

LU-TP 22-34
May 2022

Assignment of heteronuclear methyl-NMR spectrum of the 44 kDa protein MALT1 Casp-IgL₃

Maria Levkovets

Department of Astronomy and Theoretical Physics, Lund University

Master thesis supervised by:
Vladislav Orekhov - Gothenburg University
Carl Troein - Lund University



LUND
UNIVERSITY

Abstract

Mucosa-associated lymphoid tissue lymphoma translocation protein 1 (MALT1) plays an important role in the immune pathway that controls the activation and proliferation of B and T cells. Dysregularization of this pathway leads to the development of highly aggressive lymphomas. The proteolytic activity of MALT1 can be pharmacologically regulated making this protein an attractive anti-cancer drug target.

In this thesis, I study two domains of MALT1 (Casp and IgL₃) responsible for its proteolytic activity with nuclear magnetic resonance spectroscopy (NMR). I perform the ¹H, ¹³C Ile, Val, Leu-methyl assignment of monomeric MALT1 Casp-IgL₃ domains using methyl out-and-back experiments based on the existing backbone assignment and the information from the protein structure. The protein structures that were used for this thesis are the x-ray dimer structure of MALT1, and the structure predicted by the AlphaFold protein prediction software. I have compared AlphaFold structures with the x-ray structure. The methyl assignment that I have performed provides an opportunity for further studies of monomeric MALT1 structure and dynamics which is the key to anti-cancer drug design.

Populärvetenskaplig beskrivning

Ett protein som kallas MALT1 reglerar en kedja av kemiska reaktioner som kan orsaka aggressiva typer av cancer såsom B-cellslymfom. Att hitta ett läkemedel som kan blockera MALT1 såväl som kedjan av kemiska reaktioner är ett möjligt sätt att bota denna typ av cancer. För att designa det läkemedlet måste själva strukturen av MALT1 avslöjas. I denna masteruppsats studerar jag en del av MALT1 med en teknik som kallas Nuclear Magnetic Resonance (NMR) spektroskopi.

B-celler, även kallade lymfocyter, är vår kropps försvarare mot olika patogener: bakterier och virus. De blir aktiva och neutraliserar patogener med hjälp av en uppsättning mycket komplicerade inre kemiska reaktioner. Hos patienter med B-cellslymfom är lymfocyter aktiva och delar sig utan närvaro av något "hot" vilket orsakar cancer.

Ett protein med det långa namnet mucosa-associated lymphoid tissue lymphoma translocation protein, som förkortas till MALT1, spelar en viktig roll i den kemiska reaktionskaskaden som aktiverar lymfocyterna. Denna kaskad är en uppsättning kemiska reaktioner som följer varandra och MALT1 deltar i en av dem. Om vi hindrar MALT1 från att reagera med andra molekyler genom att fästa ett läkemedel till det, kan vi blockera hela kaskaden och stoppa cancer. Man kan jämföra det med en dominoeffekt. Att eliminera en domino i raden kommer att stoppa alla andra från att falla.

För att kunna delta i kaskaden måste två MALT1-molekyler fästa vid varandra och bilda ett par. De flesta läkemedel designades för att fästa vid MALT1 paret för att blockera den från att reagera med andra molekyler i kaskaden. Flera läkemedel har upptäckts, men alla har antingen för många biverkningar eller fungerar inte alls. Ett möjlig alternativ lösning skulle vara att designa ett läkemedel som binder till en enskild molekyl av MALT1 och hindra den från att bilda ett par. För att designa ett sådant läkemedel måste strukturen och framför allt dynamiken, alltså hur MALT1 interagerar med andra molekyler studeras. I dagsläget känner vi endast till strukturen av parvisa MALT1. Därför måste studier av en enskild molekyl av MALT1 utföras.

I denna masteruppsats använder jag NMR för att studera MALT1. Denna teknik möjliggör studier av en enskild molekyl. Eftersom MALT1 är ett relativt stort protein är det utmanande att studera det med vanliga NMR-metoder. Av den anledningen använder vi metylgrupper för att studera MALT1. Metylgrupper i proteinet består av en kolatom omgiven av tre väteatomer. Det första steget till eventuella studier av MALT1 med metylgrupper kallas tillordning. Att utföra en tillordning med enkla ord innebär att matcha varje NMR-signal till en motsvarande metyl i proteinstrukturen. Huvudmålet med mitt examensarbete är att utföra metyltillordning av MALT1. Resultatet kommer att möjliggöra framtida studier av MALT1 och bidra till upptäckt av en läkemedel mot MALT1-beroende cancer.

Contents

1	Introduction	6
2	Theory background	8
2.1	NMR	8
2.1.1	The basics of NMR	9
2.1.2	NMR experiments	11
2.1.3	Protein NMR	13
2.1.4	Multidimensional NMR	14
2.1.5	Methyl-based protein NMR	16
3	Materials and methods	18
4	Results and discussion	20
4.1	Assignment of $^1\text{H}/^{13}\text{C}$ HMQC methyl spectra	20
4.2	AlphaFold	27
4.3	Discussion	29
4.4	Conclusion	31

List of acronyms

BCL10 - B-cell lymphoma 10
BMRB - Biological Magnetic Resonance Bank
CARMA1 - Caspase-recruitment domain (CARD) membrane-associated guanylate kinase (MAGUK) protein 1
CASP - Caspase
CBM complex - CARMA, BCL10, MALT1
DD - Death Domain
DLBCL - Diffuse Large B-cell Lymphoma
DSS - Sodium trimethylsilylpropanesulfonate
FID - Free Induction Decay
HMQC - Heteronuclear Multiple Quantum Coherence
ILV - Ile, Leu, Val
IgL - Immunoglobulin-Like
Ile - Isoleucine

Leu - Leucine
 MALT1 - Mucosa-associated lymphoid tissue lymphoma translocation protein 1
 MSA - Multiple sequence alignment NF- κ B - Nuclear Factor kappa B
 NMR - Nuclear Magnetic Resonance
 NOESY - Nuclear Overhauser Enhancement Spectroscopy
 PDB - Protein Data Bank
 RF - Radiofrequency
 TMS - Tetramethylsilane
 Val - Valine
 ppm - parts per million

List of Figures

1	Structure of MALT1	7
2	2D $^1\text{H}/^{13}\text{C}$ HMQC methyl spectra	8
3	Precession of a spinning top and a nucleus	10
4	A simple pulse-acquire NMR experiment	11
5	The Fourier transform	12
6	Polypeptide chain	13
7	Comparison of 1D and 2D NMR spectra	14
8	A general scheme of a pulse sequence for a 2D experiment	14
9	The Fourier transform for 2D experiments	15
10	The magnetization transfer pathways in some of the multidimensional NMR experiments	16
11	Ile, Val, and Leu amino acids	17
12	Out-and-back magnetization transfer experiment	17
13	Labeling scheme of Ile, Val, and Leu amino acids used in this thesis	19
14	3D HNCA spectra	20
15	3D HMCM(CB)CA spectra	21
16	3D HMCM(CGCB)CA spectra	22
17	3D HNHM NOESY and 3D HMCM(CB)CA spectra	23
18	4D HMCM NOESY spectra	25
19	Methyl groups of ILV residues on MALT1 protein structure	26

20	Hydrophobic clusters produced by methyl groups of Ile, Val, Leu	27
21	MALT1 Casp-IgL ₃ structure predicted by AlphaFold	28
22	The output of AlphaFold	29
23	¹ H- ¹³ C HMQC spectrum of MALT1 Casp-IgL ₃ with assigned resonances . .	34

List of Tables

1	Summary of the performed assignment	30
2	Non-stereospecific methyl assignment of MALT Casp-IgL ₃ domains	35

1 Introduction

In this thesis, I am studying a part of a protein called mucosa-associated lymphoid tissue lymphoma translocation protein 1 (MALT1). This protein regulates the Nuclear factor-kappa B (NF- κ B) signaling pathway¹ which is responsible for B and T-lymphocytes activation and proliferation. Overexpression of MALT1 leads to dysregulation of the NF- κ B pathway causing multiple diseases including MALT1 lymphoma, diffuse large B-cell lymphoma (DLBCL), and several other immune-related diseases.[1]

MALT1 regulates the NF- κ B pathway by two routes. First, MALT1 acts as a scaffolding protein² by generating an oligomeric CBM complex that consists of CARMA, BCL10, and MALT1 proteins and triggers the downstream activation of the NF- κ B signaling pathway. Second, MALT1 acts as a protease³ on several proteins that are included in the NF- κ B pathway. [1]

The proteolytic activity of MALT1 can be pharmacologically regulated with the help of inhibitors⁴, therefore, making MALT1 a promising anti-cancer drug target. However, the development of MALT1 inhibitors is still in the early stages, and currently available inhibitors have various defects which limit their application in clinical settings. The activation of the proteolytic activity of MALT1 starts with its homodimerization⁵. [1] Thus, it can be beneficial to develop an inhibitor, which does not inhibit the proteolytic activity of MALT1 directly but does not allow MALT1 to form a dimer.

To design such an inhibitor, the monomeric structure of MALT1 is required. Currently, all the existing structures of MALT1 that were revealed with the x-ray crystallography technique, are dimers. Although up-to-date protein structure prediction software AlphaFold [2] can be a source of reliable information about the possible structure of monomeric MALT1, this is, as for all crystal structures, just a snapshot of energy minimized states which may provide adequate information.

The molecular mass of MALT1 is 93kDa. MALT1 consists of five functional units or domains: death domain (DD), three immunoglobulin-like domains (IgL₃, IgL₂, and IgL₃), and a caspase-like domain (Casp) followed by an unstructured C-terminal tail. (Figure 1) Casp and IgL₃ domains maintain protease activity of MALT1. The dimerization arises between the two Casp domains. [1]

In this thesis, I am studying the 44 kDa MALT1 Casp and IgL₃ domains as those are the minimum unit of protease activity of MALT1 with Nuclear Magnetic Resonance (NMR)

¹A signaling pathway is a set of chemical reactions which control a cell's function.

²A scaffolding protein is a type of protein which can bring several other proteins together and form a stable configuration.

³Proteases are enzymes that can break down proteins.

⁴Inhibitors are chemicals that can slow down or stop a chemical reaction.

⁵Homodimerization is a process in which two molecules of the same chemical (in this case a protein) react with each other and form a complex.

spectroscopy. This technique is capable of providing information about the structure of monomeric MALT1 in solution. In addition to that, NMR, in contrast to x-ray, can provide valuable knowledge of the dynamics of MALT1 which is connected to its biological functions.

However, the studies of large systems, such as MALT1, with conventional solution-NMR methods remain a challenge, as there is significant signal broadening. The use of methyl groups (CH_3 groups) in the side-chains of amino acids as spectroscopic probes helps to achieve a more resolved signal and thus enables the studies of MALT1. Methyl groups are located at the hydrophobic core, catalytic active sites, proteinprotein interfaces, and flexible loops and they serve as sensitive probes of the protein structure and dynamics.

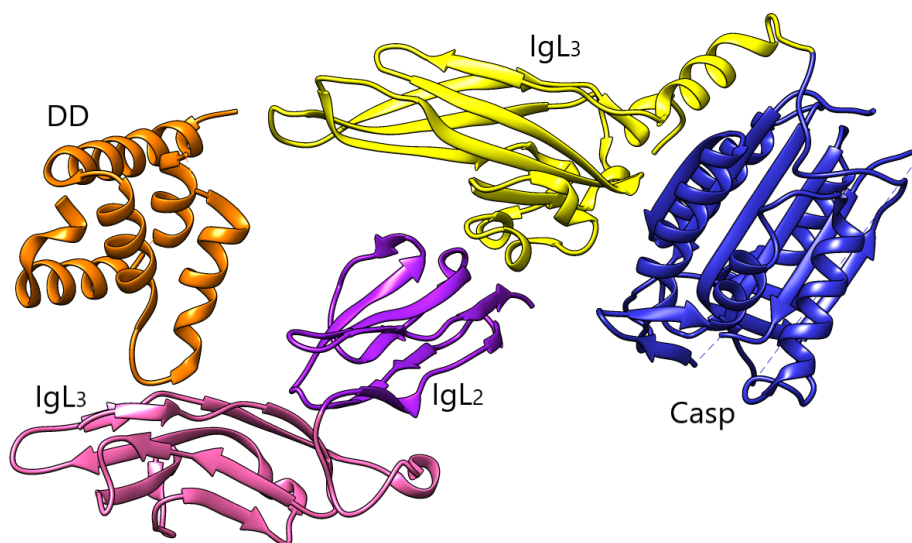


Figure 1: Structure of MALT1. The figures of individual domains were acquired from Protein Data Bank (PDB) [3]. Death domain - DD (PDB ID:2G7R), IgL_1 and IgL_2 domains (PDB: 3K0W), C - Ig_3 and Casp domains (PDB ID:3V55).

The major goal of this master thesis and the first step to the studies of MALT1 is to perform methyl resonances assignment of MALT1 Casp- IgL_3 domains. This means that each signal in the 2D $^1\text{H}/^{13}\text{C}$ HMQC methyl NMR spectra has to be related to a specific methyl group in an amino acid in the protein sequence. For MALT1 we have used a specific ILV labeling scheme meaning that only methyl groups of Ile, Val, and Leu can be observed. In Figure 2 there are the two regions that correspond to Ile and Val/Leu residues. Peaks in this spectrum must be assigned.

In addition to that I have worked with the protein structure prediction software AlphaFold[2] which was beneficial for this project as it provided the structural information about the flexible strand which contained several ILV residues and was missing in x-ray MALT1 Casp- IgL_3 dimer structure.

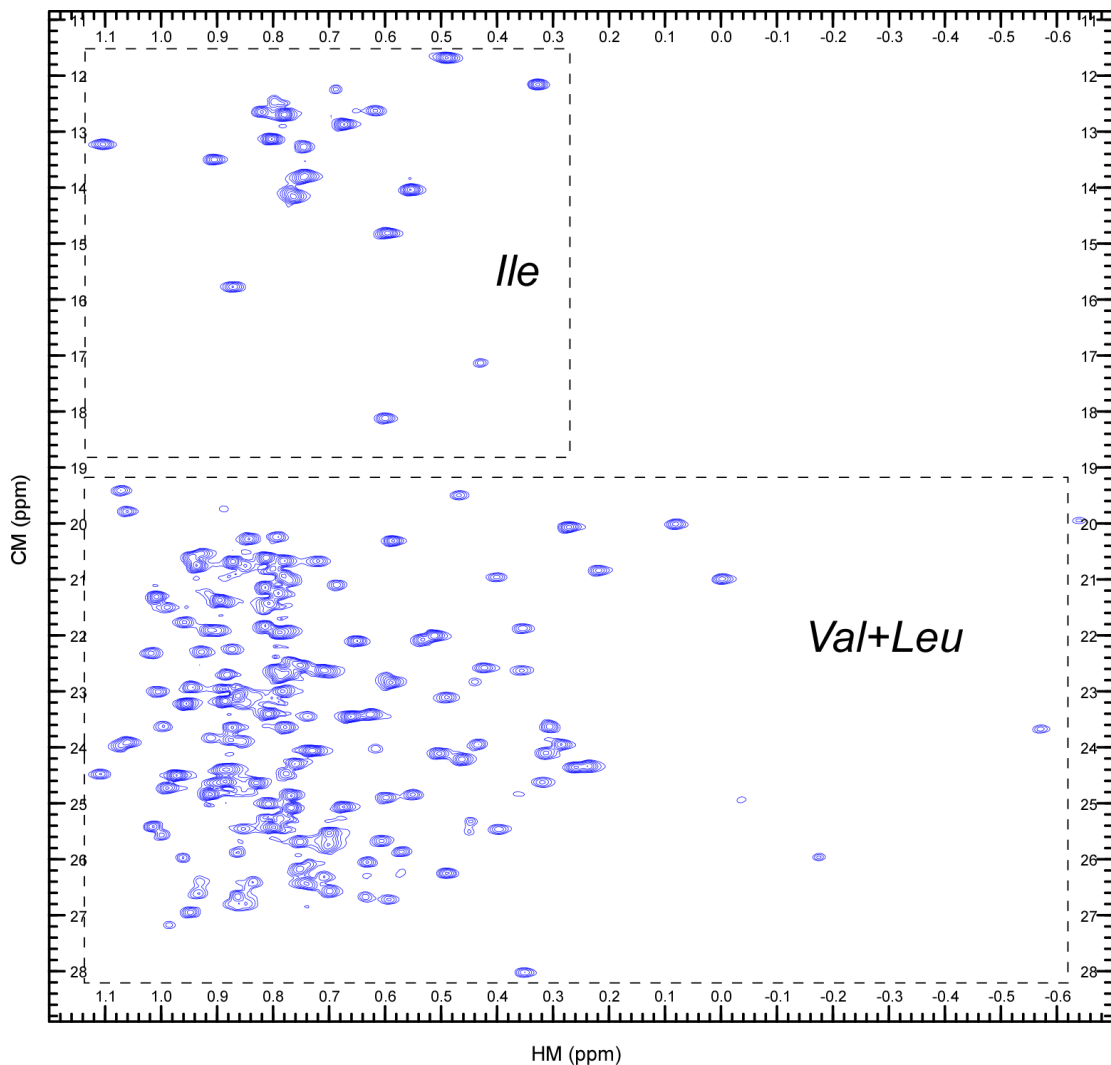


Figure 2: 2D $^1\text{H}/^{13}\text{C}$ HMQC methyl spectra. Axes display the ^1H and ^{13}C chemical shifts of methyl groups. The two regions correspond to the resonances from the methyl groups of Ile and Leu+Val respectively.

2 Theory background

2.1 NMR

NMR spectroscopy is an analytical method used to obtain quantitative information about liquid, solid, and gas samples. In NMR, a nucleus is placed in a very strong magnetic field and irradiated with a radio-frequency field which is in resonance with the frequency of the magnetic field around the nucleus. The nuclei in a sample respond to that by producing an electromagnetic signal which is directly related to the molecular structure

and dynamics. Thus, NMR is capable of providing exhaustive information about the chemistry of molecules. This technique is widely used in organic and inorganic chemistry as well as in medicine. [4]

2.1.1 The basics of NMR

The NMR phenomenon relies on the quantum mechanical property of the intrinsic angular momentum also called spin ($\mathbf{I}$) which can only take integer and half-integer values. When the non-zero spin $\mathbf{I}$ is placed in a strong magnetic field, it can take a limited number of orientations corresponding to $(2I+1)$ energy eigenstates. This is called Zeeman splitting. The transition between those states stimulated by an applied radiofrequency (RF) radiation is the basis of NMR experiments.

In addition to the quantum mechanical approach, the NMR method is frequently described in terms of classical mechanics. In this approach, the motion of spins in the magnetic field is described as similar to the motion of a spinning top in the gravitational field. The nucleus with a non-zero spin has a nuclear magnetic moment $\boldsymbol{\mu} = \gamma \mathbf{I} \hbar$ and the nuclear angular momentum $\mathbf{J}$. γ is the gyromagnetic ratio. In the presence of an applied magnetic field $\mathbf{B}_0$ the spins will experience a torque which according to Newtonian mechanics is equal to:

$$\frac{d\mathbf{J}}{dt} = \boldsymbol{\mu} \times \mathbf{B}_0. \quad (2.1)$$

As $\mathbf{J} = \mathbf{I} \hbar$ and $\boldsymbol{\mu} = \gamma \mathbf{I} \hbar = \gamma \mathbf{J}$, the equation becomes:

$$\frac{d\boldsymbol{\mu}}{dt} = \gamma \boldsymbol{\mu} \times \mathbf{B}_0. \quad (2.2)$$

This equation is similar to the equation of rotational motion of the body with mass m and angular momentum $\mathbf{L}$ at a distance $\mathbf{r}$ from the rotational axis in the gravitational field $\mathbf{g}$:

$$\frac{d\mathbf{L}}{dt} = \mathbf{r} \times m\mathbf{g}. \quad (2.3)$$

So as the spinning top will precess around the axis of the gravitational field, the same type of motion is observed when the spins are placed in the magnetic field. The energy difference of the two Zeeman states with $I = 1/2$ is equal to:

$$\Delta E = \gamma \hbar B_0. \quad (2.4)$$

From Planck's law:

$$\Delta E = \hbar \nu. \quad (2.5)$$

Combining equations 2.4 and 2.5 we get the Larmor frequency of rotation for a given nucleus in the static magnetic field B_0 :

$$\nu_0 = \frac{\gamma B_0}{2\pi} \quad (2.6)$$

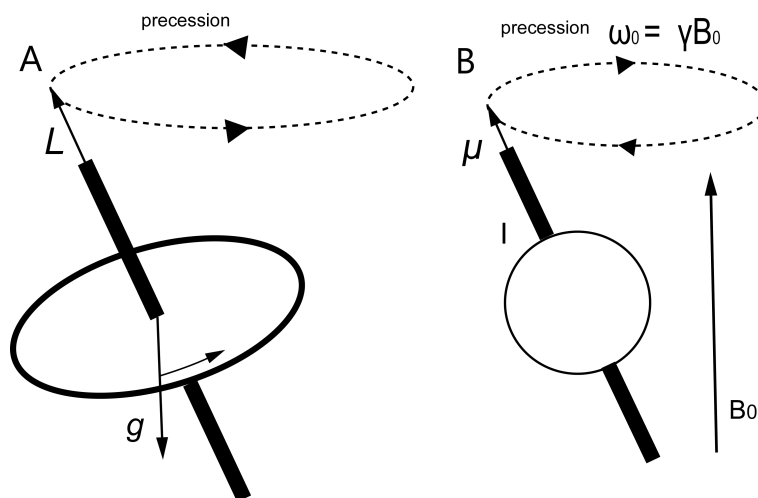


Figure 3: A - precession of a spinning top with angular momentum L in a gravitational field. B - precession of a nucleus with spin I in the magnetic field B_0 .

or

$$\omega_0 = \gamma B_0. \quad (2.7)$$

The frequency of rotation of a nuclear spin is proportional to the strength of the static magnetic field B_0 . However, in a chemical compound, the magnetic field around a single nucleus is slightly changed due to the currents caused by the motions of electrons. Thus the local field for a single nucleus is:

$$B_{loc} = (1 - \sigma)B_0 \quad (2.8)$$

where σ is the shielding factor. For that reason the frequency of each nucleus in the chemical compound will be slightly different from the others:

$$\omega = (1 - \sigma)\omega_0. \quad (2.9)$$

The precession of spins can be detected with a radiofrequency (RF) coil. This will make it possible to obtain spectra in which the individual peak will correspond to an individual group of nuclei. However, in practice, the frequency domain is rarely used in NMR spectroscopy. The frequencies are usually transformed into the chemical shifts measured in parts per million (ppm) by:

$$\delta = \frac{\nu - \nu_{ref}}{\nu_{ref}}. \quad (2.10)$$

Here ν_{ref} is the frequency of some known components such as TMS (tetramethylsilane) used for organic solvents or DSS (sodium trimethylsilylpropanesulfonate) used for protein NMR experiments as DSS is soluble in water. Sources for the paragraph: [5, p.5-8], [6, p.14].

2.1.2 NMR experiments

In the absence of a magnetic field, spins have random orientations, but when the magnetic field B_0 is turned on, the individual magnetic moments of spins tend to align along the direction of the field (Z-axis) forming a so-called bulk magnetization M_0 . When the radiofrequency radiation produced by the RF coil is applied, its magnetic component B_1 which is oscillating with the Larmor frequency can affect the bulk magnetization. This is the phenomenon of resonance that the NMR method was named after. [5, p.9-10]

The basis of all NMR experiments is a series of RF pulses and delays. This is called a pulse sequence. In the simplest NMR experiment called the pulse-acquire experiment, the RF pulse of a certain length is applied to flip the bulk magnetization vector from the Z to XY plane. The magnetization slowly returns to the equilibrium performing the motion of precession simultaneously. This is called relaxation. The precession of bulk magnetization in the XY plane induces a current in the RF coil. That signal is called free induction decay (FID). (See Figure 4.) [7, p.147-148]

In general, the pulse sequence of any one-dimensional (1D) NMR experiment can be divided into two major sections - preparation and detection periods. During the preparation period, a series of RF pulses and delays are applied to achieve the desired state of magnetization before detection of the FID. In the case of the pulse-acquire experiment, the preparation period is a 90° RF pulse after which the FID is detected.

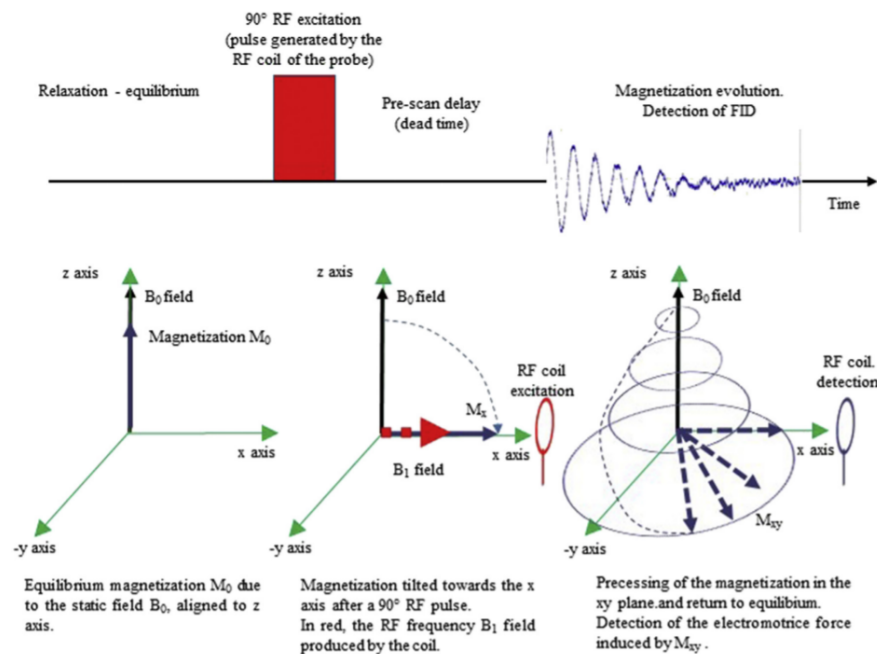


Figure 4: A simple pulse-acquire NMR experiment. Reprint from [8] under creative common licence and with acknowledgment to the authors.

To produce the NMR spectra from the FID, Fourier transformation and data processing must be performed. (See Figure 5) Mathematically, the Fourier transform can be written as:

$$S_{spectrum}(f) = \int_0^{+\infty} S_{FID}(t) \cos(2\pi ft) dt \quad (2.11)$$

where f is the frequency of a trial $\cos$ function, $S_{FID}(t)$ is the amplitude of the FID at time t and $S_{spectrum}(f)$ is the intensity at frequency f . In the spectrometer the smooth function of FID is represented as a set of data points (N points), so mathematically the integral of Fourier transform becomes a sum:

$$S_{spectrum}(f) = \sum_{i=1}^{i=N} S_{FID}(t_i) \cos(2\pi ft_i). \quad (2.12)$$

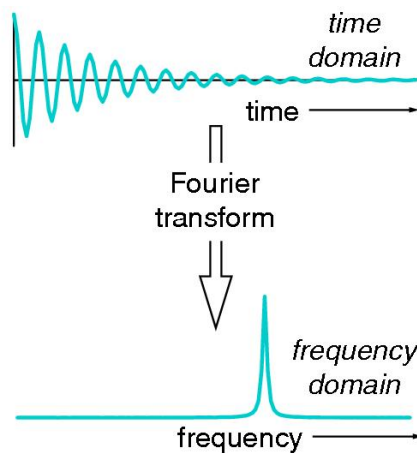


Figure 5: The Fourier transform turns the time-domain signal (FID) into the frequency domain signal. Reprint from [7, Figure 5.1] with permission from the author. This way a simple 1D spectra is acquired. [7, p.81-85]

2.1.3 Protein NMR

Proteins are biological macromolecules that are made of amino acids. NMR is a powerful tool to study protein structure, dynamics and kinetics. The major advantage of this technique is that it gives the opportunity to study proteins in an environment that is close to natural conditions. For the proteins that can not be crystallized or the crystallization process is known to alternate the protein structure, NMR is a method of choice. The drawback of NMR is that it is limited to proteins with a molecular mass of less than 40-60 kDa.

Proteins are made of one or several polypeptide chains that are arranged in space in a certain way. A polypeptide chain consists of amino acids that are connected via a peptide bond. An amino acid unit in a peptide chain is called a residue. A polypeptide chain consists of a main chain called a backbone chain and variable side chains. (See Figure 6.) Amide nitrogen (N), the amide hydrogen (H), the backbone α carbon (C_α also sometimes written as CA), and the carbon connected to oxygen (CO) are called the backbone atoms. The first carbon in the side chain is called C_β . Protein NMR studies usually start with finding chemical shifts of N , H , C_α , and CO . For each residue, those will be slightly different as the chemical environment of an individual residue is unique. This is called the backbone assignment.

To study proteins, the common isotopes 1H , ^{13}C , ^{15}N with spin 1/2 are used. However, the low natural abundance of ^{13}C (1.11%) and of ^{15}N is (0.37%) [9, p.14], would give low signal-to-noise ratio. In order to increase it, the protein is enriched (labelled) with those isotopes.

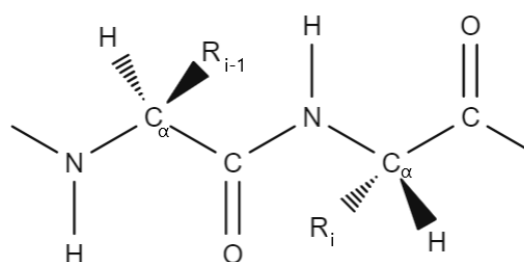


Figure 6: A schematic view of a polypeptide chain with the backbone and variable side chains marked as R_i .

2.1.4 Multidimensional NMR

To study proteins, so-called multidimensional NMR experiments are performed. In contrast to the conventional one-dimensional (1D) experiments, where we plot the intensity against frequency, here we plot the intensity against several frequencies. In Figure 7 you can see a comparison between 1D spectra and 2D spectra. The pulse sequence for a multidimensional experiment is more complicated than for a 1D experiment. A 2D experiment, for example, consists of preparation, evolution, mixing, and detection periods. (See Figure 8.) The preparation period is designed to prepare the desired state of magnetization. That magnetization evolves during the evolution period t_1 . t_1 here is not a fixed time. During the series of experiments, t_1 is gradually increased. The mixing period is designed to transfer the magnetization that exists at the end of t_1 into a detectable signal. The mixing period is frequently used for the transfer of magnetization from one spin to another. The FID is detected during the time t_2 . To conclude, to obtain a 2D spectra we perform a series of 1D experiments with various t_1 and detect FID. The resulting signal can be represented as a 2D time-domain matrix. Two separate Fourier transforms are then used to obtain frequency-domain 2D NMR spectra. (See Figure 9.) A similar approach is used to obtain NMR spectra of higher dimensions. [7, Chapter 8]

During the mixing time, the magnetization may be transferred to another spin through the spin-spin coupling (also called J-coupling or scalar coupling). J-coupling arises between the spins which are connected through one, two, or three chemical bonds. Classical NMR experiments connect the backbone atoms such as N, H, C_α , CO to obtain 2D and 3D spectra. The magnetization transfer pathways are visualized in Figure 10. [10, p.139]

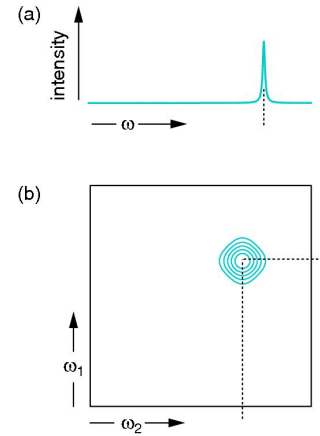


Figure 7: A: Conventional 1D NMR spectra. Peak has only one frequency. B: 2D NMR spectra. The peak has two frequencies. The intensity is shown as round lines. Reprint from [7, Figure 8.1] with the permission from the author.

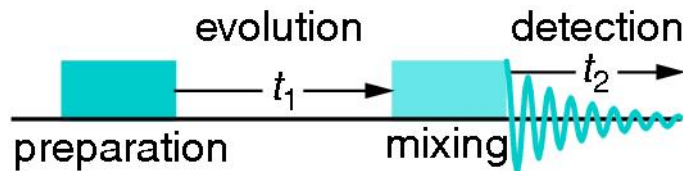


Figure 8: A general scheme of a pulse sequence for a 2D experiment. Reprint from [7, Figure 8.2] with the permission from the author.

The magnetization transfer can also occur via cross-relaxation. Cross-relaxation is the transfer of magnetization through dipolar couplings between nuclei. Such transfer of

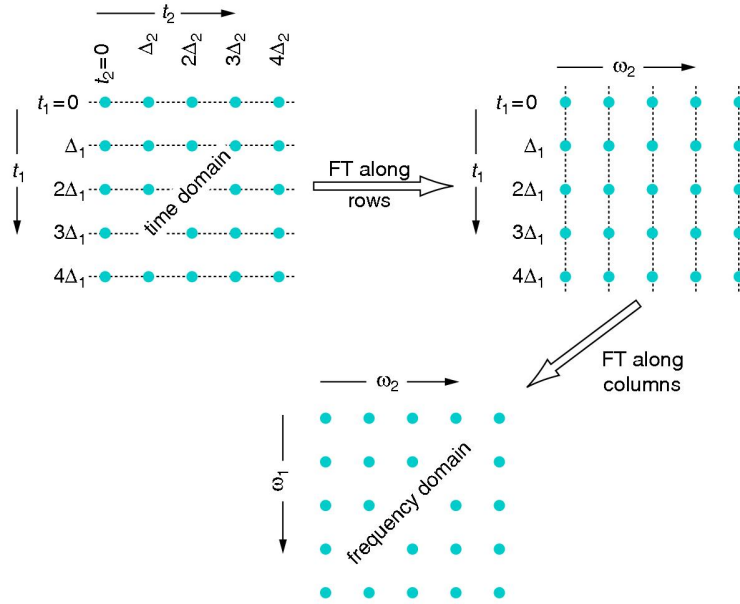


Figure 9: The recorded FID is presented as a 2D time-domain matrix. Two Fourier transforms are performed to obtain the frequency domain matrix and eventually the spectra. Reprint from [7, Figure 8.1] with the permission from the author.

magnetization is the basis for Nuclear Overhauser Effect spectroscopy (NOESY). NOESY is capable of detecting spins that are close in space but not linked through a chemical bond. (See Figure 10 ^1H - ^{15}N NOESY.) [10, p.358-360]

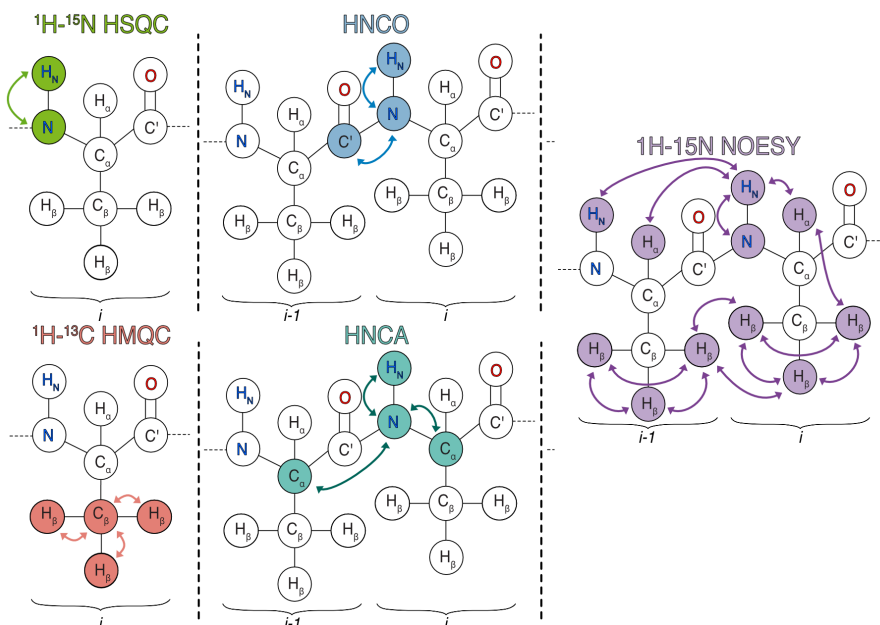


Figure 10: The magnetization transfer pathways in some of the multidimensional NMR experiments. Those or similar experiments were used in this master thesis. Adapted from [11, p.17] with the permission from the author. HSQC - heteronuclear single quantum coherence experiment, HMQC - heteronuclear multiple quantum coherence experiment.

2.1.5 Methyl-based protein NMR

NMR has historically been limited to proteins smaller than ~ 30 kDa⁶. Above that limit, the linewidths of resonances are significantly increased which decreases the spectral resolution. This problem can be solved with the help of methyl groups. They are capable of giving a higher intensity and more resolved NMR signal. Methyl groups, $-\text{CH}_3$, are present in the side chains of 6 out of 20 amino acids. In Figure 11 three amino acids - Isoleucine (Ile), Valine (Val), and Leucine (Leu) which were used in this master thesis are presented. Methyl groups are usually located at the hydrophobic core of proteins, catalytic active sites, and protein-protein interfaces and thus they serve as valuable probes of protein structure and dynamics. [12] With selectively protonated ^{13}C -labeled methyl groups it is possible to acquire high-resolution $^1\text{H}/^{13}\text{C}$ heteronuclear multiple quantum coherence (HMQC) NMR spectra for proteins with mass up to 1 MDa. The major obvious drawback of methyl-based NMR is that the parts of a protein, which do not contain methyl side-chains, cannot be observed. [13]

Methyl resonances can be connected to the backbone resonances through out-and-back correlation experiments developed by Tugarinov et al. [14], which rely on out-and-back magnetization transfer from methyl atoms (HM and CM) to C_α or to CO. The

⁶A unit of molecular mass which is equal to 1/12 of the mass of the isotope of carbon ^{12}C .

names of the experiments, such as HMCM(CGCB)CA for Leu (see Figure 12), represent the magnetization transfer path in the experiment. Spins in parenthesis participate in the magnetization transfer, but their frequency in NMR spectra is not observed.

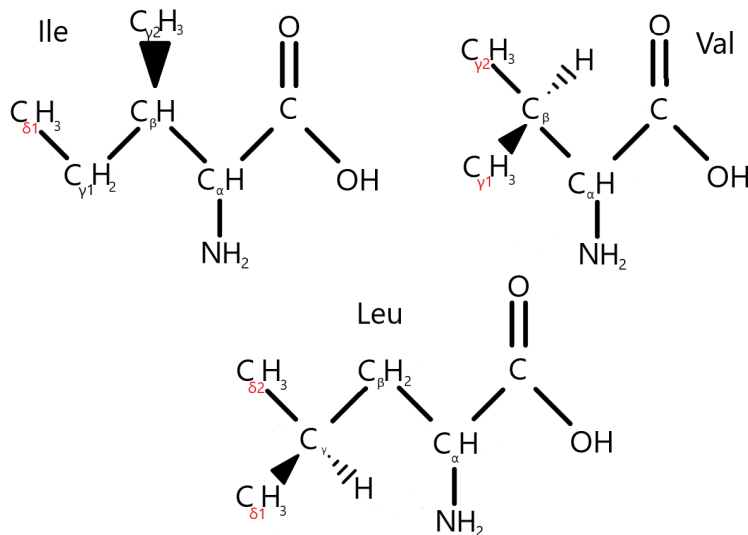


Figure 11: Amino acids that contain methyl groups: Isoleucine (Ile), Valine (Val), Leucine (Leu). The methyl groups that were specifically labeled to be observed in NMR spectra in this master thesis are $\text{C}_{\delta 1}\text{H}_3$ for Ile, $\text{C}_{\gamma 1}\text{H}_3$ and $\text{C}_{\gamma 2}\text{H}_3$ for Val, $\text{C}_{\delta 1}\text{H}_3$ and $\text{C}_{\delta 2}\text{H}_3$ for Leu.

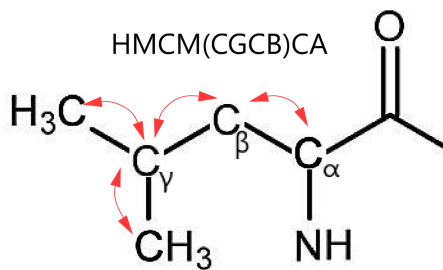


Figure 12: Magnetization transfer for out-and-back HMCM(CGCB)CA experiment for Leu amino acid. Note that C_{β} , C_{γ} , and C_{δ} atoms are usually written as CB, CG and CD in the NMR experiments.

3 Materials and methods

NMR spectra were recorded on an 800 MHz and 900 MHz Bruker ADVANCE III-HD spectrometer. Conventional 3D HNCO and HNCA experiments were recorded. The methyl-based NMR experiments that were performed are $^1\text{H}/^{13}\text{C}$ HMQC, out-and-back HMCM(CGCB)CA for Ile/Leu, and HMCM(CB)CA for Val residues at 298 K, and HMCM(CGBCA)CO for Ile, Val, and Leu residues performed at 308 K. In addition to those experiments the NH-methyl interactions were recorded with 3D $^1\text{H}-^{15}\text{N}$ SF TROSY NOESY experiment. Methyl-methyl group interactions were obtained with 4D ^{13}C , ^{13}C SF HMQC NOESY experiment. Further, we will refer to the spectra produced by 3D $^1\text{H}-^{15}\text{N}$ TROSY NOESY and 4D ^{13}C , ^{13}C HMQC NOESY experiments as 3D HNHM NOESY and 4D HMCM NOESY respectively. The names refer to the axes of those spectra.

The isotopic non-stereospecific labeling of Ile, Leu, and Val for mentioned NMR experiments is presented in Figure 13. The isotopic labeling of methyl groups of Val and Leu is not stereospecific which results in only one of the two methyl groups labelled with ^{13}C and ^1H , while the other methyl group is labelled with ^{12}C and D, hence not detectable in the experiments. The backbone chemical shifts of MALT1 have been previously assigned [15] and for the current thesis acquired from Biological Magnetic Resonance Data Bank (BMRB) [16] with access code 25674. The spectra were displayed and analyzed in CcpNmr Analysis Version 3.4.0. [17]. The x-ray MALT1 three-dimensional structure of Casp and IgL₃ domains were acquired from PDB [3] (PDB DOI: 10.2210/pdb3V55/pdb, PDB ID: 3v55). Original article: [18]. Protein structure prediction software AlphaFold [2] was used to produce the structure of monomeric MALT1 Casp-IgL₃. The input includes protein sequence of MALT1 Casp-IgL₃. All the other parameters were left as default [2].

Python scripts were written to automatize some of the procedures of the assignment. Python script “NH-methyl distances” calculates the distances between the given N atom and all other CM atoms based on the x-ray 3v55 structure of MALT1 and on the acquired AlphaFold structures, and returns the numbers of residues of the corresponding N and CM if they are closer than 5Å. The alteration of that script - “methyl-methyl distances” calculates all the distances between a given CM atom and all other CM atoms, and returns the residue numbers of those that are close in space (6Å). Additional macros code for UCSF Chimera package [19] was written to visualize the results of the methyl assignment on the acquired AlphaFold structure. The alignment of AlphaFold and 3v55 structures was performed with Needleman-Wunsch algorithm in UCSF Chimera and from that the RMSD was calculated.

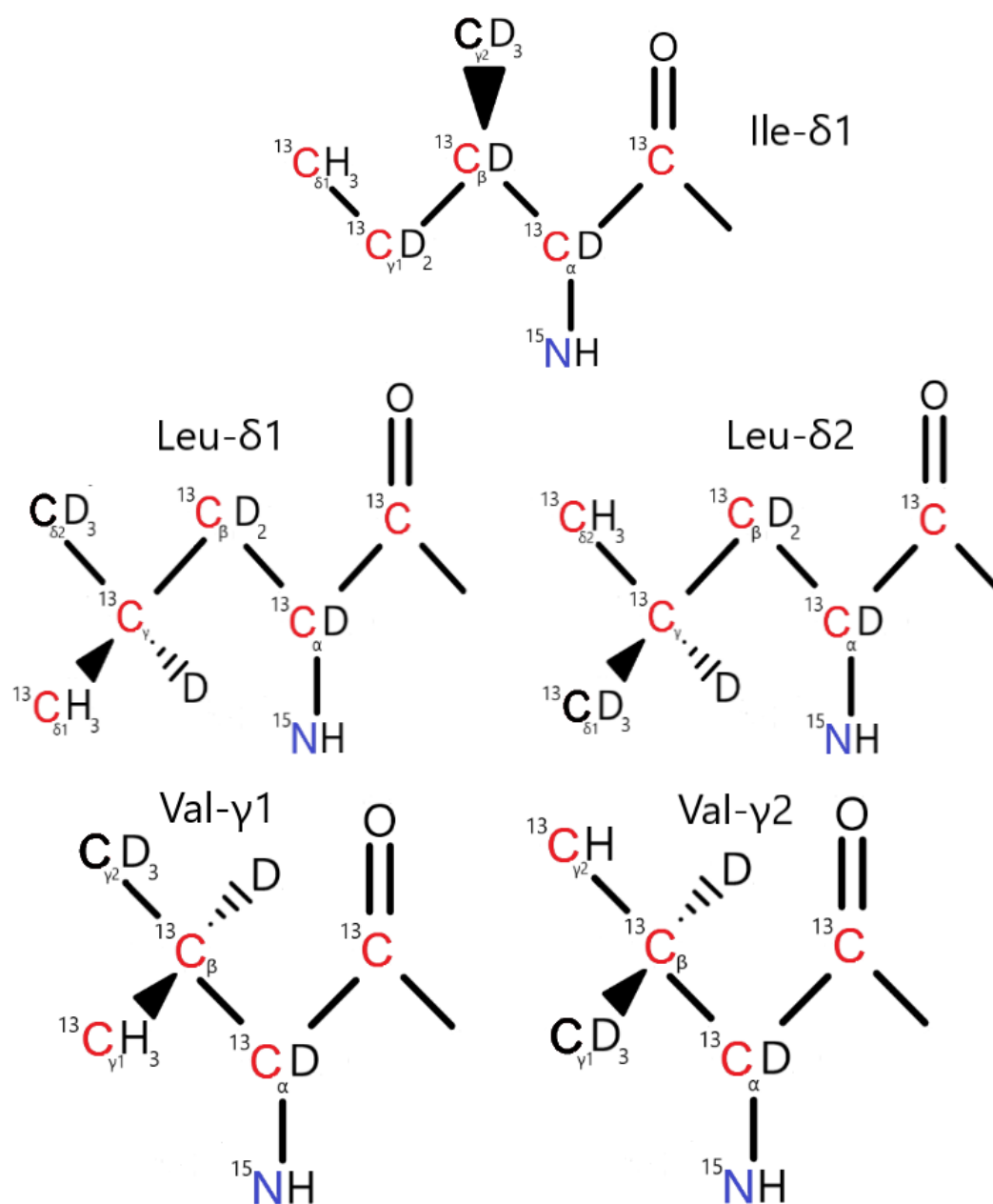


Figure 13: Labeling scheme of Ile, Val, and Leu amino acids used in this thesis. ^{13}C is marked with red. ^{12}C is marked in black.

4 Results and discussion

4.1 Assignment of $^1\text{H}/^{13}\text{C}$ HMQC methyl spectra

As mentioned previously, the major goal of this master thesis is to perform the assignment of methyl resonances of MALT1 Casp-IgL₃ domains. To assign methyl groups in MALT1, I have developed an approach that is based on an existing backbone assignment. In this section, I will describe the assignment procedure taking residue Val 705 as an example.

In the first step of the assignment procedure, a $^1\text{H}/^{15}\text{N}/^{13}\text{C}\alpha$ resonance of Val 705 was found in the HNCA spectra. The dimensions of these 3D spectra correspond to the backbone atoms - N, H, and $\text{C}\alpha$ chemical shifts (Figure 14). The position of a peak corresponding to Val 705 was established using a known assignment. As can be seen from Figure 14, the $\text{C}\alpha$ chemical shift of Val 705 is equal to 60.11 ppm.

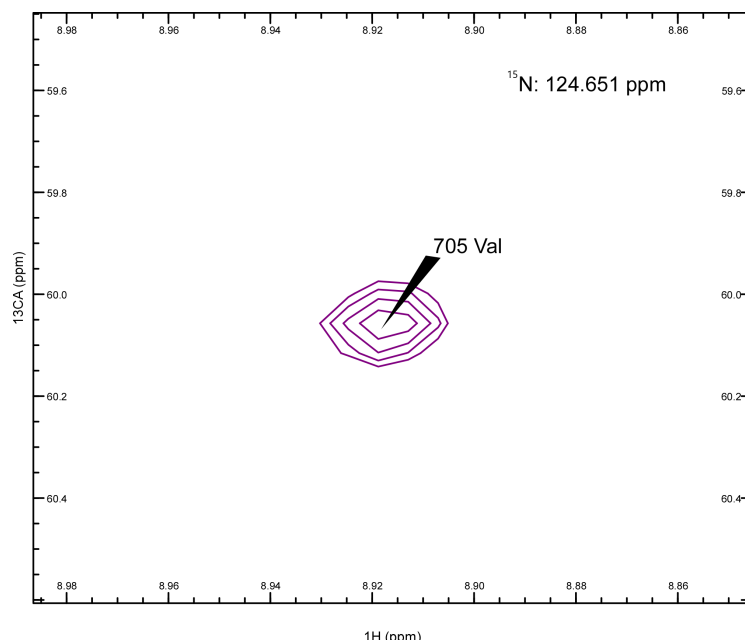


Figure 14: 3D HNCA spectra. Axes display chemical shifts of the backbone ^1H , ^{15}N and $^{13}\text{C}\alpha$ atoms. Plane $\text{N} = 124.651$ ppm is visible. Resonance from Val 705 can be observed.

In the second step the corresponding $\text{C}\alpha$ plane is found in the 3D HMCM(CB)CA spectra (HMCM(CGCB)CA spectra in the case of Leu or Ile). In those 3D spectra, the three dimensions correspond to methyl carbon (CM), methyl hydrogen (HM), and $\text{C}\alpha$ chemical shifts. In Figure 15 the HM/CM plane at $\text{C}\alpha = 60.11$ ppm is presented. That plane contains several HM/CM peaks and the choice between them has to be made. In fact, out of all peaks that can be observed in Figure 15 only two, marked as X1 and X2, have the maximum on the corresponding $\text{C}\alpha$ plane. This means that those two peaks are

probably the two methyl groups of Val 705. This can be validated through other methods. I used several approaches to determine and validate the correct resonances. As an example one of the peaks shown in Figure 15 and marked as X1 can be taken. X1 has coordinates: (0.509, 22.03).

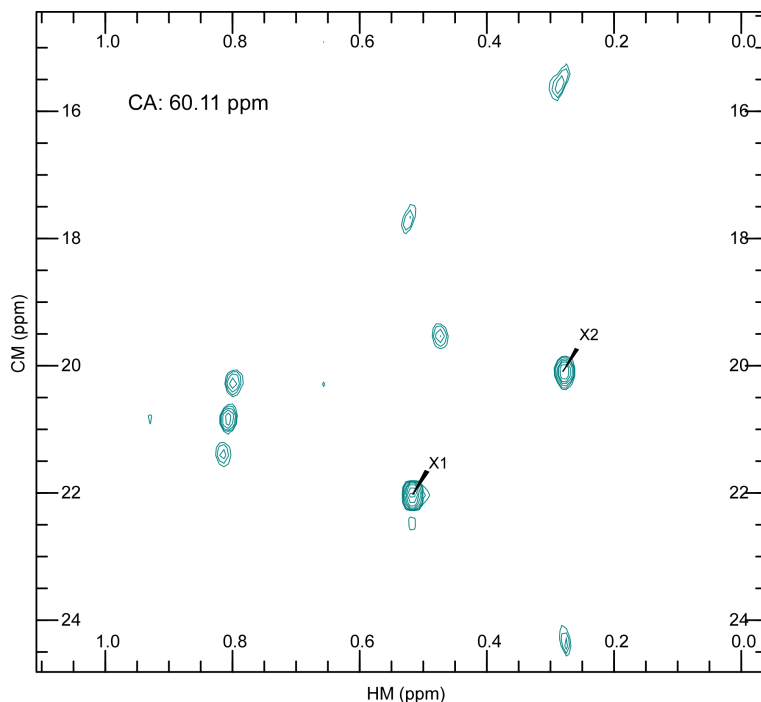


Figure 15: 3D HMCM(CB)CA spectra. Axes display the ^1H and ^{13}C chemical shifts of methyl groups and the backbone $^{13}\text{C}\alpha$ chemical shift. Plane $\text{C}\alpha = 60.11$ ppm is displayed.

One way to determine if the X1 peak originates from residue Val 705 is with the help of 3D HMCM(CGCBBCA)CO spectra which displays all three types of residues - Val, Leu, and Ile. Here, the three axes correspond to the HM, CM, and CO chemical shifts. Knowing the HM and CM shifts from the previous step, CO shift can be found. In Figure 16 you can see the HM/CO plane of 3D HMCM(CGCBBCA)CO spectra with $\text{CM} = 22.02$ ppm. The CO shift is found to be 172.279 ppm. That CO shift matches with the known assignment for Val 705. This means that the HM/CM peak we have chosen has the $\text{C}\alpha$ and CO of the Val 705. However, the peaks in HMCM(CGCBBCA)CO frequently have low signal-to-noise ratio and sometimes they are not visible in the spectra. Another issue is that the CO chemical shift may not be known.

An additional and important validation of the HM/CM resonance can be done with the 3D HN²HM NOESY spectra that displays backbone NH-methyl interactions. In this 3D spectra, three dimensions correspond to the backbone N and H chemical shifts and additionally, there is an axis with a HM chemical shift. A peak in this spectra indicates that some NH group of the backbone is close in space with a methyl group. In the spectra,

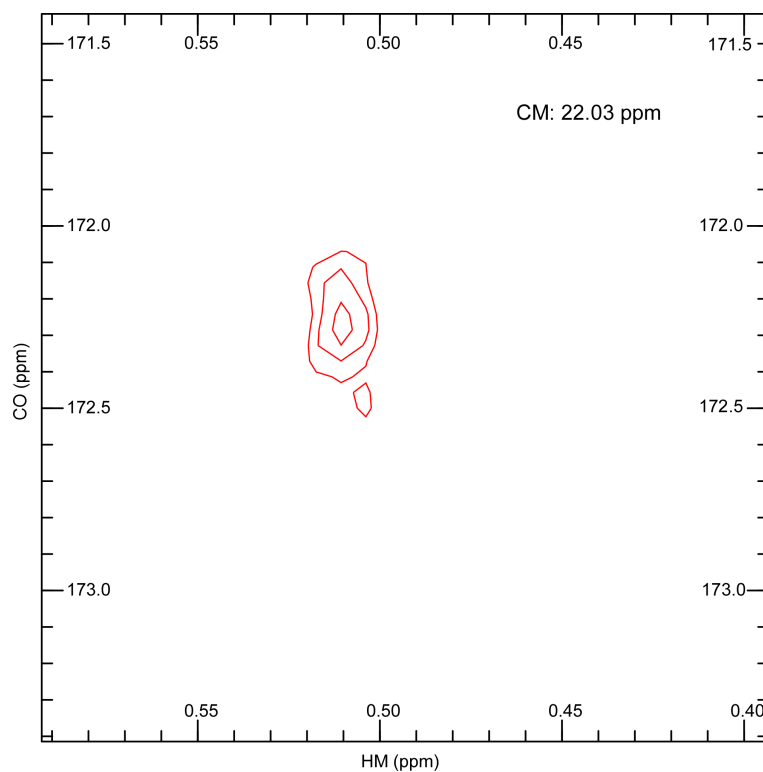


Figure 16: 3D HMCM(CGCBCA)CO spectra. Axes display the ^1H and ^{13}C chemical shifts of methyl groups and the backbone ^{13}CO chemical shift. Plane CM: 22.03 ppm.

the most intensive cross-peaks typically originate from the interaction within one amino acid. Sometimes methyls from other residues, which are close in space (I have used $5\text{\AA}$ as cutoff), give cross-peaks with the NH group.

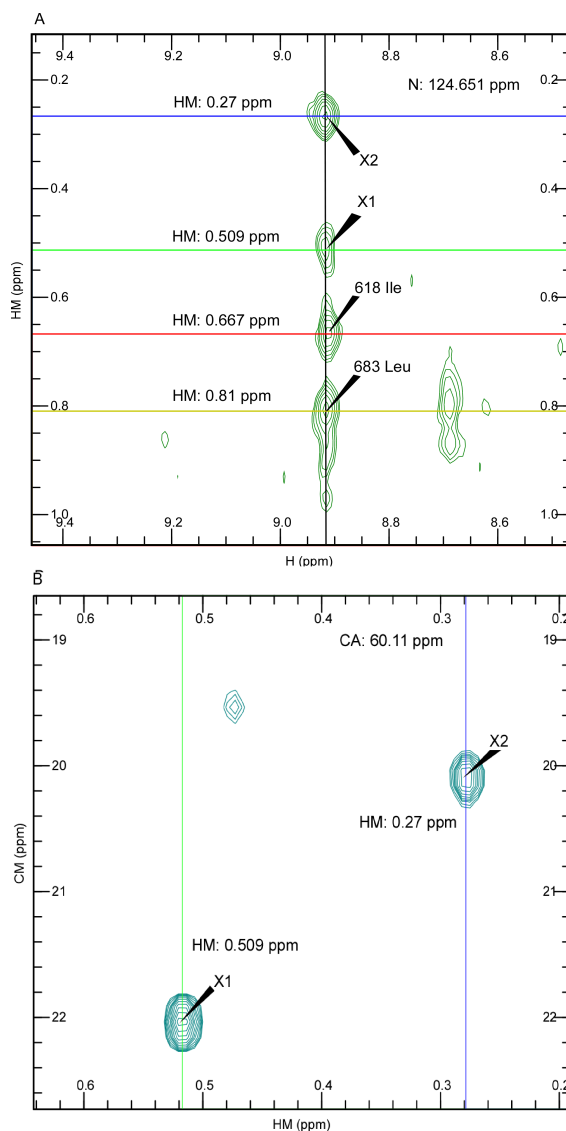


Figure 17: A: 3D HNHM NOESY spectra. Axes display the backbone ^1H , ^{15}N and the ^1H methyl group chemical shifts. A plane with $\text{N} = 124.651$ ppm, which corresponds to Val 705, is displayed. Three peaks from the interaction of NH Val 705 and the methyl groups are observed. The fourth peak has a slightly different H frequency and originates from the interaction of ^{15}N of 618 Ile with the methyl group of 618 Ile. B: 3D HMCM(CB)CA spectra. Axes display the ^1H and ^{13}C chemical shifts of methyl groups and the backbone $^{13}\text{C}\alpha$ chemical shift. Plane $\text{CA} = 60.11$ ppm, which corresponds to Val 705, is displayed. The two peaks have the same HM chemical shifts as the two peaks in Figure A.

From the output of “NH-methyl distances” script N of Val 705 is close in space with its own 705 $\text{C}\gamma 2$, $\text{C}\gamma 1$, and Leu 683 $\text{C}\gamma 1$. In Figure 17A the H/HM plane of the 3D HNHM

NOESY spectra with $N = 124.65$ ppm is presented. Three peaks corresponding to Val 705 H (black line) are observed. Two out of those four peaks have the same HM as peaks X1 and X2 (Figure 15 or Figure 17B). From that, we can conclude that X2 is 705 C γ 2 as it is more intensive and closer in space and X1 is 705 C γ 1. With that step, we have performed a stereospecific assignment of Val 705 although experiments were not designed to give us that information. Additionally, the peak from the interaction between NH and 683 C γ 1 is observed. A peak with HM = 0.667 ppm marked with a red line is slightly offset from the black line and that peak originates from Ile 618 which has N and H chemical shifts similar to Val 705. To conclude, with that step the two HM/CM peaks are validated through H, N, C α and CO.

Additional information can be acquired from the 4D HMCM NOESY spectra with two HM and two CM dimensions. With that spectrum, we can see the met-met interactions. In Figure 18 the HM/CM plane of 705 C γ 2 is presented. An intensive diagonal peak originates from 705 C γ 2 itself. The smaller peaks (cross-peaks) originate from the interaction of 705 C γ 2 with other methyls if they are close in space. The sensitivity of that interaction is 5-7Å. From the output of “methyl-methyl distances” 705 C γ 1 is close (within 5.5 Å) to 707 Val C γ 1, 602 Leu C γ 1, 592 Leu C γ 1, C γ 2, and 686 Val C γ 2. In Figure 18 you can observe the cross-peak from those assigned residues. The rest of the peaks have a maximum on a different plane meaning that those originate from a different methyl group than Val 705. One of those peaks might be 686 Val, but it is impossible to determine as this residue lacks the backbone assignment.

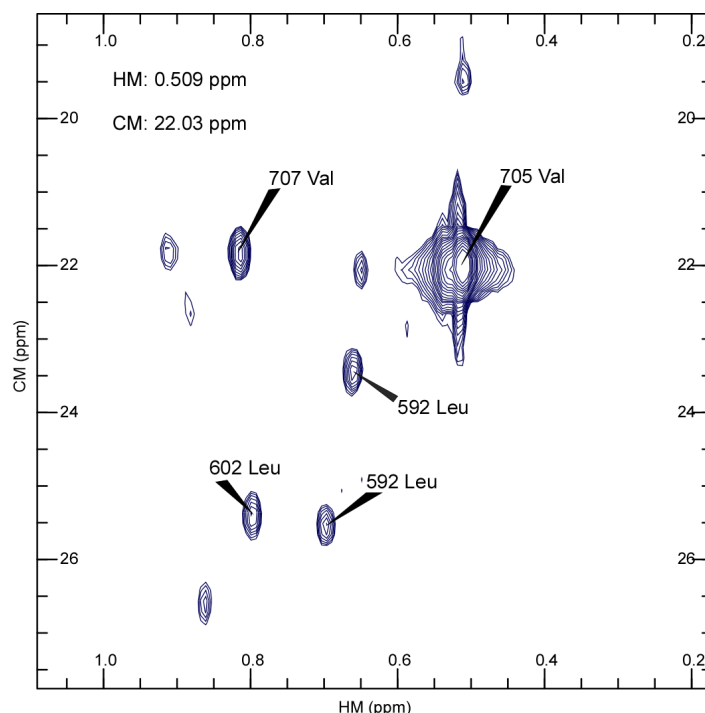


Figure 18: 4D HMCM NOESY spectra. Axes display two ^1H and two ^{13}C chemical shifts of methyl groups. Plane $\text{CM} = 22.03$ ppm and $\text{HM} = 0.509$ is displayed which potentially corresponds to one of the methyl groups of Val 705. The cross-peaks originate from the met-met interaction of Val 705 with the methyl groups of 707 Val, 592 Leu, and 602 Leu. Those are in agreement with the information acquired from the protein structure.

To conclude, we have assigned the two peaks marked with X1 and X2 on Figure 15 as Val 705 $\text{C}\gamma 1$ and $\text{C}\gamma 2$ based on the backbone H, N, $\text{C}\alpha$ and CO chemical shifts and the information from the protein structure. Knowing their CM and HM chemical shifts from the 3D HMCM(CB)CA spectra, the corresponding peak can be found and assigned in the 2D $^1\text{H}/^{13}\text{C}$ HMQC methyl spectra. Here I went from the known backbone assignment to finding the methyl peaks, assigning and validating them. The case which was described is a relatively simple one. Along with that method, I have used other approaches to assign the methyl groups. In the case when the peak was missing in the 3D spectra from methyl-and-back experiments, the assignment procedure started from the 3D HNHM NOESY spectra. With the help of a python script, I have found all possible NH-met interactions of the methyl in question and with that information, I went to the NH plane of 3D HNHM NOESY spectra and observed the cross-peaks. From that, the HM chemical shift was acquired. Then I found and checked all possible peaks with the same HM in 2D $^1\text{H}/^{13}\text{C}$ HMQC spectra through the 4D HMCM NOESY based on the protein structure. In some rare cases, when no backbone assignment exists, the assignment of methyl groups could be done based on 4D HMCM NOESY only.

MALT1 Casp-IgL₃ domains in total contain 98 ILV residues - 18 Ile, 26 Val, and 54 Leu. Of those 14 Ile, 17 Val, and 19 Leu were assigned. The ¹H-¹³C HMQC spectrum with the performed assignment is presented on Figure 23. The chemical shift table is located in the Appendix. The assigned residues are visualized on the structure of MALT1 Casp-IgL₃ on Figure 19. As can be seen from the structure, the majority of the assigned methyl groups are located in the IgL₃ domain.

In the 3D protein structure of MALT1 Casp-IgL₃, the methyl groups of Ile, Val, and Leu form four hydrophobic clusters. They are presented in Figure 20. Cluster I is located in the IgL₃ domain and is colored in yellow. Cluster II is identified between the IgL₃ and Casp domains and colored in violet. Clusters III and IV are located on both sides of the beta-sheets. They are colored in green and red (pink) respectively. Residues colored in blue can not be associated with any cluster. The methyl groups of residues that are colored in darker colors were assigned. The majority of assigned residues belong to clusters I and II.

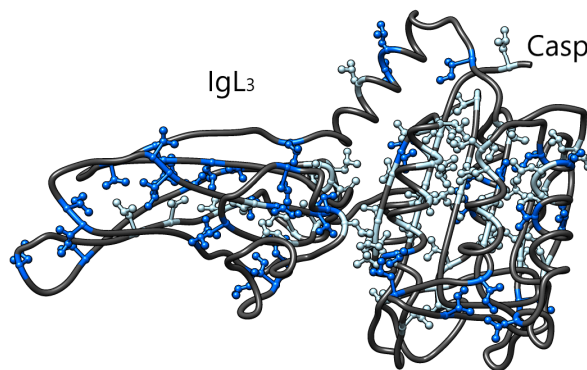


Figure 19: Methyl groups of ILV residues presented on MALT1 protein structure. Assigned residues in dark blue color, while unassigned in light blue color. The model of MALT1 that was used for this figure is predicted by AlphaFold.

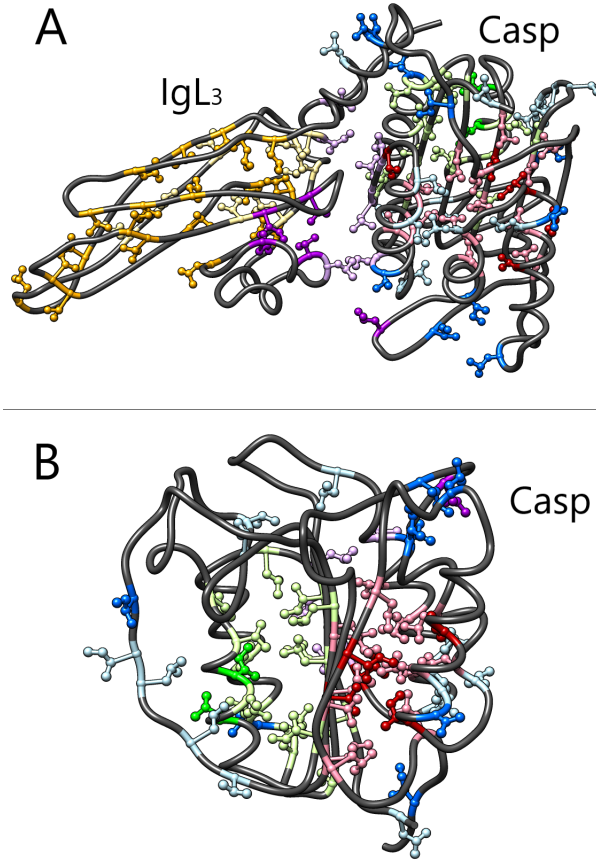


Figure 20: A: Hydrophobic clusters produced by methyl groups of Ile, Val, Leu. Cluster I - yellow, Cluster II - violet, Cluster III - green, Cluster IV - red (pink). Blue - not associated with any cluster. The assigned methyl groups are marked with darker colors and unassigned by light colors. B: The Casp domain showing Clusters III and IV between the beta sheets.

4.2 AlphaFold

AlphaFold[2] is a protein structure prediction software that won the Critical Assessment of protein Structure Prediction 14 (CASP14) in 2020. The protein structures, predicted by AlphaFold, had mean backbone accuracy of 0.96 Å RMSD⁷ which is near the experimental accuracy. AlphaFold uses novel neural network architectures in combination with evolutionary, physical, and geometrical constraints to obtain highly accurate protein structures. AlphaFold network is trained on the x-ray structures meaning that there might be a difference with the protein structure in the solution. For that reason, in my master thesis, I have obtained five structures of MALT1 Casp-IgL₃ with AlphaFold software and partially

⁷Root-mean-square deviation of atomic positions.

compared them with the structural information about the methyl groups obtained from 4D HMCM NOESY spectrum. AlphaFold structures were also valuable for this project as they contained the flexible strand (residues 468-481 in the Casp domain) which is not present in the x-ray structure. This loop contains two Ile and two Leu amino acids.

AlphaFold mean pLDDT⁸ score for obtained MALT1 Casp-IgL₃ structures is 91.98 out of 100 which is considered to be a very high confidence score. The regions of less confident structure prediction are 501-507 residues and 445-474 flexible strands in the Casp domain, the position of the alpha-helix in the IgL₃ domain (residues 575-566), and residues 654-661, 694-696 in the IgL₃ domain. The level of confidence of the structure prediction by AlphaFold is presented as a color scheme in Figure 21. Blue colors represent high confidence while yellow and red colors signal low and very low confidence respectively. In Figure 22 A there is more detailed information about the predicted pLDDT confidence score for different MALT1 Casp-IgL₃ models. AlphaFold takes the target protein sequence as an input and searches similar sequences in the databases to perform multiple sequence alignment (MSA)⁹. A schematic view of MSA for MALT1 Casp-IgL₃ is presented in Figure 22 B. The IgL₃ domain is predicted with lower confidence as there are fewer similar sequences in the databases while Casp domain is predicted with higher confidence. The two regions of Casp with low pLDDT scores are due to the fact that many sequences have deletions¹⁰ in those regions.

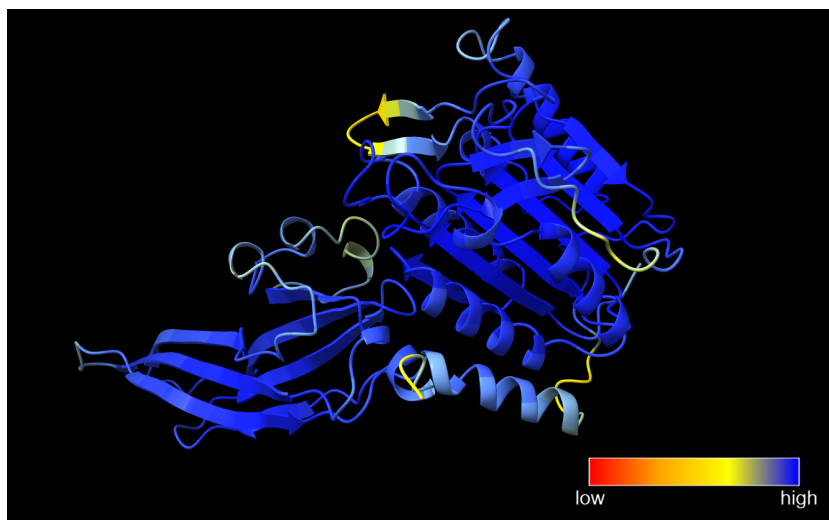


Figure 21: One of the MALT1 Casp-IgL₃ structure predicted by AlphaFold. Colors represent pLDDT structure prediction confidence score.

⁸per-residue accuracy of the structure

⁹MSA - alignment of several biological sequences. Typically used to find similarities between sequences.

¹⁰Mutation which leads to the loss of nucleotides in the DNA sequence.

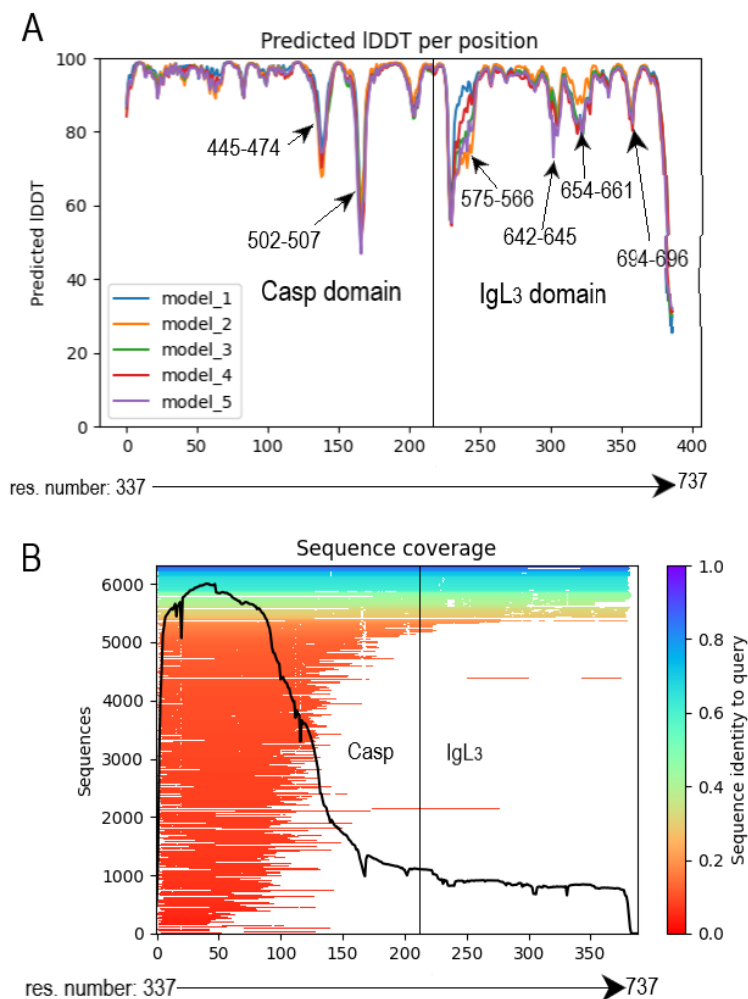


Figure 22: The output of AlphaFold. A: Detailed graph of per residue predicted pLDDT structure prediction confidence score. pLDDT = 90-100 - very high accuracy expected, 70-90 - backbone predicted well, 50-70 - low confidence, 0-50 - not confident, should not be interpreted. B: Multiple sequence alignment.

4.3 Discussion

A summary of the results of the performed assignment of methyl groups of ILV type of residues of MALT1 Casp-IgL₃ is presented in Table 1. My method of assignment to a high degree depended on the existing backbone assignment, in particular on C α and CO chemical shifts. Very few methyls were assigned based on 3D and 4D NOESY. As can be seen from Table 1 most of the methyl groups of Ile and Val amino acids, that had a backbone assignment, were assigned in this thesis. The methyl groups that were not assigned, had a resonance in the 2D ^1H - ^{13}C HMQC spectra, but no resonances in the 3D

methyl out-and-back experiments were observed. The reason for that is that during the magnetization transfer from the backbone to the methyl groups, the loss of magnetization occurs and that leads to a low signal-to-noise ratio. This is especially noticeable for Leu residues as the magnetization transfer pathway is longer for that amino acid than for Val. Ile and Leu have a similar magnetization pathway, but Ile has only one labelled residue. For that reason, the signal is higher and almost all Ile were assigned. The signal-to-noise ratio could be improved if the time of experiment is increased, but MALT1 is an unstable protein which induces limitations on the length of the experiments.

From Figure 1 it is obvious that most methyl groups that were assigned are located in the IgL₃ domain, particularly hydrophobic clusters I and II. The reason is that the Casp domain is largely involved in slow dynamical processes, which leads to the line broadening and, as a result, a weak signal.

The assignment could be improved if an additional 3D HNCM NOESY with H, N, and CM dimensions was recorded. Along with the HM chemical shift from 3D HNHM NOESY with H, N, and HM dimensions, I would have had CM chemical shift, which would give a more precise location of the peak in 2D ¹H-¹³C HMQC spectra. The low signal-to-noise ratio of methyl out-and-back spectra could be improved with stereo-specific labeling.

The confidence score of AlphaFold structures is relatively high except the regions with low sequence coverage. Those regions are the flexible loops of the protein and for that reason their structure is undefined. I have compared the AlphaFold structures of MALT1 with the x-ray 3v55 structure. Per-residue accuracy of the structures (RMSD) range from 2.28 to 2.68Å which is a relatively high number. I could not make a direct comparison of NMR data with the structures, but I have observed that met-met interactions in the 4D HMMCM NOESY in some cases are in better agreement with the x-ray structure rather than with the AlphaFold structure. If NMR constraints could be embedded in the AlphaFold algorithm, that would possibly improve the performance of AlphaFold.

Table 1: Summary of the performed assignment

Residue type	Ile	Val	Leu
Total num. of residues	18	26	54
Assigned num. of residues ¹	14	17	19
Num. of residues with H, N, Cα ²	15	20	41
Num. of residue with H, N, Cα, CO ³	13	18	39

¹ Residues that have one or both of their methyl groups assigned.

² Residues that had H, N, Cα chemical shifts.

³ Residues that had H, N, Cα, CO chemical shifts.

4.4 Conclusion

In this master thesis I have performed the assignment of the methyl groups of ILV residues of MALT1 Casp-IgL₃ domains. I used several pathways to assign the methyl groups based on the backbone assignment and on the data from the protein structure. I have assigned 14 (of 18) Ile, 17 (of 26) Val, and (19 of 54) Leu methyl groups. Although the assignment is not complete, it gives the opportunity to investigate the structure and dynamics of the hydrophobic clusters of the system. This information could further improve the development of MALT1 inhibitors. The assignment can be expanded with a different type of labeling, if supplementary experiments are performed, or if more backbone atoms are assigned.

Furthermore, in this master thesis I have worked with AlphaFold protein structure prediction software. The confidence of prediction was relatively high for the whole protein except the regions of flexible loops. I have compared the AlphaFold structures with the x-ray MALT1 Casp-IgL₃ structure. The RMSD ranged from 2.28 to 2.68 Å which is relatively high. I could not directly compare the AlphaFold structures with the NMR data, but I have observed that the x-ray in some cases is in better agreement with the NMR data. The performance of AlphaFold could be improved in the future by introducing NMR constraints into AlphaFold algorithm.

Throughout this master thesis I have acquired in-depth understanding of protein NMR spectroscopy. I have learned the theory of NMR, the related software, working with the NMR and biological data, and performing methyl side-chains assignment.

References

- [1] Xuewu Liang, YiChun Cao, Chunpu Li, Haolan Yu, Chenghua Yang, and Hong Liu. MALT1 as a promising target to treat lymphoma and other diseases related to MALT1 anomalies. *Medicinal Research Reviews*, 41(4):2388–2422, 2021.
- [2] John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589, 2021.
- [3] Z. Feng G. Gilliland T.N. Bhat H. Weissig I.N. Shindyalov P.E. Bourne H.M. Berman, J. Westbrook. Protein Data Bank. *The Protein Data Bank Nucleic Acids Research*, 28: 235-242, 2000.
- [4] Alberto Tampieri, Márk Szabó, Francesc Medina, and Henrik Gulyás. A brief introduction to the basics of NMR spectroscopy and selected examples of its applications to materials characterization. *Physical Sciences Reviews*, 6(1), 2021.

- [5] Alastair Ian Scott. Biomolecular NMR Spectroscopy By Jeremy NS Evans (Washington State University). Oxford University Press: Oxford and New York. 1995. xvi+444 pp. ISBN 0-19-85467-6., 1996.
- [6] Leif Schröder and Cornelius Faber. *In vivo NMR Imaging*. Springer, 2011.
- [7] James Keeler. *Understanding NMR spectroscopy*. John Wiley & Sons, 2011.
- [8] Alain Louis-Joseph and Philippe Lesot. Designing and building a low-cost portable FT-NMR spectrometer in 2019: A modern challenge. *Comptes Rendus Chimie*, 22(9-10):695–711, 2019.
- [9] Andrew Travers. An NMR Primer for Life Scientists. *Trends in Biochemical Sciences*, 21(3):119–119, 1996.
- [10] Gordon S Rule and T Kevin Hitchens. *Fundamentals of protein NMR spectroscopy*, volume 5. Springer Science & Business Media, 2006.
- [11] Darius Šulskis. *The Structural Modularity and Inherent Dynamics of the DegP Protease together with its Intertwined Role in the Bacterial Periplasm*. 2021.
- [12] Zachary K Boswell and Michael P Latham. Methyl-based NMR spectroscopy methods for uncovering structural dynamics in large proteins and protein complexes. *Biochemistry*, 58(3):144–155, 2018.
- [13] Iva Pritišanac, T Reid Alderson, and Peter Güntert. Automated assignment of methyl NMR spectra from large proteins. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 118:54–73, 2020.
- [14] Vitali Tugarinov, Voula Kanelis, and Lewis E Kay. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nature protocols*, 1(2):749–754, 2006.
- [15] Sofia Unnerståle, Michal Nowakowski, Vera Baraznenok, Gun Stenberg, Jimmy Lindberg, Maxim Mayzel, Vladislav Orekhov, and Tatiana Agback. Backbone assignment of the MALT1 paracaspase by solution NMR. *Plos one*, 11(1):e0146496, 2016.
- [16] CF Schulte, DE Tolmie, D Maziuk, E Nakatani, H Akutsu, H Yao, JL Markley, J Lin, JF Doreleijers, M Livny, et al. BioMagResBank. *Nucleic Acids Research*, 36(suppl_1):D402–D8, 2007.
- [17] Wim F Vranken, Wayne Boucher, Tim J Stevens, Rasmus H Fogh, Anne Pajon, Miguel Llinas, Eldon L Ulrich, John L Markley, John Ionides, and Ernest D Laue. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins: structure, function, and bioinformatics*, 59(4):687–696, 2005.

- [18] Christian Wiesmann, Lukas Leder, Jutta Blank, Anna Bernardi, Samu Melkko, Arnaud Decock, Allan D’Arcy, Frederic Villard, Paulus Erbel, Nicola Hughes, et al. Structural determinants of MALT1 protease activity. *Journal of molecular biology*, 419(1-2):4–21, 2012.
- [19] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.

Acknowledgments

I would like to thank Vladislav Orekhov, Johan Wallerstein, and Tatiana Agback for supervision and help during the project.

Appendix

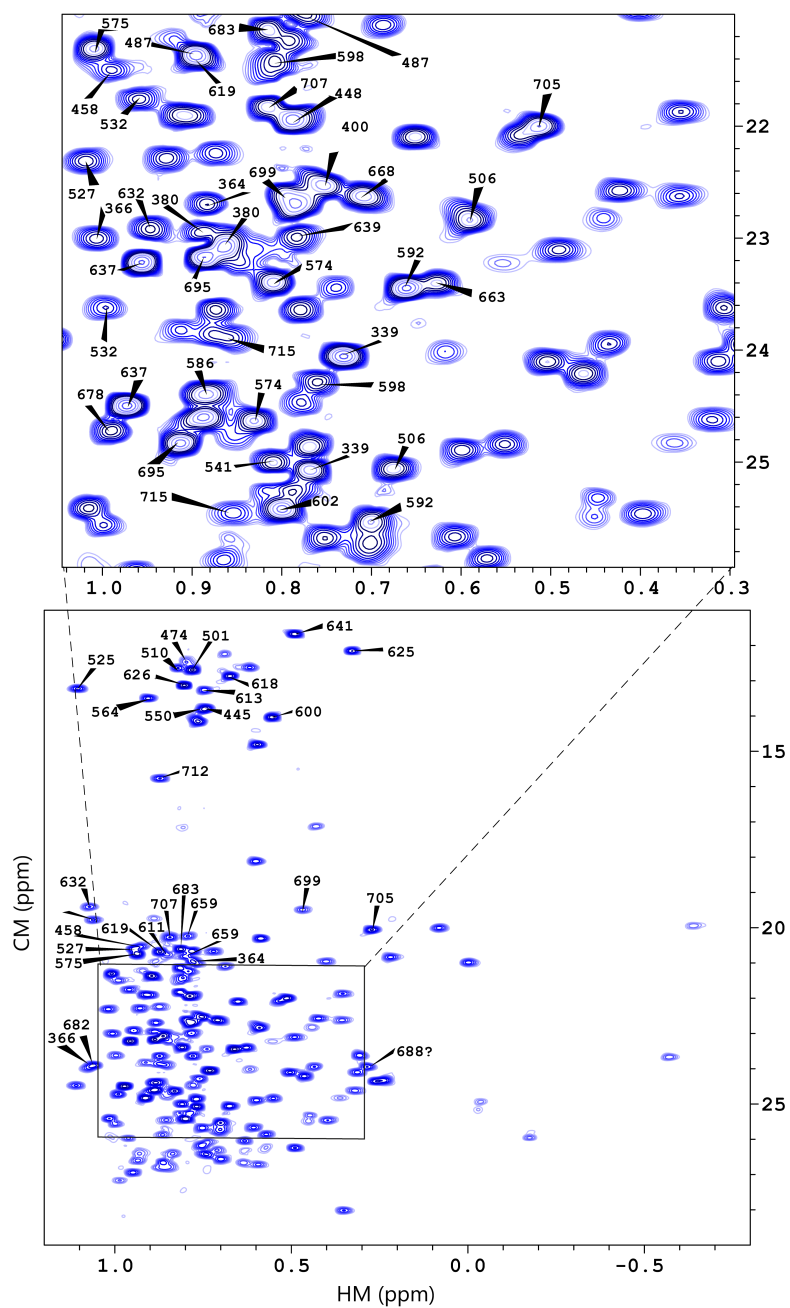


Figure 23: ^1H - ^{13}C HMQC spectrum of MALT1 Casp-IgL₃ with assigned resonances

Table 2: Non-stereospecific methyl assignment of MALT
Casp-IgL₃ domains

Residue number	¹ H (ppm)	¹³ C (ppm)
364	0.88	22.7
364	0.77	20.9
380	0.88	22.9
380	0.85	23.0
400	0.74	22.5
445	0.74	13.8
448	0.78	21.9
458	0.98	21.5
458	0.93	20.6
474	0.79	12.5
501	0.77	12.7
506	0.67	25.1
510	0.81	12.6
525	1.10	13.2
527	0.94	20.6
527	1.01	22.3
532	0.99	23.6
532	0.95	21.8
541	0.80	25.0
550	0.74	13.9
564	0.90	13.5
574	0.80	23.4
574	0.82	24.6
575	0.93	20.7
575	1.00	21.3
586	0.88	24.4
592	0.65	23.5
592	0.69	25.5
598	0.75	24.3
600	0.55	14.0
602	0.79	25.4
611	0.84	20.8
613	0.74	13.3
618	0.67	12.9
619	0.86	20.7
625	0.32	12.2
626	0.80	13.1
Continued on next page		

Table 2 – continued from previous page

First column	Second column	Third column
632	1.06	19.4
632	0.94	22.9
637	0.95	23.2
637	0.96	24.5
639	0.78	23.0
641	0.48	11.7
659	0.79	20.8
659	0.79	20.2
663	0.48	26.2
663	0.62	23.4
668	0.70	22.6
682	1.05	23.9
695	0.88	23.2
695	0.91	24.8
699	0.79	22.6
699	0.58	20.3
705	0.50	22.0
705	0.26	20.1
707	0.81	21.8
707	0.84	20.3