gly 116. There are no significant differences in the N-terminal domain (1-26 amino acids) between free and membrane-anchored GABARAP.

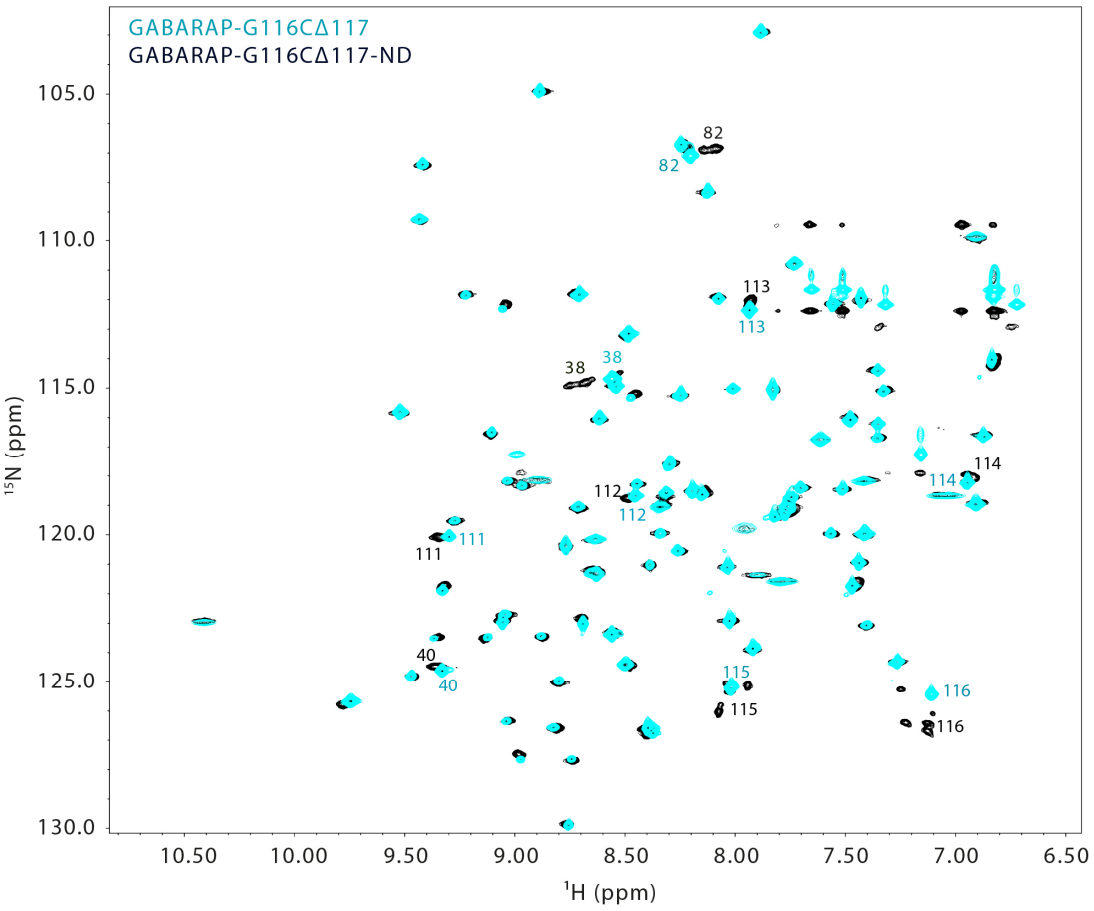


Figure 30. 2D [^{1}H - ^{15}N] HSQC spectrum of GABARAP-G116CΔ117 and 2D [^{1}H - ^{15}N] TROSY-HSQC spectrum of GABARAP-G116CΔ117 anchored to nanodisc. 2D [^{1}H ^{15}N] HSQC spectrum of free GABARAP-G116CΔ117 (blue) recorded 40°C at 700 MHz in NMR buffer (Table 13). The overlapped corrected 2D [^{1}H ^{15}N] TROSY-HSQC spectrum of GABARAP-G116CΔ117 (black) was recorded at 40°C at 600 MHz in NMR buffer (Table 13). The chemical shifts were observed for few amino acids in the C-terminal domain.

No significant changes were observed for the whole structure of GABARAP after anchoring to ND.

3.8. Differences between free wt GABARAP and GABARAP-G116CΔ117 anchored to ND by NMR spectroscopy

To see better the differences between free wild-type GABARAP and GABARAP-G116CΔ117 anchored to ND, 3D ^{15}N ^{13}C TROSY HNCO experiments were recorded (Table 17). The chemical shifts in three dimensions for each amino acid were calculated using equation 16 as described in section 2.4.5. Figure 31 shows the differences in chemical shifts of free GABARAP and GABARAP-G116CΔ117 anchored to ND in three dimensions (^{1}H, ^{15}N, ^{13}CO). As observed in the column graph, the residues with the biggest chemical shifts are close to C-terminal Ser 113, Val 114, Tyr 115, Gly 116 and two amino acids Lys 38, Arg 40 and Asn 82 in loop regions. No significant differences in the chemical shifts were observed for the N-terminal domain (1-26 amino acids) between free GABARAP and GABARAP anchored to ND. The average chemical shift for all amino acids is 0.035 ppm.

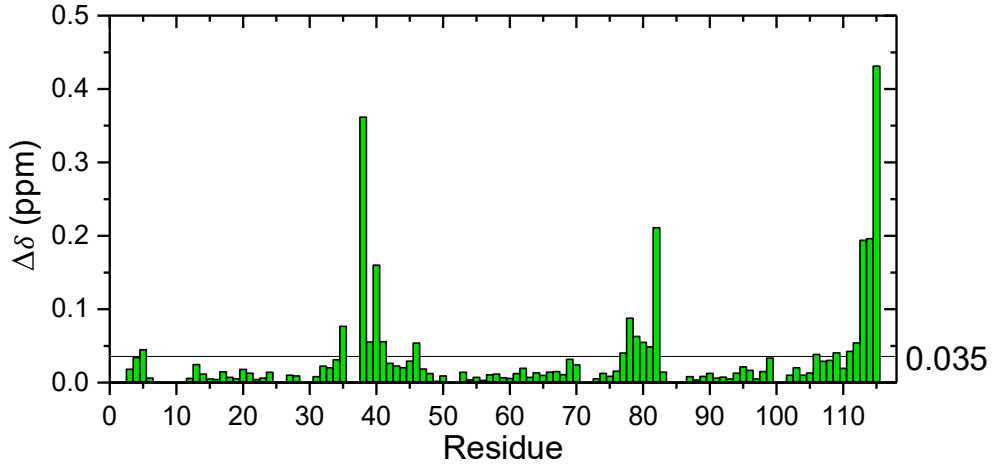


Figure 31. Weighted chemical shift (Δδ) perturbation analysis between free molecule GABARAP and GABARAP-PE anchored to ND. The Δδ values per residue between wt GABARAP and mutant GABARAP anchored to ND shows no significant changes, except for the following residues 38, 40, 82, 113, 114, 115, that are significantly above the mean value of 0.035 ppm.

All amino acids with significant chemical shifts in Figure 31 are localized on the loop region close to C-terminal, as shown in Figure 32.

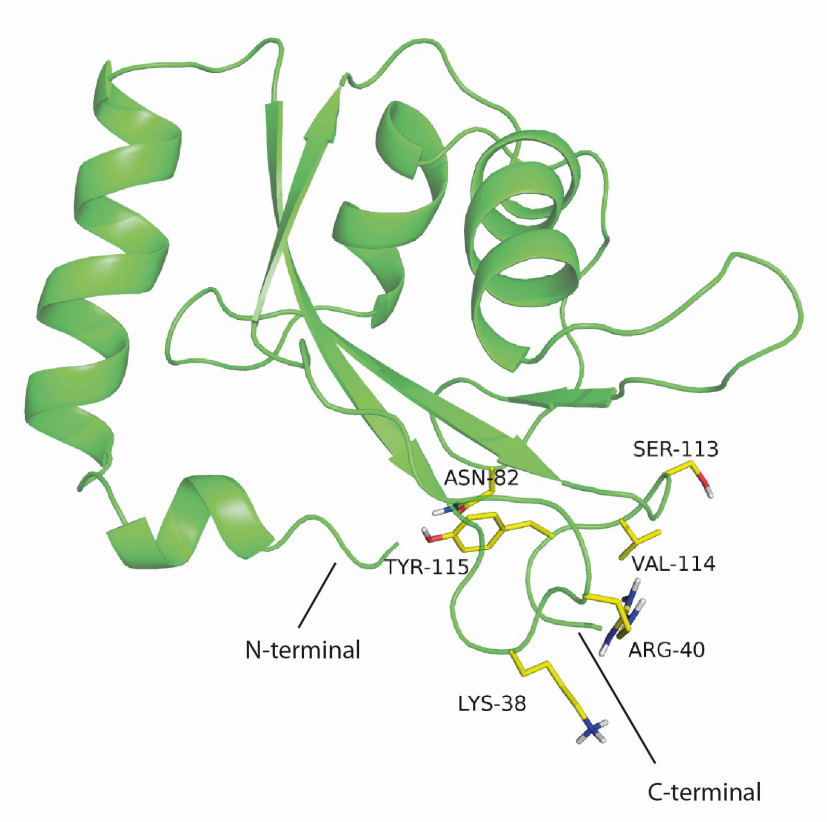


Figure 32. 3D model of GABARAP molecule. Amino acids that have the most significant chemical shifts are labelled.

3.9. Stability of GABARAP anchored to nanodisc

Relaxation experiments take a few weeks of measurements time. The stability of GABARAP-G116CΔ117 anchored to nanodiscs was checked after one month at 40°C in the spectrometer. Figure 33 shows the expanded overlapped ^{1}H-^{15}N TROSY-HSQC and ^{1}H-^{15}N HSQC spectra of the last 116 Cys amino acid anchored to DMPC nanodisc. After one month, the peak is detected at the same position, and no new peak has appeared on the position corresponding to the free GABARAP-G116CΔ117.

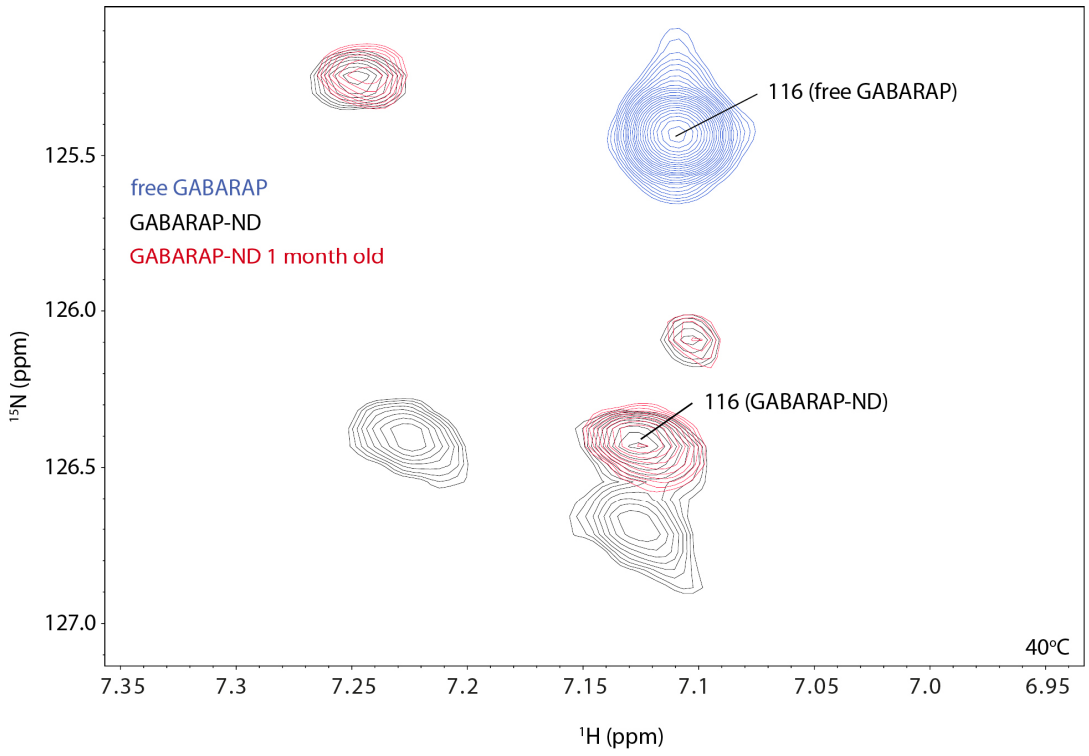


Figure 33. Expanded corrected 2D [^{1}H - ^{15}N] TROSY-HSQC and 2D [^{1}H - ^{15}N] HSQC spectra of GABARAP mutant and GABARAP mutant-ND. Resonance of Cys 116 in the overlapped NMR spectra of free GABARAP-G116CΔ117 (blue), GABARAP-G116CΔ117 anchored to ND (black) and GABARAP-G116CΔ117 anchored to ND after one month at 40°C (red) are shown here. TROSY spectra were recorded at 40°C at 600 MHz, and the HSQC spectrum was recorded at 40°C at 700 MHz.

In conclusion, GABARAP is stable on the DMPC nanodisc for one month at 40°C.

3.10. Pico- to nanosecond dynamics

Analysis of ^{15}N relaxation data recorded as described in 2.4.6 is the first step for understanding the dynamics of N-H bonds on the pico- to nanosecond time scale. The ^{15}N relaxation experiments were recorded at 40°C for two different magnetic fields. The relaxation data set contains longitudinal relaxation rates R_1, transverse relaxation rates R_2 and {^{1}H} –^{15}N NOE values at proton Larmor frequencies of 600 and 900 MHz. The data were evaluated for all amino acids in the sequence that are not overlapped or had a low signal-to-noise ratio. The determined relaxation rates R_1 and R_2 provide the information about the global rotation correlation time, while the heteronuclear {^{1}H} –^{15}N NOE data reflect the very fast internal dynamics of individual amino acid residues.

Figure 34 shows the longitudinal and transverse relaxation rates R_1 and R_2 at different magnetic field strengths. Relaxation rates were received for all available amino acids, except for overlapped or low

signal-to-noise ratio resonances and Prolines at the following positions 10, 26, 30, 37, 52, 85 and 86. Relaxation data for both magnetic fields show similar behaviour for the C-terminal part (114, 115, 116) in R_1 relaxation rates. R_2 relaxation rates show a similar trend for all amino acids. For both relaxation rates, the data of Tyr 5 and Lys 6 at 900 MHz field were absent due to a low signal-to-noise ratio. The average relaxation rate R_1 is 1.19 ± 0.01 s^{-1} for 600 MHz and 0.77 ± 0.01 s^{-1} for 900 MHz. The overall values for the R2 relaxation rate were 21.21 ± 0.35 s^{-1} for 600 MHz and 28.41 ± 0.57 s^{-1} for 900 MHz.

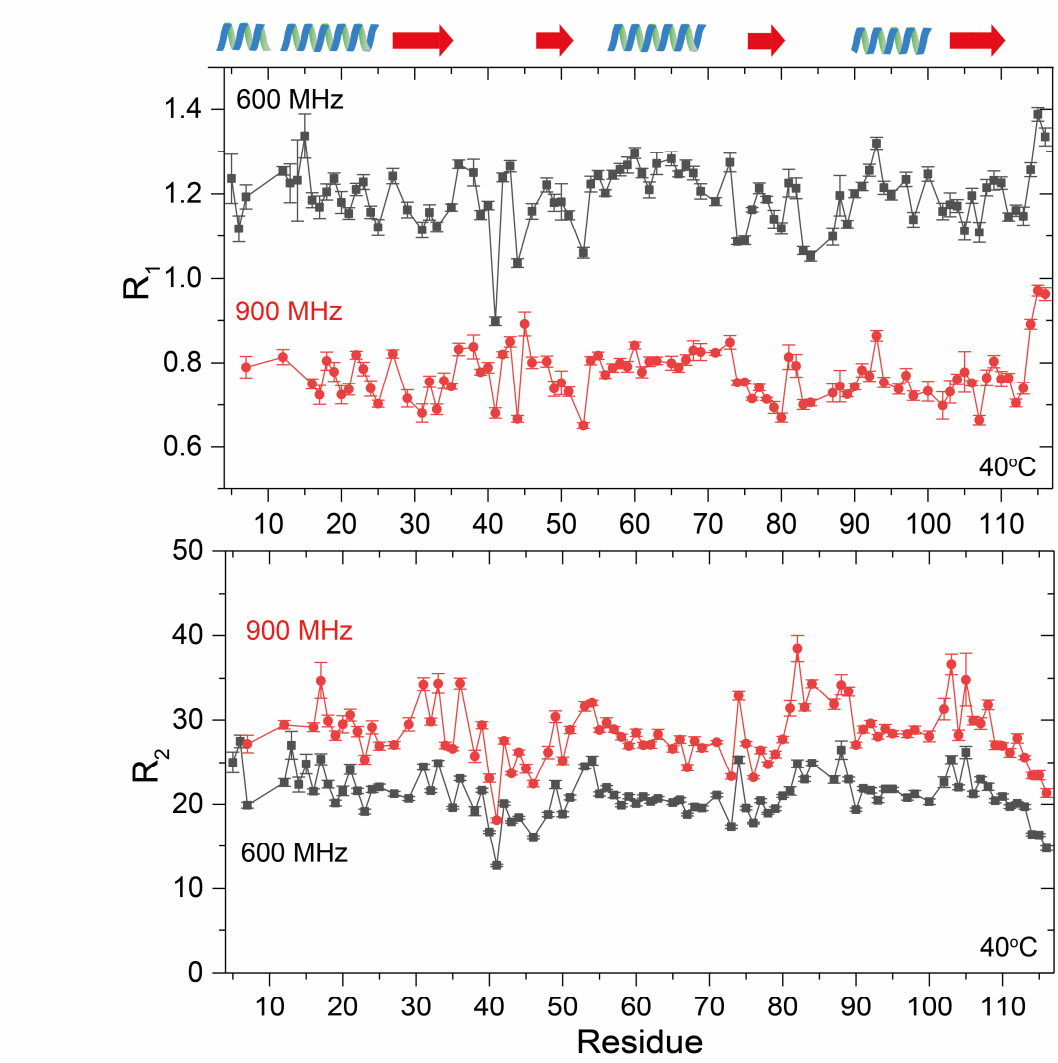


Figure 34. ^{15}N relaxation rates R_1 and R_2 of GABARAP-ND complex. Both experiments were done at 40°C for two different magnetic fields of 600 MHz (black) and 900 MHz (red). The blue spirals are related to the α-helical, and the red arrows are related β-strands of the secondary structure of the GABARAP protein.

Figure 35 shows $\{^1H\}$–^{15}N NOE for two different magnetic fields. The heteronuclear $\{^1H\}$ –^{15}N NOE values less than 0.65 is considered flexible [114]. The maximum value of NOE is 1.0, and all amino acids close to this value show the significant restrictions of the movement. C-terminal and one amino acid in a loop region around Ile 41 have the lowest NOE values, which indicates increased internal dynamics in these regions. The average NOE value is 0.72 ± 0.01 for 600 MHz and 0.80 ± 0.01 for 900 MHz.

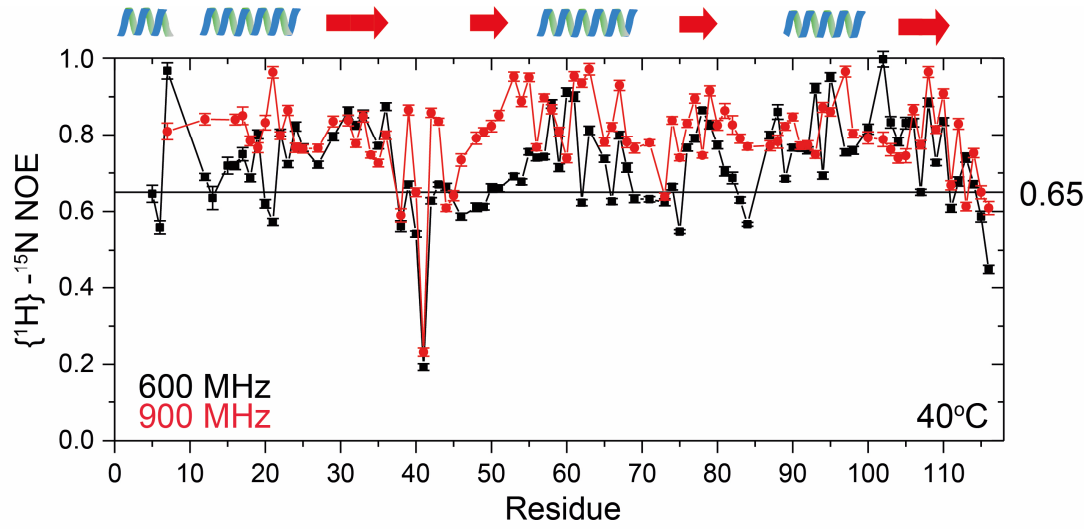


Figure 35. {^{1}H} –^{15}N NOE values of GABARAP-ND complex. The experiments were done at 40°C for two different magnetic fields, 600 MHz (black) and 900 MHz (red). The blue spirals are related to the α-helical, and the red arrows are related to β-strands of the secondary structure of GABARAP protein. The horizontal black line and the number corresponding to the value 0.65.

These relaxation data were optimized before the evaluation of backbone and sidechain dynamics. First, the {^{1}H} –^{15}N NOE values less than 0.65 were removed, also R_1 and R_2 relaxation rates deviated more than 20% from the average number were cut off. Then, the relaxation data were used for model-free analysis in the Tensor2 (Table 8). The rotation correlation time was determined for both magnetic fields and had values $12.9 \cdot 10^{-9} \pm 1.86 \cdot 10^{-11}$ s for 600 MHz and $12.7 \cdot 10^{-9} \pm 1.88 \cdot 10^{-11}$ s for 900 MHz. The values for both magnetic fields are in good agreement. With the help of equation 20 from section 2.4.7, the hydrodynamical radius of GABARAP anchored to nanodisc was calculated. Assuming a viscosity of 90% H_2O/10% D_2O mixture equal to η=0.665 mPa [115] corresponds to a hydrodynamical radius of r_H= 27.2 ± 1.1 Å.

The average order parameter S^2 = 0.89 ± 0.02 (Chapter I, 2.4.2) for 600 MHz indicates a high motion restriction of the amide bond of all amino acids in the ordered secondary structure of the protein. Figure 36 shows the order parameters for all available residue in the complex GABARAP anchored to ND. Due to the complexity of the studied GABARAP-ND complex, the case of isotropic tumbling was observed.

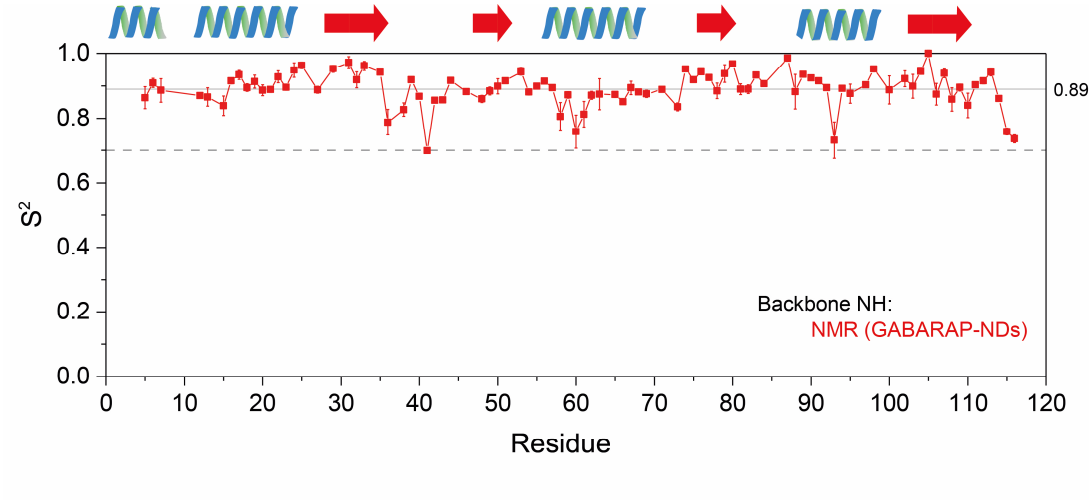


Figure 36. Order parameter S^2 was determined from ^{15}N relaxation data for 600 MHz at 40 ° C. The blue spirals are related to the α-helical, and the red arrows are related to β-strands of the secondary structure of GABARAP protein. The horizontal grey line and the number corresponding to the average value of 0.89.

4. Discussion

4.1. High yield expression of GABARAP

To demonstrate the efficiency of the new high yield protocol expression in the M9+ medium [97], we compared the yields of two different expressions using different M9 mediums. Figure 37 shows the comparison of two anion exchange chromatograms for GABARAP-G116CΔ117, which was expressed in 125 ml of new M9+ medium and 500 ml of regular M9/D_2O medium.

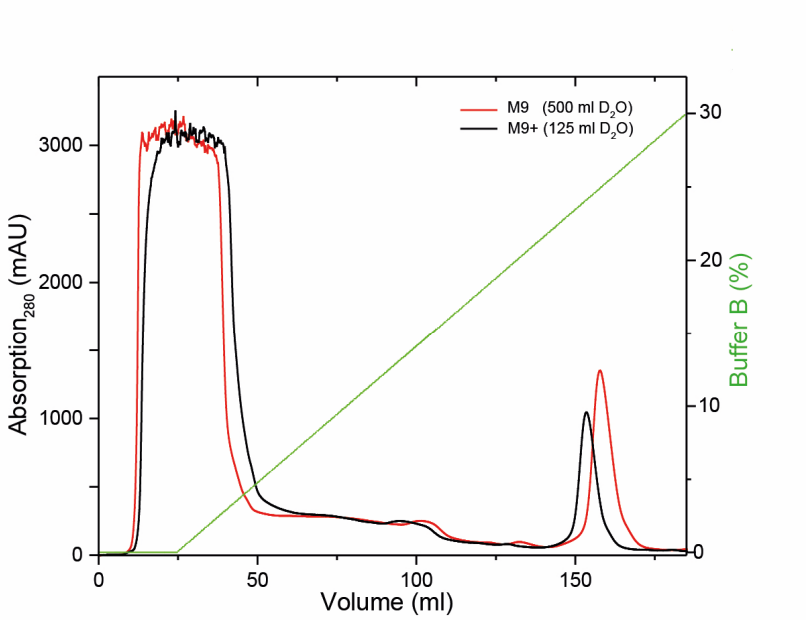


Figure 37. Anion Exchange Chromatography of deuterated GABARAP-G116CΔ117 from 125 ml of M9+ and 500 ml of M9 mediums. Black absorption line: elution profile of GABARAP-G116CΔ117 protein from 125 ml of new M9+ medium eluted from HiLoad SP Sepharose 16/10 column with 0-1 M KCl gradient over 9 CV. Red absorption line: elution profile of GABARAP-G116CΔ117 protein from 500 ml of regular M9 medium. Green line: Concentration gradient of Buffer B in %.

Figure 38 shows the Size Exclusion Chromatogram of GABARAP-G116CΔ117 absorption at 280 nm for both M9 mediums. The maximum absorption was 1700 mAU and 1100 mAU for the new M9+ medium. The amount of protein in the new M9+ medium was less by 35%.

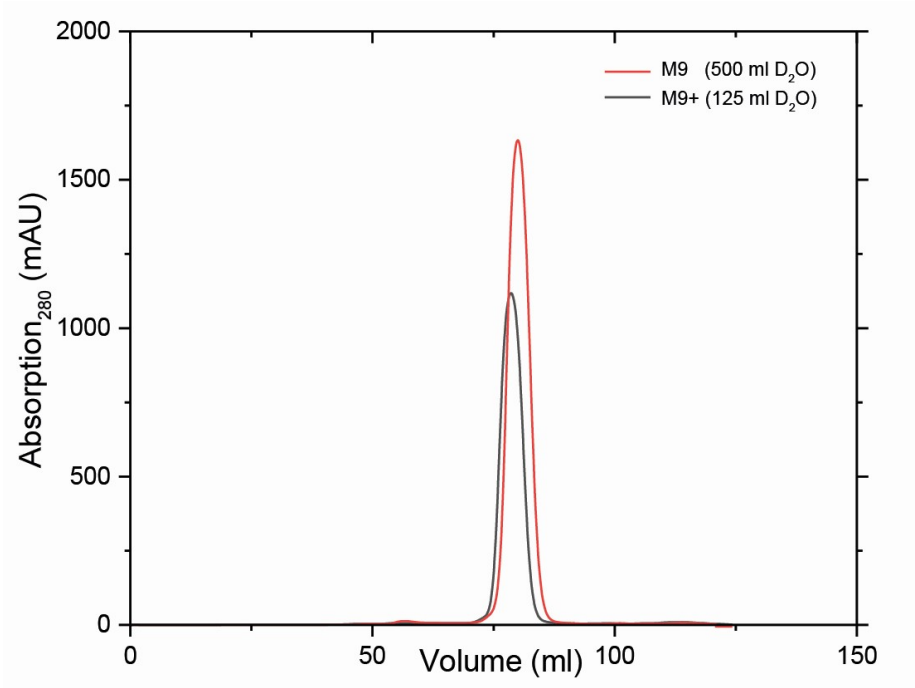


Figure 38. ***Size Exclusion Chromatography of deuterated GABARAP-G116CΔ117 from M9 and M9+ mediums.*** Red absorption line: elution profile of GABARAP-G116CΔ117 protein from 500 ml of M9 medium using HiLoad 16/600 Superdex 75 pg column. Black absorption line: elution profile of GABARAP-G116CΔ117 protein from 125 ml of deuterated new M9+ medium using HiLoad 16/600 Superdex 75 pg column.

After SEC in 10 ml buffer was purified 0.83 mg/ml for 500 ml D_2O M9 and 0.58 mg/ml for 125 ml new M9+ D_2O. M9+ medium needs four times less expensive compounds as D_2O and the same amount of D-Glucose-$^{13}C_6$ then the standard M9/D_2O medium. Based on the price calculation and quality and amount of the final deuterated protein sample, this protein expression method makes the production of fully deuterated GABARAP protein roughly five times cheaper.

Thus, the deuterated GABARAP-G116CΔ117 was purified after expression in 250 ml of the M9+ medium. This amount of protein was enough for three identical NMR samples of GABARAP anchored to ND with a protein concentration ~390 µM.

4.2. Structural comparison of GABARAP, GABARAP mutant and GABARAP mutant anchored to nanodisc

GABARAP protein has two domains: a C-terminal domain (residues 27–117) and a small N-terminal subdomain (residues 1–26) that contains two α-helixes [2]. C-terminal shows structural similarity to ubiquitin and includes a central four-stranded β sheet with two helices and loops that connects them. N-terminal is located on the opposite side from C-terminal.

Changes in chemical shifts of GABARAP and GABARAP-G116CΔ117 was described in [27], and the following chemical shift changes were observed: Lys 38, Arg 40, Val 114, Tyr 115. In the deuterated proteins, GABARAP and GABARAP-G116CΔ117 chemical shifts were observed for the same amino acids and additionally for Asn 82. These amino acids were affected upon C-terminal mutation.

To identify the nanodisc interaction site of lipidated GABARAP-G116CΔ117, chemical shifts between free GABARAP-G116CΔ117 and GABARAP-G116CΔ117 anchored to ND were compared. Some

previous results for the non-deuterated sample were described in [27]. The residues Lys 2, Lys 35, Lys 38, Val 80, Asn 82, Ser 113, Val 114 and Tyr 115 were affected upon lipidation and the following anchoring to nanodisc. In current work, chemical shifts differences of deuterated proteins were observed only for Lys 38, Asn 82, Ser 113, Val 114 and Tyr 115. Lys 2 was absent in the spectra due to low signal-to-noise. Lys 35 shows no chemical shift; the same was observed for Val 80. Additionally, chemical shifts differences were observed for Arg 40, Asp 111, Glu 112 and Gly 116. Seemingly these amino acids were affected by the presence of lipid nanodisc. The presence of chemical shift of C-terminal and loop regions around Lys 38 and Asn 82 indicates that the C-terminal and these loop regions are close to the membrane. Since GABARAP was anchored to nanodisc by lipidated C-terminus, the absence of significant chemical shift differences (apart from those mentioned above) suggests that GABARAP does not show strong interactions with the lipid nanodisc. Also, the N-terminal region was not affected by the presence of nanodisc.

GABARAP protein can exist in two distinct conformations [2]. The monomeric confirmation of free GABARAP was observed for low protein concentrations up to 100 µM when not associated with other proteins or membranes. In contrast, the oligomeric conformation of GABARAP was found in the crystal structure at high salt conditions (2.4 M ammonium sulphate). In the crystallography data [2], the first six amino acids from N-terminal bind a β2 strand of neighbouring GABARAP molecule in the crystal structure, which can explain how GABARAP participate in the autophagosome formation. In the current work, based on the chemical shifts differences of N-terminal domain between free protein and anchored to ND, GABARAP oligomerisation was not induced and stabilized by interactions with the membrane (ND) in the liquid state.

Also, note that the monomeric size of free GABARAP is equal to 21.5 - 21.8 Å at 5°C - 35°C [113]. In this work, the hydrodynamic radius of the GABARAP anchored to ND at 40°C was bigger than monomeric: 27.2 Å. This effect of increasing r_H was not discussed in the current work and needed more additional information for understanding.

4.3. Changes in pico- to nanosecond dynamics of GABARAP anchored to ND

The analysis of the ^{15}N relaxation rates R_1 and R_2 and the $\{^1H\}$-^{15}N heteronuclear NOE values were used for determining the dynamics of GABARAP anchored to ND from pico- to nanosecond time scale. The measurements were done at two magnetic fields and showed similar results between them.

Backbone ^{15}N spin relaxation rates R_1, R_2 and the $\{^1H\}$-^{15}N hetNOE of 1 mM free GABARAP at 600 and 900 MHz measured at 25.0 °C were published in [78] and are used in this work as a reference.

The overall values of $\{^1H\}$-^{15}N hetNOE values of free GABARAP protein were found around 0.70 ± 0.06 at 600 MHz and 0.80 ± 0.06 at 900 MHz that reveals a stable tertiary fold at 25°C. The average $\{^1H\}$-^{15}N hetNOE values of GABARAP anchored to ND is 0.72 ± 0.01 for 600 MHz and 0.80 ± 0.01 for 900 MHz are close to the values found for the free GABARAP. C-terminal of free GABARAP shows decreasing values for both magnetic fields suggesting that the backbone of the C-terminal region is more flexible.

The current work shows that NOE in the C-terminal region has the same values (in error range) at higher temperatures as free GABARAP. Thus, no significant internal mobility changes were observed after anchoring GABARAP to ND. Similar to free GBARAP, Ile 41 in a loop region of GABARAP-ND shows high internal mobility.

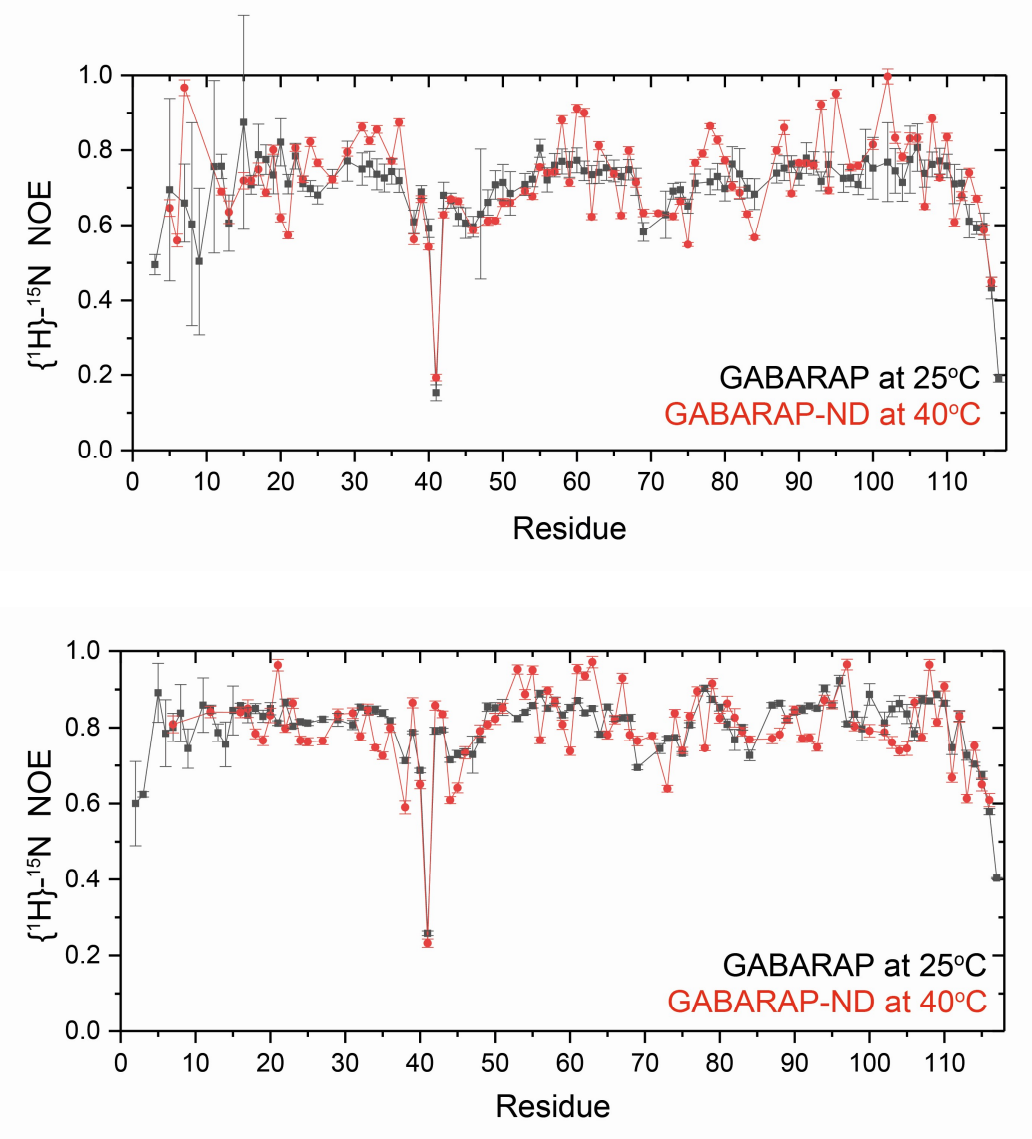


Figure 39. {^{1}H}-^{15}N heteronuclear NOE values of GABARAP (black) and GABARAP anchored to ND (red) for 600 MHz (top) and 900 MHz (bottom).

{^{1}H}-^{15}N HetNOE data of GABARAP anchored to nanodisc shows the same behaviour as free GABARAP. Changes in the temperature don't influence the very fast internal mobility of NOE values.

The average relaxation rates for different temperatures were compared between free GABARAP studied earlier [113] and GABARAP anchored to ND. The measured average relaxation rate was lower than expected for 40°C due to the presence of nanodiscs. Table 18 shows the average relaxation rated for four different temperatures of free GABARAP and one temperature of GABARAP anchored to ND.

Table 18: Dependence of the average Relaxation rates R_1 from different temperatures for free GABARAP and GABARAP anchored to ND at 600 MHz.

R_1 of free GABARAP at 600 MHz		R_1 of GABARAP anchored to ND at 600 MHz	
5°C	0.7 s^{-1}		
15°C	1.1 s^{-1}		
25°C	1.3 s^{-1}		
35°C	1.6 s^{-1}		
		40°C	1.2 s^{-1}

The flexible C-terminal region also displays significantly elevated longitudinal ^{15}N relaxation rates, R_1 the same as it was detected for free GABARAP [78].

Table 19 shows the average relaxation rate R_2. With increasing the temperature, the relaxation rate R_2 decreases, but for the GABARAP anchored to ND, R_2 is faster than expected due to the presence of the nanodisc.

Table 19: Dependence of the average relaxation rates R_2 from different temperatures for free GABARAP and GABARAP anchored to ND at 600 MHz.

R_2 of free GABARAP at 600 MHz		R_2 of GABARAP anchored to ND at 600 MHz	
5°C	25 s^{-1}		
15°C	19 s^{-1}		
25°C	15 s^{-1}		
35°C	11 s^{-1}		
		40°C	21 s^{-1}

Relaxation rates for different amino acids are shown in Figures 40 for 600 MHz data and 41 for 900 MHz data. Both relaxation rates show similar behaviour for almost all amino acids. R_1 relaxation rate didn't show any changes in C-terminal, while the R_2, which is more related to the protein's global motion, shows the significant changes for 113, 114, 115 amino acids in the C-terminal. Especially Val 114, which has a much less R_2 relaxation rate (slowly T_2 relaxation time) in free GABARAP. The same effect was observed for the 900 MHz magnetic field.

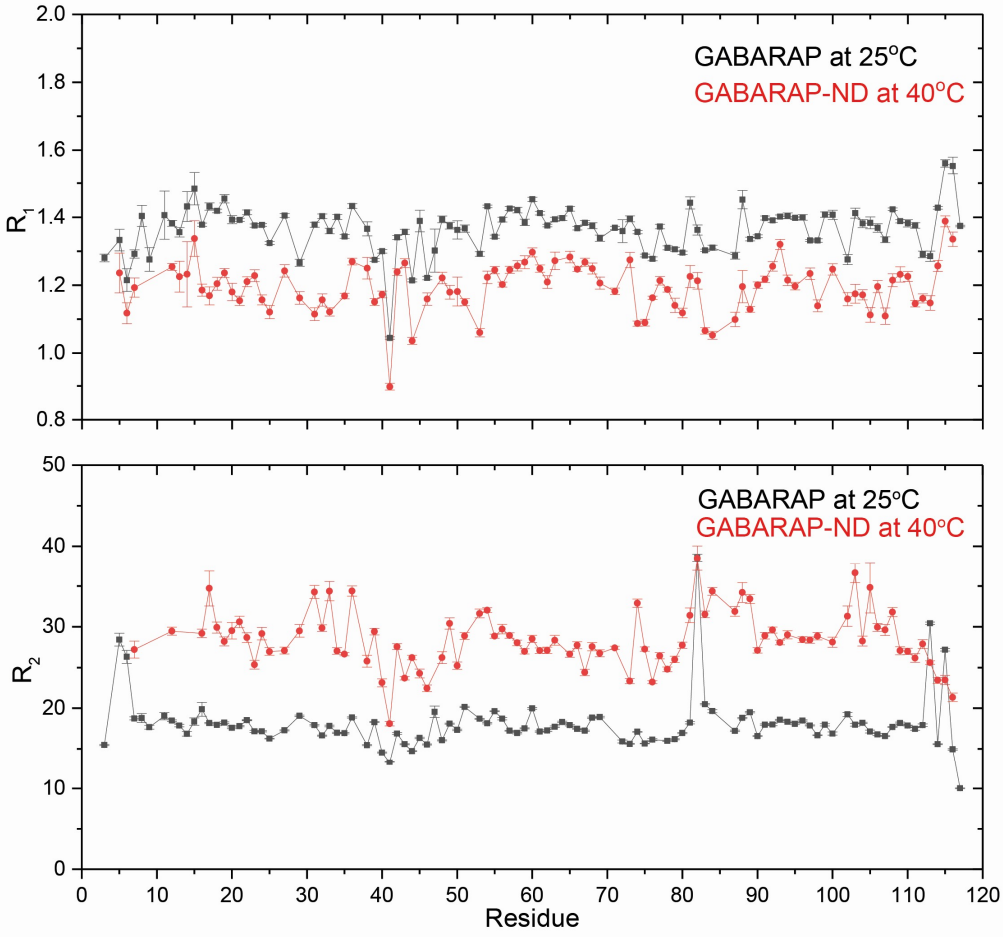


Figure 40. R_1 and R_2 relaxation rates of GABARAP (black) and GABARAP anchored to ND (red) at 600 MHz.

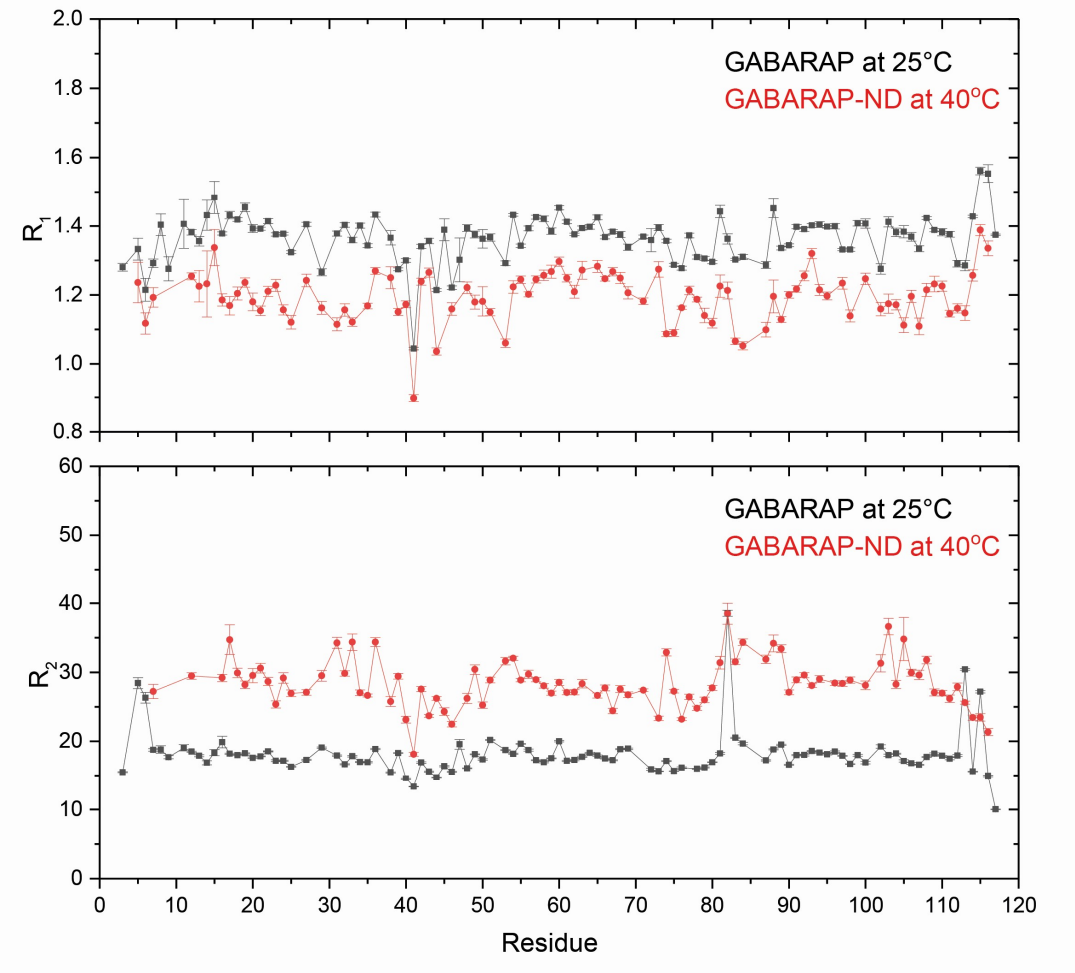


Figure 41. R_1 and R_2 relaxation rates of GABARAP (black) and GABARAP anchored to ND (red) at 900 MHz.

In the "original" model-free formalism (Chapter I, 2.4.6, equation 6*)*, the overall motions are unrestricted with the correlation function that decays exponentially to zero with a single characteristic time scale τ_m. Internal motions are assumed to be restricted because the correlation function for internal motion decays faster. Then the resulting total correlation function (Chapter I, 2.4.6, equation 5) in this model has a double exponent with the fast and slow phases representing internal and global motion. The amplitude of the global (slow) phase is characterized as the square of the order parameter (S^2) and represents the degree of spatial restriction of the backbone H-N internal motions. Reviewed in this work, order parameter was calculated with the help of only isotropic tumbling "simplified" model-free formalism from Chapter I, 2.4.6 equation 14. Nevertheless, the order parameter was compared with previously published data for free GABARAP [78] received from the anisotropic model.

In the current work, the dynamics of GABARAP anchored to ND were studied together with the research group of Prof. Dr B. Strodel (IBI-7). The NMR data were compared with MD simulation data performed by Xue Wang. Figure 42 shows the S^2 for previously published free GABARAP and new NMR and MD order parameters of the GABARAP-ND complex. The backbone order parameters S^2 from NMR spectroscopy and MD simulations primarily reveal internal motions in loop regions of Ile 41 and at the N and C-terminal, but the regular secondary structure of GABARAP show high rigidity (higher

than 0.7). GABARAP anchored to ND shows less mobility for Ile 41 and C-terminal amino acids in both experiments NMR and MD simulation.

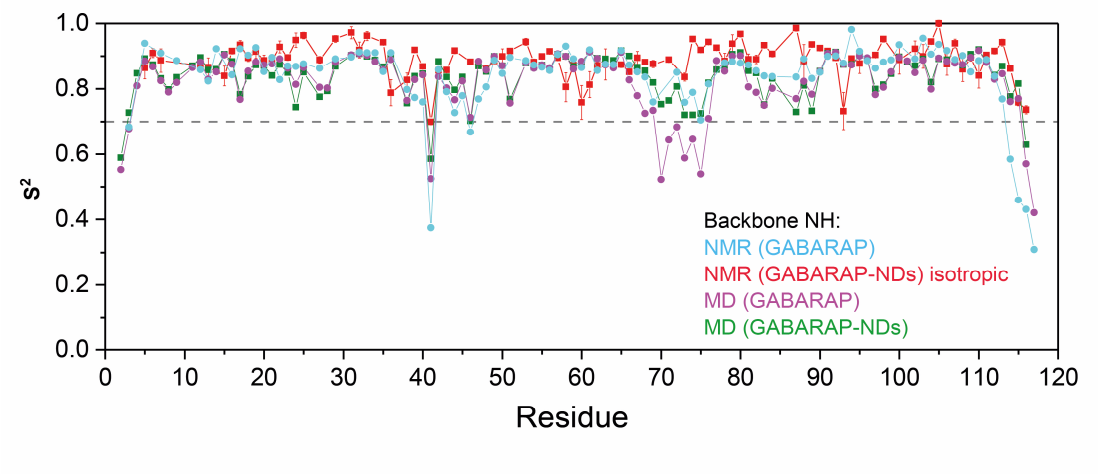


Figure 42. Backbone flexibility of free GABARAP and GABARAP anchored to ND. Comparison of S^2 free GABARAP from 15N relaxation data (blue), free GABARAP from MD simulation (violet), GABARAP-ND from 15N relaxation data in this thesis (red) and GABARAP-ND from MD simulation (green).

One more parameter from relaxation data is correlation time τ_c, which was received from the TENSOR2 program (Table 8) was compared with a correlation time of free GABARAP at different temperatures [113]. Table 20 shows the dependence of correlation time on the temperature.

Table 20: Dependence of the correlation time τ_c from different temperatures for free GABARAP and GABARAP anchored to ND at 600 MHz.

τ_c of free GABARAP at 600 MHz		τ_c of GABARAP anchored to ND at 600 MHz	
5°C	16.9 ± 0.1 ns		
15°C	12.3 ± 0.1 ns		
25°C	9.5 ± 0.1 ns		
35°C	7.5 ± 0.1 ns		
		40°C	12.9 ± 0.1 ns

The correlation time τ_c of free GABARAP decreased with increasing the temperature due to faster rotation at high temperature, but τ_c of GABARAP anchored to ND has a slow value because it is tumbling together with the ND.

The general conclusion for the GABARAP project are observed in Chapter IV.

Chapter III Lipoprotein

Structure Calculation of Lipoprotein CD1348 from *Clostridium difficile* by Solution NMR

Lipoprotein is a membrane protein encoded directly in front of the CprABC operon of the resistance system in bacteria *Clostridium difficile*. The structure of this protein is scarcely known and will be considered in this chapter. NMR is one of the main techniques that can provide a complete description of the protein structure at the atomic level. This chapter describes methods for NMR spectra assignments of fully labelled $^{15}N,^{13}C$ lipoprotein, and structure calculation algorithms. Sequence-specific assignments for backbone resonance is obtained from TROSY 2D and 3D triple-resonance experiments (HNCA, HN(CO)CA, HNCACB, HN(CO)CACB, H(CCO)NH, C(CO)NH). NMR structure determination was based on distances between ^{1}H-^{1}H and dihedral angles. The distances were determined from NOE intensities of peaks from 3D ^{15}N or ^{13}C NOESY spectra. The additional geometrical information about distance restraints was derived from the 4D $^{13}C,^{13}C$ NOESY experiment, avoiding several chemical shifts overlap. Also, titration experiments with lantibiotic gallidermin show chemical shifts differences in the loop region of lipoprotein. This loop region of lipoprotein might play a role of a binding site and thereby lowers the concentration of the lantibiotic reaching the membrane. The observed interaction with lantibiotics can help understand the function of the lipoprotein in the resistance machinery of *Clostridium difficile*.

1. Materials

1.1. NMR spectrometers

Liquid state NMR spectrometers operating at different magnetic fields were used for obtained NMR data for assignment, structure calculation and titration experiments.

Table 21: NMR spectrometers were used for the Lipoprotein project.

Spectrometers	Manufacturer
Bruker Avance III HD NMR (600 MHz)	Bruker, Billerica, USA
Bruker Avance III HD NMR (600 MHz)	Bruker, Billerica, USA
Bruker Avance III HD NMR (700 MHz)	Bruker, Billerica, USA
Variant (800 MHz)	Agilent (Varian), Santa Clara, USA
VNMRS (900 MHz)	Agilent (Varian), Santa Clara, USA

1.2. Software and databases

Table 22 lists all software and databases used for assignment, structure calculation, and visualization of the lipoprotein structure.

Table 22: Software and databases used for the Lipoprotein project.

Software/database	Usage	Reference/distribution
BMRB Databank	Open databank of chemical shifts of biomolecules	http://www.bmrb.wisc.edu [82]
Expasy-ProtParam	Analysis of proteins based on their primary sequence	http://web.expasy.org/protparam [83]
NMRPipe	Processing of NMR spectra	[84]
NMRDraw	Visualisation of processed MNR spectra	[84]
NMRViewJ 8.0.3	Visualisation and analysis of NMR spectra	[85]
PDB Database	Structures of biological macromolecules	http://rcsb.org
PyMOL	Visualisation of protein structures	
RasMol 2.7.5	Visualisation of protein structures	[88]
TALOS-N	Prediction of torsion angles based on chemical shifts of the protein	[70, 116]
AssignNOE	Assign and determination the distance between NOE	[117]
Xplor-NIH	Determination of molecular structure based on NMR data and crystallography	[118]
PROCHECK-NMR	Checking the quality of predicted NMR	[119]

	structures	
TopSpin	Basic software of Bruker NMR spectrometers	
VnmrJ	Basic software of Bruker NMR spectrometers	
Adobe Illustrator CS5	Generation of figures	
Origin 2017	Chemical shifts calculation and visualisation of relaxation data	

2. Methods

2.1. Nuclear magnetic resonance spectroscopy

All NMR measurements were carried out using NMR spectrometers from Varian and Bruker. Detailed specifications of the spectrometers can be found in Table 23.

Table 23: NMR spectrometers are used in this work.

Spectrometer	Frequency	NMR CRYO Probe
Bruker (Magnex)	600 MHz	5 mm H-P/C/N-D
Bruker (Oxford)	600 MHz	5 mm H-/C/N-D
Bruker	700 MHz	5 mm H-/C/N-D
Varian	900 MHz	5 mm H-/C/N-D
Varian	800 MHz	5 mm H-/C/N-D

2.1.1. NMR sample

The studied NMR samples were prepared at the University of Düsseldorf by Dr Rebecca Clemens, Dr Lothar Gremer and Dr Sander Smits. The samples contained 630 µM [U-^{15}N] or 630 µM [U-^{13}C,^{15}N] (His)10-lipoprotein in NMR buffer (100 mM NaCl, 5 mM NaN_3, 25 mM MES (pH 6.5)). The sample is contained 10% (v/v) D_2O to lock the magnetic field/frequency. The samples were used as received.

For studying the binding of the lantibiotic gallidermin with lipoprotein, the 630 µM [U-^{15}N] (His)10-lipoprotein in NMR buffer and 10% of D_2O was titrated with five different rations of concentration gallidermin to lipoprotein: 0.125:1, 0.25:1, 0.5:1, 0.75:1, 1:1. The stock solution of gallidermin was provided by Dr Jens Reiners.

2.1.2. NMR spectra

NMR experiments were performed at 30°C on the spectrometers operating at ^{1}H frequencies of 600, 700, 800 and 900 MHz (Table 23). The sample temperature was calibrated using a deuterated methanol sample methanol-d_4 as describes in [102]. Sequence-specific assignments for the backbone resonances were obtained from TROSY [103, 104, 120] versions of the following 2D, 3D and 4D triple resonance experiments [68, 121]. Dr Philipp Neudecker did the optimization of pulse sequences and the measurements. The resolution of the recorded data was increased by the processing of the spectra with the program NMRPipe (Table 22). The data were visualised and analysed in NMRDraw and converted to NMRViewJ format by Dr Philipp Neudecker. All spectra were analysed with the NMRViewJ program (Table 22).

Some selected acquisition parameters such as the spectral width (SW) for all dimensions, number of points T (complex), which is twice the value of the corresponding time-domain size, offset is a proper chemical shift referencing in all dimensions and number of scans (NS) are listed in Table 24.

Table 24: *NMR experiments. The first column is the number of the experiment; the second is the name of the experiment, then the selected acquisition parameters and references for used pulse programs.*

N	Experiment	Spectrometer	SW_1 (ppm)	T_1 (complex)	offset ^1H (ppm)	SW_2 (ppm)	T_2 (complex)	$offset_2$ (ppm)	SW_3 (ppm)	T_3 (complex)	$offset_3$ (ppm)	NS	References
1	2D ^{15}N-TROSY-HSQC	700 Bruker	16	1536	4.7	30	256	117.0 (^{15}N)	-	-	-	24	[103], [104] [122], [123], [124], [125], [126]
2	2D ^{15}N-HSQC	700 Bruker	16	1024	4.7	30	256	117.0 (^{15}N)	-	-	-	24	
3	2D ^{13}C CT-HSQC	600 Bruker	16	1024	4.7	40	82	128 (^{13}C)	-	-	-	160	
4	3D ^{15}N,^{13}C-TROSY-HNCO	700 Bruker	16	1024	4.7	30	48	117.0 (^{15}N)	12	32	176.0 (^{13}C)	12	
5	3D ^{15}N,^{13}C-TROSY-HN(CO)CA	700 Bruker	16	1024	4.7	30	48	117.0 (^{15}N)	26	46	56.0 (^{13}C)	16	
6	3D ^{15}N,^{13}C-TROSY-HN(CO)CACB	700 Bruker	16	1024	4.7	30	48	117.0 (^{15}N)	52	65	43.0 (^{13}C)	16	
7	3D ^{15}N,^{13}C-TROSY-HNCA	700 Bruker	16	1024	4.7	30	48	117.0 (^{15}N)	26	46	56.0 (^{13}C)	16	
8	3D ^{15}N,^{13}C-TROSY-HN(CA)CO	700 Bruker	16	1024	4.7	30	48	117.0 (^{15}N)	12	32	176.0 (^{13}C)	16	
9	3D ^{15}N,^{13}C-TROSY-HNCACB	900 Varian	16	1024	4.7	30	64	43		48	117.0 (^{15}N)	16	
10	3D ^{15}N,^{13}C-H(CCO)NH-TROSY	600 Bruker	16	1024	4.7	30	44	117.0 (^{15}N)	154	128	82.0 (^{13}C)	16	
11	3D ^{15}N,^{13}C-C(CO)NH-TROSY	600 Bruker	16	1024	4.7	30	44	117.0 (^{15}N)	154	128	82.0 (^{13}C)	16	
12	3D ^{15}N ^{13}C – H(C)CHTOCSY	600 Bruker	16	512	4.7	30	96	4.7	12	32	77.0 (^{13}C)	8	
13	3D ^{15}N-TOCSY-HSQC	700 Bruker	16	1024	4.7	30	48	117.0 (^{15}N)	14	128	4.7 (^{1}H)	16	
14	3D ^{15}N-NOESY-TROSY	800 Varian	16	1024	4.7	30	128	4.7 (^{1}H)		48	117.0 (^{15}N)	16	
15	3D ^{15}N-NOESY-HSQC	600 Bruker	16	1024	4.7	30	48	117.0	13	128	4.7 (^{1}H)	16	
16	3D ^{15}N, ^{13}C-NOESY-HSQC	700 Bruker	16	1024	4.7	30	48	117.0 (^{15}N)	70	64	43.0 (^{13}C)	16	
17	3D ^{13}C-NOESY-HSQC	700 Bruker	16	1024	4.7	30	46	77.0 (^{13}C)	14	128	4.7 (^{1}H)	16	
18	4D ^{13}C ^{13}C NOESY HSQC	600 Bruker	16	1536	4.7	30	36	77.0 (^{13}C)	12	128	4.7 (^{1}H)*	4	

* One more additional plane for 4D experiment

SW_4 (ppm)	T_4 (complex)	$offset_4$ (ppm)
30	36	77.0 (^{13}C)

2.1.3. Assignment of resonances

One of the most useful functions of solution NMR in structural biology is its ability to determine the three-dimensional structures of proteins, especially which one is hard to crystallize, as in the case of lipoprotein CD1348. The assignment of resonances in the NMR spectra is one of the first steps in obtaining the protein structure. This step involves few methods of sequential assignment of systems of spin-spin coupled resonances. The sequence-specific assignment was obtained from J-correlated 3D/4D resonance experiments (Table 24). This process involves the assignment of one amino acid with its sequential neighbours by linking backbone nuclei. Experiments 4-9 from the Table 24 were used to measure HNCO, HN(CO)CA, HNCA, HN(CA)CO, CACB(CO)NH and HNCACB spectra to assign the peaks in (^{1}HN, ^{15}N, $^{1}H_{\alpha}$, $^{13}C_{\alpha}$, $^{13}C_{\beta}$, ^{13}C') in the backbone. The C(CO)NH and H(CCO)NH spectra (experiments 10-11 in Table 24) were used to assign the nuclei in the aliphatic side chains (^{1}H, ^{13}C), which Bettina Wagner did during her bachelor work [127]. The 3D HCCH-TOCSY spectrum (experiment 12, Table 24) was used to assign the chemical shifts of the aromatic rings ($^{1}H_{x}$, $^{13}C_{x}$). The 2D ^{15}N HSQC spectrum (experiment 2, Table 24) was used to assign the peaks for the sidechain HN_2 groups of Asn and Gln.

The resonances in the [U-^{13}C,^{15}N] (His)10-lipoprotein spectrum were assigned manually using the program NMRViewJ (Table 22).

The 2D ^{15}N-TROSY-HSQC spectrum generally represents a “fingerprint” of the protein. It contains information about all backbone amides of each residue, except for Pro. Each amino acid's 1H and 15N chemical shifts were assigned using triple resonance multidimensional NMR experiments and then labelled on the 2D spectrum. Another 2D spectrum, the ^{1}H-^{13}C-CT-HSQC spectrum, is a “fingerprint” of the carbon groups of the protein. It shows the peaks for the $^{13}C_{\alpha}$, $^{13}C_{\beta}$ backbone and the directly bound protons $^{1}H_{\alpha}$ and $^{1}H_{\beta}$. Spectral assignments for this spectrum were also obtained by analysing the 3D experiments, as described in the following.

The graphical representation of all triple-resonance experiments is shown in Chapter I, 2.2, Figure 11. The experiments such as HNCA and HN(CO)CA correlate a backbone ^{1}HN and ^{15}N chemical shifts with one or more $^{13}C_{\alpha}$ chemical shifts. The HNCA experiment correlates proton and nitrogen chemical shifts with two $^{13}C_{\alpha}$ chemical shifts of its residue *i* and the previous one *i-1*. The HN(CO)CA experiment transfers magnetisation between the HN group to the $^{13}C_{\alpha}$ (*i-1*) residue with coherence transfer via the preceding ^{13}CO. Combing these two experiments, therefore, allows identifying the $^{13}C_{\alpha}$ (*i*) and $^{13}C_{\alpha}$ (*i-1*) peaks.

Additionally, HNCACB and CACB(CO)NH were used to identify $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts of the intra (*i*) and inter (*i-1*) nuclei in the backbone. The HNCACB experiment correlates the ^{1}HN and ^{15}N chemical shifts of its residue (*i*) with $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts of *i* and *i-1* residues. In contrast, the CACB(CO)NH experiment provides correlations only between amino group chemical shifts of a residue *i* with $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ shifts of the previous residue (*i-1*). The resonances received from $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ nuclei have a 180° phase difference on the spectrum where $^{13}C_{\alpha}$ resonances have positive peak

intensity, and the other ($^{13}C_\beta$) is negative at the same spectrum. Since the $^{13}C_\beta$ is recorded alongside the $^{13}C_\alpha$ nucleus, it reduces the ambiguities in the assignment of $^{13}C_\alpha$ in the HNCA spectrum. The intensity of peaks depends on the ^3J coupling constant (Chapter I, 2.2, Figure 10); therefore, the signal obtained from the inter-residue correlations between the amide group and the carbons of the previous amino acid (*i-1*) has a weaker intensity than the signal from the intramolecular correlation (*i*).

The HNCO and HN(CA)CO spectra were used to assign the backbone carbonyl ^{13}CO group. The 3D HNCO experiment correlates ^1HN and ^{15}N chemical shifts of residue *i* with the carbonyl group of residue *i-1*. The HN(CA)CO experiment was used to detect the ^{13}CO chemical shift of both residues *i* and *i-1* via transfer magnetization through $^{13}C_\alpha$ of *i* and *i-1* residues.

The 3D C(CO)NH and H(CCO)NH experiments are used to assign the carbon and proton nuclei in the side chains. In the 3D triple resonance C(CO)NH experiment, all aliphatic carbons correlate with the amide group of the preceding residue *i–1* via transferred magnetisation of the ^{13}CO group. In the H(CCO)NH experiment, magnetisation is transferred from the side-chain hydrogen to their attached carbon group and then via the carbonyl group to the amide group of the previous residue *i-1*.

The 3D HCCH-TOCSY NMR experiment was used to complete ^1H_x and $^{13}C_x$ assignments of the aromatic rings. The magnetisation is transferred from the side chains protons to the nuclei of the $^{13}C_x$ atoms they are attached to. In the 3D HCCH-TOCSY spectrum, ^1H_x-^1H_x cross-peaks are spread out in the third dimension according to their $^{13}C_x$ chemical shifts. For assignment of side chains on this spectrum, assignments of $^{13}C_\alpha/^{13}C_\beta$ and $^1H_\alpha/^1H_\beta$ obtained from backbone assignment experiments were used as the starting point.

The 2D HSQC spectrum has been used to assign the Trp side chain H_ε-N_ε groups, Asn $H_{\delta2}$-$N_{\delta2}$ and Gln $H_{\varepsilon2}$-$N_{\varepsilon2}$ side-chain groups. The Asn and Gln side chain NH_2 groups correspond to double peaks in the N dimension but for two different proton shifts. The Trp side chain H_ε-N_ε peaks usually have specific proton (around 10 ppm) chemical shifts and are located separately on the spectrum.

Statistically calculated chemical shifts from atoms in all amino acids were used as a reference for assigned lipoprotein. These chemical shifts are available from the biological magnetic resonance data bank (BMRB) (Table 22).

2.1.4. Determination of torsion angle restraints

The NMR spectra contain all necessary information about interatomic distances and angular geometries used for the structure determination of the protein [128]. The dependence of the isotropic chemical shifts on the local backbone geometry determines protein torsion (dihedral) angle restraints [129]. Backbone and sidechain torsion angle restraints used for the structure calculation were determined using TALOS-N (Torsion Angle Likelihood Obtained from Shifts) program (Table 22). This program predicts protein backbone torsion angles ϕ and ψ and sidechain torsion angle χ1 (Chapter I, 2.3, Figure 13) from NMR chemical shifts of ^1HN, ^{15}N, $^{13}C'$, $^{13}C\alpha$, $^{13}C_\beta$ and $^1H_\alpha$ nuclei [116]. TALOS-N is an artificial neural network trained with curated data on chemical shifts for specific amino acids and PDB coordinates from high-resolution X-ray structures with torsion angle information. The

program classifies chemical shifts and amino acid sequence data of a particular secondary structure unit (helix, sheet, and coil).

For the backbone torsion angle prediction, TALOS-N uses prediction of 324-state ϕ/ψ distribution of each residue based on the chemical shift, which amino acid and place in a sequence. This 324-state ϕ/ψ distribution corresponds to the residue on the Ramachandran map [130]. Then the program searches for 1000 heptapeptide fragments from the database with ϕ/ψ angles that best match the 324-state ϕ/ψ distribution. The 25 best-matched database heptapeptides and the averages of ϕ/ψ angles of their centre residues are assigned as the backbone torsion angles of the centre residue of the query heptapeptide [116]. The prediction is shown in the program directly.

Besides prediction of ϕ/ψ angles, TALOS-N also was used for prediction of the sidechain χ1 torsion angles. First, TALOS-N searches the database for the 1000 best-matched heptapeptides from ϕ/ψ torsion angles and residue types [70]. There are three possible χ1 rotameric states of all sidechain types (except Pro): *gauche (+), trans or gauche (-).*

2.1.5. NOE resonance assignment and extraction of distance restraints from NOE chemical shifts

The interproton distance obtained from NOE data is based on a comparison of NOE peak intensities for pairs of spins in NOESY experiments [131]. The through-space interactions were observed using 3D NOESY experiments which were made for both [U-^{15}N] or [U-^{13}C,^{15}N] (His)10-lipoprotein samples and listed in Table 24, № 14-18. The 3D ^{15}N and ^{13}C NOESY-TROSY spectra were recorded to assign side-chain protons of intra and intermolecular residues. The mixing time used for the ^{15}N NOESY-TROSY experiment was 120 ms. Additionally, 3D-^{13}C,^{15}N HSQC-NOESY-HSQC experiment to obtain the intermolecular distance information was used. The mixing time was set to 150 ms.

Verena Kienapfel did the automatic assignment of NOEs and NOE-derived distances from these experiments during her bachelor work [132]. AssignNOE program was used for chemical shifts assignment and determination of the distances between possible coupling partners [117].

The 4D NOESY experiment was used to solve the problem of overlapping some cross-peaks and less resolution. The 4D ^{13}C, ^{13}C HSQC-NOESY-HSQC experiment includes one additional ^{13}C dimension (Table 24, number 18). NOEs between side-chain protons were spread out by the chemical shifts of the directly bounded ^{13}C atoms. Peaks were assigned manually using NMRViewJ (Table 22).

2.2. Protein structure determination

Protein structure determination is based on energy minimization methods. Native protein structures correspond to a system at thermodynamic equilibrium with minimum free energy. The 3D structure determination from assigned NOE data was performed in the structure generation software program X-PLOR-NIH (Table 2). X-PLOR-NIH program uses the NOE statement sets for determination and refines solution NMR structures based on throw space distance estimates, coupling constant measurements and others. The following information about molecular structure, NOE-derived

distance bounds, and coupling-constant-derived dihedral angle restraints was used for the NMR structure determination. The program was run with a target radius of gyration of 13.8 Å [133].

The input files with the experimental data such as NOE-derived distances, torsion angles and residual dipolar couplings (RDCs) were run 22 times. Every next run was updated with some new data in input files. Distance restraint files were added separately from 4 different sources:

Table 25: Assigned spectra that were used for 3D structure calculation

3D ^{15}N NOESY-TROSY	3D ^{15}N NOESY-TROSY	3D ^{13}C NOESY-HSQC	4D ^{13}C ^{13}C HSQC NOESY HSQC
Automatic + manual assignment (AssignNOE + NMRViewJ) by Verena	Manual assignment (NMRViewJ)	Manual assignment (NMRViewJ)	Manual assignment (NMRViewJ)

Since the NOE assignment and structure calculation require around 90% completeness of the chemical shift assignment, which is not straightforward to achieve by automated peak picking and automated resonance assignments [134], a manual assignment was added for completing these data.

The quality of the calculated 3D structure of the lipoprotein was validated using dedicated software as PROCHECK_NMR (Table 22). Backbone and sidechain torsion angles were analyzed using the program PROCHECK_NMR for calculating the "stereochemical quality" of a given protein structure. This program was used for ten models with the lowest energy structures. The RMSD of atomic positions measured the average distance of superimposed ten models with the lowest energy structures.

2.3. Chemical shift perturbation analysis

After the resonances are assigned and the lipoprotein structure is identified, the interaction strength between the nuclei can be studied. NMR spectroscopy provides a powerful tool to study interactions between proteins or protein and ligand at atomic resolution. In particular, the localisation of the binding sites on the protein surface can be determined based on the chemical shift perturbation (CSP) analysis. It is performed analogously to the case of GABARAP as described in Chapter 2, 2.4.5. In the current research, the CSP was used to calculate chemical shifts difference in the lipoprotein in the presence of lantibiotic gallidermin at five different concentrations. The processed spectra were superposed to identify the ^{1}H–^{15}N correlation peaks subjected to the most significant frequency shifts upon increasing the concentration of lantibiotic. A peak-picking on each spectrum was done by using the program NMRViewJ (Table 22). The chemical shifts differences for each amino acid were calculated using:

$$\Delta\delta_{ave} = \sqrt{{\Delta\delta_{HN}}^2 + \left(\frac{\Delta\delta_N}{10}\right)^2} \qquad (22)$$

Where $\Delta\delta_{HN}$ and $\Delta\delta N$ are the difference between the proton and nitrogen chemical shifts, which were measured with a given amount of gallidermin and without it respectively.

3. Results

3.1. Assignment of backbone resonances of lipoprotein

The sequential assignment of the following nuclei ^{1}HN, ^{15}N, $^{13}C'$, $^{13}C_{\alpha}$, $^{1}H_{\alpha}$ and $^{13}C_{\beta}$ of lipoprotein was performed in a sequence-specific manner using the resonance experiments described in section 2.1.3. and Chapter I, 2.2. Two- and three-dimensional NMR data sets were recorded for two NMR samples 630 µM [U-15N] and 630 µM [U-13C,15N] (His)10-lipoprotein. Figure 43 represents a part of the sequential assignment of lipoprotein (Met 21 - Leu 30) using the 3D HNCACB spectrum. The observable peaks are $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ of intra- and inter-residual resonances. The preceding and succeeding resonances are connected by matching chemical shifts at the same positions at ^{1}H-^{15}N correlations.

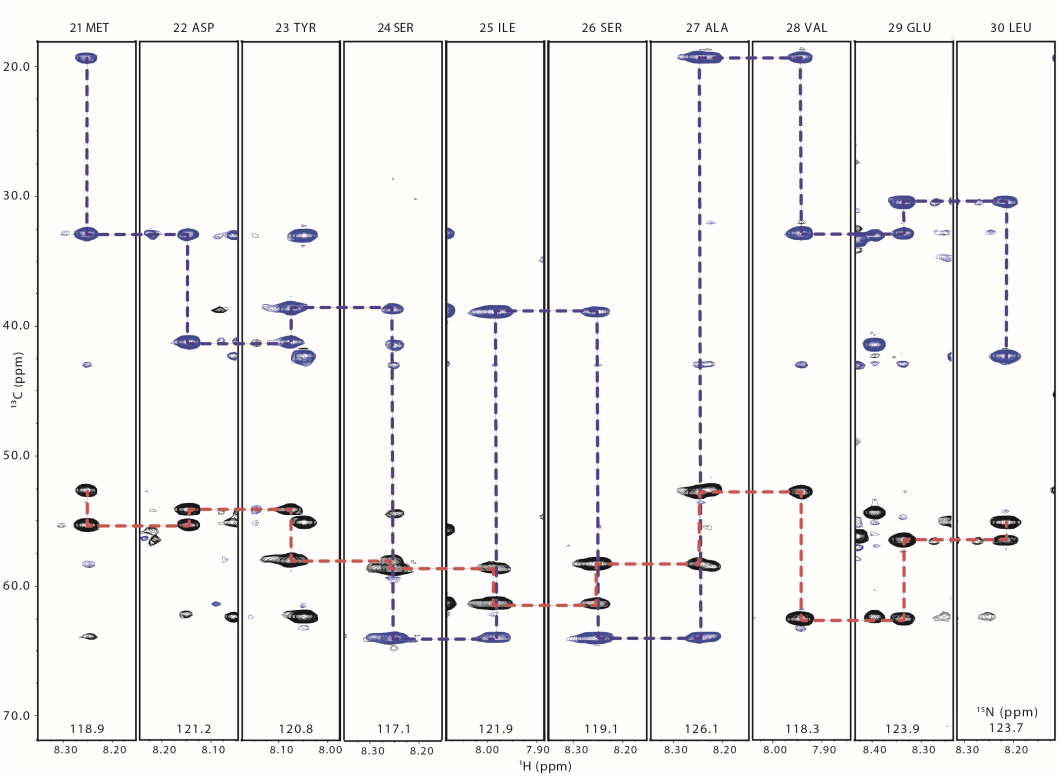


Figure 43. Sequential assignment of backbone nuclei resonances of lipoprotein residues M21-L30 from the 3D HNCACB spectrum. Strip plots from the 3D ^{15}N,^{13}C-TROSY-HNCACB spectrum, recorded on the sample described in 3.1.1. are shown for ten connected residues. Each strip is extracted from the ^{15}N plane and contains two different resonances from $^{13}C\alpha$ (black) and $^{13}C\beta$ (blue) nuclei. For each ^{1}H the strip contains four cross-peaks ($^{13}C\alpha_{i}$, $^{13}C\alpha_{i-1}$ and $^{13}C\beta_{i}$,$^{13}C\beta_{i-1}$). The dotted lines indicate the intra- and inter-residue connections of cross-peaks of the preceding and succeeding residues: red for Cα and blue for Cβ.

The chemical shifts of the amide group were assigned entirely in the 2D ^{1}H, ^{15}N TROSY-HSQC spectrum [61]. Figure 44 shows the complete assignment of all available amino acids in the protein (162 amino acids). The lipoprotein contains two Pro residues at positions 63 and 124, excluded from the analysis because they do not have amino group NH_2. Also, ^{1}H-^{15}N correlations of Gln 155 and Glu 156 amino

acids were not found and assigned on the spectrum. The expected number of peaks in the spectrum was 160 (without Pro), which corresponds to all lipoprotein amino acids. The area around 115-108 ppm of N-dimensional and 6.5-7.8 ppm of H dimension contains additional resonances from the specific TROSY experiment. Also, the side chain guanidino group NHε of Arg and $NH\varepsilon_1$ of Trp rings were picked and assigned on the spectrum. 98.75% of the visible peaks were assigned on the spectrum.

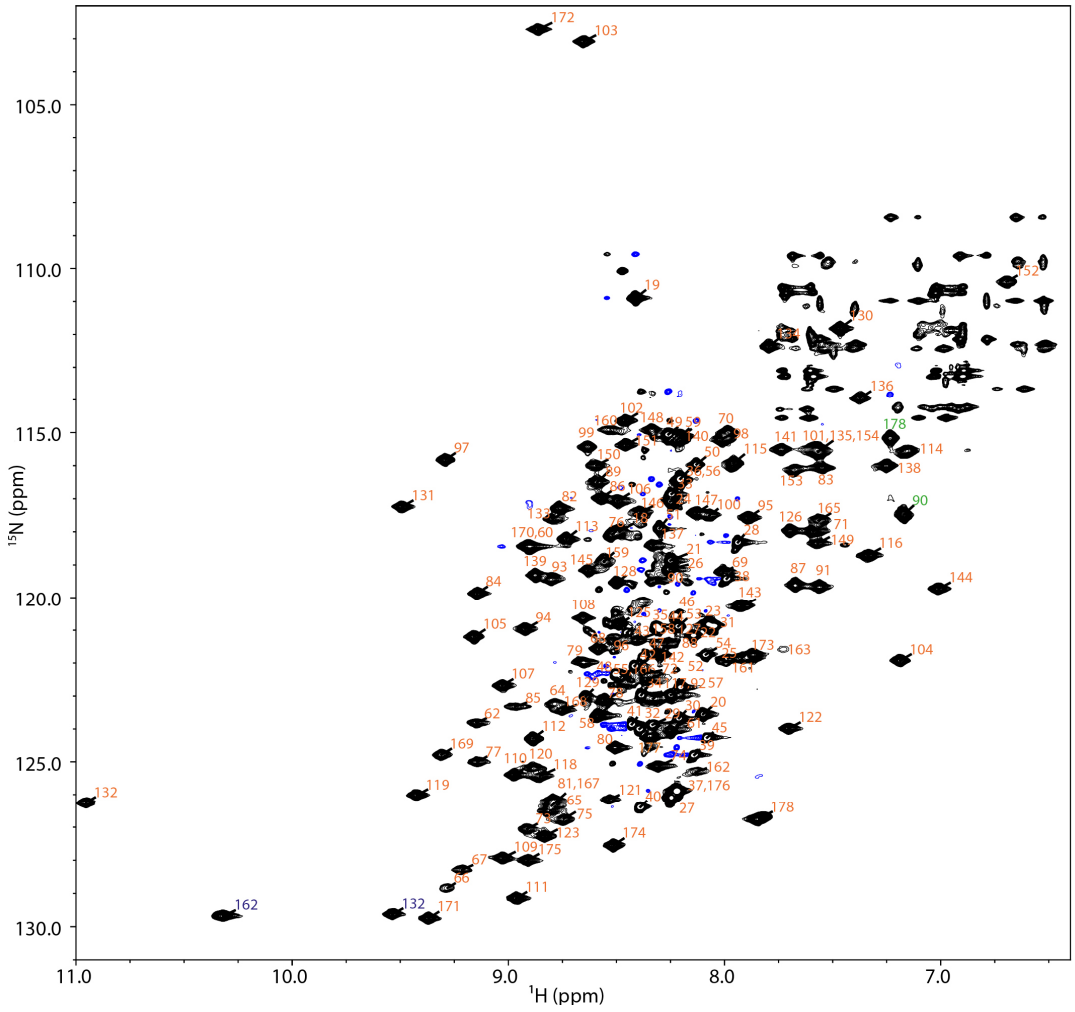


Figure 44. Fully assigned 2D [^{1}H, ^{15}N]-TROSY-HSQC spectrum of lipoprotein. 2D [^{1}H, ^{15}N]-TROSY-HSQC spectrum of lipoprotein were recorded at 30°C at 700 MHz. The numbers in the label correspond to the position of the respective amino acid in the sequence. The spectrum contains peaks from backbone amides (orange numbers). Trp indole $^1H\varepsilon_1$-$^{15}N\varepsilon_1$ peaks show with blue atom numbers. The side chain guanidino group, NHε of Arg, have the green colour of labels.

The Asn and Gln side chain amide groups assignment was completed for 13 amino acids out of 14. Figure 45 represents the part of the fully assigned ^{1}H, ^{15}N HSQC spectrum. Pairs of peaks correlating an NH_2 group of Asn or Gln shifts with single chemical shifts for N and double chemical shifts for two H. The side-chain amide groups assignment information for Asn and Gln was obtained by comparing the chemical shifts of the side chain correlated carbon resonances $^{13}C\alpha$ and $^{13}C\beta$ using 3D TROSY-HN(CO)CACB or 3D TROSY-HNCACB experiments. Also, 3D ^{15}N NOESY HSQC and 3D TROSY HNCO experiments were used to obtain the final assignment of the backbone NH_2 group. Only one NH_2 group of Asn 152 was not found and assigned.

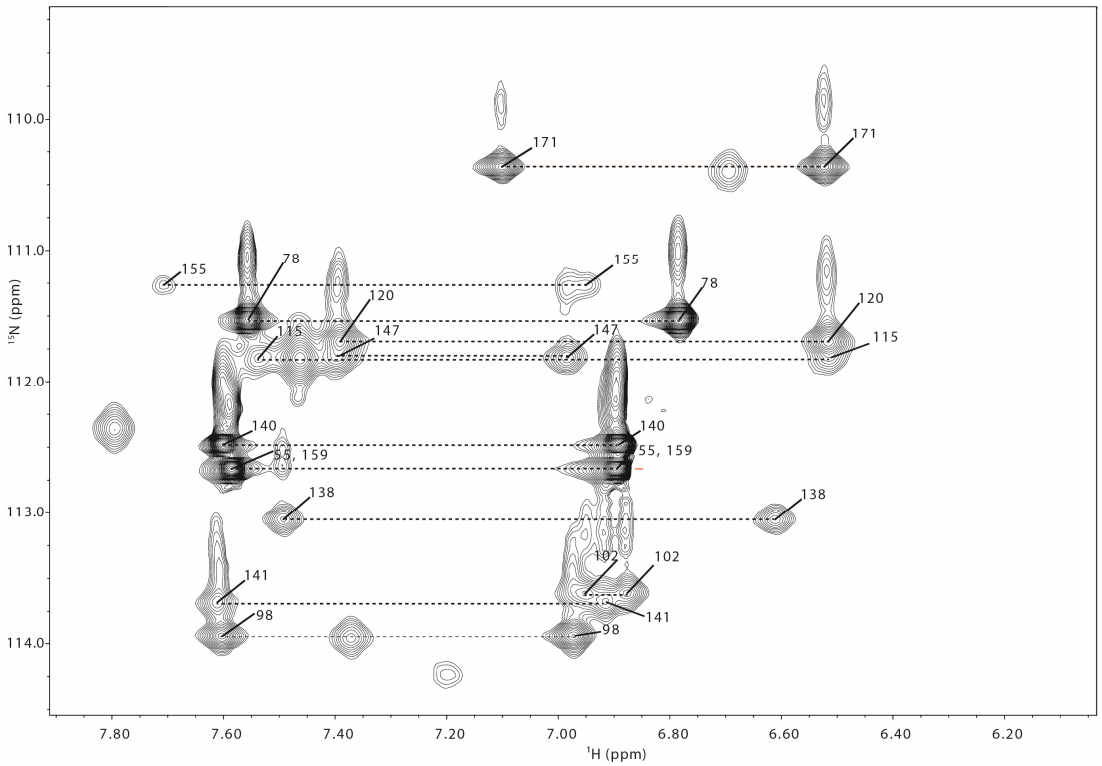


Figure 45. Part of the [^{1}H, ^{15}N] HSQC spectrum of lipoprotein. The expanded region of the ^{1}H, ^{15}N HSQC spectrum at 30°C at 700 MHz represents peaks from backbone amides and side-chain NH_2 of Asn and Gln. Horizontal dashed lines connect pairs of peaks. The numbers in the label correspond to the position of the respective Asn or Gln in the sequence.

The rest of the side-chain assignment of $^{13}C_x$ carbons and connected to them $^{1}H_x$ protons were done using 3D C(CO)NH and 3D H(CCO)NH experiments. Additionally, TOCSY-HSQC and NOESY-TROSY experiments were used to determine all possible proton resonances of the spin system. Figure 46 shows the expanded region's ^{1}H, ^{13}C CT-HSQC NMR spectrum of aromatic rings assignment for the residues of nine amino acids. The assignment was made using 3D ^{13}C NOESY HSQC, 3D ^{15}N NOESY-HSQC and 4D ^{13}C,^{13}C NOESY HSQC spectra.

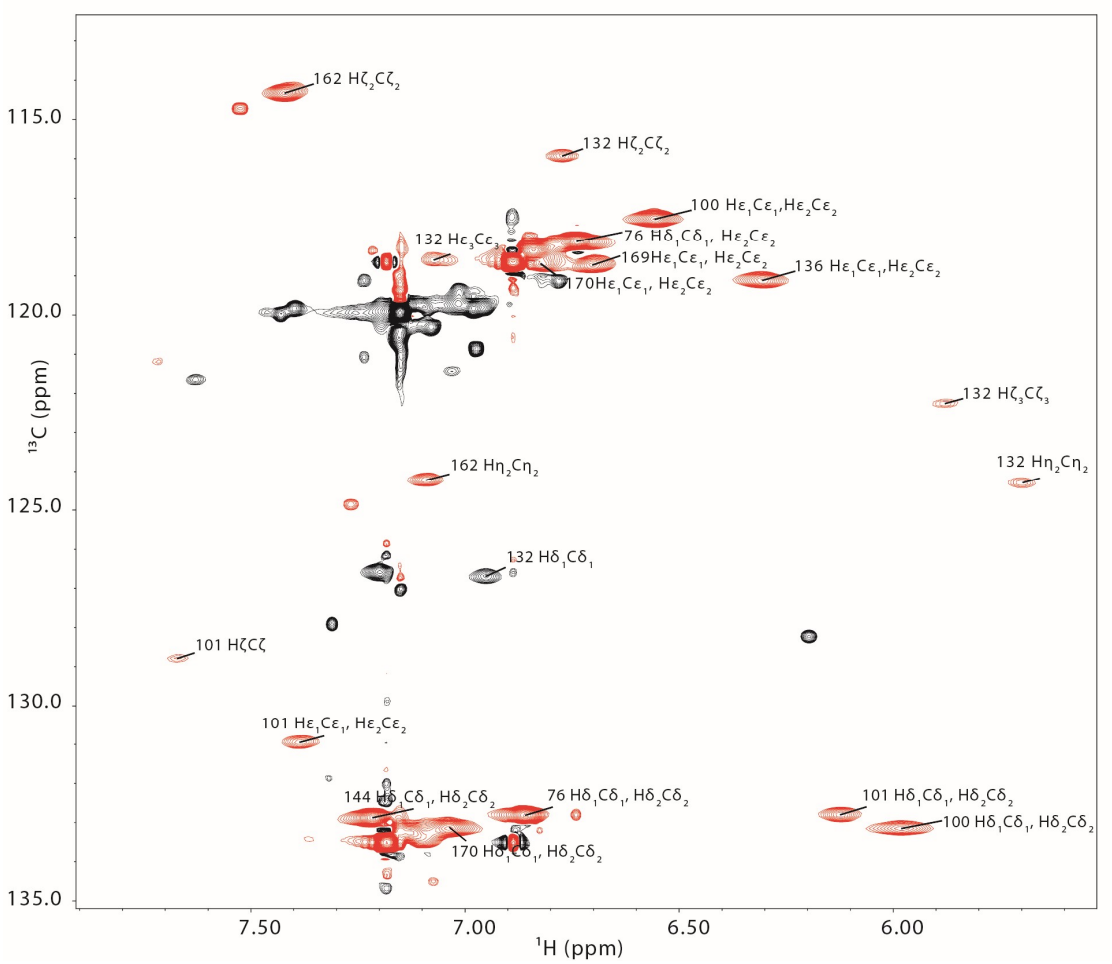


Figure 46. Part of [^{1}H, ^{13}C] CT-HSQC spectrum of lipoprotein. Red peaks are negative, while black is positive. The expanded region of the [^{1}H, ^{13}C] CT-HSQC spectrum at 30°C at 700 MHz represents peaks from aromatic side-chains of Phe, Trp and Tyr. The numbers in the label correspond to the position of the respective Phe, Trp and Tyr in the sequence.

All chemical shifts of completely assigned lipoprotein might be found in the table in the Appendix part.

3.2. Determination of structural restraints from NMR data for lipoprotein secondary structure calculation

The secondary structure, torsion angles ϕ and ψ of the lipoprotein were predicted using chemical shifts assignment in the TALOS-N program (Section 2.1.5). The program calculated the predicted backbone rigidity for each predictable residue with input chemical shifts as RCI-S^2 order parameter (see Chapter I, 2.4.2). An input file with chemical shifts data was generated automatically from NMRViewJ software.

Reliable torsion angle prediction was made for amino acids from Thr 59 to Leu 178, where Gln 155 and Glu 156 had no data prediction due to the absence of chemical shifts. The N terminal amino acids from Gly 19 to Lys 58 were dynamic because the residues were below the threshold RCI-S^2 ≤ 0.6. Figure 47 shows the predicted structure conformation for lipoprotein using the TALOS-N program.

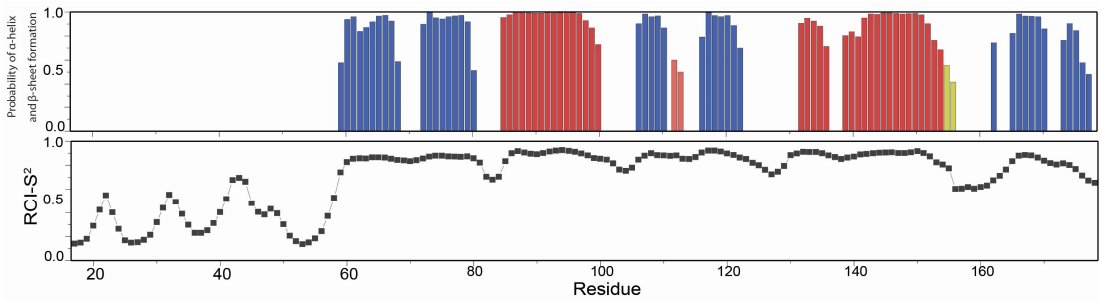


Figure 47. Secondary structure restraints were obtained from TALOS-N torsion angle prediction. The upper picture is the probability of α helix (red) and β strands (blue) formation calculated from NMR chemical shifts. The bottom picture is the predicted backbone rigidity RCI-S^2 for each residue.

Based on the TALOS-N prediction, the final secondary structure of the lipoprotein is shown in Figure 48, which contains the loop region in the N-terminal, six β strands and three α helixes.

Figure 48. Lipoprotein secondary structure prediction from TALOS-N with the help of backbone chemical shifts. The black line corresponds to the chemical shifts of the random coil structure (loop), the blue arrow is the β strand, and the red spiral is the α helix.

3.3. Tertiary structure calculation

The tertiary structure of the lipoprotein was determined from NMR chemical shifts assignment, and NOE derived distance restraints data using XPLOR-NIH NMR software (section 2.1.6). Backbone and side-chain torsion angles were received from the TALOS-N program. The quality of the ensemble of the ten lowest energy structures was evaluated regarding the number and quality of the violations of the experimentally derived restraints. Based on the secondary structure prediction, the loop in the N-terminal region has random coordinates for all ten models. The main structural parts (C-terminal domain) of the ten models of lipoprotein have a good agreement in structure ensemble. The program was run with a target radius of gyration of 13.8 Å [133]. Figure 49 shows ten models with the lowest energies after structure calculation in XPLOR-NIH. Most of the calculations in XPLOR-NIH were done together with Dr Neudecker. The figure was plotted in PyMOL (Table 22). Except for the dynamic N-terminal residues from 1-58 and the loop of 155-160, lipoprotein shows a well-defined structure in solution with an average atomic RMSD from the average structure of 0.65 Å for the backbone and 1.03 Å for all heavy atoms.

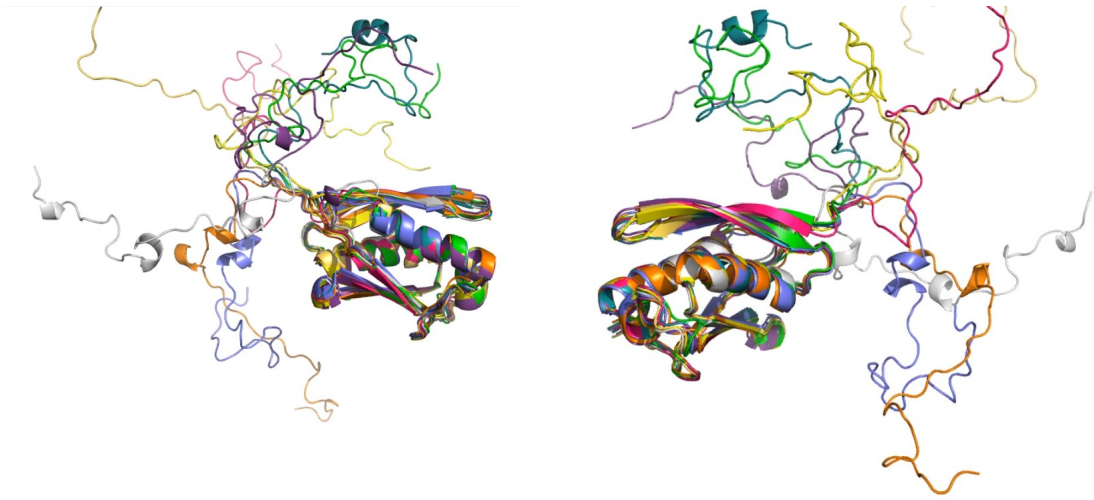


Figure 49. Superposition of the ten lowest energy structures calculated from NMR data from two points of view. The superposition of the ten lowest energy models in the structural ensemble. The N-terminal loops are disordered, while the ordered parts of the protein are in good agreement for all models. Different colours correspond to the different ten models.

The covalent geometry of the 3D structure was assessed using the program PROCHECK_NMR (Table 22). This program sorts the backbone torsion angles in energetically favoured, allowed, and disallowed regions in the Ramachandran plot [135]. PROCHECK_NMR allows seeing the Ramachandran plot for all selected models. The Ramachandran plot (Figure 50) shows ϕ and ψ torsion angles for all residues in the lipoprotein structure for the ten lowest energy models. Gly residues are separately identified by triangles Δ as these are not restricted to the plot regions appropriate to the other sidechain types. 93,2% of the backbone torsion angles of the lipoprotein is located in the most favoured regions of the Ramachandran plot, 6% in the additionally allowed regions and 0.8% in both, the generously allowed and disallowed regions. Only two amino acids that were not assigned (Gln 155 and Glu 156) are located in the generously allowed and disallowed regions of the Ramachandran plot.

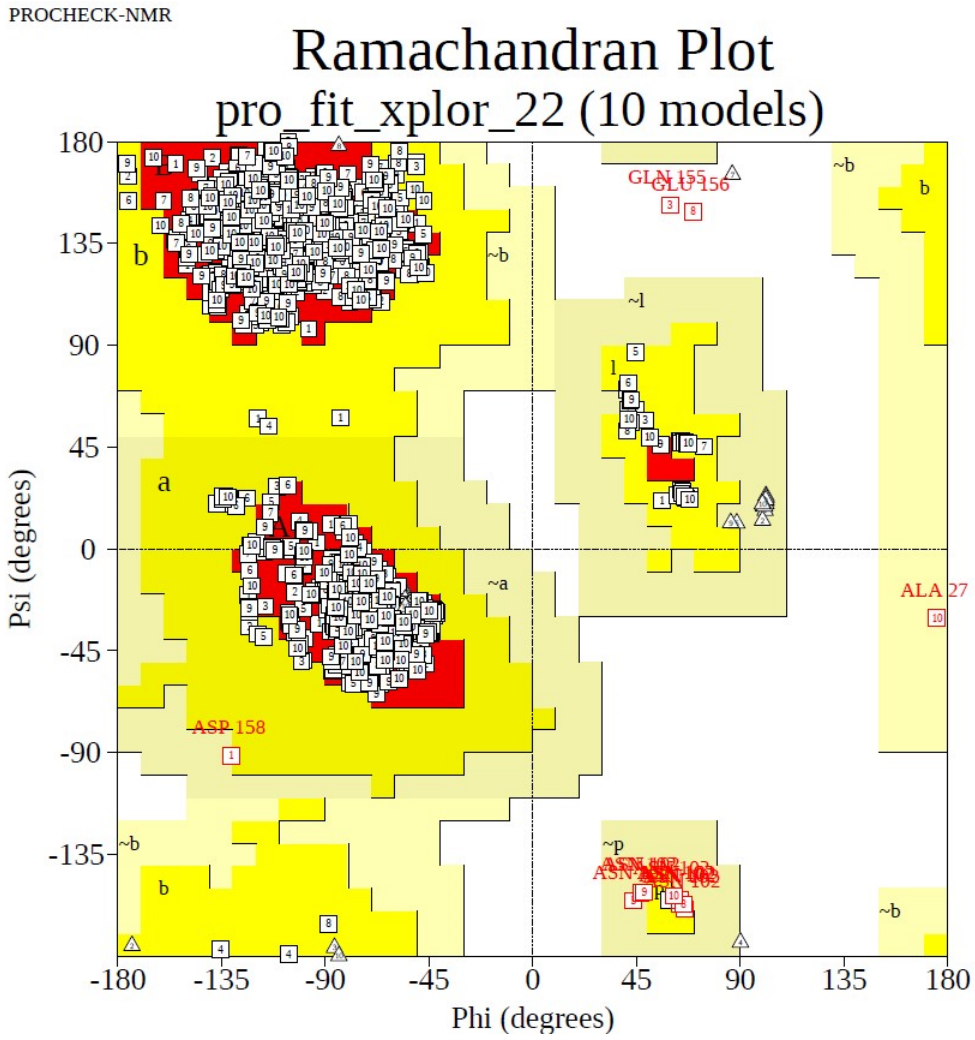


Figure 50. Ramachandran plot of the backbone torsion angles of the lipoprotein ten lowest energy models. The plot shows the correlation of ϕ and ψ torsion angles for all residues of the ten lowest energy structures after lipoprotein structure calculation. All models with their numbers are depicted as white squares, except for Gly residues, shown here as triangles. The colouring on the plot represents the different regions described in [135]: the red areas represent the most favourable combinations of phi-psi values, the yellow colour is additionally allowed regions, the light yellow areas are the generously allowed regions, and the white regions are disallowed regions. Two amino acids Gln 155 and Glu 156, were found in the disallowed region because they both were not the assignment. The letter b is the β strand region, a is the α helix, and l is the loop region in each model structure.

The plotted statistic based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Table 26: Plot statistics of Ramachandran plot of the backbone torsion angles of the lipoprotein ten lowest energy models. A-α helix, B-β strand, L-loop.

Residues in most favoured regions [A,B,L]	1444	93.2%
Residues in additional allowed regions [a,b,l,p]	93	6.0%
Residues in generously allowed regions [~a,~b,~l,~p]	11	0.7%
Residues in disallowed regions	2	0.1%
Number of non-glycine and non-proline residues	1550	100%
Number of end-residues (excl. Gly and Pro)	20	
Number of glycine residues (shown as triangles)	30	
Number of proline residues	20	
Total number of residues	1620	

Figure 51 represents the one picked lipoprotein structure from Figure 49 with the lowest energy 37.5 kcal/mol. It contains the loop region in N-terminal, and compact C-terminal with antiparallel β1 and β2 strands, antiparallel β3 and β4 strands, parallel β4 and β5 strands, and three α helixes.

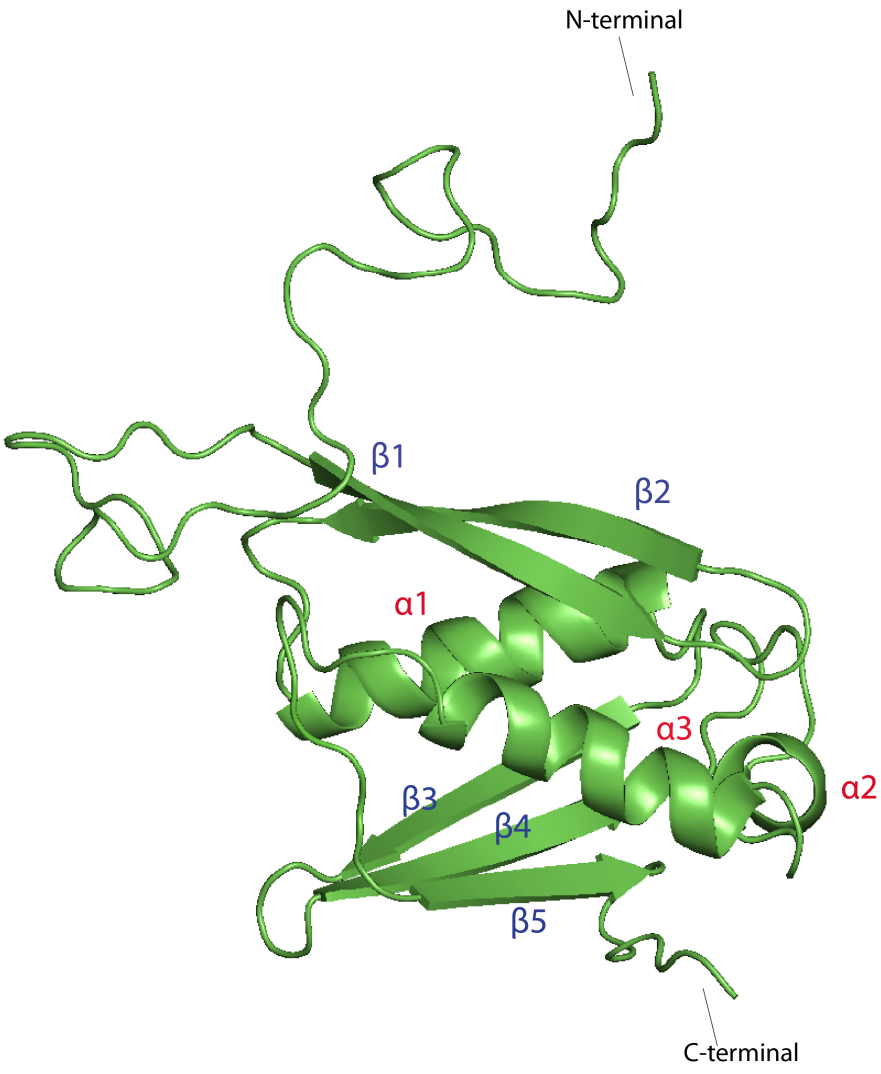


Figure 51. The final 3D model of lipoprotein from NMR data. The final model was calculated using XPLOR-NIH NMR software.

3.4. Titration experiments of lipoprotein and lantibiotic gallidermin

The titration experiments were done for a 630 µM [U-15N] (His)10-lipoprotein sample (see section 2.1.1). The protein concentration was the same for all NMR experiments. For the titration experiments, a stock solution of gallidermin was prepared. Gallidermin was added to lipoprotein sample in 5 different concentrations with ration between gallidermin : lipoprotein as 0.125:1, 0.25:1, 0.5:1, 0.75:1, 1:1. Chemical shift differences (in ppm) for each amino acid were calculated by equation 1. The chemical shift was calculated between samples without gallidermin. The biggest chemical shifts were observed for Ser 160, Ile 161 and Ile 163. The average chemical shift for all amino acids for a 1:1 ratio is 0.01047 ppm. Figure 52 shows the average chemical shifts for each assignment amino acid of lipoprotein and lipoprotein with different concentrations of gallidermin.

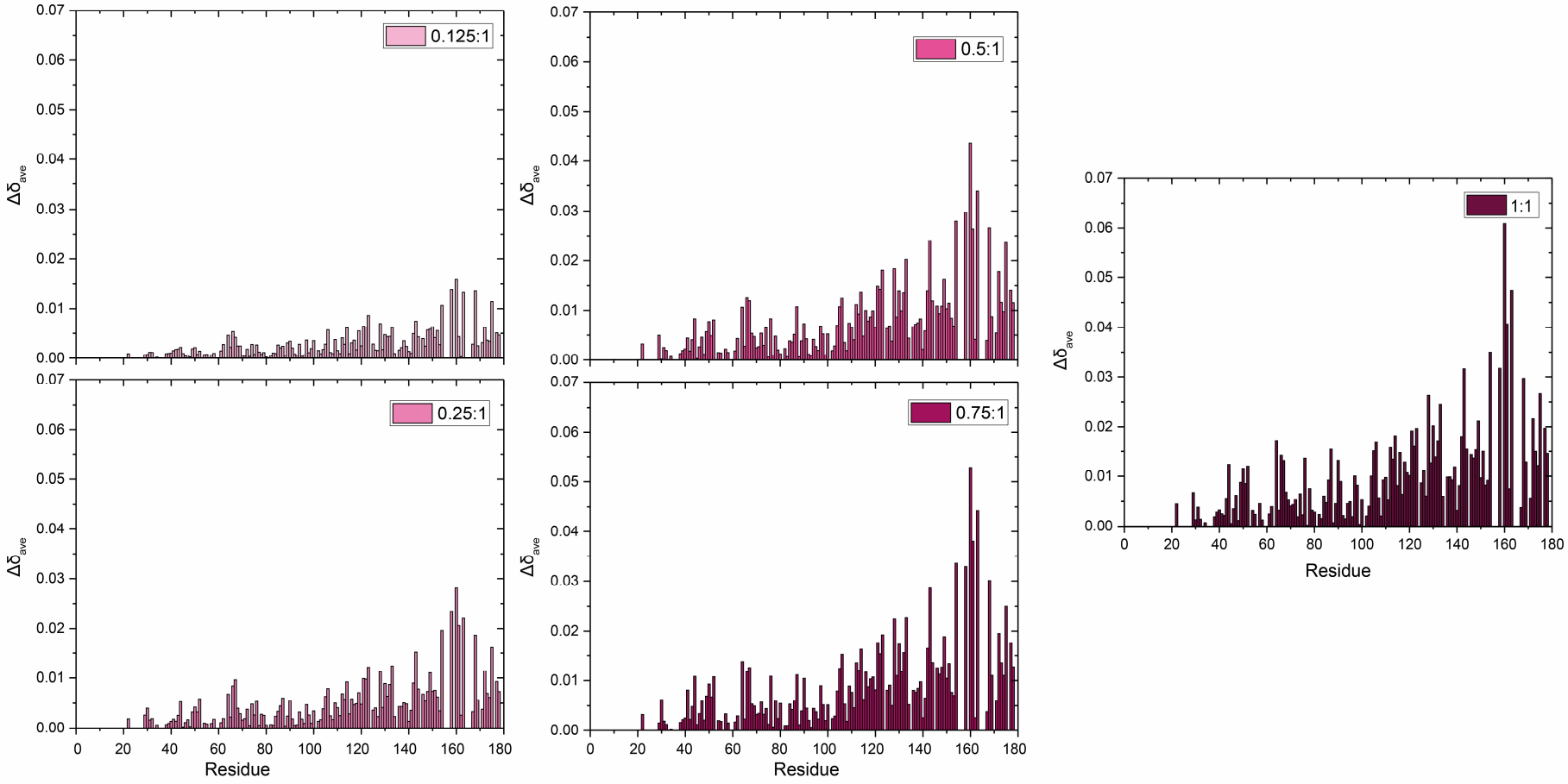


Figure 52. Weighted chemical shift ($\Delta\delta_{ave}$) perturbation analysis between 630 µM lipoprotein and lipoprotein with gallidermin. The $\Delta\delta_{ave}$ values per residue between lipoprotein sample and different rations lipoprotein. The most significant chemical shifts are observed for Ser 160, Ile 161 and Ile 163.

4. Discussion

4.1. Comparison of the theoretically predicted lipoprotein model with secondary structure from NMR data.

Lipoprotein is the membrane resistance protein that is encoded in the CprABC system of *Clostridium difficile*. This protein is likely involved in lantibiotic resistance, but its exact function is still not completely understood. Knowledge of protein tertiary structure (3D) is an essential criterium for understanding its function. The most popular techniques used to determine 3D structures of small proteins are X-ray crystallography and nuclear magnetic resonance. The advantage of solution NMR over X-ray crystallography is that the protein structure can be determined in a solution that is more closely to its actual physiological state. Some proteins, including the lipoprotein, are hard to crystallize; solution NMR was used as the main experimental method for structure calculation. One more additional method to solve the 3D structure is computational structure prediction. The first visualized model of lipoprotein was achieved by H. Gohlke's research group using the fully automated meta-method "TopModel" for protein structure prediction. The description of the method might be found in [61]. From NMR experiments, the secondary structure and b-strand pairing together with SAXS experiments were performed to estimate the shape and radius of gyration (RG). The secondary structure experimental data were superimposed on the theoretical model. The 3D lipoprotein model from TopModel has a good agreement with NMR secondary structure assignment, but there are still some discrepancies. Figure 54 shows the initial 3D prospective model of lipoprotein (the disordered tail was removed from the structure) with superimposed NMR secondary structure data (Figure 48). The predicted model has a good agreement with the experimental data. The expected model has a local TopScore where 0.1 indicating about 10% error in the model, 0.3 is 30%, etc. The highest number shows that the model may not be highly confident.

Four discrepancies between theoretical and experimental models can be found in Figure 53. The first (1) shows that the first a-helix of TopModel is eight residues shorter than in NMR predicted structure, which also indicates low TopScore. The second (2) difference is the third β-strand shifted by two residues and the loop on TopModel was identified as β-strand in NMR, the loop indicated by TopScore to contain high errors. The third (3) is the difference in the C-terminus, on the NMR structure prediction is folded into a β6-strand. The fourth (4) difference is a different length of second α-helix 2, which also scores poorly according to TopScore.

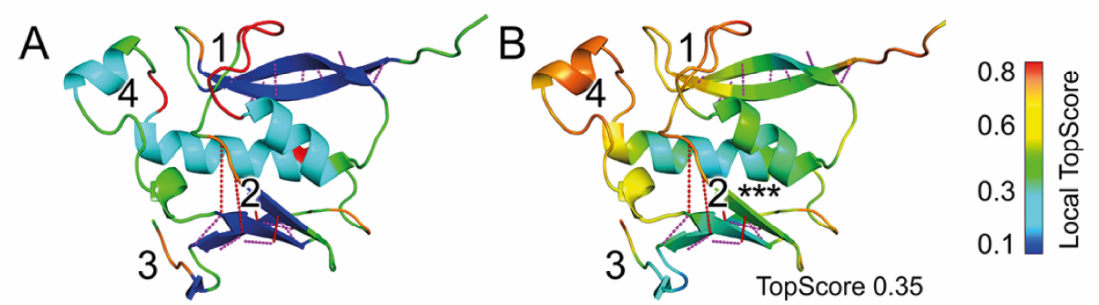
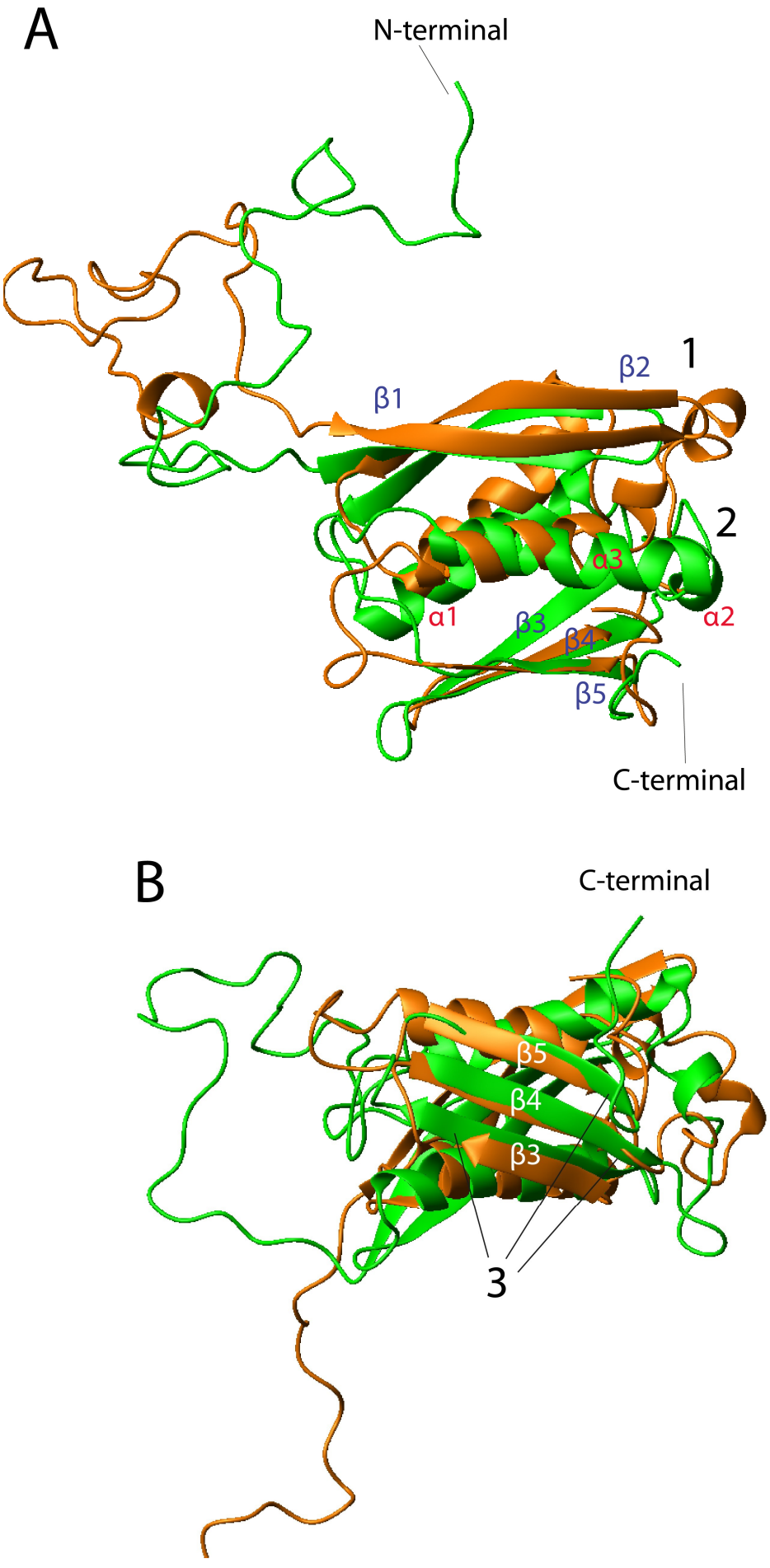
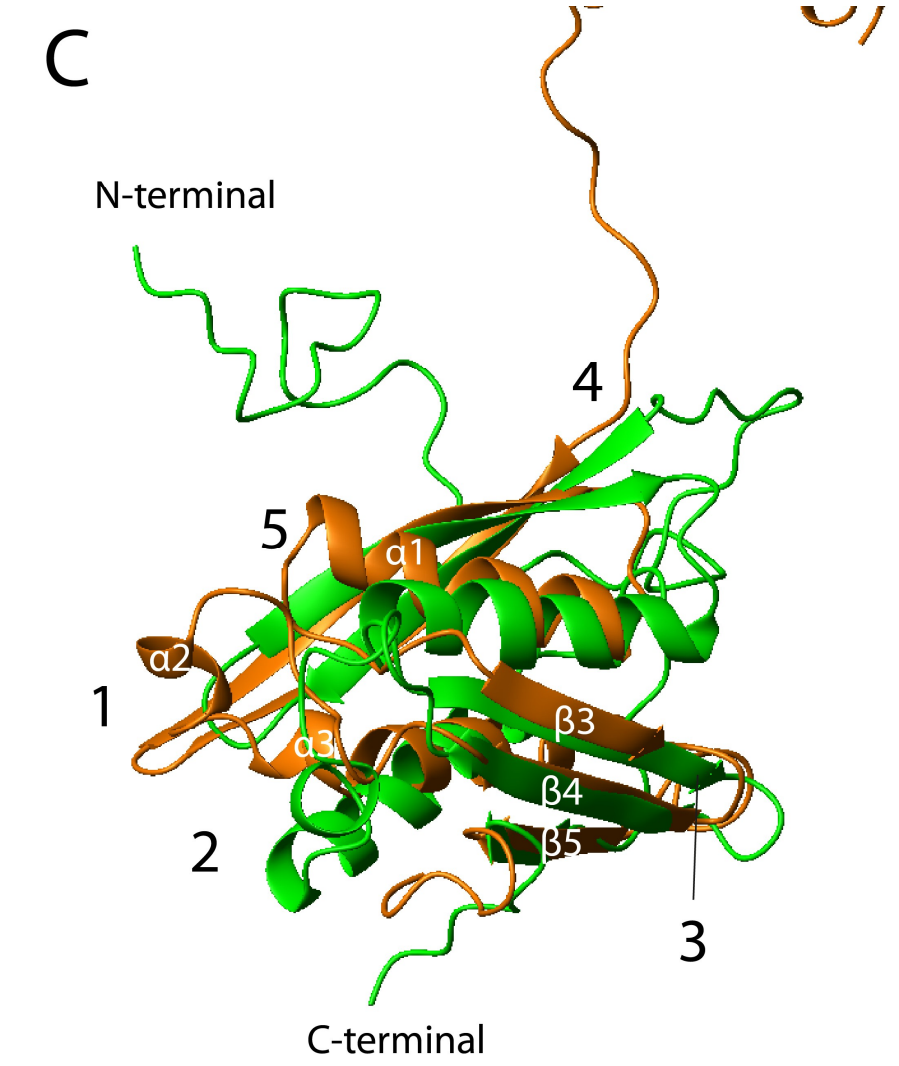


Figure 53. Prospective modelling of lipoprotein from *C. difficile* (without disordered N-terminal tail). (A) The initial TopModel with superimposed NMR secondary structure data. The dark blue colour corresponds to the good agreement of β-sheets. Orange colour: residues identified as a β-strand in NMR but not found so in the theoretical model. Cyan colour is α-helixes in good agreement. Red: residues identified as α-helical in NMR but not found so in the theoretical model. Violet lines: experimental β-sheet NOE restraints are showing agreement with the model. Red dotted lines: experimental β-sheet NOE restraints show a shift of two residue positions of the third β-strand. (B) The initial TopModel is coloured according to Local TopScore. Blue regions indicated the best residue-wise error (~30%). Yellow/red regions indicate regions with a high residue-wise error (>50%). The numbers correspond to four differences between TopModel and NMR data. The picture was published in the joint article with H. Gohlke's research group [61].

4.2. Comparison of the theoretically predicted lipoprotein model with tertiary structure from NMR data.

The tertiary structure of the lipoprotein was determined from NMR chemical shifts assignment, and NOE derived distance restraints data using XPLOR-NIH NMR software with a target radius of gyration of 13.8 Å. Figure 54 A-D shows the overlapped 3D structure of lipoprotein from NMR data with the theoretical model from [61]. Both models are shown from different angles of vision. The 3D lipoprotein model from experimental NMR data has a good agreement with the theoretical model, but there are still some discrepancies. There are five differences: (1) the turn from β1 to β2 is shifted for few residues; (2) α3 from the NMR structure is closer to β3, β4, β5 strands; (3) shifts for three β3, β4, β5 strands which were also detected with the secondary structure comparison; (4) shift of β1 and β2 which is a consequence of the first difference; (5) shift for one turn in α1. N-terminal was not compared because it has random coordinates due to its fast dynamic.


A
N-terminal
β2
1
β1
2
α3
α1
β3
β4
α2
β5
C-terminal
B
C-terminal
β5
β4
β3
3


C
N-terminal
4
5
α1
α2
1
β3
α3
β4
β5
2
3
C-terminal

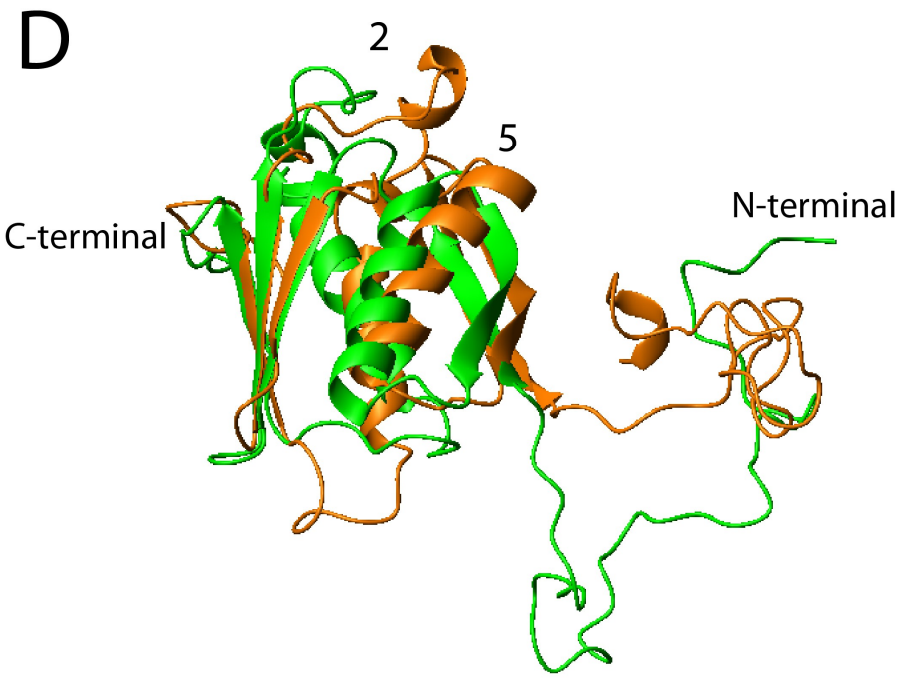

D
2
5
C-terminal
N-terminal

Figure 54. Overlapped two 3D models of lipoprotein from *C. difficile* from NMR experimental data and theoretical prediction. Comparison of two models of lipoprotein from 3D NMR data (green) and computational model (orange). The models represent from four different angles of view. The number corresponds to the differences described here. (A) shows all β strands (blue) and all α helixes (red). (B) shows the labelled β strands that are involved in the third difference. (C,D) is β strands and α helixes are involved in the differences between models.

Searching the database characterized lipoproteins from bacteria showed the similarity in the structure with some bacterial protein involved in cell-cell adhesion. The structure similarity with other proteins which might be involved in lantibiotic resistance was not found so far. The determined structure of lipoprotein CD1348 is rare and poorly characterized in the literature.

4.3. Titration experiments with lantibiotic gallidermin

The role of lipoprotein in the resistance machinery of *C.difficile* is still not completely understood. The solved 3D structure of the lipoprotein can help to understand if this protein is involved or not in lantibiotic resistance. Some previous work was performed for lipoprotein CD1348 and lantibiotics nisin and gallidermin [60]. The interaction between lipoprotein and lantibiotics was tested *in vivo* and *in vitro* experiments. To see the interactions, *in vitro* analytical co-elution and tryptophan fluorescence experiments were provided. After co-elution studies, the chromatogram of the lipoprotein with nisin does not show any differences, which means there is no interaction between them. Another method is tryptophan fluorescence spectroscopy is more sensitive to detect interactions. The 1 µM of purified lipoprotein was measured with different concentrations of the lantibiotics nisin or gallidermin. If the lantibiotic binds to the lipoprotein, a tryptophan quenching should be observed. The decreasing fluorescence signal of tryptophan quenching was observed for lantibiotic gallidermin. With increasing gallidermin concentration (0-47.85 µM), the calculated K_D value of 2.4 µM gallidermin shows an interaction. The sequence of lipoprotein contains two Trp at 132 and 162 positions. Trp 132 is located in the α2 helix, and Trp 162 is located in the loop region between α3-helix and β5-strand. The additional fluorescence experiments were performed to see which Trp is more affected by interaction with gallidermin. Figure 56 showed the recent unpublished fluorescence data from Dr Smits research group for lipoprotein wildtype mutant when Trp 132 was replaced with Ala 132 and mutant with Ala 162 instead of Trp 162. The wildtype data shows the increasing fluorescence signal of tryptophan quenching with increasing the concentration of gallidermin. For the mutant Trp 132 to Ala, the increase of tryptophan quenching was also observed, which means that Trp 162 has interactions with gallidermin. The K_D values were measure for fluorescent experiments. K_D of wildtype is 4.3 µM, mutant Trp132Ala is 4.9 µM and mutant Trp162Ala is 21.3 µM.

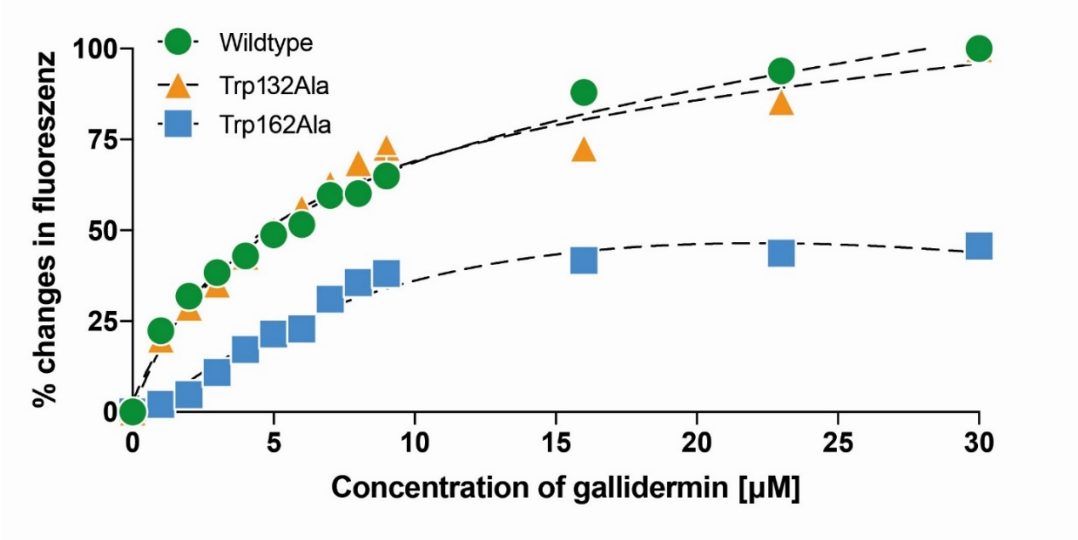


Figure 55. ***Fluorescence of samples in the presence of different concentrations of gallidermin.*** Fluorescence data were measured for three different samples of lipoprotein wildtype (green circles), lipoprotein mutant Ala 132 instead of Trp (orange triangles) and lipoprotein mutant Ala 162 instead of Trp (blue squares). The picture is published here with the allowance of Dr Smits.

From the fluorescence data, Trp 162 is more affected by interaction with gallidermin. With the help of titration experiments, the most significant chemical shift changes were observed for Ser 160, Ile 161 and Ile 163, which is a dynamic loop close to where the disordered N-terminal is attached to the globular structure. The NMR data is similar to fluorescence data and might indicate possible interactions between lipoprotein and lantibiotic gallidermin. Figure 57 shows the location of the amino acids in which chemical shifts were observed.

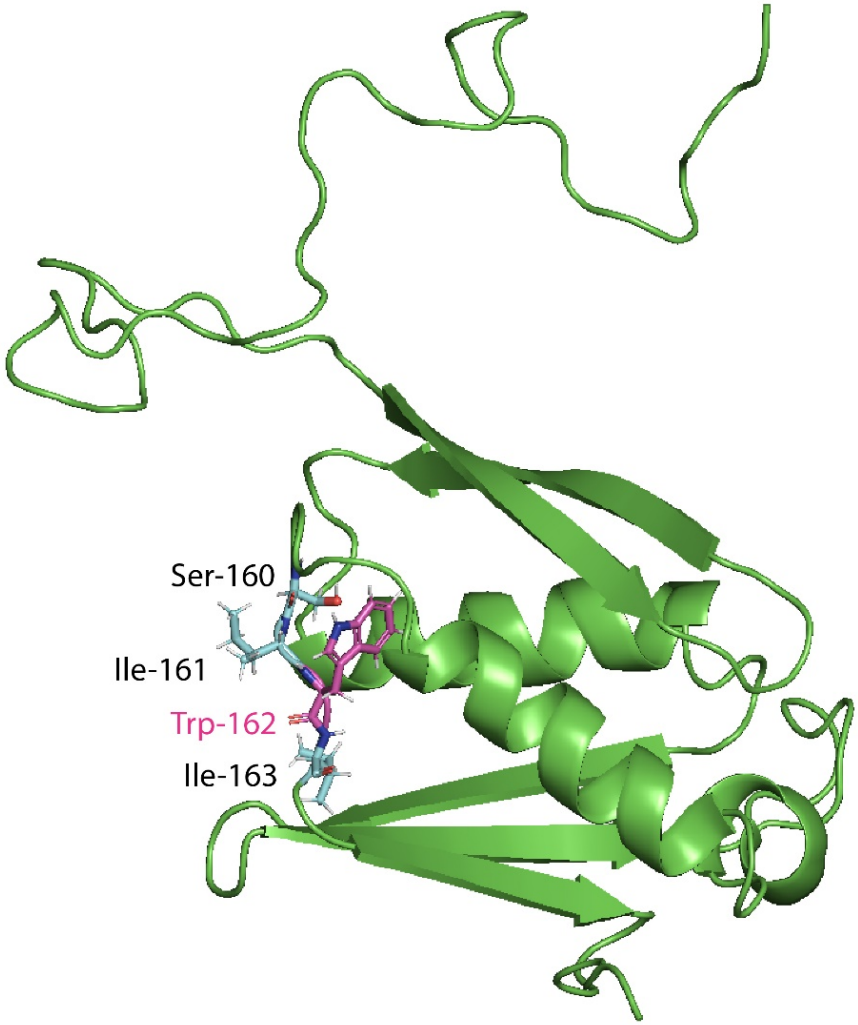


Figure 56. 3D model of lipoprotein molecule. Amino acids that have the most significant chemical shifts are labelled black. The Tryptophan quenching of 162 Trp residue, which was observed by fluorescence spectroscopy, is labelled pink.

The general conclusion for the lipoprotein project is observed in Chapter IV.

Chapter IV Conclusion and Outlook

Nuclear Magnetic Resonance Spectroscopy was used for structure calculation and dynamic determination of two membrane proteins: GABARAP and lipoprotein CD1348. Molecular dynamics simulation, SAXS and fluorescence spectroscopy were used to complement the NMR data analysis.

The characterization of the structural conformation changes of GABARAP anchored to nanodiscs was performed using solution NMR. Based on the chemical shift differences for the N-terminal domains between free protein and protein anchored to nanodiscs (ND), GABARAP oligomerization was not induced and stabilized by interactions with the ND membrane in the solution. Anchoring to NDs leads to significant chemical shift changes of the backbone amide groups of Lys 38, Asn 82, Ser 113, Val 114 and Tyr 115. Free GABARAP protein was predominantly monomeric in the temperature range from 5°C to 35°C, as indicated by the hydrodynamic radius between 21.5 - 21.8 Å. In this work, the hydrodynamic radius of the GABARAP anchored to ND at 40°C was bigger than monomeric: 27.2 Å. The increased hydrodynamic radius was not discussed in the current work and needs additional information for understanding. Further studies using MD simulation, small-angle scattering, or fluorescence spectroscopy for the GABARAP-ND complex can provide the necessary information.

In this thesis, a combination of NMR spectroscopy and MD simulations were used to characterize the internal dynamics of GABARAP anchored to nanodiscs at various time scales. In particular, the order parameter S^2 of free GABARAP obtained by MD simulations is in good agreement with the experimental data presented in this work. Also, S^2 values for the GABARAP-ND complex were compared with MD simulation data. The data shows that the backbone order parameters S^2 from NMR spectroscopy and MD simulations primarily reveal internal motions in the loop regions of Ile 41 and at the N and C-terminal. At the same time, the regular secondary structure of GABARAP shows high rigidity (higher than 0.7). GABARAP anchored to a ND shows less mobility for the Ile 41 and C-terminal amino acids in NMR experiments and MD simulation. The correlation time τ_c of GABARAP in the ND complex is much slower than for the free GABARAP protein. It has been explained by the tumbling of GABARAP together with the ND. This data is a good starting point for the next steps in studying the protein-nanodisc dynamics. A full model-free analysis can be done if the data are available for two or more magnetic fields.

The second part of the thesis was dedicated to the structure determination of another membrane protein. Lipoprotein from *C. difficile* is encoded directly before a known lantibiotic resistance ABC-transporter CprABC. This resistance operon consists of the ABC-transporter and a two-component system, including histidine kinase (HK) and a response regulator (RR). The lipoprotein function is still not completely understood, but it has been hypothesised that this membrane-associated protein also performs resistance against lantibiotics. Therefore, determining the 3D structure of the lipoprotein should provide insight into its role in the context of lantibiotic resistance of *C.difficile*.

The structure of the lipoprotein was solved by solution NMR, SAXS and computer simulations using the “TopModel” model. Standard 2D and 3D NMR experiments were successfully recorded to gain

assignment and structural information of the lipoprotein. The assignment of the amino acids in the 178-residue lipoprotein was done manually with the help of sequential backbone assignment. The tertiary structure of lipoprotein CD1348 was received using multidimensional 3D and 4D NOESY experiments and represented a so far undescribed model. In particular, the N-terminal of the final model is disordered, while the C-terminal is well ordered with antiparallel β1 and β2 strands, antiparallel β3 and β4 strands, parallel β4 and β5 strands, and three α helixes. The simulated TopModel structure showed good agreement with the experimental structure of the lipoprotein. The structural shift of the 3rd β-strand is registered in the NMR structure.

Titration experiments of lipoprotein with lantibiotic gallidermin show chemical shifts in a loop region of Ser 160, Ile 161 and Ile 163 residues. The data is in good agreement with fluorescence experiments. The presence of gallidermin and changing chemical shifts can indicate possible interactions between lipoprotein and the lantibiotic gallidermin. This exciting result is key for future investigation of lipoprotein function.

Literature

1. Mohrluder, J., M. Schwarten, and D. Willbold, *Structure and potential function of gamma-aminobutyrate type A receptor-associated protein.* FEBS J, 2009. **276**(18): p. 4989-5005.
2. Coyle, J.E., et al., *Structure of GABARAP in two conformations: implications for GABA(A) receptor localization and tubulin binding.* Neuron, 2002. **33**(1): p. 63-74.
3. von Heijne, G., *The membrane protein universe: what's out there and why bother?* J Intern Med, 2007. **261**(6): p. 543-57.
4. Krogh, A., et al., *Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.* J Mol Biol, 2001. **305**(3): p. 567-80.
5. Wimley, W.C., *Toward genomic identification of beta-barrel membrane proteins: composition and architecture of known structures.* Protein Sci, 2002. **11**(2): p. 301-12.
6. Alberts, B.J., A.; Lewis, J.; et al., *Molecular Biology of the Cell. 4th edition.* New York: Garland Science, 2002.
7. Levine, B. and D.J. Klionsky, *Development by self-digestion: molecular mechanisms and biological functions of autophagy.* Dev Cell, 2004. **6**(4): p. 463-77.
8. Cuervo, A.M., *Autophagy: in sickness and in health.* Trends Cell Biol, 2004. **14**(2): p. 70-7.
9. Yun, C.W. and S.H. Lee, *The Roles of Autophagy in Cancer.* Int J Mol Sci, 2018. **19**(11).
10. White, E., *Deconvoluting the context-dependent role for autophagy in cancer.* Nat Rev Cancer, 2012. **12**(6): p. 401-10.
11. Lorin, S., et al., *Autophagy regulation and its role in cancer.* Semin Cancer Biol, 2013. **23**(5): p. 361-79.
12. Grumati, P., et al., *Autophagy is defective in collagen VI muscular dystrophies, and its reactivation rescues myofiber degeneration.* Nat Med, 2010. **16**(11): p. 1313-20.
13. Nah, J., J. Yuan, and Y.K. Jung, *Autophagy in neurodegenerative diseases: from mechanism to therapeutic approach.* Mol Cells, 2015. **38**(5): p. 381-9.
14. Saha, S., et al., *Autophagy in health and disease: A comprehensive review.* Biomed Pharmacother, 2018. **104**: p. 485-495.
15. Parzych, K.R. and D.J. Klionsky, *An overview of autophagy: morphology, mechanism, and regulation.* Antioxid Redox Signal, 2014. **20**(3): p. 460-73.
16. Chen, Y. and D.J. Klionsky, *The regulation of autophagy - unanswered questions.* J Cell Sci, 2011. **124**(Pt 2): p. 161-70.
17. Hayashi-Nishino, M., et al., *A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation.* Nature Cell Biology, 2009. **11**(12): p. 1433-1437.
18. Axe, E.L., et al., *Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum.* J Cell Biol, 2008. **182**(4): p. 685-701.
19. Xie, Z. and D.J. Klionsky, *Autophagosome formation: core machinery and adaptations.* Nat Cell Biol, 2007. **9**(10): p. 1102-9.

20. Young, A.R., et al., *Starvation and ULK1-dependent cycling of mammalian Atg9 between the TGN and endosomes.* J Cell Sci, 2006. **119**(Pt 18): p. 3888-900.
21. Walker, S., et al., *Making autophagosomes: Localized synthesis of phosphatidylinositol 3-phosphate holds the clue.* Autophagy, 2008. **4**(8): p. 1093-1096.
22. Mizushima, N., T. Yoshimori, and Y. Ohsumi, *The role of Atg proteins in autophagosome formation.* Annu Rev Cell Dev Biol, 2011. **27**: p. 107-32.
23. Slobodkin, M.R. and Z. Elazar, *The Atg8 family: multifunctional ubiquitin-like key regulators of autophagy.* Essays Biochem, 2013. **55**: p. 51-64.
24. Kabeya, Y., et al., *LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing.* Embo j, 2000. **19**(21): p. 5720-8.
25. Kabeya, Y., et al., *LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation.* Journal of Cell Science, 2004. **117**(13): p. 2805-2812.
26. Shpilka, T., et al., *Atg8: an autophagy-related ubiquitin-like protein family.* Genome Biology, 2011. **12**(7): p. 226.
27. Ma, P., et al., *Preparation of a functional GABARAP-lipid conjugate in nanodiscs and its investigation by solution NMR spectroscopy.* Chembiochem, 2010. **11**(14): p. 1967-70.
28. Tanida, I., et al., *GATE-16 and GABARAP are authentic modifiers mediated by Apg7 and Apg3.* Biochem Biophys Res Commun, 2003. **300**(3): p. 637-44.
29. Ritchie, T.K., et al., *Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer nanodiscs.* Methods Enzymol, 2009. **464**: p. 211-31.
30. Borch, J.H., T., *The nanodisc: a novel tool for membrane protein studies.* 2009. **390**(8): p. 805-814.
31. Viegas, A., T. Viennet, and M. Etzkorn, *The power, pitfalls and potential of the nanodisc system for NMR-based studies.* Biol Chem, 2016. **397**(12): p. 1335-1354.
32. Jonas, A., *Reconstitution of high-density lipoproteins.* Methods Enzymol, 1986. **128**: p. 553-82.
33. Bayburt, T.H.G.Y.V.a.S.S.G., *Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins.* Nano Letters 2002. **2(8)**: p. 853-856.
34. Denisov, I.G., et al., *Directed self-assembly of monodisperse phospholipid bilayer Nanodiscs with controlled size.* J Am Chem Soc, 2004. **126**(11): p. 3477-87.
35. Hagn, F., et al., *Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins.* J Am Chem Soc, 2013. **135**(5): p. 1919-25.
36. Nath, A., W.M. Atkins, and S.G. Sligar, *Applications of phospholipid bilayer nanodiscs in the study of membranes and membrane proteins.* Biochemistry, 2007. **46**(8): p. 2059-69.
37. Kohanski, M.A., D.J. Dwyer, and J.J. Collins, *How antibiotics kill bacteria: from targets to networks.* Nat Rev Microbiol, 2010. **8**(6): p. 423-35.
38. Müller, A., A. Klöckner, and T. Schneider, *Targeting a cell wall biosynthesis hot spot.* Nat Prod Rep, 2017. **34**(7): p. 909-932.

39. Kapoor, G., S. Saigal, and A. Elongavan, *Action and resistance mechanisms of antibiotics: A guide for clinicians.* J Anaesthesiol Clin Pharmacol, 2017. **33**(3): p. 300-305.
40. Breukink, E. and B. de Kruijff, *Lipid II as a target for antibiotics.* Nat Rev Drug Discov, 2006. **5**(4): p. 321-32.
41. Bierbaum, G. and H.G. Sahl, *Lantibiotics: mode of action, biosynthesis and bioengineering.* Curr Pharm Biotechnol, 2009. **10**(1): p. 2-18.
42. Jung, G., *Lantibiotics-ribosomally sythesized biologically acve polypeptides containing sulfide bridges and a,b-didehyroamino acids.* Angew. Chem. Int. Edn. English, 1991. **30**: p. 1051–1068.
43. Willey, J.M. and W.A. van der Donk, *Lantibiotics: peptides of diverse structure and function.* Annu Rev Microbiol, 2007. **61**: p. 477-501.
44. Knerr, P.J. and W.A. van der Donk, *Discovery, biosynthesis, and engineering of lantipeptides.* Annu Rev Biochem, 2012. **81**: p. 479-505.
45. Reiners, J., et al., *Stoichiometry and structure of a lantibiotic maturation complex.* Scientific Reports, 2017. **7**(1): p. 42163.
46. Clemens, R., et al., *Insight into Two ABC Transporter Families Involved in Lantibiotic Resistance.* Front Mol Biosci, 2017. **4**: p. 91.
47. Wiedemann, I., et al., *Specific binding of nisin to the peptidoglycan precursor lipid II combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity.* J Biol Chem, 2001. **276**(3): p. 1772-9.
48. Hasper, H.E., B. de Kruijff, and E. Breukink, *Assembly and stability of nisin-lipid II pores.* Biochemistry, 2004. **43**(36): p. 11567-75.
49. Götz, F., et al., *Epidermin and gallidermin: Staphylococcal lantibiotics.* International Journal of Medical Microbiology, 2014. **304**(1): p. 63-71.
50. Parisot, J., et al., *Molecular mechanism of target recognition by subtilin, a class I lanthionine antibiotic.* Antimicrob Agents Chemother, 2008. **52**(2): p. 612-8.
51. Barbut, F., et al., *Clinical features of Clostridium difficile-associated diarrhoea due to binary toxin (actin-specific ADP-ribosyltransferase)-producing strains.* Journal of medical microbiology, 2005. **54**(Pt 2): p. 181-185.
52. Vollaard, E.J. and H.A. Clasener, *Colonization resistance.* Antimicrob Agents Chemother, 1994. **38**(3): p. 409-14.
53. Nolan, N.P., et al., *An epidemic of pseudomembranous colitis: importance of person to person spread.* Gut, 1987. **28**(11): p. 1467-73.
54. Peschel, A. and H.G. Sahl, *The co-evolution of host cationic antimicrobial peptides and microbial resistance.* Nat Rev Microbiol, 2006. **4**(7): p. 529-36.
55. Hancock, R.E. and G. Diamond, *The role of cationic antimicrobial peptides in innate host defences.* Trends Microbiol, 2000. **8**(9): p. 402-10.

56. Suárez, J.M., A.N. Edwards, and S.M. McBride, *The Clostridium difficile cpr locus is regulated by a noncontiguous two-component system in response to type A and B lantibiotics.* J Bacteriol, 2013. **195**(11): p. 2621-31.
57. McBride, S.M. and A.L. Sonenshein, *The dlt operon confers resistance to cationic antimicrobial peptides in Clostridium difficile.* Microbiology (Reading), 2011. **157**(Pt 5): p. 1457-1465.
58. Schmitt, L. and R. Tampé, *Structure and mechanism of ABC transporters.* Current Opinion in Structural Biology, 2002. **12**(6): p. 754-760.
59. Pipatthana, M., et al., *The repertoire of ABC proteins in Clostridioides difficile.* Comput Struct Biotechnol J, 2021. **19**: p. 2905-2920.
60. Clemens, R., *Characterization of a lipoprotein CD1348 from Clostridium difficile and the viral infectivity factor of HIV-1.* Inaugural-Dissertation, 2018.
61. Mulnaes, D., et al., *TopModel: Template-Based Protein Structure Prediction at Low Sequence Identity Using Top-Down Consensus and Deep Neural Networks.* J Chem Theory Comput, 2020. **16**(3): p. 1953-1967.
62. Foster, M.P., C.A. McElroy, and C.D. Amero, *Solution NMR of large molecules and assemblies.* Biochemistry, 2007. **46**(2): p. 331-40.
63. Jeener, J. and G. Alewaeters, *"Pulse pair technique in high resolution NMR" a reprint of the historical 1971 lecture notes on two-dimensional spectroscopy.* Prog Nucl Magn Reson Spectrosc, 2016. **94-95**: p. 75-80.
64. Bodenhausen, G. and D.J. Ruben, *Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy.* Chemical Physics Letters, 1980. **69**: p. 185.
65. Kay, L.E., et al., *Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1 beta in solution.* Science, 1990. **249**(4967): p. 411-414.
66. Motáčková, V., et al., *Strategy for complete NMR assignment of disordered proteins with highly repetitive sequences based on resolution-enhanced 5D experiments.* Journal of Biomolecular NMR, 2010. **48**(3): p. 169-177.
67. Karplus, M., *Vicinal Proton Coupling in Nuclear Magnetic Resonance*. Journal of the American Chemical Society, 1963. **85**(18): p. 2870-2871.
68. Sattler, M., Schleucher, J., & Griesinger, C., *Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution.* Prog. Nucl. Magn. Reson. Spectrosc, 1999. **34**: p. 93-158.
69. Wüthrich, K., *NMR of Proteins and Nucleic Acids.* Wiley: New York, 1986.
70. Shen, Y. and A. Bax, *Protein structural information derived from NMR chemical shift with the neural network program TALOS-N.* Methods Mol Biol, 2015. **1260**: p. 17-32.
71. Case, D.A., H.J. Dyson, and P.E. Wright, *[13] Use of chemical shifts and coupling constants in nuclear magnetic resonance structural studies on peptides and proteins*, in *Methods in Enzymology*. 1994, Academic Press. p. 392-416.
72. Neuhaus, D., *Nuclear Overhauser Effect.* Encyclopedia of Magnetic Resonance, 2011.

73. Reddy, J.G., et al., *Simultaneous determination of fast and slow dynamics in molecules using extreme CPMG relaxation dispersion experiments.* J Biomol NMR, 2018. **70**(1): p. 1-9.
74. Palmer, A.G., 3rd, *Chemical exchange in biomacromolecules: past, present, and future.* J Magn Reson, 2014. **241**: p. 3-17.
75. Vallurupalli, P., G. Bouvignies, and L.E. Kay, *Studying "invisible" excited protein states in slow exchange with a major state conformation.* J Am Chem Soc, 2012. **134**(19): p. 8148-61.
76. Kleckner, I.R. and M.P. Foster, *An introduction to NMR-based approaches for measuring protein dynamics.* Biochim Biophys Acta, 2011. **1814**(8): p. 942-68.
77. Zeeb, M. and J. Balbach, *Protein folding studied by real-time NMR spectroscopy.* Methods, 2004. **34**(1): p. 65-74.
78. Möckel, C., et al., *Integrated NMR, Fluorescence, and Molecular Dynamics Benchmark Study of Protein Mechanics and Hydrodynamics.* The Journal of Physical Chemistry B, 2019. **123**(7): p. 1453-1480.
79. Jarymowycz, V.A. and M.J. Stone, *Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences.* Chemical Reviews, 2006. **106**(5): p. 1624-1671.
80. Farrow, N.A., et al., *Characterization of the backbone dynamics of folded and denatured states of an SH3 domain.* Biochemistry, 1997. **36**(9): p. 2390-402.
81. Rule G. S., H.T.K., *Fundamentals of Protein NMR Spectroscopy.* Springer Netherlands, 2006.
82. Seavey, B.R., et al., *A relational database for sequence-specific protein NMR data.* J Biomol NMR, 1991. **1**(3): p. 217-36.
83. Gasteiger, E., et al., *Protein Identification and Analysis Tools on the ExPASy Server*, in *The Proteomics Protocols Handbook*, J.M. Walker, Editor. 2005, Humana Press: Totowa, NJ. p. 571-607.
84. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.
85. Johnson, B.A. and R.A. Blevins, *NMR View: A computer program for the visualization and analysis of NMR data.* J Biomol NMR, 1994. **4**(5): p. 603-14.
86. Orekhov, V.Y., I.V. Ibraghimov, and M. Billeter, *MUNIN: a new approach to multi-dimensional NMR spectra interpretation.* J Biomol NMR, 2001. **20**(1): p. 49-60.
87. Korzhneva, D.M., et al., *MUNIN: application of three-way decomposition to the analysis of heteronuclear NMR relaxation data.* J Biomol NMR, 2001. **21**(3): p. 263-8.
88. Bernstein, H.J., *Recent changes to RasMol, recombining the variants.* Trends Biochem Sci, 2000. **25**(9): p. 453-5.
89. Dosset, P., et al., *Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data.* Journal of Biomolecular NMR, 2000. **16**(1): p. 23-28.
90. Tropea, J.E., S. Cherry, and D.S. Waugh, *Expression and Purification of Soluble His6-Tagged TEV Protease*, in *High Throughput Protein Expression and Purification: Methods and Protocols*, S.A. Doyle, Editor. 2009, Humana Press: Totowa, NJ. p. 297-307.

91. Bayburt, T.H., Y.V. Grinkova, and S.G. Sligar, *Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins.* Nano Letters, 2002. **2**(8): p. 853-856.

92. Farmer, B.T. and R.A. Venters, *Assignment of Side-Chain C-13 Resonances in Perdeuterated Proteins.* Journal of the American Chemical Society, 1995. **117**(14): p. 4187-4188.

93. Grzesiek, S. and A. Bax, *The Importance of Not Saturating H2o in Protein Nmr - Application to Sensitivity Enhancement and Noe Measurements.* Journal of the American Chemical Society, 1993. **115**(26): p. 12593-12594.

94. Nietlispach, D., et al., *An approach to the structure determination of larger proteins using triple resonance NMR experiments in conjunction with random fractional deuteration.* Journal of the American Chemical Society, 1996. **118**(2): p. 407-415.

95. Yamazaki, T., et al., *A Suite of Triple-Resonance Nmr Experiments for the Backbone Assignment of N-15, C-13, H-2 Labeled Proteins with High-Sensitivity.* Journal of the American Chemical Society, 1994. **116**(26): p. 11655-11666.

96. Yamazaki, T., et al., *An Hnca Pulse Scheme for the Backbone Assignment of N-15,C-13,H-2-Labeled Proteins - Application to a 37-Kda Trp Repressor DNA Complex.* Journal of the American Chemical Society, 1994. **116**(14): p. 6464-6465.

97. Cai, M., et al., *A simple and robust protocol for high-yield expression of perdeuterated proteins in Escherichia coli grown in shaker flasks.* J Biomol NMR, 2016. **66**(2): p. 85-91.

98. Timothy H. Bayburt, Y.V.G., and Stephen G. Sligar, *Self-Assembly of Discoidal Phospholipid Bilayer Nanoparticles with Membrane Scaffold Proteins.* 2002.

99. Pavlidou, M., et al., *Nanodiscs allow phage display selection for ligands to non-linear epitopes on membrane proteins.* PLoS One, 2013. **8**(9): p. e72272.

100. Takeda, M., et al., *Construction and performance of an NMR tube with a sample cavity formed within magnetic susceptibility-matched glass.* J Magn Reson, 2011. **209**(2): p. 167-73.

101. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

102. Findeisen, M.B., T.; Berger, S.,, *A 1HNMR thermometer suitable for cryoprobes. .* Magnetic Resonance in Chemistry, 2007. **45 (2)**: p. 175-178.

103. Nietlispach, D., *Suppression of anti-TROSY lines in a sensitivity enhanced gradient selection TROSY scheme.* Journal of Biomolecular NMR, 2005. **31**(2): p. 161-166.

104. Yang, D. and L.E. Kay, *Improved 1HN-detected triple resonance TROSY-based experiments.* J Biomol NMR, 1999. **13**(1): p. 3-10.

105. Farrow, N.A., et al., *Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation.* Biochemistry, 1994. **33**(19): p. 5984-6003.

106. Lakomek, N.A., J. Ying, and A. Bax, *Measurement of (1)(5)N relaxation rates in perdeuterated proteins by TROSY-based methods.* J Biomol NMR, 2012. **53**(3): p. 209-21.

107. Korzhnev, D.M., et al., *An NMR experiment for the accurate measurement of heteronuclear spin-lock relaxation rates.* J Am Chem Soc, 2002. **124**(36): p. 10743-53.

108. Gardner, K.H., M.K. Rosen, and L.E. Kay, *Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR.* Biochemistry, 1997. **36**(6): p. 1389-401.

109. Kharchenko, V., et al., *Dynamic 15N{1H} NOE measurements: a tool for studying protein dynamics.* Journal of Biomolecular NMR, 2020. **74**(12): p. 707-716.

110. Idiyatullin, D., V.A. Daragan, and K.H. Mayo, *Improved measurement of (15)N-[(1)H] NOEs in the presence of H(N)-water proton chemical exchange.* J Magn Reson, 2001. **153**(1): p. 138-43.

111. Gong, Q. and R. Ishima, *15N–{1H} NOE experiment at high magnetic field strengths.* Journal of Biomolecular NMR, 2007. **37**(2): p. 147-157.

112. Noggle, J.H. and R.E. Schirmer, *The nuclear Overhauser effect. Chemical application.* Academic Press, New York, 1971.

113. Möller, C., *NMR-spektroskopische Charakterisierung der Dynamik des Autophagie-relevanten Proteins GABARAP auf verschiedenen Zeitskalen.* Inaugural-Dissertation, 2015.

114. Pawley, N.H., et al., *An improved method for distinguishing between anisotropic tumbling and chemical exchange in analysis of 15N relaxation parameters.* J Biomol NMR, 2001. **20**(2): p. 149-65.

115. Hardy, R. and R.L. Cottington, *Viscosity of deuterium oxide and water in the range 5 to 125 C.* Journal of research of the National Bureau of Standards, 1949. **42**: p. 573.

116. Shen, Y. and A. Bax, *Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks.* Journal of biomolecular NMR, 2013. **56**(3): p. 227-241.

117. Schweimer, J.K., *Mehrdimensionale heteronukleare NMR Spektroskopie zur Bestimmung der Strukturen des Birkenpollenallergens Bet v 1, des Guillardia theta Rubredoxins und des [2Fe-2S] Ferredoxins aus Halobacterium salinarum.* Dissertation zur Erlangung des Doktorgrades der Fakultät für Biologie, Chemie und Geowissenschaften der Universität Bayreuth, 2000.

118. Schwieters, C.D., et al., *The Xplor-NIH NMR molecular structure determination package.* J Magn Reson, 2003. **160**(1): p. 65-73.

119. Laskowski, R.A., et al., *AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR.* J Biomol NMR, 1996. **8**(4): p. 477-86.

120. Salzmann, M., et al., *Improved sensitivity and coherence selection for [15N,1H]-TROSY elements in triple resonance experiments.* J Biomol NMR, 1999. **15**(2): p. 181-4.

121. Cavanagh, J.F., W. J.; Palmer Iii, A. G.; Skelton, N. J., *Protein NMR spectroscopy: principles and practice.* Elsevier, 1995.

122. Palmer, A.G., et al., *Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy.* Journal of Magnetic Resonance (1969), 1991. **93**(1): p. 151-170.

123. Kay, L., P. Keifer, and T. Saarinen, *Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity.* Journal of the American Chemical Society, 1992. **114**(26): p. 10663-10665.

124. Schleucher, J., et al., *A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients.* Journal of Biomolecular NMR, 1994. **4**(2): p. 301-306.

125. Cavanagh, J. and M. Rance, *Suppression of cross-relaxation effects in TOCSY spectra via a modified DIPSI-2 mixing sequence.* Journal of Magnetic Resonance (1969), 1992. **96**(3): p. 670-678.

126. Zhang, O., et al., *Backbone 1H and 15N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques.* Journal of Biomolecular NMR, 1994. **4**(6): p. 845-858.

127. Wagner, B., *Sequenzspezifische Zuordnung der NMR-Seitenkettenresonanzen des Lipoproteins aus C. difficile.* Bachelorarbeit, 2020.

128. Griesinger C., H.M., Marino J.P., Reft B., Richter C., Schwalbe H. , *Methods for the Determination of Torsion Angle Restraints in Biomacromolecules.* Biological Magnetic Resonance, 2002. **16**.

129. Mielke, S.P. and V.V. Krishnan, *Characterization of protein secondary structure from NMR chemical shifts.* Prog Nucl Magn Reson Spectrosc, 2009. **54**(3-4): p. 141-165.

130. Ramachandran, G.N., C. Ramakrishnan, and V. Sasisekharan, *Stereochemistry of polypeptide chain configurations.* J Mol Biol, 1963. **7**: p. 95-9.

131. Butts, C.P., et al., *Interproton distance determinations by NOE--surprising accuracy and precision in a rigid organic molecule.* Org Biomol Chem, 2011. **9**(1): p. 177-84.

132. Kienapfel, V., *Strukturelle Charakterisierung des Lipoproteins aus C. difficile mittels NMR-Spektroskopie.* Bachelorarbeit, 2019.

133. Huang, X. and R. Powers, *Validity of Using the Radius of Gyration as a Restraint in NMR Protein Structure Determination.* Journal of the American Chemical Society, 2001. **123**(16): p. 3834-3835.

134. López-Méndez, B. and P. Güntert, *Automated Protein Structure Determination from NMR Spectra.* Journal of the American Chemical Society, 2006. **128**(40): p. 13112-13122.

135. Morris, A.L., et al., *Stereochemical quality of protein structure coordinates.* Proteins, 1992. **12**(4): p. 345-64.

Abbreviations

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
4D	Four-dimensional
ABC transporter	The ATP-Binding Cassette transporter
Atg	Autophagy-related
BMRB	Biological magnetic resonance data bank
CAMP	Cationic antimicrobial peptides
CMA	Chaperone-mediated autophagy
Cpr	Cationic antimicrobial peptide resistance
CSA	Chemical shift anisotropy
CV	Column volume
C-terminus	Carboxy-terminus
DD	Dipole-dipole
Dha	Didehydroalanine
Dhb	Didehydrobutyrine
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
DTT	Dithiothreitol
ε	Molar extinction coefficient
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic reticulum
FID	Free Induction Decay
GABARAP	Gamma-aminobutyric acid receptor-associated protein
HSQC	Heteronuclear Single Quantum Correlation
IPTG	Isopropyl-β-D-thiogalaktopyranosid
LB	Lysogeny broth
MD	Molekulardynamik
MPB-PE	11,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-N-[4-(pmaleimidophenyl)butyramide]
MSP	Membrane Scaffold Protein
MSP1D1Δ5	Membrane Scaffold Protein delta 5 α-helix
MWCO	Molecular weight cut-off
ND	Nanodisc

Abbreviations

N-Terminus	Amino-Terminus
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
PAS	Phagophore assembly site
PDB	Protein Data Bank
PE	Phosphatidylethanolamine
RF	Radio Frequency
RMSD	Root Mean Square Deviation
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
TALOS	Torsion angle likelihood obtained from shift and sequence similarity
TROSY	Transverse relaxation optimized spectroscopy
v/v	Volume per volume
UV	Ultraviolet
wt	Wildtype

The one- and three-letter code was used for amino acids:

Alanine	Ala A
Arginine	Arg R
Asparagine	Asn N
Aspartic acid	Asp D
Cysteine	Cys C
Glutamic acid	Glu E
Glutamine	Gln Q
Glycine	Gly G
Histidine	His H
Isoleucine	Ile I
Leucine	Leu L
Lysine	Lys K
Methionine	Met M
Phenylalanine	Phe F
Proline	Pro P
Serine	Ser S
Threonine	Thr T
Tryptophan	Trp W
Tyrosine	Tyr Y
Valine	Val V

Appendix

Chemical shifts in ppm of the lipoprotein CD1348 assignment:

0	17	SER	CA	58.41	1086	101	PHE	N	115.53
1	17	SER	HA	4.485	1087	101	PHE	HN	7.567
2	17	SER	CB	63.9	1088	101	PHE	CA	57.77
3	17	SER	HB2	3.856	1089	101	PHE	HA	4.196
4	17	SER	HB1	3.856	1090	101	PHE	CB	38.55
5	17	SER	C	174.61	1091	101	PHE	HB2	2.629
6	18	SER	N	117.99	1092	101	PHE	HB1	2.956
7	18	SER	HN	8.486	1093	101	PHE	CD1	132.8
8	18	SER	CA	58.7	1094	101	PHE	HD1	6.118
9	18	SER	HA	4.461	1095	101	PHE	CE1	130.876
10	18	SER	CB	63.89	1096	101	PHE	HE1	7.376
11	18	SER	HB2	3.897	1097	101	PHE	CZ	128.7
12	18	SER	HB1	3.897	1098	101	PHE	HZ	7.665
13	18	SER	C	175.01	1099	101	PHE	CE2	130.876
14	19	GLY	N	110.9	1100	101	PHE	HE2	7.376
15	19	GLY	HN	8.411	1101	101	PHE	CD2	132.8
16	19	GLY	CA	45.34	1102	101	PHE	HD2	6.118
17	19	GLY	HA2	3.961	1103	101	PHE	C	173.85
18	19	GLY	HA1	3.961	1104	102	ASN	N	114.634
19	19	GLY	C	173.84	1105	102	ASN	HN	8.458
20	20	ALA	N	123.55	1106	102	ASN	CA	54.95
21	20	ALA	HN	8.099	1107	102	ASN	HA	4.215
22	20	ALA	CA	52.61	1108	102	ASN	CB	39.1
23	20	ALA	HA	4.29	1109	102	ASN	HB2	2.583
24	20	ALA	CB	19.32	1110	102	ASN	HB1	3.075
25	20	ALA	HB1	1.353	1111	102	ASN	CG	179.447
26	20	ALA	HB2	1.353	1112	102	ASN	ND2	113.611
27	20	ALA	HB3	1.353	1113	102	ASN	HD21	6.952
28	20	ALA	C	177.74	1114	102	ASN	HD22	6.876
29	21	MET	N	118.87	1115	102	ASN	C	174.88
30	21	MET	HN	8.25	1116	103	GLY	N	103.077
31	21	MET	CA	55.34	1117	103	GLY	HN	8.651
32	21	MET	HA	4.394	1118	103	GLY	CA	45.45
33	21	MET	CB	32.87	1119	103	GLY	HA2	3.458
34	21	MET	HB2	1.925	1120	103	GLY	HA1	3.987
35	21	MET	HB1	1.925	1121	103	GLY	C	175.15
36	21	MET	CG	31.85	1122	104	LEU	N	121.92
37	21	MET	HG2	2.442	1123	104	LEU	HN	7.187
38	21	MET	HG1	2.442	1124	104	LEU	CA	55.63
39	21	MET	C	175.57	1125	104	LEU	HA	4.079
40	22	ASP	N	121.16	1126	104	LEU	CB	41.42
41	22	ASP	HN	8.143	1127	104	LEU	HB2	0.937
42	22	ASP	CA	54.14	1128	104	LEU	HB1	0.969
43	22	ASP	HA	4.569	1129	104	LEU	CG	25.69
44	22	ASP	CB	41.2	1130	104	LEU	HG	1.424
45	22	ASP	HB2	2.603	1131	104	LEU	CD1	25.35
46	22	ASP	HB1	2.706	1132	104	LEU	HD11	-0.375
47	22	ASP	C	175.92	1133	104	LEU	HD12	-0.375
48	23	TYR	N	120.83	1134	104	LEU	HD13	-0.375
49	23	TYR	HN	8.072	1135	104	LEU	CD2	22.8
50	23	TYR	CA	58.01	1136	104	LEU	HD21	0.652
51	23	TYR	HA	4.563	1137	104	LEU	HD22	0.652
52	23	TYR	CB	38.6	1138	104	LEU	HD23	0.652
53	23	TYR	HB2	2.904	1139	104	LEU	C	177.4
54	23	TYR	HB1	3.083	1140	105	ASP	N	121.2
55	23	TYR	CD1	133.212	1141	105	ASP	HN	9.159
56	23	TYR	HD1	7.084	1142	105	ASP	CA	55.23
57	23	TYR	CE1	118.329	1143	105	ASP	HA	4.861
58	23	TYR	HE1	6.829	1144	105	ASP	CB	43.39
59	23	TYR	CE2	118.329	1145	105	ASP	HB2	2.68
60	23	TYR	HE2	6.829	1146	105	ASP	HB1	2.741
61	23	TYR	CD2	133.212	1147	105	ASP	C	175.67
62	23	TYR	HD2	7.084	1148	106	MET	N	117.065
63	23	TYR	C	175.79	1149	106	MET	HN	8.496
64	24	SER	N	117.08	1150	106	MET	CA	54.62
65	24	SER	HN	8.251	1151	106	MET	HA	5.115
66	24	SER	CA	58.62	1152	106	MET	CB	38.54
67	24	SER	HA	4.405	1153	106	MET	HB2	1.778
68	24	SER	CB	63.92	1154	106	MET	HB1	1.844
69	24	SER	HB2	3.838	1155	106	MET	CG	33.63
70	24	SER	HB1	3.838	1156	106	MET	HG2	2.377
71	24	SER	C	174.41	1157	106	MET	HG1	2.404
72	25	ILE	N	121.865	1158	106	MET	CE	19.2
73	25	ILE	HN	7.983	1159	106	MET	HE1	2.093
74	25	ILE	CA	61.32	1160	106	MET	HE2	2.093
75	25	ILE	HA	4.208	1161	106	MET	HE3	2.093
76	25	ILE	CB	38.89	1162	106	MET	C	175.17
77	25	ILE	HB	1.869	1163	107	GLU	N	122.68
78	25	ILE	CG1	27.31	1164	107	GLU	HN	9.029
79	25	ILE	HG12	1.177	1165	107	GLU	CA	54.73
80	25	ILE	HG11	1.456	1166	107	GLU	HA	4.626
81	25	ILE	CD1	13.06	1167	107	GLU	CB	33.09
82	25	ILE	HD11	0.846	1168	107	GLU	HB2	1.889
83	25	ILE	HD12	0.846	1169	107	GLU	HB1	2.074
84	25	ILE	HD13	0.846	1170	107	GLU	CG	36.07
85	25	ILE	CG2	17.54	1171	107	GLU	HG2	2.252
86	25	ILE	HG21	0.897	1172	107	GLU	HG1	2.252
87	25	ILE	HG22	0.897	1173	107	GLU	C	174.83
88	25	ILE	HG23	0.897	1174	108	VAL	N	120.67
89	25	ILE	C	176.19	1175	108	VAL	HN	8.653
90	26	SER	N	119.13	1176	108	VAL	CA	59.15
91	26	SER	HN	8.254	1177	108	VAL	HA	5.223
92	26	SER	CA	58.4	1178	108	VAL	CB	35.24
93	26	SER	HA	4.452	1179	108	VAL	HB	1.785
94	26	SER	CB	63.89	1180	108	VAL	CG2	19.18
95	26	SER	HB2	3.848	1181	108	VAL	HG21	0.834
96	26	SER	HB1	3.848	1182	108	VAL	HG22	0.834
97	26	SER	C	174.44	1183	108	VAL	HG23	0.834
98	27	ALA	N	126.09	1184	108	VAL	CG1	23.09
99	27	ALA	HN	8.246	1185	108	VAL	HG11	0.797
100	27	ALA	CA	52.72	1186	108	VAL	HG12	0.797
101	27	ALA	HA	4.346	1187	108	VAL	HG13	0.797
102	27	ALA	CB	19.33	1188	108	VAL	C	172.815
103	27	ALA	HB1	1.38	1189	109	GLU	N	127.93
104	27	ALA	HB2	1.38	1190	109	GLU	HN	9.027
105	27	ALA	HB3	1.38	1191	109	GLU	CA	54.87
106	27	ALA	C	177.67	1192	109	GLU	HA	4.629
107	28	VAL	N	118.32	1193	109	GLU	CB	33.16
108	28	VAL	HN	7.938	1194	109	GLU	HB2	1.838

109	28	VAL	CA	62.48	1195	109	GLU	HB1	1.945
110	28	VAL	HA	4.079	1196	109	GLU	CG	36.122
111	28	VAL	CB	32.73	1197	109	GLU	HG2	2.182
112	28	VAL	HB	2.037	1198	109	GLU	HG1	2.182
113	28	VAL	CG2	20.76	1199	109	GLU	C	173.11
114	28	VAL	HG21	0.917	1200	110	VAL	N	125.4
115	28	VAL	HG22	0.917	1201	110	VAL	HN	8.973
116	28	VAL	HG23	0.917	1202	110	VAL	CA	61.79
117	28	VAL	CG1	21.16	1203	110	VAL	HA	4.356
118	28	VAL	HG11	0.901	1204	110	VAL	CB	33.21
119	28	VAL	HG12	0.901	1205	110	VAL	HB	2.089
120	28	VAL	HG13	0.901	1206	110	VAL	CG2	21.59
121	28	VAL	C	176.07	1207	110	VAL	HG21	0.855
122	29	GLU	N	123.87	1208	110	VAL	HG22	0.855
123	29	GLU	HN	8.333	1209	110	VAL	HG23	0.855
124	29	GLU	CA	56.53	1210	110	VAL	CG1	21.62
125	29	GLU	HA	4.291	1211	110	VAL	HG11	0.855
126	29	GLU	CB	30.34	1212	110	VAL	HG12	0.855
127	29	GLU	HB2	1.94	1213	110	VAL	HG13	0.855
128	29	GLU	HB1	2.059	1214	110	VAL	C	175.23
129	29	GLU	CG	36.24	1215	111	LYS	N	129.15
130	29	GLU	HG2	2.258	1216	111	LYS	HN	8.963
131	29	GLU	HG1	2.258	1217	111	LYS	CA	54.09
132	29	GLU	C	176.17	1218	111	LYS	HA	4.727
133	30	LEU	N	123.67	1219	111	LYS	CB	30.84
134	30	LEU	HN	8.211	1220	111	LYS	HB2	1.898
135	30	LEU	CA	55.05	1221	111	LYS	HB1	1.97
136	30	LEU	HA	4.387	1222	111	LYS	CG	24.09
137	30	LEU	CB	42.3	1223	111	LYS	HG2	1.401
138	30	LEU	HB2	1.58	1224	111	LYS	HG1	1.303
139	30	LEU	HB1	1.641	1225	111	LYS	CD	28.66
140	30	LEU	CG	27.01	1226	111	LYS	HD2	1.638
141	30	LEU	HG	1.597	1227	111	LYS	HD1	1.638
142	30	LEU	CD1	25.1	1228	111	LYS	CE	41.92
143	30	LEU	HD11	0.894	1229	111	LYS	HE2	2.912
144	30	LEU	HD12	0.894	1230	111	LYS	HE1	2.912
145	30	LEU	HD13	0.894	1231	111	LYS	C	176.73
146	30	LEU	CD2	24.57	1232	112	GLU	N	124.29
147	30	LEU	HD21	0.833	1233	112	GLU	HN	8.886
148	30	LEU	HD22	0.833	1234	112	GLU	CA	60.52
149	30	LEU	HD23	0.833	1235	112	GLU	HA	3.849
150	30	LEU	C	177.225	1236	112	GLU	CB	29.61
151	31	VAL	N	120.87	1237	112	GLU	HB2	2.049
152	31	VAL	HN	8.048	1238	112	GLU	HB1	2.049
153	31	VAL	CA	62.34	1239	112	GLU	CG	36.84
154	31	VAL	HA	4.104	1240	112	GLU	HG2	2.256
155	31	VAL	CB	33.02	1241	112	GLU	HG1	2.299
156	31	VAL	HB	2.058	1242	112	GLU	C	179.18
157	31	VAL	CG2	20.61	1243	113	LYS	N	118.21
158	31	VAL	HG21	0.975	1244	113	LYS	HN	8.732
159	31	VAL	HG22	0.975	1245	113	LYS	CA	58.42
160	31	VAL	HG23	0.975	1246	113	LYS	HA	4.119
161	31	VAL	CG1	21.16	1247	113	LYS	CB	31.72
162	31	VAL	HG11	0.901	1248	113	LYS	HB2	1.801
163	31	VAL	HG12	0.901	1249	113	LYS	HB1	1.844
164	31	VAL	HG13	0.901	1250	113	LYS	CG	24.8
165	31	VAL	C	175.8	1251	113	LYS	HG2	1.451
166	32	ASP	N	124	1252	113	LYS	HG1	1.451
167	32	ASP	HN	8.393	1253	113	LYS	CD	29.06
168	32	ASP	CA	54.34	1254	113	LYS	HD2	1.68
169	32	ASP	HA	4.628	1255	113	LYS	HD1	1.68
170	32	ASP	CB	41.31	1256	113	LYS	CE	42.16
171	32	ASP	HB2	2.607	1257	113	LYS	HE2	2.995
172	32	ASP	HB1	2.709	1258	113	LYS	HE1	2.995
173	32	ASP	C	176.35	1259	113	LYS	C	176.51
174	33	SER	N	116.86	1260	114	ASP	N	115.56
175	33	SER	HN	8.245	1261	114	ASP	HN	7.156
176	33	SER	CA	58.58	1262	114	ASP	CA	53.75
177	33	SER	HA	4.408	1263	114	ASP	HA	4.761
178	33	SER	CB	63.84	1264	114	ASP	CB	42.7
179	33	SER	HB2	3.863	1265	114	ASP	HB2	2.306
180	33	SER	HB1	3.909	1266	114	ASP	HB1	2.881
181	33	SER	C	174.99	1267	114	ASP	C	174.59
182	34	LYS	N	122.97	1268	115	ASN	N	116.01
183	34	LYS	HN	8.385	1269	115	ASN	HN	7.966
184	34	LYS	CA	56.84	1270	115	ASN	CA	54.4
185	34	LYS	HA	4.294	1271	115	ASN	HA	4.279
186	34	LYS	CB	32.73	1272	115	ASN	CB	37.84
187	34	LYS	HB2	1.748	1273	115	ASN	HB2	2.887
188	34	LYS	HB1	1.844	1274	115	ASN	HB1	2.939
189	34	LYS	CG	24.74	1275	115	ASN	CG	177.429
190	34	LYS	HG2	1.435	1276	115	ASN	ND2	111.966
191	34	LYS	HG1	1.435	1277	115	ASN	HD21	7.564
192	34	LYS	CD	29.07	1278	115	ASN	HD22	6.552
193	34	LYS	HD2	1.678	1279	115	ASN	C	173.28
194	34	LYS	HD1	1.678	1280	116	LEU	N	118.71
195	34	LYS	CE	42.15	1281	116	LEU	HN	7.337
196	34	LYS	HE2	2.997	1282	116	LEU	CA	54.27
197	34	LYS	HE1	2.997	1283	116	LEU	HA	5.22
198	34	LYS	C	176.96	1284	116	LEU	CB	46.85
199	35	GLU	N	120.89	1285	116	LEU	HB2	1.409
200	35	GLU	HN	8.313	1286	116	LEU	HB1	1.475
201	35	GLU	CA	56.82	1287	116	LEU	CG	27.46
202	35	GLU	HA	4.284	1288	116	LEU	HG	1.296
203	35	GLU	CB	30.2	1289	116	LEU	CD1	26.18
204	35	GLU	HB2	1.942	1290	116	LEU	HD11	0.76
205	35	GLU	HB1	2.058	1291	116	LEU	HD12	0.76
206	35	GLU	CG	36.24	1292	116	LEU	HD13	0.76
207	35	GLU	HG2	2.261	1293	116	LEU	CD2	24.47
208	35	GLU	HG1	2.261	1294	116	LEU	HD21	0.763
209	35	GLU	C	176.72	1295	116	LEU	HD22	0.763
210	36	SER	N	116.4	1296	116	LEU	HD23	0.763
211	36	SER	HN	8.204	1297	116	LEU	C	174.62
212	36	SER	CA	58.49	1298	117	VAL	N	123.08
213	36	SER	HA	4.412	1299	117	VAL	HN	8.33
214	36	SER	CB	63.84	1300	117	VAL	CA	60.44
215	36	SER	HB2	3.859	1301	117	VAL	HA	4.724
216	36	SER	HB1	3.859	1302	117	VAL	CB	34.86
217	36	SER	C	174.27	1303	117	VAL	HB	1.874
218	37	ALA	N	125.89	1304	117	VAL	CG2	22.095
219	37	ALA	HN	8.22	1305	117	VAL	HG21	0.874
220	37	ALA	CA	52.56	1306	117	VAL	HG22	0.874
221	37	ALA	HA	4.354	1307	117	VAL	HG23	0.874
222	37	ALA	CB	19.31	1308	117	VAL	CG1	22.936
223	37	ALA	HB1	1.383	1309	117	VAL	HG11	0.9
224	37	ALA	HB2	1.383	1310	117	VAL	HG12	0.9
225	37	ALA	HB3	1.383	1311	117	VAL	HG13	0.9
226	37	ALA	C	177.56	1312	117	VAL	C	173.11

Appendix

227	38	VAL	N	119.42	1313	118	LYS	N	125.43
228	38	VAL	HN	7.99	1314	118	LYS	HN	8.854
229	38	VAL	CA	62.49	1315	118	LYS	CA	54.64
230	38	VAL	HA	4.076	1316	118	LYS	HA	5.371
231	38	VAL	CB	32.72	1317	118	LYS	CB	35.1
232	38	VAL	HB	2.03	1318	118	LYS	HB2	1.677
233	38	VAL	CG2	20.76	1319	118	LYS	HB1	1.677
234	38	VAL	HG21	0.933	1320	118	LYS	CG	25.28
235	38	VAL	HG22	0.933	1321	118	LYS	HG2	1.342
236	38	VAL	HG23	0.933	1322	118	LYS	HG1	1.426
237	38	VAL	CG1	21.16	1323	118	LYS	CD	29.76
238	38	VAL	HG11	0.901	1324	118	LYS	HD2	1.449
239	38	VAL	HG12	0.901	1325	118	LYS	HD1	1.546
240	38	VAL	HG13	0.901	1326	118	LYS	CE	41.79
241	38	VAL	C	176.1	1327	118	LYS	HE2	2.588
242	39	VAL	N	124.79	1328	118	LYS	HE1	2.687
243	39	VAL	HN	8.14	1329	118	LYS	C	175.28
244	39	VAL	CA	62.36	1330	119	ILE	N	126.02
245	39	VAL	HA	4.08	1331	119	ILE	HN	9.423
246	39	VAL	CB	32.77	1332	119	ILE	CA	59.12
247	39	VAL	HB	2.028	1333	119	ILE	HA	4.953
248	39	VAL	CG2	20.61	1334	119	ILE	CB	40.52
249	39	VAL	HG21	0.915	1335	119	ILE	HB	1.712
250	39	VAL	HG22	0.915	1336	119	ILE	CG1	28.18
251	39	VAL	HG23	0.915	1337	119	ILE	HG12	0.995
252	39	VAL	CG1	21.16	1338	119	ILE	HG11	1.436
253	39	VAL	HG11	0.901	1339	119	ILE	CD1	14.24
254	39	VAL	HG12	0.901	1340	119	ILE	HD11	0.702
255	39	VAL	HG13	0.901	1341	119	ILE	HD12	0.702
256	39	VAL	C	175.85	1342	119	ILE	HD13	0.702
257	40	LYS	N	126.39	1343	119	ILE	CG2	17.44
258	40	LYS	HN	8.389	1344	119	ILE	HG21	0.724
259	40	LYS	CA	56.15	1345	119	ILE	HG22	0.724
260	40	LYS	HA	4.33	1346	119	ILE	HG23	0.724
261	40	LYS	CB	33.19	1347	119	ILE	C	173.78
262	40	LYS	HB2	1.746	1348	120	ASN	N	125.2
263	40	LYS	HB1	1.835	1349	120	ASN	HN	8.882
264	40	LYS	CG	24.7	1350	120	ASN	CA	51.35
265	40	LYS	HG2	1.415	1351	120	ASN	HA	5.363
266	40	LYS	HG1	1.415	1352	120	ASN	CB	41.17
267	40	LYS	CD	29.07	1353	120	ASN	HB2	2.512
268	40	LYS	HD2	1.679	1354	120	ASN	HB1	2.512
269	40	LYS	HD1	1.679	1355	120	ASN	CG	175.596
270	40	LYS	CE	42.13	1356	120	ASN	ND2	111.693
271	40	LYS	HE2	2.985	1357	120	ASN	HD21	7.391
272	40	LYS	HE1	2.985	1358	120	ASN	HD22	6.516
273	40	LYS	C	176.164	1359	120	ASN	C	174.01
274	41	LYS	N	123.86	1360	121	LEU	N	126.16
275	41	LYS	HN	8.429	1361	121	LEU	HN	8.536
276	41	LYS	CA	56.24	1362	121	LEU	CA	54.72
277	41	LYS	HA	4.329	1363	121	LEU	HA	3.85
278	41	LYS	CB	33.24	1364	121	LEU	CB	38.62
279	41	LYS	HB2	1.746	1365	121	LEU	HB2	-0.897
280	41	LYS	HB1	1.835	1366	121	LEU	HB1	1.423
281	41	LYS	CG	24.7	1367	121	LEU	CG	25.44
282	41	LYS	HG2	1.423	1368	121	LEU	HG	1.134
283	41	LYS	HG1	1.423	1369	121	LEU	CD1	24.98
284	41	LYS	CD	29.08	1370	121	LEU	HD11	0.024
285	41	LYS	HD2	1.678	1371	121	LEU	HD12	0.024
286	41	LYS	HD1	1.678	1372	121	LEU	HD13	0.024
287	41	LYS	CE	42.11	1373	121	LEU	CD2	21.86
288	41	LYS	HE2	2.993	1374	121	LEU	HD21	-0.393
289	41	LYS	HE1	2.993	1375	121	LEU	HD22	-0.393
290	41	LYS	C	176.27	1376	121	LEU	HD23	-0.393
291	42	ASP	N	122.09	1377	121	LEU	C	175.02
292	42	ASP	HN	8.415	1378	122	ILE	N	123.99
293	42	ASP	CA	54.63	1379	122	ILE	HN	7.703
294	42	ASP	HA	4.571	1380	122	ILE	CA	59.39
295	42	ASP	CB	41.16	1381	122	ILE	HA	4.404
296	42	ASP	HB2	2.605	1382	122	ILE	CB	39.31
297	42	ASP	HB1	2.705	1383	122	ILE	HB	1.839
298	42	ASP	C	176.47	1384	122	ILE	CG1	27.33
299	43	GLU	N	121.07	1385	122	ILE	HG12	1.167
300	43	GLU	HN	8.45	1386	122	ILE	HG11	1.38
301	43	GLU	CA	56.97	1387	122	ILE	CD1	11.71
302	43	GLU	HA	4.235	1388	122	ILE	HD11	0.652
303	43	GLU	CB	30.28	1389	122	ILE	HD12	0.652
304	43	GLU	HB2	1.932	1390	122	ILE	HD13	0.652
305	43	GLU	HB1	2.069	1391	122	ILE	CG2	17.52
306	43	GLU	CG	36.24	1392	122	ILE	HG21	0.935
307	43	GLU	HG2	2.258	1393	122	ILE	HG22	0.935
308	43	GLU	HG1	2.258	1394	122	ILE	HG23	0.935
309	43	GLU	C	176.435	1395	122	ILE	C	176.74
310	44	ASP	N	120.96	1396	123	GLU	N	127.28
311	44	ASP	HN	8.301	1397	123	GLU	HN	8.834
312	44	ASP	CA	54.59	1398	123	GLU	CA	54.01
313	44	ASP	HA	4.572	1399	123	GLU	HA	5.806
314	44	ASP	CB	41.18	1400	123	GLU	CB	29.176
315	44	ASP	HB2	2.602	1401	123	GLU	HB2	1.949
316	44	ASP	HB1	2.706	1402	123	GLU	HB1	2.368
317	44	ASP	C	176.13	1403	123	GLU	CG	34.2
318	45	ALA	N	124.26	1404	123	GLU	HG2	2.724
319	45	ALA	HN	8.08	1405	123	GLU	HG1	2.819
320	45	ALA	CA	52.61	1406	124	PRO	CA	62.71
321	45	ALA	HA	4.298	1407	124	PRO	HA	4.638
322	45	ALA	CB	19.15	1408	124	PRO	CB	32.69
323	45	ALA	HB1	1.387	1409	124	PRO	HB2	2.192
324	45	ALA	HB2	1.387	1410	124	PRO	HB1	2.444
325	45	ALA	HB3	1.387	1411	124	PRO	CG	27.36
326	45	ALA	C	177.75	1412	124	PRO	HG2	2.156
327	46	LYS	N	120.56	1413	124	PRO	HG1	2.184
328	46	LYS	HN	8.228	1414	124	PRO	CD	50.99
329	46	LYS	CA	56.45	1415	124	PRO	HD2	4.105
330	46	LYS	HA	4.293	1416	124	PRO	HD1	4.247
331	46	LYS	CB	33.06	1417	124	PRO	C	176.74
332	46	LYS	HB2	1.751	1418	125	ASP	N	120.82
333	46	LYS	HB1	1.839	1419	125	ASP	HN	8.491
334	46	LYS	CG	24.7	1420	125	ASP	CA	55.11
335	46	LYS	HG2	1.428	1421	125	ASP	HA	4.498
336	46	LYS	HG1	1.428	1422	125	ASP	CB	41.19
337	46	LYS	CD	29.05	1423	125	ASP	HB2	2.699
338	46	LYS	HD2	1.674	1424	125	ASP	HB1	2.785
339	46	LYS	HD1	1.674	1425	125	ASP	C	176.25
340	46	LYS	CE	42.15	1426	126	LYS	N	117.96
341	46	LYS	HE2	2.996	1427	126	LYS	HN	7.697
342	46	LYS	HE1	2.996	1428	126	LYS	CA	55.24
343	46	LYS	C	176.77	1429	126	LYS	HA	4.422
344	47	GLU	N	121.82	1430	126	LYS	CB	33.31

345	47	GLU	HN	8.376
346	47	GLU	CA	56.69
347	47	GLU	HA	4.285
348	47	GLU	CB	30.3
349	47	GLU	HB2	1.942
350	47	GLU	HB1	2.056
351	47	GLU	CG	36.25
352	47	GLU	HG2	2.258
353	47	GLU	HG1	2.258
354	47	GLU	C	176.66
355	48	GLU	N	122.32
356	48	GLU	HN	8.503
357	48	GLU	CA	56.82
358	48	GLU	HA	4.358
359	48	GLU	CB	30.3
360	48	GLU	HB2	1.952
361	48	GLU	HB1	2.069
362	48	GLU	CG	36.25
363	48	GLU	HG2	2.266
364	48	GLU	HG1	2.266
365	48	GLU	C	176.94
366	49	THR	N	115.05
367	49	THR	HN	8.259
368	49	THR	CA	62.37
369	49	THR	HA	4.377
370	49	THR	CB	69.7
371	49	THR	HB	4.271
372	49	THR	CG2	21.71
373	49	THR	HG21	1.226
374	49	THR	HG22	1.226
375	49	THR	HG23	1.226
376	49	THR	C	175.15
377	50	THR	N	115.965
378	50	THR	HN	8.13
379	50	THR	CA	62.43
380	50	THR	HA	4.348
381	50	THR	CB	69.7
382	50	THR	HB	4.271
383	50	THR	CG2	21.72
384	50	THR	HG21	1.229
385	50	THR	HG22	1.229
386	50	THR	HG23	1.229
387	50	THR	C	175
388	51	SER	N	117.89
389	51	SER	HN	8.301
390	51	SER	CA	58.83
391	51	SER	HA	4.415
392	51	SER	CB	63.68
393	51	SER	HB2	3.861
394	51	SER	HB1	3.916
395	51	SER	C	174.76
396	52	LYS	N	122.73
397	52	LYS	HN	8.203
398	52	LYS	CA	56.52
399	52	LYS	HA	4.309
400	52	LYS	CB	33.06
401	52	LYS	HB2	1.746
402	52	LYS	HB1	1.835
403	52	LYS	CG	24.7
404	52	LYS	HG2	1.436
405	52	LYS	HG1	1.928
406	52	LYS	CD	29.05
407	52	LYS	HD2	1.674
408	52	LYS	HD1	1.674
409	52	LYS	CE	42.16
410	52	LYS	HE2	2.989
411	52	LYS	HE1	2.989
412	52	LYS	C	176.58
413	53	MET	N	120.88
414	53	MET	HN	8.214
415	53	MET	CA	55.68
416	53	MET	HA	4.46
417	53	MET	CB	32.77
418	53	MET	HB2	2.011
419	53	MET	HB1	2.083
420	53	MET	CG	32.09
421	53	MET	HG2	2.517
422	53	MET	HG1	2.592
423	53	MET	C	176.28
424	54	ILE	N	121.74
425	54	ILE	HN	8.086
426	54	ILE	CA	61.33
427	54	ILE	HA	4.145
428	54	ILE	CB	38.82
429	54	ILE	HB	1.863
430	54	ILE	CG1	27.32
431	54	ILE	HG12	1.182
432	54	ILE	HG11	1.459
433	54	ILE	CD1	12.95
434	54	ILE	HD11	0.848
435	54	ILE	HD12	0.848
436	54	ILE	HD13	0.848
437	54	ILE	CG2	17.52
438	54	ILE	HG21	0.894
439	54	ILE	HG22	0.894
440	54	ILE	HG23	0.894
441	54	ILE	C	176.02
442	55	ASN	N	122.42
443	55	ASN	HN	8.48
444	55	ASN	CA	53.27
445	55	ASN	HA	4.762
446	55	ASN	CB	38.97
447	55	ASN	HB2	2.76
448	55	ASN	HB1	2.866
449	55	ASN	CG	177.112
450	55	ASN	ND2	112.664
451	55	ASN	HD21	7.585
452	55	ASN	HD22	6.895
453	55	ASN	C	175.11
454	56	SER	N	116.46
455	56	SER	HN	8.206
456	56	SER	CA	58.53
457	56	SER	HA	4.42
458	56	SER	CB	63.85
459	56	SER	HB2	3.863
460	56	SER	HB1	3.916
461	56	SER	C	174.32
462	57	LYS	N	122.97
1431	126	LYS	HB2	1.745
1432	126	LYS	HB1	1.835
1433	126	LYS	CG	24.31
1434	126	LYS	HG2	1.407
1435	126	LYS	HG1	1.407
1436	126	LYS	CD	29.08
1437	126	LYS	HD2	1.68
1438	126	LYS	HD1	1.68
1439	126	LYS	CE	42.13
1440	126	LYS	HE2	3.002
1441	126	LYS	HE1	3.002
1442	126	LYS	C	175.78
1443	127	LYS	N	121.71
1444	127	LYS	HN	8.301
1445	127	LYS	CA	57.66
1446	127	LYS	HA	4.115
1447	127	LYS	CB	32.09
1448	127	LYS	HB2	1.797
1449	127	LYS	HB1	1.797
1450	127	LYS	CG	24.74
1451	127	LYS	HG2	1.428
1452	127	LYS	HG1	1.428
1453	127	LYS	CD	29.07
1454	127	LYS	HD2	1.68
1455	127	LYS	HD1	1.68
1456	127	LYS	CE	42.16
1457	127	LYS	HE2	2.993
1458	127	LYS	HE1	2.993
1459	127	LYS	C	176.75
1460	128	SER	N	119.53
1461	128	SER	HN	8.498
1462	128	SER	CA	57.12
1463	128	SER	HA	4.66
1464	128	SER	CB	64.28
1465	128	SER	HB2	3.877
1466	128	SER	HB1	3.935
1467	128	SER	C	174.66
1468	129	ARG	N	123.02
1469	129	ARG	HN	8.642
1470	129	ARG	CA	57.99
1471	129	ARG	HA	4.244
1472	129	ARG	CB	30.8
1473	129	ARG	HB2	1.886
1474	129	ARG	HB1	1.886
1475	129	ARG	CG	27.4
1476	129	ARG	HG2	1.667
1477	129	ARG	HG1	1.667
1478	129	ARG	CD	43.25
1479	129	ARG	HD2	3.205
1480	129	ARG	HD1	3.205
1481	129	ARG	C	175.88
1482	130	VAL	N	111.82
1483	130	VAL	HN	7.466
1484	130	VAL	CA	59.16
1485	130	VAL	HA	4.829
1486	130	VAL	CB	35.91
1487	130	VAL	HB	1.935
1488	130	VAL	CG2	19.44
1489	130	VAL	HG21	0.825
1490	130	VAL	HG22	0.825
1491	130	VAL	HG23	0.825
1492	130	VAL	CG1	21.69
1493	130	VAL	HG11	0.818
1494	130	VAL	HG12	0.818
1495	130	VAL	HG13	0.818
1496	130	VAL	C	174.73
1497	131	SER	N	117.24
1498	131	SER	HN	9.492
1499	131	SER	CA	56.8
1500	131	SER	HA	4.761
1501	131	SER	CB	66.8
1502	131	SER	HB2	3.883
1503	131	SER	HB1	4.492
1504	131	SER	C	175.04
1505	132	TRP	N	126.25
1506	132	TRP	HN	10.955
1507	132	TRP	CA	61.62
1508	132	TRP	HA	4.445
1509	132	TRP	CB	29.18
1510	132	TRP	HB2	2.79
1511	132	TRP	HB1	3.351
1512	132	TRP	CD1	126.691
1513	132	TRP	HD1	6.951
1514	132	TRP	NE1	129.63
1515	132	TRP	HE1	9.537
1516	132	TRP	CZ2	115.923
1517	132	TRP	HZ2	6.772
1518	132	TRP	CH2	124.299
1519	132	TRP	HH2	5.695
1520	132	TRP	CZ3	122.256
1521	132	TRP	HZ3	5.865
1522	132	TRP	CE3	118.552
1523	132	TRP	HE3	7.068
1524	132	TRP	C	180.07
1525	133	LYS	N	117.6
1526	133	LYS	HN	8.795
1527	133	LYS	CA	60.11
1528	133	LYS	HA	3.62
1529	133	LYS	CB	32.23
1530	133	LYS	HB2	1.753
1531	133	LYS	HB1	2.148
1532	133	LYS	CG	23.96
1533	133	LYS	HG2	1.376
1534	133	LYS	HG1	1.404
1535	133	LYS	CD	29.15
1536	133	LYS	HD2	1.757
1537	133	LYS	HD1	1.757
1538	133	LYS	CE	42.16
1539	133	LYS	HE2	3.073
1540	133	LYS	HE1	3.073
1541	133	LYS	C	177.03
1542	134	ASP	N	112.36
1543	134	ASP	HN	7.796
1544	134	ASP	CA	56.02
1545	134	ASP	HA	4.647
1546	134	ASP	CB	41.45
1547	134	ASP	HB2	2.477
1548	134	ASP	HB1	2.588

463	57	LYS	HN	8.248	1549	134	ASP	C	177.49
464	57	LYS	CA	56.26	1550	135	ASP	N	115.58
465	57	LYS	HA	4.355	1551	135	ASP	HN	7.567
466	57	LYS	CB	33.11	1552	135	ASP	CA	56.29
467	57	LYS	HB2	1.747	1553	135	ASP	HA	4.661
468	57	LYS	HB1	1.834	1554	135	ASP	CB	42.89
469	57	LYS	CG	24.72	1555	135	ASP	HB2	1.797
470	57	LYS	HG2	1.43	1556	135	ASP	HB1	2.734
471	57	LYS	HG1	1.43	1557	135	ASP	C	176.125
472	57	LYS	CD	29.05	1558	136	TYR	N	113.95
473	57	LYS	HD2	1.676	1559	136	TYR	HN	7.374
474	57	LYS	HD1	1.676	1560	136	TYR	CA	60.68
475	57	LYS	CE	42.1	1561	136	TYR	HA	4.888
476	57	LYS	HE2	2.987	1562	136	TYR	CB	41.12
477	57	LYS	HE1	2.987	1563	136	TYR	HB2	2.994
478	57	LYS	C	176.36	1564	136	TYR	HB1	3.951
479	58	LYS	N	123.71	1565	136	TYR	CD1	133.438
480	58	LYS	HN	8.591	1566	136	TYR	HD1	7.076
481	58	LYS	CA	56.26	1567	136	TYR	CE1	119.094
482	58	LYS	HA	4.456	1568	136	TYR	HE1	6.304
483	58	LYS	CB	33.13	1569	136	TYR	CE2	119.094
484	58	LYS	HB2	1.758	1570	136	TYR	HE2	6.304
485	58	LYS	HB1	1.84	1571	136	TYR	CD2	133.438
486	58	LYS	CG	24.72	1572	136	TYR	HD2	7.076
487	58	LYS	HG2	1.426	1573	136	TYR	C	176.86
488	58	LYS	HG1	1.426	1574	137	LEU	N	118.41
489	58	LYS	CD	29.06	1575	137	LEU	HN	8.332
490	58	LYS	HD2	1.681	1576	137	LEU	CA	53.54
491	58	LYS	HD1	1.681	1577	137	LEU	HA	4.67
492	58	LYS	CE	42.16	1578	137	LEU	CB	41
493	58	LYS	HE2	2.998	1579	137	LEU	HB2	1.659
494	58	LYS	HE1	2.998	1580	137	LEU	HB1	1.721
495	58	LYS	C	176.34	1581	137	LEU	CG	26.57
496	59	THR	N	115.12	1582	137	LEU	HG	1.553
497	59	THR	HN	8.203	1583	137	LEU	CD1	25.51
498	59	THR	CA	60.73	1584	137	LEU	HD11	0.191
499	59	THR	HA	4.888	1585	137	LEU	HD12	0.191
500	59	THR	CB	71.41	1586	137	LEU	HD13	0.191
501	59	THR	HB	4.076	1587	137	LEU	CD2	21.79
502	59	THR	CG2	21.63	1588	137	LEU	HD21	0.707
503	59	THR	HG21	1.118	1589	137	LEU	HD22	0.707
504	59	THR	HG22	1.118	1590	137	LEU	HD23	0.707
505	59	THR	HG23	1.118	1591	137	LEU	C	175.7
506	59	THR	C	174.1	1592	138	ASN	N	116.01
507	60	THR	N	118.43	1593	138	ASN	HN	7.253
508	60	THR	HN	8.915	1594	138	ASN	CA	52.48
509	60	THR	CA	60.96	1595	138	ASN	HA	4.667
510	60	THR	HA	4.574	1596	138	ASN	CB	38.14
511	60	THR	CB	70.674	1597	138	ASN	HB2	2.806
512	60	THR	HB	4.057	1598	138	ASN	HB1	3.248
513	60	THR	CG2	21.223	1599	138	ASN	CG	176.105
514	60	THR	HG21	1.159	1600	138	ASN	ND2	113.052
515	60	THR	HG22	1.159	1601	138	ASN	HD21	7.492
516	60	THR	HG23	1.159	1602	138	ASN	HD22	6.611
517	60	THR	C	172.55	1603	138	ASN	C	174.9
518	61	LYS	N	124.03	1604	139	GLU	N	119.31
519	61	LYS	HN	8.246	1605	139	GLU	HN	8.876
520	61	LYS	CA	55	1606	139	GLU	CA	60.11
521	61	LYS	HA	5.242	1607	139	GLU	HA	3.975
522	61	LYS	CB	34.6	1608	139	GLU	CB	29.66
523	61	LYS	HB2	1.637	1609	139	GLU	HB2	2.083
524	61	LYS	HB1	1.674	1610	139	GLU	HB1	2.083
525	61	LYS	CG	25.24	1611	139	GLU	CG	36.41
526	61	LYS	HG2	1.193	1612	139	GLU	HG2	2.376
527	61	LYS	HG1	1.376	1613	139	GLU	HG1	2.376
528	61	LYS	CD	29.42	1614	139	GLU	C	178.16
529	61	LYS	HD2	1.409	1615	140	GLN	N	115.23
530	61	LYS	HD1	1.409	1616	140	GLN	HN	8.223
531	61	LYS	CE	41.77	1617	140	GLN	CA	58.33
532	61	LYS	HE2	2.663	1618	140	GLN	HA	4.259
533	61	LYS	HE1	2.663	1619	140	GLN	CB	28.9
534	61	LYS	C	176.03	1620	140	GLN	HB2	2.088
535	62	ILE	N	123.81	1621	140	GLN	HB1	2.088
536	62	ILE	HN	9.148	1622	140	GLN	CG	34.46
537	62	ILE	CA	57.44	1623	140	GLN	HG2	2.451
538	62	ILE	HA	4.855	1624	140	GLN	HG1	2.451
539	62	ILE	CB	40.29	1625	140	GLN	CD	180.69
540	62	ILE	HB	1.872	1626	140	GLN	NE2	112.482
541	62	ILE	CG1	26.36	1627	140	GLN	HE21	7.601
542	62	ILE	HG12	1.177	1628	140	GLN	HE22	6.892
543	62	ILE	HG11	1.514	1629	140	GLN	C	177.3
544	62	ILE	CD1	13.5	1630	141	ASN	N	115.5
545	62	ILE	HD11	0.795	1631	141	ASN	HN	7.739
546	62	ILE	HD12	0.795	1632	141	ASN	CA	53.96
547	62	ILE	HD13	0.795	1633	141	ASN	HA	5.379
548	62	ILE	CG2	17.47	1634	141	ASN	CB	41.5
549	62	ILE	HG21	0.861	1635	141	ASN	HB2	2.611
550	62	ILE	HG22	0.861	1636	141	ASN	HB1	3.048
551	62	ILE	HG23	0.861	1637	141	ASN	CG	176.725
552	63	PRO	CA	61.9	1638	141	ASN	ND2	113.675
553	63	PRO	HA	4.461	1639	141	ASN	HD21	7.611
554	63	PRO	CB	31.92	1640	141	ASN	HD22	6.915
555	63	PRO	HB2	1.737	1641	141	ASN	C	177.44
556	63	PRO	HB1	1.883	1642	142	ILE	N	122.325
557	63	PRO	CG	26.98	1643	142	ILE	HN	8.369
558	63	PRO	HG2	2.091	1644	142	ILE	CA	66.07
559	63	PRO	HG1	2.333	1645	142	ILE	HA	4.403
560	63	PRO	CD	50.84	1646	142	ILE	CB	37.19
561	63	PRO	HD2	3.892	1647	142	ILE	HB	1.93
562	63	PRO	HD1	4.093	1648	142	ILE	CG1	29.75
563	63	PRO	C	175.86	1649	142	ILE	HG12	1.177
564	64	ILE	N	123.24	1650	142	ILE	HG11	1.501
565	64	ILE	HN	8.786	1651	142	ILE	CD1	14.27
566	64	ILE	CA	58.06	1652	142	ILE	HD11	0.918
567	64	ILE	HA	4.428	1653	142	ILE	HD12	0.918
568	64	ILE	CB	38.02	1654	142	ILE	HD13	0.918
569	64	ILE	HB	1.733	1655	142	ILE	CG2	16.08
570	64	ILE	CG1	26.63	1656	142	ILE	HG21	0.976
571	64	ILE	HG12	1.219	1657	142	ILE	HG22	0.976
572	64	ILE	HG11	1.299	1658	142	ILE	HG23	0.976
573	64	ILE	CD1	12.12	1659	142	ILE	C	175.34
574	64	ILE	HD11	0.727	1660	143	ILE	N	120.25
575	64	ILE	HD12	0.727	1661	143	ILE	HN	7.925
576	64	ILE	HD13	0.727	1662	143	ILE	CA	64.27
577	64	ILE	CG2	16.86	1663	143	ILE	HA	3.751
578	64	ILE	HG21	0.563	1664	143	ILE	CB	37
579	64	ILE	HG22	0.563	1665	143	ILE	HB	1.557
580	64	ILE	HG23	0.563	1666	143	ILE	CG1	28.68

581	64	ILE	C	174.265	1667	143	ILE	HG12	1.075
582	65	GLU	N	126.49	1668	143	ILE	HG11	1.297
583	65	GLU	HN	8.807	1669	143	ILE	CD1	12.71
584	65	GLU	CA	54.65	1670	143	ILE	HD11	0.742
585	65	GLU	HA	4.986	1671	143	ILE	HD12	0.742
586	65	GLU	CB	32.2	1672	143	ILE	HD13	0.742
587	65	GLU	HB2	1.819	1673	143	ILE	CG2	17.31
588	65	GLU	HB1	1.819	1674	143	ILE	HG21	0.529
589	65	GLU	CG	36.9	1675	143	ILE	HG22	0.529
590	65	GLU	HG2	1.835	1676	143	ILE	HG23	0.529
591	65	GLU	HG1	2.024	1677	143	ILE	C	177.86
592	65	GLU	C	175.365	1678	144	TYR	N	119.72
593	66	ILE	N	128.83	1679	144	TYR	HN	7.009
594	66	ILE	HN	9.28	1680	144	TYR	CA	60.9
595	66	ILE	CA	59.49	1681	144	TYR	HA	4.28
596	66	ILE	HA	4.071	1682	144	TYR	CB	37.49
597	66	ILE	CB	39.16	1683	144	TYR	HB2	3.34
598	66	ILE	HB	1.567	1684	144	TYR	HB1	3.303
599	66	ILE	CG1	26.94	1685	144	TYR	CD1	132.79
600	66	ILE	HG12	0.454	1686	144	TYR	HD1	7.215
601	66	ILE	HG11	1.035	1687	144	TYR	CE1	118.543
602	66	ILE	CD1	12.91	1688	144	TYR	HE1	6.83
603	66	ILE	HD11	0.46	1689	144	TYR	CE2	118.543
604	66	ILE	HD12	0.46	1690	144	TYR	HE2	6.83
605	66	ILE	HD13	0.46	1691	144	TYR	CD2	132.79
606	66	ILE	CG2	17.93	1692	144	TYR	HD2	7.215
607	66	ILE	HG21	0.357	1693	144	TYR	C	178.75
608	66	ILE	HG22	0.357	1694	145	THR	N	119.16
609	66	ILE	HG23	0.357	1695	145	THR	HN	8.633
610	66	ILE	C	175.05	1696	145	THR	CA	67.35
611	67	ILE	N	128.29	1697	145	THR	HA	4.242
612	67	ILE	HN	9.215	1698	145	THR	CB	69.42
613	67	ILE	CA	60.35	1699	145	THR	HB	4.728
614	67	ILE	HA	4.684	1700	145	THR	CG2	22.22
615	67	ILE	CB	39.78	1701	145	THR	HG21	1.402
616	67	ILE	HB	1.874	1702	145	THR	HG22	1.402
617	67	ILE	CG1	27.8	1703	145	THR	HG23	1.402
618	67	ILE	HG12	1.097	1704	145	THR	C	174.96
619	67	ILE	HG11	1.527	1705	146	ILE	N	117.41
620	67	ILE	CD1	13.4	1706	146	ILE	HN	8.388
621	67	ILE	HD11	0.888	1707	146	ILE	CA	64.03
622	67	ILE	HD12	0.888	1708	146	ILE	HA	3.612
623	67	ILE	HD13	0.888	1709	146	ILE	CB	36.81
624	67	ILE	CG2	17.84	1710	146	ILE	HB	1.933
625	67	ILE	HG21	0.966	1711	146	ILE	CG1	28.06
626	67	ILE	HG22	0.966	1712	146	ILE	HG12	1.425
627	67	ILE	HG23	0.966	1713	146	ILE	HG11	1.49
628	67	ILE	C	175	1714	146	ILE	CD1	11.67
629	68	SER	N	121.57	1715	146	ILE	HD11	0.677
630	68	SER	HN	8.581	1716	146	ILE	HD12	0.677
631	68	SER	CA	55.91	1717	146	ILE	HD13	0.677
632	68	SER	HA	4.349	1718	146	ILE	CG2	17.58
633	68	SER	CB	65.96	1719	146	ILE	HG21	0.891
634	68	SER	HB2	3.832	1720	146	ILE	HG22	0.891
635	68	SER	HB1	4.079	1721	146	ILE	HG23	0.891
636	68	SER	C	173.5	1722	146	ILE	C	178.3
637	69	LYS	N	119.22	1723	147	ASN	N	117.4
638	69	LYS	HN	8.007	1724	147	ASN	HN	8.133
639	69	LYS	CA	59.16	1725	147	ASN	CA	57.13
640	69	LYS	HA	3.91	1726	147	ASN	HA	4.445
641	69	LYS	CB	32.06	1727	147	ASN	CB	38.57
642	69	LYS	HB2	1.673	1728	147	ASN	HB2	2.663
643	69	LYS	HB1	1.778	1729	147	ASN	HB1	2.865
644	69	LYS	CG	24.29	1730	147	ASN	CG	176.024
645	69	LYS	HG2	1.336	1731	147	ASN	ND2	111.811
646	69	LYS	HG1	1.336	1732	147	ASN	HD21	7.393
647	69	LYS	CD	29.08	1733	147	ASN	HD22	6.983
648	69	LYS	HD2	1.68	1734	147	ASN	C	177.13
649	69	LYS	HD1	1.68	1735	148	ASN	N	114.94
650	69	LYS	CE	42.15	1736	148	ASN	HN	8.336
651	69	LYS	HE2	3.02	1737	148	ASN	CA	55.75
652	69	LYS	HE1	3.02	1738	148	ASN	HA	4.619
653	69	LYS	C	176.71	1739	148	ASN	CB	38
654	70	ASP	N	115.01	1740	148	ASN	HB2	2.161
655	70	ASP	HN	7.983	1741	148	ASN	HB1	2.881
656	70	ASP	CA	54.68	1742	148	ASN	CG	175.17
657	70	ASP	HA	4.746	1743	148	ASN	ND2	110.543
658	70	ASP	CB	42.93	1744	148	ASN	HD21	7.288
659	70	ASP	HB2	2.521	1745	148	ASN	HD22	5.482
660	70	ASP	HB1	2.741	1746	148	ASN	C	177.24
661	70	ASP	C	175.38	1747	149	ILE	N	118.34
662	71	GLU	N	117.99	1748	149	ILE	HN	7.569
663	71	GLU	HN	7.59	1749	149	ILE	CA	65.98
664	71	GLU	CA	55.77	1750	149	ILE	HA	3.844
665	71	GLU	HA	4.621	1751	149	ILE	CB	38.83
666	71	GLU	CB	33	1752	149	ILE	HB	1.873
667	71	GLU	HB2	1.847	1753	149	ILE	CG1	29.64
668	71	GLU	HB1	1.949	1754	149	ILE	HG12	0.763
669	71	GLU	CG	35.58	1755	149	ILE	HG11	2.322
670	71	GLU	HG2	2.089	1756	149	ILE	CD1	14.21
671	71	GLU	HG1	2.089	1757	149	ILE	HD11	0.654
672	71	GLU	C	173.74	1758	149	ILE	HD12	0.654
673	72	LYS	N	122.49	1759	149	ILE	HD13	0.654
674	72	LYS	HN	8.357	1760	149	ILE	CG2	18.15
675	72	LYS	CA	55.68	1761	149	ILE	HG21	0.699
676	72	LYS	HA	4.988	1762	149	ILE	HG22	0.699
677	72	LYS	CB	34.61	1763	149	ILE	HG23	0.699
678	72	LYS	HB2	1.632	1764	149	ILE	C	176.58
679	72	LYS	HB1	1.674	1765	150	ILE	N	116
680	72	LYS	CG	24.97	1766	150	ILE	HN	8.594
681	72	LYS	HG2	1.113	1767	150	ILE	CA	66.72
682	72	LYS	HG1	1.228	1768	150	ILE	HA	3.441
683	72	LYS	CD	29.443	1769	150	ILE	CB	37.82
684	72	LYS	HD2	1.577	1770	150	ILE	HB	1.872
685	72	LYS	HD1	1.577	1771	150	ILE	CG1	29.06
686	72	LYS	CE	41.94	1772	150	ILE	HG12	1.063
687	72	LYS	HE2	2.806	1773	150	ILE	HG11	1.763
688	72	LYS	HE1	2.806	1774	150	ILE	CD1	13.45
689	72	LYS	C	175.04	1775	150	ILE	HD11	0.805
690	73	ILE	N	127.06	1776	150	ILE	HD12	0.805
691	73	ILE	HN	8.912	1777	150	ILE	HD13	0.805
692	73	ILE	CA	59.96	1778	150	ILE	CG2	17.51
693	73	ILE	HA	4.397	1779	150	ILE	HG21	0.872
694	73	ILE	CB	41.67	1780	150	ILE	HG22	0.872
695	73	ILE	HB	1.687	1781	150	ILE	HG23	0.872
696	73	ILE	CG1	27.47	1782	150	ILE	C	178.1
697	73	ILE	HG12	1.075	1783	151	LYS	N	115.38
698	73	ILE	HG11	1.396	1784	151	LYS	HN	8.457

699	73	ILE	CD1	13.67	1785	151	LYS	CA	59.1
700	73	ILE	HD11	0.812	1786	151	LYS	HA	3.991
701	73	ILE	HD12	0.812	1787	151	LYS	CB	32.05
702	73	ILE	HD13	0.812	1788	151	LYS	HB2	1.793
703	73	ILE	CG2	17.54	1789	151	LYS	HB1	2.012
704	73	ILE	HG21	0.812	1790	151	LYS	CG	25.3
705	73	ILE	HG22	0.812	1791	151	LYS	HG2	1.459
706	73	ILE	HG23	0.812	1792	151	LYS	HG1	1.459
707	73	ILE	C	174.17	1793	151	LYS	CD	28.68
708	74	VAL	N	125.13	1794	151	LYS	HD2	1.637
709	74	VAL	HN	8.305	1795	151	LYS	HD1	1.637
710	74	VAL	CA	60.95	1796	151	LYS	CE	42.1
711	74	VAL	HA	4.899	1797	151	LYS	HE2	2.922
712	74	VAL	CB	33.95	1798	151	LYS	HE1	2.922
713	74	VAL	HB	1.804	1799	151	LYS	C	178.15
714	74	VAL	CG2	20.71	1800	152	ASN	N	110.4
715	74	VAL	HG21	0.536	1801	152	ASN	HN	6.696
716	74	VAL	HG22	0.536	1802	152	ASN	CA	55.1
717	74	VAL	HG23	0.536	1803	152	ASN	HA	4.805
718	74	VAL	CG1	21.27	1804	152	ASN	CB	40.74
719	74	VAL	HG11	0.793	1805	152	ASN	HB2	2.453
720	74	VAL	HG12	0.793	1806	152	ASN	HB1	2.608
721	74	VAL	HG13	0.793	1807	152	ASN	C	174.95
722	74	VAL	C	175.45	1808	153	VAL	N	116.13
723	75	LYS	N	126.75	1809	153	VAL	HN	7.677
724	75	LYS	HN	8.744	1810	153	VAL	CA	64.36
725	75	LYS	CA	53.71	1811	153	VAL	HA	4.056
726	75	LYS	HA	4.547	1812	153	VAL	CB	32.16
727	75	LYS	CB	35.37	1813	153	VAL	HB	2.138
728	75	LYS	HB2	1.473	1814	153	VAL	CG2	22.09
729	75	LYS	HB1	1.75	1815	153	VAL	HG21	0.865
730	75	LYS	CG	24.63	1816	153	VAL	HG22	0.865
731	75	LYS	HG2	1.22	1817	153	VAL	HG23	0.865
732	75	LYS	HG1	1.22	1818	153	VAL	CG1	22.39
733	75	LYS	CD	29.11	1819	153	VAL	HG11	0.947
734	75	LYS	HD2	1.524	1820	153	VAL	HG12	0.947
735	75	LYS	HD1	1.575	1821	153	VAL	HG13	0.947
736	75	LYS	CE	41.94	1822	154	ILE	N	115.42
737	75	LYS	HE2	2.838	1823	154	ILE	HN	7.579
738	75	LYS	HE1	2.838	1824	154	ILE	CA	62.79
739	75	LYS	C	173.7	1825	154	ILE	HA	3.879
740	76	TYR	N	118.08	1826	154	ILE	CB	37.43
741	76	TYR	HN	8.52	1827	154	ILE	HB	1.875
742	76	TYR	CA	57.48	1828	154	ILE	CG1	27.47
743	76	TYR	HA	5.186	1829	154	ILE	HG12	1.16
744	76	TYR	CB	40.72	1830	154	ILE	HG11	1.467
745	76	TYR	HB2	2.589	1831	154	ILE	CD1	13.14
746	76	TYR	HB1	2.75	1832	154	ILE	HD11	0.698
747	76	TYR	CD1	132.737	1833	154	ILE	HD12	0.698
748	76	TYR	HD1	6.856	1834	154	ILE	HD13	0.698
749	76	TYR	CE1	118.076	1835	154	ILE	CG2	17.56
750	76	TYR	HE1	6.732	1836	154	ILE	HG21	0.859
751	76	TYR	CE2	118.076	1837	154	ILE	HG22	0.859
752	76	TYR	HE2	6.732	1838	154	ILE	HG23	0.859
753	76	TYR	CD2	132.737	1839	154	ILE	C	174.88
754	76	TYR	HD2	6.856	1840	155	GLN	CG	32.461
755	76	TYR	C	175.65	1841	155	GLN	HG2	2.13
756	77	LEU	N	125	1842	155	GLN	HG1	2.13
757	77	LEU	HN	9.144	1843	155	GLN	CD	179.688
758	77	LEU	CA	53.25	1844	155	GLN	NE2	111.265
759	77	LEU	HA	4.687	1845	155	GLN	HE21	7.707
760	77	LEU	CB	45.07	1846	155	GLN	HE22	6.951
761	77	LEU	HB2	1.519	1847	157	GLU	CA	56.7
762	77	LEU	HB1	1.604	1848	157	GLU	HA	4.371
763	77	LEU	CG	26.997	1849	157	GLU	CB	30.53
764	77	LEU	HG	1.552	1850	157	GLU	HB2	1.951
765	77	LEU	CD1	24.999	1851	157	GLU	HB1	2.065
766	77	LEU	HD11	0.872	1852	157	GLU	CG	36.22
767	77	LEU	HD12	0.872	1853	157	GLU	HG2	2.26
768	77	LEU	HD13	0.872	1854	157	GLU	HG1	2.26
769	77	LEU	CD2	24.665	1855	157	GLU	C	175.836
770	77	LEU	HD21	0.842	1856	158	ASP	N	121.3
771	77	LEU	HD22	0.842	1857	158	ASP	HN	8.404
772	77	LEU	HD23	0.842	1858	158	ASP	CA	54.42
773	77	LEU	C	175.61	1859	158	ASP	HA	4.642
774	78	GLN	N	123.14	1860	158	ASP	CB	41.53
775	78	GLN	HN	8.556	1861	158	ASP	HB2	2.606
776	78	GLN	CA	55.32	1862	158	ASP	HB1	2.712
777	78	GLN	HA	5.047	1863	158	ASP	C	176.18
778	78	GLN	CB	30.27	1864	159	ASN	N	118.9
779	78	GLN	HB2	1.947	1865	159	ASN	HN	8.558
780	78	GLN	HB1	1.947	1866	159	ASN	CA	54.33
781	78	GLN	CG	34.3	1867	159	ASN	HA	4.636
782	78	GLN	HG2	2.213	1868	159	ASN	CB	38.57
783	78	GLN	HG1	2.287	1869	159	ASN	HB2	2.891
784	78	GLN	CD	180.135	1870	159	ASN	HB1	2.916
785	78	GLN	NE2	111.526	1871	159	ASN	CG	177.409
786	78	GLN	HE21	7.555	1872	159	ASN	ND2	112.664
787	78	GLN	HE22	6.783	1873	159	ASN	HD21	7.593
788	78	GLN	C	174.85	1874	159	ASN	HD22	6.901
789	79	ILE	N	121.98	1875	159	ASN	C	175.44
790	79	ILE	HN	8.652	1876	160	SER	N	114.93
791	79	ILE	CA	59.27	1877	160	SER	HN	8.527
792	79	ILE	HA	4.552	1878	160	SER	CA	59.31
793	79	ILE	CB	41.71	1879	160	SER	HA	4.519
794	79	ILE	HB	1.763	1880	160	SER	CB	64.12
795	79	ILE	CG1	27.29	1881	160	SER	HB2	3.864
796	79	ILE	HG12	1.209	1882	160	SER	HB1	3.912
797	79	ILE	HG11	1.475	1883	160	SER	C	173.77
798	79	ILE	CD1	13.25	1884	161	ILE	N	121.86
799	79	ILE	HD11	0.838	1885	161	ILE	HN	7.92
800	79	ILE	HD12	0.838	1886	161	ILE	CA	61.05
801	79	ILE	HD13	0.838	1887	161	ILE	HA	4.235
802	79	ILE	CG2	17.11	1888	161	ILE	CB	39.43
803	79	ILE	HG21	0.819	1889	161	ILE	HB	1.784
804	79	ILE	HG22	0.819	1890	161	ILE	CG1	27.26
805	79	ILE	HG23	0.819	1891	161	ILE	HG12	1.186
806	79	ILE	C	174.125	1892	161	ILE	HG11	1.452
807	80	ASP	N	124.56	1893	161	ILE	CD1	13.54
808	80	ASP	HN	8.507	1894	161	ILE	HD11	0.792
809	80	ASP	CA	54.47	1895	161	ILE	HD12	0.792
810	80	ASP	HA	4.668	1896	161	ILE	HD13	0.792
811	80	ASP	CB	41.32	1897	161	ILE	CG2	17.56
812	80	ASP	HB2	2.598	1898	161	ILE	HG21	0.816
813	80	ASP	HB1	2.804	1899	161	ILE	HG22	0.816
814	80	ASP	C	177.75	1900	161	ILE	HG23	0.816
815	81	GLU	N	126.34	1901	161	ILE	C	174.74
816	81	GLU	HN	8.8	1902	162	TRP	N	125.31

817	81	GLU	CA	59.08	1903	162	TRP	HN	8.132
818	81	GLU	HA	4.015	1904	162	TRP	CA	56.71
819	81	GLU	CB	30.3	1905	162	TRP	HA	4.79
820	81	GLU	HB2	1.942	1906	162	TRP	CB	29.75
821	81	GLU	HB1	2.064	1907	162	TRP	HB2	3.187
822	81	GLU	CG	36.24	1908	162	TRP	HB1	3.251
823	81	GLU	HG2	2.254	1909	162	TRP	CD1	126.199
824	81	GLU	HG1	2.254	1910	162	TRP	HD1	7.181
825	81	GLU	C	177.57	1911	162	TRP	NE1	129.68
826	82	GLU	N	117.29	1912	162	TRP	HE1	10.319
827	82	GLU	HN	8.764	1913	162	TRP	CZ2	114.345
828	82	GLU	CA	56.17	1914	162	TRP	HZ2	7.41
829	82	GLU	HA	4.328	1915	162	TRP	CH2	124.27
830	82	GLU	CB	28.98	1916	162	TRP	HH2	7.079
831	82	GLU	HB2	1.973	1917	162	TRP	CZ3	121.479
832	82	GLU	HB1	2.202	1918	162	TRP	HZ3	7.024
833	82	GLU	CG	36.84	1919	162	TRP	CE3	121.564
834	82	GLU	HG2	2.264	1920	162	TRP	HE3	7.621
835	82	GLU	HG1	2.264	1921	162	TRP	C	174.66
836	82	GLU	C	176.4	1922	163	ILE	N	121.59
837	83	SER	N	116.07	1923	163	ILE	HN	7.739
838	83	SER	HN	7.549	1924	163	ILE	CA	59.92
839	83	SER	CA	59.3	1925	163	ILE	HA	3.967
840	83	SER	HA	4.449	1926	163	ILE	CB	39.48
841	83	SER	CB	64.79	1927	163	ILE	HB	1.709
842	83	SER	HB2	3.961	1928	163	ILE	CG1	27.08
843	83	SER	HB1	3.961	1929	163	ILE	HG12	1.027
844	83	SER	C	173.91	1930	163	ILE	HG11	1.286
845	84	SER	N	119.866	1931	163	ILE	CD1	12.95
846	84	SER	HN	9.145	1932	163	ILE	HD11	0.703
847	84	SER	CA	57.64	1933	163	ILE	HD12	0.703
848	84	SER	HA	4.448	1934	163	ILE	HD13	0.703
849	84	SER	CB	65.19	1935	163	ILE	CG2	18.84
850	84	SER	HB2	4.104	1936	163	ILE	HG21	0.708
851	84	SER	HB1	4.338	1937	163	ILE	HG22	0.708
852	84	SER	C	174.86	1938	163	ILE	HG23	0.708
853	85	LEU	N	123.32	1939	164	GLU	CA	57.19
854	85	LEU	HN	8.968	1940	164	GLU	HA	4.334
855	85	LEU	CA	58.79	1941	164	GLU	CB	31.91
856	85	LEU	HA	3.983	1942	164	GLU	HB2	1.953
857	85	LEU	CB	41.1	1943	164	GLU	HB1	2.148
858	85	LEU	HB2	1.661	1944	164	GLU	CG	36.4
859	85	LEU	HB1	1.746	1945	164	GLU	HG2	2.258
860	85	LEU	CG	27.14	1946	164	GLU	HG1	2.258
861	85	LEU	HG	1.428	1947	164	GLU	C	175.82
862	85	LEU	CD1	25.83	1948	165	GLU	N	117.655
863	85	LEU	HD11	0.804	1949	165	GLU	HN	7.563
864	85	LEU	HD12	0.804	1950	165	GLU	CA	55.16
865	85	LEU	HD13	0.804	1951	165	GLU	HA	4.699
866	85	LEU	CD2	23.6	1952	165	GLU	CB	33.27
867	85	LEU	HD21	0.81	1953	165	GLU	HB2	1.941
868	85	LEU	HD22	0.81	1954	165	GLU	HB1	1.953
869	85	LEU	HD23	0.81	1955	165	GLU	CG	35.69
870	85	LEU	C	178.63	1956	165	GLU	HG2	2.209
871	86	LYS	N	116.98	1957	165	GLU	HG1	2.239
872	86	LYS	HN	8.572	1958	165	GLU	C	173.53
873	86	LYS	CA	60.68	1959	166	VAL	N	122.7
874	86	LYS	HA	3.73	1960	166	VAL	HN	8.427
875	86	LYS	CB	31.93	1961	166	VAL	CA	60.14
876	86	LYS	HB2	1.766	1962	166	VAL	HA	4.95
877	86	LYS	HB1	1.833	1963	166	VAL	CB	34.61
878	86	LYS	CG	25.22	1964	166	VAL	HB	1.835
879	86	LYS	HG2	1.424	1965	166	VAL	CG2	21.75
880	86	LYS	HG1	1.424	1966	166	VAL	HG21	0.747
881	86	LYS	CD	29.09	1967	166	VAL	HG22	0.747
882	86	LYS	HD2	1.678	1968	166	VAL	HG23	0.747
883	86	LYS	HD1	1.678	1969	166	VAL	CG1	22.122
884	86	LYS	CE	42.13	1970	166	VAL	HG11	0.782
885	86	LYS	HE2	3.003	1971	166	VAL	HG12	0.782
886	86	LYS	HE1	3.003	1972	166	VAL	HG13	0.782
887	86	LYS	C	177.77	1973	166	VAL	C	174.81
888	87	ASP	N	119.62	1974	167	GLU	N	126.2
889	87	ASP	HN	7.673	1975	167	GLU	HN	8.79
890	87	ASP	CA	57.46	1976	167	GLU	CA	54.44
891	87	ASP	HA	4.438	1977	167	GLU	HA	4.466
892	87	ASP	CB	40.31	1978	167	GLU	CB	34.34
893	87	ASP	HB2	2.676	1979	167	GLU	HB2	1.763
894	87	ASP	HB1	2.929	1980	167	GLU	HB1	2.024
895	87	ASP	C	179.38	1981	167	GLU	CG	36.27
896	88	LYS	N	121.42	1982	167	GLU	HG2	2.001
897	88	LYS	HN	8.263	1983	167	GLU	HG1	2.201
898	88	LYS	CA	60.71	1984	167	GLU	C	174.13
899	88	LYS	HA	3.922	1985	168	ILE	N	123.43
900	88	LYS	CB	33.18	1986	168	ILE	HN	8.749
901	88	LYS	HB2	1.726	1987	168	ILE	CA	60.02
902	88	LYS	HB1	1.987	1988	168	ILE	HA	4.847
903	88	LYS	CG	26.47	1989	168	ILE	CB	39.16
904	88	LYS	HG2	1.194	1990	168	ILE	HB	1.567
905	88	LYS	HG1	1.688	1991	168	ILE	CG1	28.14
906	88	LYS	CD	30.21	1992	168	ILE	HG12	0.786
907	88	LYS	HD2	1.432	1993	168	ILE	HG11	1.497
908	88	LYS	HD1	1.564	1994	168	ILE	CD1	14.44
909	88	LYS	CE	41.99	1995	168	ILE	HD11	0.784
910	88	LYS	HE2	2.753	1996	168	ILE	HD12	0.784
911	88	LYS	HE1	2.753	1997	168	ILE	HD13	0.784
912	88	LYS	C	178.28	1998	168	ILE	CG2	18.1
913	89	LEU	N	116.5	1999	168	ILE	HG21	0.616
914	89	LEU	HN	8.588	2000	168	ILE	HG22	0.616
915	89	LEU	CA	57.86	2001	168	ILE	HG23	0.616
916	89	LEU	HA	4.033	2002	168	ILE	C	173.27
917	89	LEU	CB	41.9	2003	169	TYR	N	124.78
918	89	LEU	HB2	1.179	2004	169	TYR	HN	9.309
919	89	LEU	HB1	1.943	2005	169	TYR	CA	56.52
920	89	LEU	CG	27	2006	169	TYR	HA	5.041
921	89	LEU	HG	1.932	2007	169	TYR	CB	42.7
922	89	LEU	CD1	27.25	2008	169	TYR	HB2	2.417
923	89	LEU	HD11	0.777	2009	169	TYR	HB1	2.946
924	89	LEU	HD12	0.777	2010	169	TYR	CD1	132.45
925	89	LEU	HD13	0.777	2011	169	TYR	HD1	6.678
926	89	LEU	CD2	23.82	2012	169	TYR	CE1	118.672
927	89	LEU	HD21	0.972	2013	169	TYR	HE1	6.698
928	89	LEU	HD22	0.972	2014	169	TYR	CE2	118.672
929	89	LEU	HD23	0.972	2015	169	TYR	HE2	6.698
930	89	LEU	C	178.63	2016	169	TYR	CD2	132.45
931	90	ARG	N	119.47	2017	169	TYR	HD2	6.678
932	90	ARG	HN	8.316	2018	169	TYR	C	174.62
933	90	ARG	CA	59.06	2019	170	TYR	N	118.5
934	90	ARG	HA	3.962	2020	170	TYR	HN	8.896

935	90	ARG	CB	31.95	2021	170	TYR	CA	56.97
936	90	ARG	HB2	1.976	2022	170	TYR	HA	4.83
937	90	ARG	HB1	2.187	2023	170	TYR	CB	41.18
938	90	ARG	CG	26.11	2024	170	TYR	HB2	2.746
939	90	ARG	HG2	1.685	2025	170	TYR	HB1	2.698
940	90	ARG	HG1	1.92	2026	170	TYR	CD1	133.03
941	90	ARG	CD	44.38	2027	170	TYR	HD1	7.04
942	90	ARG	HD2	3.184	2028	170	TYR	CE1	118.65
943	90	ARG	HD1	3.463	2029	170	TYR	HE1	6.814
944	90	ARG	NE	87.47	2030	170	TYR	CE2	118.65
945	90	ARG	HE	7.171	2031	170	TYR	HE2	6.814
946	90	ARG	CZ	160.02	2032	170	TYR	CD2	133.03
947	90	ARG	C	177.89	2033	170	TYR	HD2	7.04
948	91	LEU	N	119.67	2034	170	TYR	C	175.74
949	91	LEU	HN	7.561	2035	171	ASN	N	129.75
950	91	LEU	CA	58.32	2036	171	ASN	HN	9.369
951	91	LEU	HA	4.316	2037	171	ASN	CA	53.33
952	91	LEU	CB	42.017	2038	171	ASN	HA	4.084
953	91	LEU	HB2	1.84	2039	171	ASN	CB	37.21
954	91	LEU	HB1	1.84	2040	171	ASN	HB2	1.948
955	91	LEU	CG	27.2	2041	171	ASN	HB1	2.932
956	91	LEU	HG	1.703	2042	171	ASN	CG	177.451
957	91	LEU	CD1	25.31	2043	171	ASN	ND2	110.363
958	91	LEU	HD11	0.975	2044	171	ASN	HD21	7.101
959	91	LEU	HD12	0.975	2045	171	ASN	HD22	6.523
960	91	LEU	HD13	0.975	2046	171	ASN	C	175.54
961	91	LEU	CD2	24.15	2047	172	GLY	N	102.71
962	91	LEU	HD21	1.033	2048	172	GLY	HN	8.861
963	91	LEU	HD22	1.033	2049	172	GLY	CA	45.39
964	91	LEU	HD23	1.033	2050	172	GLY	HA2	3.502
965	91	LEU	C	180.74	2051	172	GLY	HA1	4.099
966	92	ILE	N	123.1	2052	172	GLY	C	173.69
967	92	ILE	HN	8.311	2053	173	LYS	N	121.75
968	92	ILE	CA	66.58	2054	173	LYS	HN	7.868
969	92	ILE	HA	3.4	2055	173	LYS	CA	54.64
970	92	ILE	CB	38.86	2056	173	LYS	HA	4.766
971	92	ILE	HB	1.967	2057	173	LYS	CB	34.81
972	92	ILE	CG1	29.76	2058	173	LYS	HB2	1.849
973	92	ILE	HG12	0.829	2059	173	LYS	HB1	1.956
974	92	ILE	HG11	1.875	2060	173	LYS	CG	24.773
975	92	ILE	CD1	14.21	2061	173	LYS	HG2	1.472
976	92	ILE	HD11	0.972	2062	173	LYS	HG1	1.472
977	92	ILE	HD12	0.972	2063	173	LYS	CD	29.1
978	92	ILE	HD13	0.972	2064	173	LYS	HD2	1.758
979	92	ILE	CG2	16.45	2065	173	LYS	HD1	1.806
980	92	ILE	HG21	0.761	2066	173	LYS	CE	42.16
981	92	ILE	HG22	0.761	2067	173	LYS	HE2	3.075
982	92	ILE	HG23	0.761	2068	173	LYS	HE1	3.075
983	92	ILE	C	177.79	2069	173	LYS	C	175.12
984	93	LEU	N	119.45	2070	174	LEU	N	127.55
985	93	LEU	HN	8.803	2071	174	LEU	HN	8.513
986	93	LEU	CA	58.51	2072	174	LEU	CA	55.49
987	93	LEU	HA	4.049	2073	174	LEU	HA	3.607
988	93	LEU	CB	40.6	2074	174	LEU	CB	42.95
989	93	LEU	HB2	1.597	2075	174	LEU	HB2	1.186
990	93	LEU	HB1	1.958	2076	174	LEU	HB1	1.501
991	93	LEU	CG	26.91	2077	174	LEU	CG	26.14
992	93	LEU	HG	1.929	2078	174	LEU	HG	1.142
993	93	LEU	CD1	25.86	2079	174	LEU	CD1	25.54
994	93	LEU	HD11	0.716	2080	174	LEU	HD11	0.62
995	93	LEU	HD12	0.716	2081	174	LEU	HD12	0.62
996	93	LEU	HD13	0.716	2082	174	LEU	HD13	0.62
997	93	LEU	CD2	22.1	2083	174	LEU	CD2	23.56
998	93	LEU	HD21	0.864	2084	174	LEU	HD21	0.436
999	93	LEU	HD22	0.864	2085	174	LEU	HD22	0.436
1000	93	LEU	HD23	0.864	2086	174	LEU	HD23	0.436
1001	93	LEU	C	179.25	2087	174	LEU	C	177.27
1002	94	ASP	N	120.97	2088	175	ILE	N	128
1003	94	ASP	HN	8.922	2089	175	ILE	HN	8.912
1004	94	ASP	CA	57.8	2090	175	ILE	CA	60.64
1005	94	ASP	HA	4.452	2091	175	ILE	HA	4.106
1006	94	ASP	CB	40.36	2092	175	ILE	CB	37.37
1007	94	ASP	HB2	2.708	2093	175	ILE	HB	1.698
1008	94	ASP	HB1	2.965	2094	175	ILE	CG1	26.86
1009	94	ASP	C	179.14	2095	175	ILE	HG12	1.203
1010	95	THR	N	117.575	2096	175	ILE	HG11	1.347
1011	95	THR	HN	7.892	2097	175	ILE	CD1	12.83
1012	95	THR	CA	67.54	2098	175	ILE	HD11	0.703
1013	95	THR	HA	4.092	2099	175	ILE	HD12	0.703
1014	95	THR	CB	68.1	2100	175	ILE	HD13	0.703
1015	95	THR	HB	4.47	2101	175	ILE	CG2	17.36
1016	95	THR	CG2	20.9	2102	175	ILE	HG21	0.762
1017	95	THR	HG21	1.161	2103	175	ILE	HG22	0.762
1018	95	THR	HG22	1.161	2104	175	ILE	HG23	0.762
1019	95	THR	HG23	1.161	2105	175	ILE	C	174.93
1020	95	THR	C	176.43	2106	176	GLU	N	125.89
1021	96	LEU	N	121.33	2107	176	GLU	HN	8.22
1022	96	LEU	HN	8.515	2108	176	GLU	CA	55.55
1023	96	LEU	CA	58.66	2109	176	GLU	HA	4.402
1024	96	LEU	HA	3.921	2110	176	GLU	CB	31.94
1025	96	LEU	CB	42.36	2111	176	GLU	HB2	1.855
1026	96	LEU	HB2	1.643	2112	176	GLU	HB1	1.95
1027	96	LEU	HB1	2.194	2113	176	GLU	CG	36.65
1028	96	LEU	CG	26.87	2114	176	GLU	HG2	2.099
1029	96	LEU	HG	1.934	2115	176	GLU	HG1	2.099
1030	96	LEU	CD1	26.54	2116	176	GLU	C	175.31
1031	96	LEU	HD11	0.976	2117	177	LEU	N	124.22
1032	96	LEU	HD12	0.976	2118	177	LEU	HN	8.347
1033	96	LEU	HD13	0.976	2119	177	LEU	CA	54.98
1034	96	LEU	CD2	23.46	2120	177	LEU	HA	4.384
1035	96	LEU	HD21	0.722	2121	177	LEU	CB	42.3
1036	96	LEU	HD22	0.722	2122	177	LEU	HB2	1.575
1037	96	LEU	HD23	0.722	2123	177	LEU	HB1	1.575
1038	96	LEU	C	179.59	2124	177	LEU	CG	27.09
1039	97	SER	N	115.82	2125	177	LEU	HG	1.587
1040	97	SER	HN	9.29	2126	177	LEU	CD1	25.37
1041	97	SER	CA	63.47	2127	177	LEU	HD11	0.877
1042	97	SER	HA	4.339	2128	177	LEU	HD12	0.877
1043	97	SER	CB	63.09	2129	177	LEU	HD13	0.877
1044	97	SER	HB2	4.038	2130	177	LEU	CD2	24
1045	97	SER	HB1	4.209	2131	177	LEU	HD21	0.801
1046	97	SER	C	176.38	2132	177	LEU	HD22	0.801
1047	98	ASN	N	115.21	2133	177	LEU	HD23	0.801
1048	98	ASN	HN	8.011	2134	177	LEU	C	175.89
1049	98	ASN	CA	56.35	2135	178	ARG	N	126.75
1050	98	ASN	HA	4.413	2136	178	ARG	HN	7.841
1051	98	ASN	CB	39.64	2137	178	ARG	CA	57.1
1052	98	ASN	HB2	2.788	2138	178	ARG	HA	4.171

1053	98	ASN	HB1	2.926
1054	98	ASN	CG	176.128
1055	98	ASN	ND2	113.939
1056	98	ASN	HD21	7.603
1057	98	ASN	HD22	6.973
1058	98	ASN	C	177.72
1059	99	GLU	N	115.43
1060	99	GLU	HN	8.633
1061	99	GLU	CA	57.57
1062	99	GLU	HA	4.279
1063	99	GLU	CB	30.37
1064	99	GLU	HB2	1.931
1065	99	GLU	HB1	1.689
1066	99	GLU	CG	36.26
1067	99	GLU	HG2	2.116
1068	99	GLU	HG1	2.339
1069	99	GLU	C	176.72
1070	100	TYR	N	117.5
1071	100	TYR	HN	8.07
1072	100	TYR	CA	59.59
1073	100	TYR	HA	4.629
1074	100	TYR	CB	41.93
1075	100	TYR	HB2	2.279
1076	100	TYR	HB1	2.461
1077	100	TYR	CD1	133.144
1078	100	TYR	HD1	5.978
1079	100	TYR	CE1	117.5
1080	100	TYR	HE1	6.553
1081	100	TYR	CE2	117.5
1082	100	TYR	HE2	6.553
1083	100	TYR	CD2	133.144
1084	100	TYR	HD2	5.978
1085	100	TYR	C	174.74

2139	178	ARG	CB	32.05
2140	178	ARG	HB2	1.673
2141	178	ARG	HB1	1.854
2142	178	ARG	CG	27.07
2143	178	ARG	HG2	1.544
2144	178	ARG	HG1	1.544
2145	178	ARG	CD	43.46
2146	178	ARG	HD2	3.138
2147	178	ARG	HD1	3.201
2148	178	ARG	NE	85.16
2149	178	ARG	HE	7.234
2150	178	ARG	CZ	159.61
2151	178	ARG	C	180.61

Acknowledgements

I want to thank everyone who has contributed to this work:

Firstly, I would like to thank my supervisor Dieter Willbold for his consistent support and guidance during the running of both projects and for the opportunity to work in the excellent, fully equipped laboratory and NMR facilities.

I would also like to thank my mentor and co-supervisor, Birgit Strodel, for her support and scientific discussions about current results during my PhD.

Many thanks to Philipp Neudecker, who helps me a lot with NMR experiments and some data evaluation. Also, thank you for this opportunity to work here and for helping me with a German bureaucracy.

My sincere thanks go to all NMR people at the institute IBI-7. In particular, I would like to thank Nils Lakomek for his support in my last year of the PhD, constructive discussions during thesis writing and great ideas about the data evaluation. Also, many thanks to Henrike Heise and all her groups' members, Baran, Nina, Luis, and Anna, for your assistance, advice and friendly atmosphere. Thank you so much to Rudi and Kevin, who helped with my questions during the experiment settings, and, also, without you, I would never move the Dewar with liquid nitrogen even for a centimetre. And one more special thanks to my dear friend Baran for her support, help, care, and tasty food she brought to me when I worked over weekends.

Thank you to all people who work at the biological laboratory of IBI-7 and Düsseldorf University lab. And especially to people who introduced me to biology lab work and taught me how to express and purify the best protein samples. In particular, thank you to Vineet, Marianne, Serife, Kai and Thibault. Also, for people who work at the "Russian lab" at IBI-7, it was always nice to know that I can count on your help if I need something urgent during the lab work.

Many thanks to all my colleges at the IBI-7 for exciting discussions and an enjoyable work environment. Special thanks to my best office mates Line, Tim and Marc for a very friendly office atmosphere and many funny moments. Thank you for listening, support and help in challenging times.

I would further like to thank Xue Wang, Verena Kienapfel and Bettina Wagner for the excellent collaboration and many scientific discussions for both projects.

I thank my colleagues from Düsseldorf University, Sander Smits, Lothar Gremer and Manuel Etzkorn, for instructive collaboration and helpful advice.

My sincere thanks to my colleagues and friends from other research institutes at FZJ Thorsten Auth, Margarita Kruteva and my friend Natalia for the continuous support with thesis writing, discussions, motivation and for cheering me up in tough times.

Acknowledgements

I would like to acknowledge the financial support from the Collaborative Research Center 1208 (SFB 1208) from the Heinrich-Heine-University Düsseldorf.

Last but not least, I thank my family, my mum, my sister, my lovely grandparents V and V, and all of my friends from Russia for their constant support during my education and moving to another country. Spasibo that you never doubt in me.

Publications and Presentations

Publications

1. Irina Apanasenko, Rebeca Clemens, Jens Reiners, Lothar Gremer, Sander Smits, and Philipp Neudecker: Structure Calculation of Lipoprotein CD 1348 from *Clostridium difficile* by Solution NMR. In progress

2. Daniel Mulnaes, Nicola Porta, Rebecca Clemens, Irina Apanasenko, Jens Reiners, Lothar Gremer, Philipp Neudecker, Sander H. J. Smits, and Holger Gohlke: TopModel: Template-Based Protein Structure Prediction at Low Sequence Identity Using Top-Down Consensus and Deep Neural Networks, Journal of Chemical Theory and Computation 2020 16 (3), 1953-1967, DOI: 10.1021/acs.jctc.9b00825

3. B. Uluca-Yazgi, L. Siemons, M. Sevenich, K. Schmitz, I. Apanasenko, N. Becker, A. S. König, W. Hoyer, P. Neudecker, D. F. Hansen, H. Heise: Isoleucine Side Chains as a Reporter of Conformational Freedom. In progress

Oral presentations

1. Irina Apanasenko, Jülich. "Elucidation of the dynamics of the autophagosomal membrane-associated protein GABARAP by NMR spectroscopy". Annual meeting 2019 of the Bio-NMR-Network-NRW (bio-N^3MR), Jülich, Germany

Eidesstattliche Erklärung

Ich versichere an Eides statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der „Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf“ erstellt worden ist.

_____ _____

Ort, Datum Unterschrift

Eidesstattliche Erklärung