UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

The role of the MyD88 intermediate domain in the formation of the Myddosome assembly in innate immunity

By

Takudzwa Carlton Mervyn Jeranyama

Student number: 16087713

Supervisor: Dr. Precious Gugulethu Motshwene

Submitted in fulfilment for the degree Master of Science in Biochemistry

Department of Biochemistry, Genetics and Microbiology

Faculty of Natural and Agricultural Sciences

University of Pretoria

12 September 2023

Contents

Declaration of originality

University of Pretoria

The Department of **Biochemistry, Genetics and Microbiology** places great emphasis upon integrity and ethical conduct in the preparation of all written work submitted for academic evaluation.

While academic staff teach you about referencing techniques and how to avoid plagiarism, you too have a responsibility in this regard. If you are at any stage uncertain as to what is required, you should speak to your lecturer before any written work is submitted.

You are guilty of plagiarism if you copy something from another author's work (for example, a book, an article or a website) without acknowledging the source and pass it off as your own. In effect you are stealing something that belongs to someone else. This is not only the case when you copy work word-for-word (verbatim), but also when you submit someone else's work in a slightly altered form (paraphrase) or use a line of argument without acknowledging it. You are not allowed to use work previously produced by another student. You are also not allowed to let anybody copy your work with the intention of passing it off as his/her work.

Students who commit plagiarism will not be given any credit for plagiarised work. The matter may also be referred to the Disciplinary Committee (Students) for a ruling. Plagiarism is regarded as a serious contravention of the University's rules and can lead to expulsion from the University.

The declaration which follows must accompany all written work submitted while you are a student of the Department of **Biochemistry, Genetics and Microbiology**. No written work will be accepted unless the declaration has been completed and attached.

Full names of student: Takudzwa Carlton Mervyn Jeranyama

Student number: 16087713

Topic of work - The role of the MyD88 intermediate domain in the formation of the Myddosome assembly in innate immunity

(Division of Biochemistry)

Declaration

1. I understand what plagiarism is and am aware of the University's policy in this regard.

2. I declare that this project proposal (for example, essay, report, project, assignment, dissertation, thesis, etc) is my own original work. Where other people's work has been used (either from a printed source, internet or any other source), this has been properly acknowledged and referenced in accordance with departmental requirements.

3. I have not used work previously produced by another student or any other person to hand in as my own.

4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

Signature:  **Date:** 12 September 2023

Acknowledgements

Firstly, I would like to thank God for all that I have been blessed with. I may have forgotten HIM several times during my struggles on this journey, but clearly, HE never forgot me. Next, I would like to extend my gratitude to the following groups and people:

- I am immensely grateful for the supervision I received from Dr. Precious Gugulethu Motshwene. Your guidance and mentorship, beyond just academic obligations, will always be appreciated on my journey as a scientist.
- I would also like to thank the Council of Scientific and Industrial Research (CSIR) for the use of their facilities and support during my studies.
- I would also like to thank the University of Pretoria for the environment and opportunity to do my tertiary studies, (as well as COMESA) from the very beginning.
- In the laboratory, a lot of people have been pivotal in the progress I made during the course of my experiments. At the forefront however, I would like to thank David Thipa, Clifford Ntui, Vukosi Munyai, Court Chiramba and Linience Maposa for their assistance and constructive criticism. All your doctorates will be well deserved.
- Outside the lab, I am grateful for all my friends, especially the founding members of the Tuks Relief Committee (TRC). Namely Sam, Kevin, Nomti, Tumi, Court and Jonathan.

Last, but certainly not least, I would like to thank my family. It goes without saying that the covid pandemic was a difficult time for the whole world, but I don't think I would have made it through without these 3 people at the centre of my heart: My father, Gordon Madyiwa Jeranyama, whose presence built the moral compass for the person I am. My mother, Sarudzayi Jeranyama, whose support emphasized the value of family. My brother, Tafadzwa Huggins Jeranyama, whose optimism and positive energy inspires me to overcome any challenge I face.

Summary

Myeloid differentiation primary response protein-88 (MyD88) and Interleukin-1 receptor associated kinase 4 (IRAK-4) are two proteins that are involved in post-receptor signalling in innate immunity. These two proteins have multiple domains. However, they both share a death domain which they use for interacting with each other through homotypic death domain interactions. Although the death domains are essential for the interaction of these two proteins, the MyD88 intermediate domain is also essential for the interaction with IRAK-4. A splice variant of MyD88 that lacks the intermediate domain fails to recruit IRAK-4 during signalling.

Death domains of both MyD88 and IRAK-4 interact with each other *in vitro* to form an undecameric complex known as the Myddosome complex. This complex has a unique stoichiometry of seven MyD88 death domain molecules to four IRAK-4 death domain molecules. It was reconstituted using MyD88 that had both the death and intermediate domains. However, there was a research group that reconstituted the Myddosome without the MyD88 intermediate domain. Their finding was contrary to existing literature.

This project set out to investigate the role of the MyD88 intermediate domain in the assembly of the Myddosome complex. We therefore expressed two variants of MyD88 in addition to the IRAK-4 death domain to reconstitute the Myddosome complex. The two variants were the MyD88 death domain and MyD88 that contained both the death and intermediate domains. MyD88 that lacked the intermediate domain was found to be an oligomer using size exclusion chromatography and dynamic light scattering. However, MyD88 that contained only the death domain was monomeric.

MyD88 that contained both the death and intermediate domain was mixed with IRAK-4 death domain to reconstitute the Myddosome complex. As expected, we managed to form the Myddosome complex. The presence of this complex was confirmed by size exclusion chromatography and SDS-PAGE. On the contrary, the MyD88 death domain that lacked the intermediate domain failed to form the Myddosome complex when mixed with IRAK-4 death domain. This finding showed that the MyD88 intermediate domain was essential for the Myddosome assembly. It was also in support of existing literature that MyD88 does not interact with IRAK-4 in the absence of its intermediate domain.

List of abbreviations

Abbreviation	Full nomenclature
DD	Death domain
DNA	Deoxyribonucleic acid
KD	Kinase domain
ID	Intermediate domain
IFN	Interferon
IL-1/6	Interleukin-1/6
IL-1R	Interleukin-1 receptor
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRAK-1/2/4/M	Interleukin-1 receptor associated kinase 1/2/4/M
LPS	Lipopolysaccharide
MAPK	Mitogen associated protein kinase
MyD88	Myeloid differentiation primary response protein-88
NF-κB	Nuclear factor kappa B
OD_{600}	Optical density at 600 nm
PAMP	Pathogen associated molecular pattern
PRR	Pathogen recognition receptor
SDM	Site directed mutagenesis
SDS-PAGE	Sodium dodecyl-sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
TIR	Toll/Interleukin-1 receptor
TIRAP	TIR-adaptor protein
TLR	Toll-like receptor
TRAF6	TNF receptor associated factor 6

List of figures

Chapter 6: Expression, purification and analysis of IRAK-4-DD protein

Chapter 7: Protein to protein interaction studies in Myddosome assembly

List of tables

Chapter 1: Literature review

1.1 Innate immunity

The immune system is divided into two, namely the adaptive immune system and the innate immune system. The adaptive immune system uses specialised cells known as B and T lymphocytes to mount an immune response against an antigen (den Haan, Arens, & van Zelm, 2014). An adaptive immune response can take days or even weeks to produce antibodies that are targeted towards an invading antigen (Iwasaki & Medzhitov, 2015). The innate immune system on the other hand, is present from birth – hence the name "innate". This system is the first line of defence and it uses a limited number of receptors known as pathogen recognition receptors (PRRs) to elicit an innate immune response (Amarante-Mendes et al., 2018). Unlike the adaptive immune system, the innate immune system responds very quickly to antigens (Riera Romo, Pérez-Martínez, & Castillo Ferrer, 2016). The focus of this dissertation is primarily on the innate immune system.

PRRs are distinct proteins that recognize molecules found in pathogens to elicit a signalling response that leads to the production of pro-inflammatory cytokines and type I interferons (Konno et al., 2018). While these receptors are limited in number, they have a crucial role in detecting evolutionarily conserved molecules that are commonly present in pathogens (Akira, Uematsu, & Takeuchi, 2006). These evolutionarily conserved molecules are known as pathogen associated molecular patterns (PAMPs).

Another important role of PRRs is in the detection of endogenous cell and tissue damage. This can occur when an organism has sustained some form of physical or mechanical trauma, such as falling to the ground. The impact of the incurred injury leads to cell and tissue damage. In such a case, PRRs will detect specific molecules released by the damaged cells to elicit an immediate response (Santoni et al., 2015). These molecules are known as damage associated molecular patterns (DAMPs) and their detection is important to maintain the well-being of an organism (Zindel & Kubes, 2020).

PRR proteins are classified into five groups (Amarante-Mendes et al., 2018). These are the Toll-like receptors (TLRs), retinoic acid-inducible gene I (RIG-I) -like receptors (RLRs), nucleotide-binding oligomerisation and domain (NOD) -like receptors (NLRs), AIM2-like receptors and C-type lectin receptors (CLRs). However, our focus will only be on TLRs in this dissertation.

1.2 Toll-like receptors (TLRs)

Toll is a protein found in the fruit fly, *Drosophila melanogaster*, where its primary role is dorsoventral patterning during embryogenesis (Anderson, Jürgens, & Nüsslein-Volhard, 1985). It was later found to have another function of protecting the fruit flies against fungal infections (Lemaitre, Nicolas, Michaut, Reichhart, & Hoffmann, 1996). This finding was an important discovery as it led scientists to search expressed sequence tag (EST) databases for mammalian homologues of Toll. Thirteen were found in

mammals and were named Toll-like receptors (TLRs) due to their resemblance of the protein Toll (Takeda & Akira, 2015). Out of the thirteen TLRs, human beings only have ten as shown in Table 1 below (Fitzgerald & Kagan, 2020). The discovery of TLRs led scientists to conclude that mammals also possess an innate immune system (Müller, Vogel, Alber, & Schaub, 2008).

TLRs are type 1 transmembrane proteins that recognise pathogen associated molecular patterns to induce an innate immunity response (Federico et al., 2020). They have a cytoplasmic domain that is homologous to interleukin-1 receptors (IL-1Rs). Due to its homology with interleukin-1 receptors (IL-1Rs), it was named the Toll/IL-1R domain (Kawasaki & Kawai, 2014). TLRs also possess a unique extracellular domain (ECD) that comprises of a large concentration of leucine-rich repeats (Bryant, Symmons, & Gay, 2015).

The ten human TLRs shown in Table 1 are highly specific for the ligands they bind. For example, TLR4 binds lipopolysaccharide (LPS) also known as endotoxin. However, other TLRs work in pairs as heterodimers for ligand binding. An example of these includes TLR1 and TLR2 which are both required to bind lipoproteins.

Table 1. Toll-like receptors in humans. Current set of known human TLRs and corresponding ligands detected. Adapted from figure 1 of (Fitzgerald & Kagan, 2020)

Name	Ligands detected
TLR1/TLR2	Lipoproteins
TLR2/TLR6	Lipoproteins
TLR3	Double stranded viral RNA
TLR4	Lipopolysaccharide (LPS) (endotoxin)
TLR5	Flagellin
TLR7	ssRNA
TLR8	ssRNA
TLR9	CpG DNA
TLR10	Double stranded RNA

1.3 TLR - Signalling

When TLRs bind to their appropriate ligands, they undergo dimerization to become activated and initiate a signalling response (Fitzgerald & Kagan, 2020). There are various signalling pathways that occur for the different TLRs. However, for this dissertation, we will only focus on TLR4 signalling pathways which are activated by the detection of lipopolysaccharides (LPS). The LPS molecules are typically found in the cell wall of gram-negative bacteria and studies have shown that humans are hypersensitive to this ligand (Brinkworth & Valizadegan, 2021). TLR4 detection of LPS, leads to a signalling cascade that results in the production of pro-inflammatory cytokines (Płóciennikowska, Hromada-Judycka, Borzęcka, & Kwiatkowska, 2015). However, due to LPS hypersensitivity in humans, the pro-inflammatory response can at times be excessive. Consequently, this can unfortunately cause septic

shock as well as chronic inflammatory disorders (Brinkworth & Valizadegan, 2021). Therefore, the TLR4 signalling pathway is a potential target for drug design aimed at developing anti-inflammatories.

There are four adaptor proteins that are used in TLR4 signalling. These adaptor proteins mediate two different signalling pathways in response to LPS detection. One of the signalling pathways is known as the MyD88 dependent signalling pathway. It uses an adaptor protein called Myeloid differentiation primary response protein-88 (MyD88) in conjunction with another adaptor protein named TIR-domain containing adaptor protein (TIRAP) (Chan et al., 2018). The other signalling pathway that is triggered by TLR4 activation, is the MyD88 independent signalling pathway. It utilizes the two remaining adaptor proteins which are the TIR domain-containing adapter-inducing interferon-β (TRIF) and the TRIF-related adaptor molecule (TRAM).

In the MyD88 dependent signalling pathway, LPS detection is facilitated by lipopolysaccharide binding molecules (LBP) as illustrated in Figure 1. The LPS-loaded LBP molecules are assisted by a special transmembrane receptor called, CD14, to effectively shuttle the LPS molecules to the TLR4 ectodomain (Ciesielska et al., 2021). Ligand detection is subsequently achieved via binding of the LPS to a critical complex of TLR4 ectodomain and the MD-2 molecule. Signal transduction then occurs along the TLR4 protein, to induce recruitment of MyD88 and TIRAP. These adaptor protein interactions are governed by C-terminal TIR-TIR domain associations. Once the signal is received by the adaptor proteins, MyD88 will then recruit a series of kinases sequentially to form a crucial signalling complex. The first kinase that is recruited to MyD88, is Interleukin-1 receptor associated kinase 4 (IRAK-4). This recruitment occurs through N-terminal homotypic death domain interactions. IRAK-4 subsequently recruits and phosphorylates IRAK-1 and IRAK-2 proteins. The complex formed by the death domain interactions of MyD88 and the IRAK proteins, is known as the Myddosome complex (Motshwene et al., 2009).

The Myddosome complex amplifies the signalling cascades in the MyD88 dependent pathway. The TNF-receptor associated factor 6 (TRAF6) protein receives the amplified signal from the phosphorylated IRAK-1 in the Myddosome complex. Thereafter, TRAF6 transmits the signal to transforming growth factor β activating kinase (TAK1) protein which in turn activates the nuclear factor kappa B (NF-κB) and the mitogen associated protein kinases (MAPK). Ultimately, NF-κB and activator protein-1 (AP-1) transcription factors induce activation of genes that upregulate the production of pro-inflammatory cytokines as shown in Figure 1.

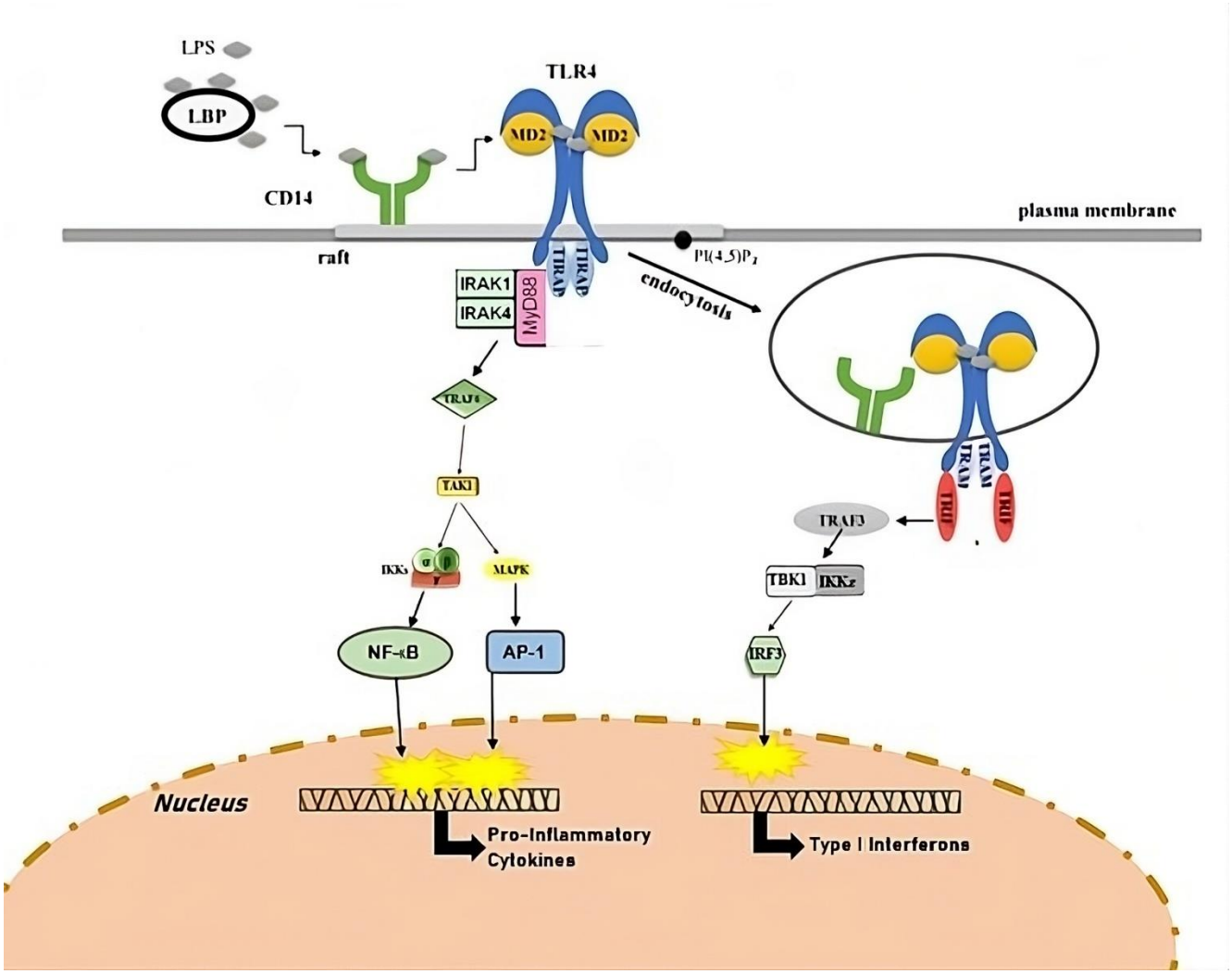


Figure 1. TLR4 signalling pathways. Illustration of the two signalling pathways that occur upon LPS detection by the TLR4/MD-2 complex. The MyD88 dependent signalling pathway cascades into the production of pro-inflammatory cytokines. The MyD88 independent signalling pathway results in the production of type I interferons (Adapted from figure 1 of (Ciesielska et al., 2021).

In the MyD88 independent signalling pathway, LBP and CD14 receptor co-operate in the trafficking of LPS to the TLR4/MD-2 binding complex. As shown in figure 1, endocytosis of the stimulated TLR4/MD-2 complex region on the plasma membrane occurs, and this becomes the trigger to initiate the MyD88 independent signalling pathway (Tsukamoto et al., 2018). The signalling cascade proceeds with the recruitment of TRAM and TRIF adaptor proteins via TIR-TIR domain interactions. The transduced signal is received by TNF-receptor associated factor 3 (TRAF3) which leads to the recruitment of TANK binding kinase 1 (TBK1) protein and the IKKε kinases. This in turn leads to the phosphorylation of the interferon regulatory factor 3 (IRF3) transcription factor (Ciesielska et al., 2021). Activation of this protein leads to transcriptional activation of genes that mediate the production of type I interferons.

1.4 Domains involved in TLR signalling

Toll-like receptors (TLRs) are composed of three domains as shown in Figure 2. These domains facilitate the signalling process during an innate immune response (Behzadi, García-Perdomo, & Karpiński, 2021). There is the ectodomain (ECD) which is located in the extracellular space. It is responsible for highly specific ligand/PAMP detection to initiate a signalling response. The next domain is known as the transmembrane domain (TD), which is in the form of a single helical structure. This domain is embedded in the cell membrane and basically anchors the TLR to it (Godfroy, Roostan, Moroz, Korendovych, Yin, 2012). The third domain is known as the Toll/Interleukin-1 receptor (TIR) domain. It is found in the cytosol, where it interacts through homotypic TIR-TIR interactions during active TLR signalling.

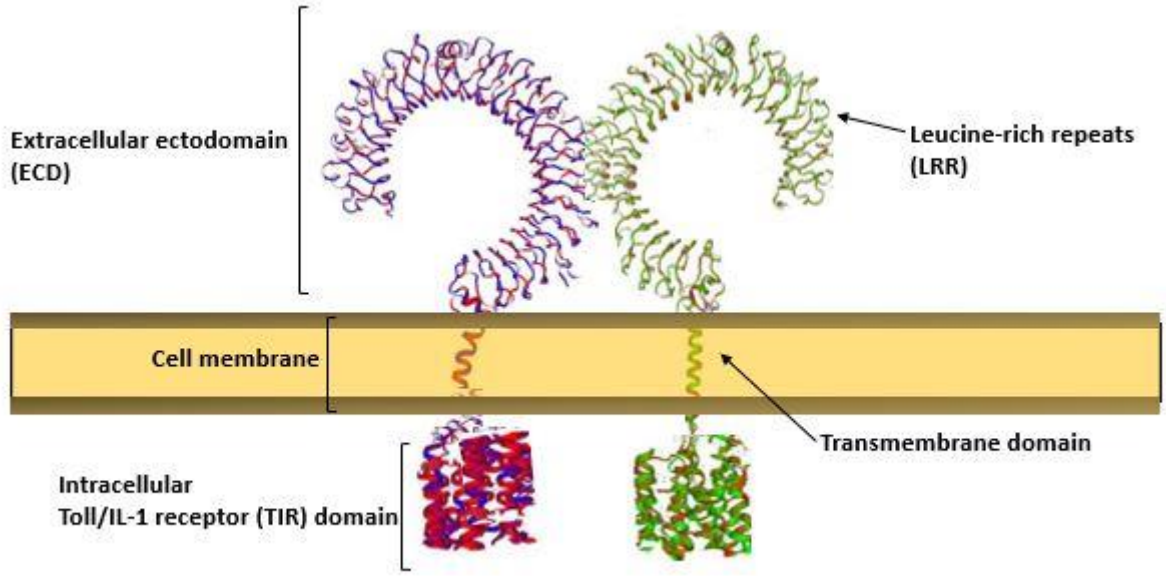


Figure 2. Domain architecture for TLRs. TLRs are structurally composed of three domains. These include an extracellular ectodomain (composed of several leucine-rich repeats), a single helix transmembrane domain and an intracellular TIR domain. (Adapted from figure 1 by Sameer & Nissar, 2021)

The ECD is composed of a large number of leucine-rich repeat (LRR) motifs (Sameer & Nissar, 2021). As a result of the dense packing of the leucine residues in these LRR motifs, the ECD forms a "horse-shoe" shape structure (Wang, Zhang, Liu, Zhao, & Yin, 2016). The LRR motifs are made up of conserved twenty-four amino acid chains (Botos et al., 2011) that contain "LxxLxLxxN/CxL" consensus sequences (whereby "x" represents any amino acid) (Botos et al., 2011). Subtle variation in these motifs allows the diversity of ligand detection that occurs when comparing different TLR ectodomains. Several crystal structures have been solved already for many ligand-bound ECDs. A few examples include the

TLR1/TLR2-lipoprotein structure (Jin et al., 2007), the TLR4-MD2-LPS complex (B. S. Park et al., 2009) and the TLR5-flagellin crystal structure (Yoon et al., 2012).

TIR domains are situated in the cytosol. They are generally composed of a five β-sheet central core which is surrounded by five alpha helices (Yin, Fu, Li, & Wu, 2015). One of the most important features in TIR domains is a loop that connects the β B-strand and the α B-helix. It is known as the BB-loop and it functions as a critical interface during TLR signalling (Nimma, Ve, Williams, & Kobe, 2017). TIR domains associate with other TIR domains at this critical interface through homotypic interactions. These interactions occur when TLRs dimerize as well as during the recruitment of post-receptor TIR-domain containing proteins. Examples of such post-receptor TIR-domain containing proteins include the adaptor proteins namely MyD88, TIRAP, TRIF and TRAM (Nanson, Kobe, & Ve, 2019).

1.5 Domains involved in TLR post-receptor signalling

In TLR post-receptor signalling, there are two prominent domains. The first domain is the Toll/Interleukin-1 receptor (TIR) domain. This intracellular domain mediates the initial signalling cascades for stimulated TLRs (Nanson, Kobe, & Ve, 2019). The TIR domain is also crucially present on cytosolic adaptor proteins (MyD88, TRIF, TIRAP, TRAM) that mediate post-receptor signalling. These adaptor proteins possess the TIR domain on their C-terminus, and their shared structural homology of this domain facilitates homotypic TIR-TIR interactions. However, proteins such as MyD88 also have another crucial domain present on their N-terminus (Avbelj et al., 2018). It is known as the death domain (DD), and it represents the second prominent domain in TLR post-receptor signalling.

Death domains (DDs) are a group of proteins that mediate post-receptor signalling in inflammation as well as apoptosis (Meng et al., 2018). These DDs belong to a large superfamily of proteins known as the Death Domain proteins which is classed into four sub-family groups. The groups are the death domains (DDs), death effector domains (DEDs), caspase recruitment domains (CARDs) and lastly the pyrin domains (PYDs) (Huoh & Hur, 2022). While the Death Domain super family has these different sub-classes, they are all important as conserved domains for innate immunity and programmed cell death processes.

Structurally, all members of the Death Domain superfamily share a unifying feature as shown in Figure 3 (Bai, Ma, Liu, Huang, & Hu, 2022). It is a conserved structural fold that is comprised of six alpha helices (Ding et al., 2019). The differences in the orientation of these helices, leads to the different sub-classes that define the Death Domain superfamily (H. H. Park et al., 2007). Nonetheless, death domains are known to self-associate and associate with other death domains through homotypic protein-protein interactions (Zhou, Kaneko, Nakagita, Takeda, & Masumoto, 2021). The result of these DD-DD interactions is the formation of supramolecular organizing centres (SMOCs) such as the PIDDosome and Myddosome complexes (Kagan et al., 2014; Tartey & Takeuchi, 2017).

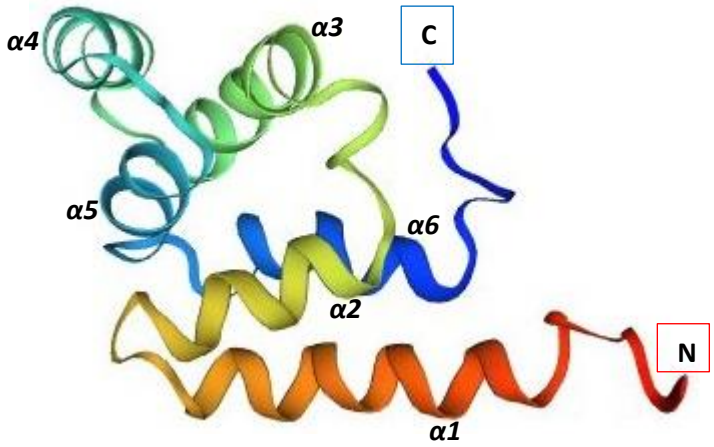


Figure 3. Unifying structural core for the Death Domain superfamily. Model crystal structure of the conserved six alpha-helical bundle in death domains. Each alpha helix is labelled (α1-α6) from the N-terminus to the C-terminus. Adapted from figure 1 of (Bai, Ma, Liu, Huang, & Hu, 2022)

1.6 MyD88 adaptor protein

Myeloid differentiation primary response protein-88 (MyD88) is an adaptor protein that mediates MyD88 dependent TLR signalling pathways in innate immunity (Kiripolsky, Romano, Kasperek, Yu, & Kramer, 2020). In these signalling cascades, activated TLRs recruit this adaptor protein through TIR-TIR interactions (Vetreno, Qin, Coleman Jr, & Crews, 2021). MyD88 is also supported by an adaptor protein called TIRAP (through TIR domain interactions) during this initial signalling process. Afterwards, MyD88 plays a key role in the co-ordinated recruitment of kinase proteins known as Interleukin-1 receptor associated kinases (IRAKs) (Dossang et al., 2016). Death domain interactions occur between MyD88 and IRAKs which subsequently leads to the formation of supramolecular organizing centres, namely the Myddosome complex (Kagan et al., 2014). Therefore, MyD88 is very important in both the initiation of cytosolic signal transduction, and the downstream post-receptor signalling processes.

The primary structure of MyD88 shows that it is composed of three domains as shown in Figure 4 (Deguine & Barton, 2014). It has a death domain on the N-terminus, that covers the first 110 amino acids. As mentioned previously, the domain is made up of a conserved six alpha-helical bundle which is a universal feature amongst members of the DD superfamily (Tran, Liu, Jakovlić, & Wang, 2015). The intermediate domain (ID) spans from amino acids 111 to 155 and it is unstructured based on available data in literature (Avbelj et al., 2018). On the C-terminus of MyD88 is the TIR domain which spans from amino acids 156 to 296. The TIR domain is made up of a central five β-sheet bundle which is surrounded by five alpha helices (Yin et al., 2015). Therefore, the DD, ID and TIR domain are the three characterizing domains for the MyD88 protein. However, there is splice variant of MyD88 that lacks the intermediate domain. This variant is known as MyD88s (Xin et al., 2016).

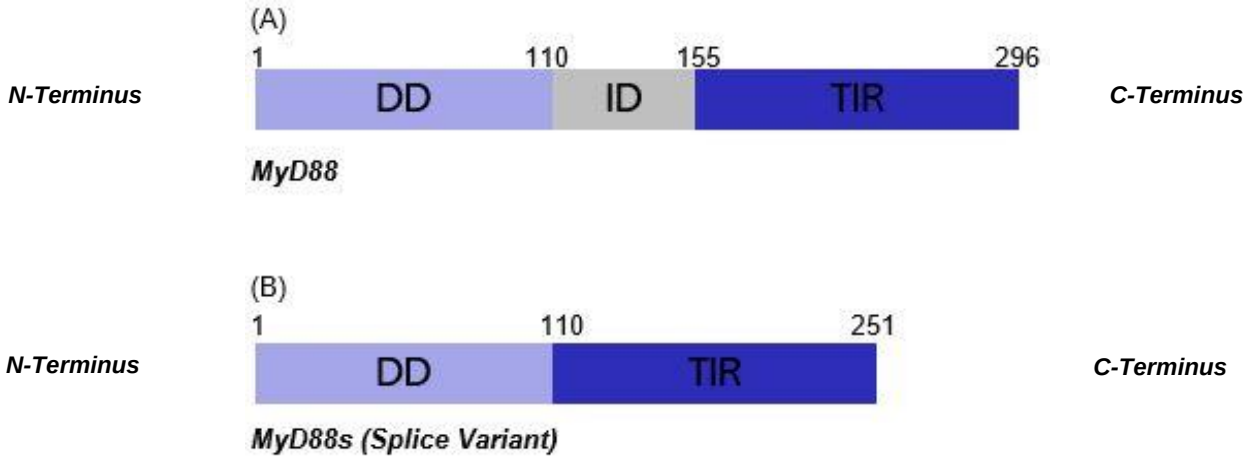


Figure 4. MyD88 domains overview. (A) The MyD88 protein is comprised of three domains: N-terminal death domain (DD), a short intermediate domain (ID) and a C-terminal Toll/interleukin-1 receptor domain. **(B)** A splice variant of MyD88 protein lacking the intermediate domain (Adapted from (Deguine & Barton, 2014).

In addition to the structural differences between MyD88 and MyD88s, there are also different physiological effects exhibited by these proteins during TLR signalling as shown in figure 5 (Janssens & Beyaert, 2003). During TLR signalling, full-length MyD88 is recruited to the TLR through homotypic TIR-TIR domain interactions. Subsequently, MyD88 recruits the IRAK-4 protein through homotypic death domain interactions leading to the activation of NF-κB transcription factors. These transcription factors translocate to the nucleus where they induce the production of pro-inflammatory cytokines. MyD88s on the other hand, is unable to recruit the IRAK-4 protein (Janssens et al., 2003). Failure to recruit the IRAK-4 protein prevents the signalling cascade from proceeding all the way to the translocation of NF-κB into the nucleus where it induces the transcription of pro-inflammatory cytokines.

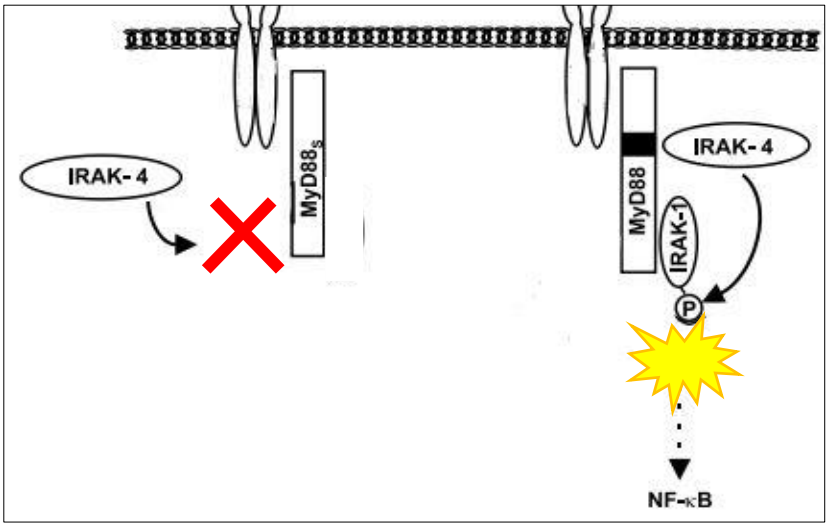


Figure 5. Comparison of MyD88s and MyD88 intracellular effects. A simplified representation of MyD88 proteins after stimulation by LPS. ID-deficient MyD88s fails to recruit IRAK-4 while MyD88, which contains the ID, is able to recruit IRAK-4 and induce NF-κB signalling pathway (Adapted from Figure 3 of (Janssens et al., 2003)

1.7 The importance of the MyD88 adaptor protein

The MyD88 adaptor protein has proven to be an essential component for a pro-inflammatory response in innate immunity (Saikh, 2021). This is evident from several studies that show the importance of MyD88 in TLR4 signalling pathways. For example, MyD88 knockout mice were used to investigate the significance of this protein in response to LPS, the only known ligand for TLR4 (Kawai et al., 1999). The results from the study showed that MyD88-deficient mice were resistant to high lethal doses of LPS (Kawai et al., 1999). All the MyD88 knockout mice survived the lethal doses of LPS. In contrast, wild-type mice that possessed MyD88 died of septic shock within four days of exposure to the high lethal doses of LPS (Kawai et al., 1999). This led to the conclusion that MyD88 is a necessary mediator of an innate immunity signalling responses. Further evidence of this conclusion was observed in a study of children with an autosomal recessive MyD88 deficiency (Israel et al., 2017). These children were highly susceptible to infections from pyogenic bacteria such as the *Pseudomonas aeruginosa* and *Staphylococcus aureus* species (Akar-Ghibril, 2022). However, they showed a poor inflammatory response due to the absence of the MyD88 protein.

The absence of MyD88 in the above-mentioned studies, led to a poor inflammatory response. This finding means that inhibiting the MyD88 protein would have an anti-inflammatory effect. Therefore, this makes the MyDD88 protein an ideal drug target for the development of new anti-inflammatories (Saikh, 2021). While there are many anti-inflammatory drugs available in the market, there are those that have adverse side effects (Haley & von Recum, 2019). As such, MyD88 is an ideal candidate to explore the development of better anti-inflammatories that target the TLR4 pathway.

1.8 IRAK-4 protein

Interleukin-1 receptor associated kinase-4 (IRAK-4) is an important mediator of post-receptor TLR signalling in innate immunity (Su, Xu, & Huang, 2020). It belongs to the IRAK family of serine/threonine kinases of which there are four known members. These are IRAK-1, IRAK-2, IRAK-3 (also referred to as IRAK-M) and IRAK-4 (Balka & De Nardo, 2019).

IRAK-4 is composed of three domains as shown in Figure 6a (Balka & De Nardo, 2019). It has an N-terminal death domain that covers the first 104 amino acids. This domain facilitates homotypic death domain interactions between IRAK-4 and MyD88 as well as IRAK-4 and other IRAK proteins that also possess a death domain. The second domain spans from amino acid 105 to 186. This domain is known as the proST domain and it is characterized by a dense concentration of proline, serine and threonine residues (Boraschi, Italiani, Weil, & Martin, 2018). The last domain of IRAK-4 is the kinase domain (KD). The KD is a large domain with a sequence ranging from amino acid 187 to 460. It is responsible for the kinase activity of the IRAK-4 protein in the phosphorylation of its substrates during post-receptor signalling (Z. Wang et al., 2009).

The crystal structure of the kinase domain (KD) was solved and it is shown in Figure 6b (Z. Wang et al., 2009). It is composed of a small N-lobe and a large C-lobe. The N-lobe is comprised of a single alpha helix (αC) that is linked to 5 twisted β-sheet strands. There is a substrate-binding cleft that is adjacent to this N-lobe. Two critical amino residues are found in this substrate-binding cleft. One of the residues is known as the invariant lysine 213 residue (K213) because it is conserved in the IRAK family. This residue is crucial because it is the ATP binding site of IRAK-4. Kinases require ATP in order to phosphorylate their substrates (Liu et al., 2020). Studies have revealed that disruption of this lysine residue abolishes kinase activity of the IRAK-4 protein (Jain, Kaczanowska, & Davila, 2014). The other important residue in the substrate binding cleft is aspartate 311 (D311). It is essential in enabling catalytic activity of IRAK-4. Mutating this residue renders an inactive IRAK-4 protein (Liu et al., 2020). Lastly, there is the large C-lobe which is made up of seven alpha helices (αD to αI). The C-lobe also possesses an important activation loop that contains a few residues that are sites for autophosphorylation of IRAK-4. These include two threonine residues (T342 and T345) as well as a serine residue (S346).

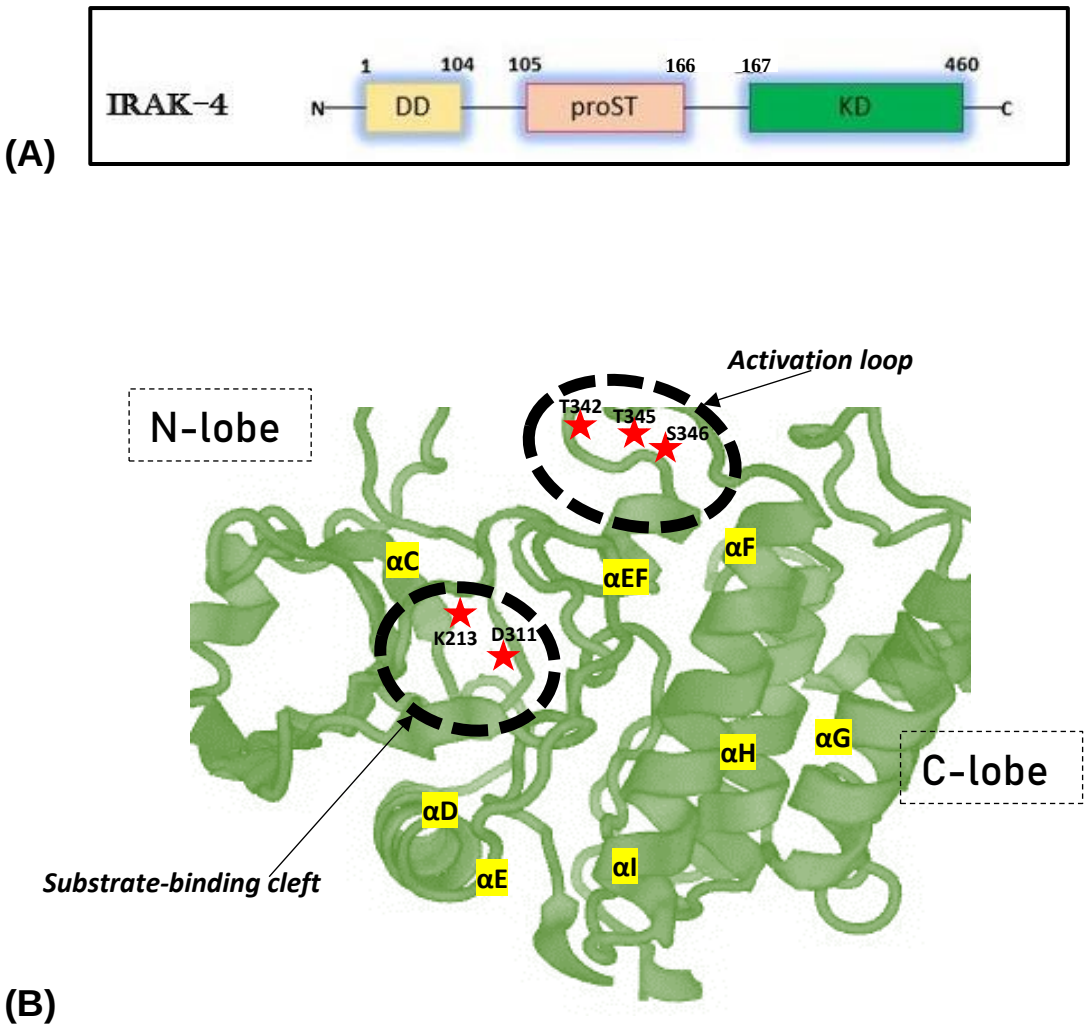


Figure 6. IRAK-4 domains and the solved structure for the kinase domain. (A) The domain architecture of the IRAK-4 protein consists of three domains. These include a death domain (DD), a proST domain and a kinase domain (KD). **(B)** The ribbon diagram illustrates the crystal structure of the kinase domain (KD). Alpha helices are annotated with yellow tags (αC to αI) while red stars denote critical residues. Adapted from figure 1 of (Balka & De Nardo, 2019) and figure 2 of (Z. Wang et al., 2009)

1.9 The importance of the IRAK-4 protein

IRAK-4 is one of the many proteins that mediate TLR4 signalling pathways in innate immunity (Saikh, 2021). However, its role as a critical protein in the TLR4 signalling cascades was shown in a study published in 2002. In that study, IRAK-4 knockout mice were carefully monitored after being injected with high lethal doses of LPS (Suzuki et al., 2002). The results showed that all the wild-type mice succumbed to death within 24 hours of exposure to the high lethal LPS doses, while all the IRAK-4 knockout mice survived (Suzuki et al., 2002). The cause of death in the wild-type mice was septic shock. Elevated levels of the IL-1 and IL 6 pro-inflammatory cytokines detected in the sera of these mice were further confirmation of septic shock (Suzuki et al., 2002). In contrast, the IRAK-4 knockout mice showed no change in their cytokine levels during the study. Furthermore, IRAK-4 knockout experiments using B-lymphocytes and bone-marrow-derived macrophages also showed no detectible responses to LPS induction (Suzuki et al., 2002). Therefore, all the results confirmed the indispensable role of IRAK-4 in inflammation.

IRAK-4 deficiency in humans also leads to a poor inflammation during an infection (Karananou, Alataki, & Papadopoulou-Alataki, 2020). One of the earliest studies in literature exemplified this observation through three unrelated children with an IRAK-4 deficiency (Picard et al., 2003). These children had recurring bacterial infections. The microorganisms that were the main cause of these infections were the gram-positive bacteria, *Streptococcus pneumoniae* and *Staphylococcus aureus* (Picard et al., 2003). Infection by these pyogenic bacteria in the three IRAK-4 deficient children was made worse by their poor inflammatory responses. Monocytes from these children were tested through LPS-induced signalling, and the results showed a diminished production of pro-inflammatory cytokines (Picard et al., 2003).

Since IRAK-4 deficiency leads to a poor inflammatory response, this then makes the IRAK-4 protein an ideal candidate that can be targeted for the design of anti-inflammatories. Molecules that inhibit IRAK-4 will result in a poor inflammatory response similar to that shown by IRAK-4 deficient humans and IRAK-4 knockout mice (Suzuki et al., 2002; Picard et al., 2003). There are at least four IRAK-4 drug inhibitors that are already in clinical trials towards treating rheumatoid arthritis, an autoimmune disease that causes inflammation (Wiese, Manning-Bennett, & Abuhelwa, 2020).

1.10 Supramolecular organizing centres - The Myddosome complex

Supramolecular organizing centres (SMOCs) are critical platforms that mediate cell signalling pathways in apoptosis and innate immunity (Kagan et al., 2014). The primary roles of SMOCs includes the localized concentrating of effector molecules to enhance and amplify signal transduction along activated signalling pathways (Sušjan-Leite, Ramuta, Boršić, Orehek, & Hafner-Bratkovič, 2022). SMOCs are formed through homotypic death domain (DD) interactions which are a universal feature in the formation of these platforms (Kagan et al., 2014). There are many different SMOC complexes that assemble due to the various DD protein-protein interactions. Examples of such complexes include the Myddosome complex, the PIDDosome complex, the FAS DISC complex and inflammasome complexes (Kagan et al., 2014). However, the focus of this dissertation will only be the Myddosome complex.

The Myddosome complex is an oligomeric platform that is assembled during post-receptor signalling in the MyD88 dependent signalling pathway (Gay, Gangloff, & O'Neill, 2011). It is formed through the interactions of the death domains between MyD88 and the IRAK proteins (Balka & De Nardo, 2019). The evidence for the existence of this complex was first published in 2009 (Motshwene et al., 2009). It revealed through *in vitro* studies that the MyD88 death domains interacted with the IRAK-4 death domains to form a heterogenous pool of Myddosome complexes. The dominant complex had a stoichiometry of 7:4 MyD88 to IRAK-4 death domain molecules (Motshwene et al., 2009). An important note however is that these Myddosome complexes are solely composed of death domains and not the full-length MyD88 and IRAK-4 proteins.

Another study was able to solve the crystal structure of a Myddosome variant as shown in Figure 7 (Lin et al., 2010). It was also composed of death domains only which assembled into a helical oligomeric structure. The complex showed a stoichiometry of six MyD88 molecules to four IRAK-4 molecules and four IRAK-2 molecules. Admittedly, the researchers that solved this structure state that they failed to produce IRAK-1 hence they used IRAK-2 when reconstituting their Myddosome variant complex (Lin et al., 2010).

Solving the Myddosome structure is a commendable milestone. However, the interactions of MyD88 and IRAK proteins *in vivo* involve the use of full-length proteins and not individual death domains. Therefore, it is not yet known if full-length MyD88 and IRAK proteins will form oligomeric complexes with unusual stoichiometries *in vivo*.

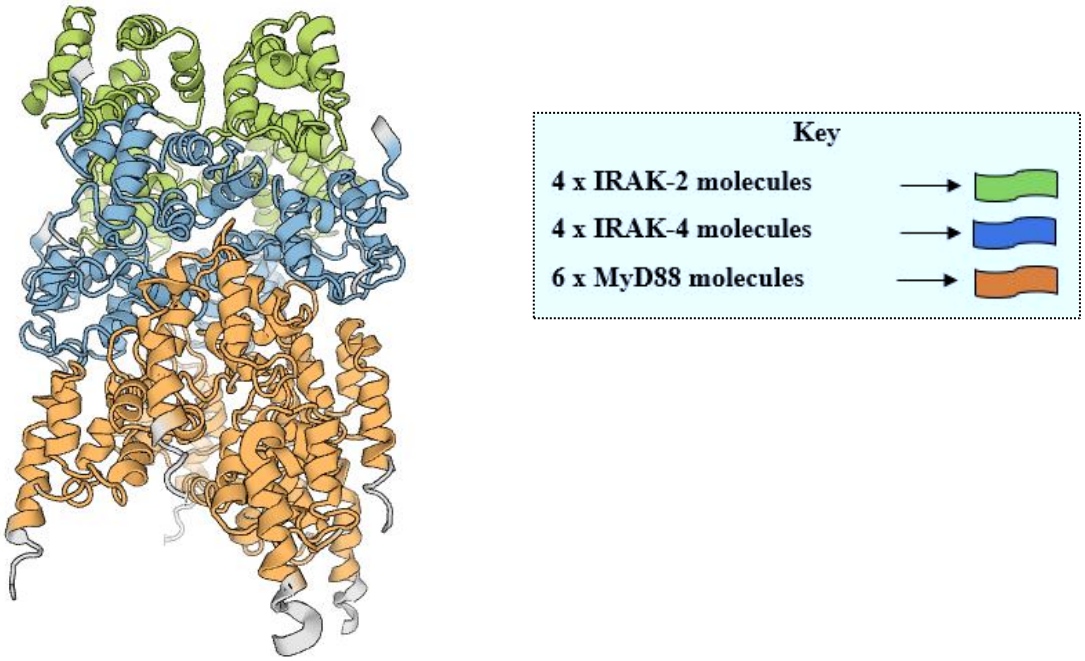


Figure 7. Myddosome complex composed of MyD88, IRAK-4 and IRAK-2 death domains. Crystal structure of a variant Myddosome complex solved by x-ray diffraction (3.4 Ångstrom) and represented as a ribbon diagram. The proteins interact through homotypic death domain interactions to assemble as follows: bottom layer of 6 MyD88 molecules (orange ribbons), middle layer of 4 IRAK-4 molecules (blue ribbons) and top layer of 4 IRAK-2 molecules (green ribbons). Adapted from figure 1 of (Lin et al., 2010)

1.11 Aim and problem identification

The intermediate domain (ID) is one of the three domains of the MyD88 adaptor protein (Moncrieffe et al., 2020). This domain has been shown to be essential for IRAK-4 binding to MyD88 (Janssens et al., 2003). A splice variant of MyD88 that lacks the intermediate domain, known as MyD88s fails to bind IRAK-4, a known binding partner of MyD88 (Janssens et al., 2003).

As mentioned earlier (Section 1.10), the Myddosome is an oligomeric complex formed *in vitro* through the interactions of death domains of MyD88 and IRAK-4. The first study that was published on the Myddosome, used a construct that expressed both the death domain and the intermediate domain of MyD88 (Motshwene et al., 2009). IRAK-4 death domain was then mixed with MyD88 which has both the death domain and the intermediate domain to reconstitute the Myddosome. A year later, another study was published where the crystal structure of a reconstituted Myddosome variant was solved (Lin et al., 2010). To our surprise, the MyD88 protein used in that study lacked the intermediate domain only possessing the death domain (Lin et al., 2010). Based on currently available literature, it is known that IRAK-4 cannot interact with MyD88 that lacks an intermediate domain (Lv et al., 2019). The big question is how did the Myddosome variant published in the 2010 study form without an intermediate domain? Formation of such a variant is contrary to available literature. Therefore, we designed a project that was aimed at investigating if the Myddosome complex forms in the absence of the MyD88 intermediate domain.

1.12 Hypothesis

The intermediate domain within MyD88 plays a pivotal role in Myddosome complex assembly and therefore, the intermediate domain deficient MyD88 death domain does not form the Myddosome complex.

1.13 Objectives

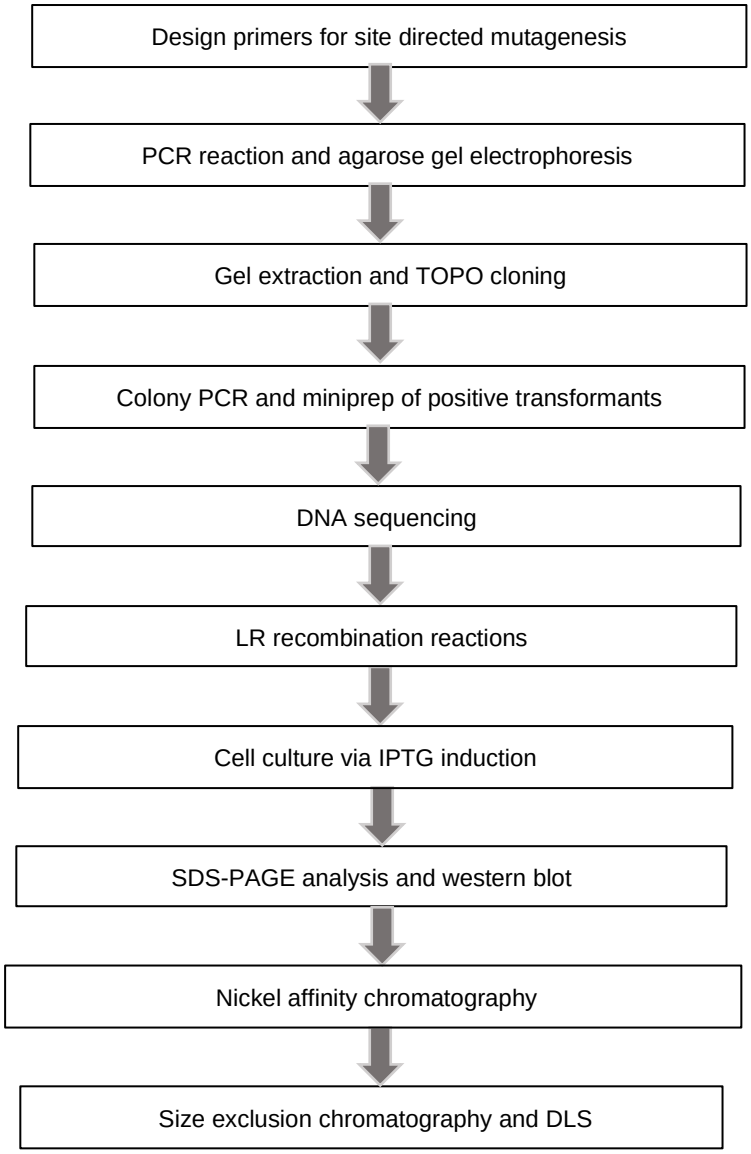
1) Express and purify ID-inclusive MyD88 death domain (MyD88-DD-ID)
2) Clone, express, and purify ID-deficient MyD88 death domain (MyD88-DD)
3) Express and purify IRAK-4 death domain (IRAK-4-DD)
4) Reconstitute the ID-inclusive and ID-deficient Myddosome complexes
5) Characterize and verify the protein-protein interactions by size exclusion chromatography

Chapter 2: Materials and methods

2.1 Introduction

This chapter outlines the materials and methods used in this project. Experiments were conducted in a biosafety level 1 (BSL-1) laboratory with adherence to set SOP and GLP standards. The structure of this chapter (excluding the introduction) is divided into six main sections. The first section shows a basic flow diagram which is a summary of all the experiments carried out. The second section gives details on how site-directed mutagenesis was carried out. The third section gives details on the background of gateway cloning. The fourth section gives details on the gateway cloning that was conducted in this project. The fifth section gives details on the protein expression, purification, and analysis in this project. The sixth section shows the table that contains the full set of instruments used in this project.

2.2 Flow diagram



2.3 Site directed mutagenesis (SDM)

2.3.1 Designing SDM primers

Table 2. Site directed mutagenesis (SDM) primers

Domain region	**Death domain** — **Intermediate domain**
Position	104 105 106 107 108 109 110 **111** 112 113 114 115 116 117 118
Amino acid sequence	E L G P S I E **E** D C Q K Y I L
MyD88 gene sequence	5' GAG CTG GGA CCC AGC ATT GAG **GAG** GAT TGC CAA AAG TAT ATC TTG 3'
	[The codon for the first amino acid of the intermediate domain: **GAG***]*
Position	104 105 106 107 108 109 110 111 112 113 114 115 116 117 118
Amino acid sequence	E L G P S I E [Stop] D C Q K Y I L
SDM forward primer	**5' GAG CTG GGA CCC AGC ATT GAG TAA GAT TGC CAA AAG TAT ATC TTG 3'**
	[The stop codon mutation for the first amino acid of the intermediate domain: **TAA**]
Position	104 105 106 107 108 109 110 111 112 113 114 115 116 117 118
SDM forward primer	5' GAG CTG GGA CCC AGC ATT GAG TAA GAT TGC CAA AAG TAT ATC TTG 3'
Reverse sequence	3' GTT CTA TAT GAA AAC CGT TAG AAT GAG TTA CGA CCC AGG GTC GAG 5'
SDM reverse compliment primer	**5' CAA GAT ATA CTT TTG GCA ATC TTA CTC AAT GCT GGG TCC CAG CTC 3'**

2.3.2 Methods for SDM

To conduct this study, we needed to have a construct that would produce an ID-deficient MyD88-DD protein (MyD88-DD). To achieve this, we decided to use site-directed mutagenesis on an MyD88-DD-ID-pETG-30A expression clone that already existed in our laboratory. It had previously shown good expression levels. Our strategy was to use SDM to mutate the nucleotide sequence for the first amino acid of the MyD88 intermediate domain into a stop codon thereby producing an ID-deficient fusion protein. The first amino acid of the intermediate domain is glutamic acid 111 (GAG) which was mutated to a stop codon (TAA). The "TAA" mutation was flanked by 21 bases upstream (which spanned only the last 21 residues of the MyD88 death domain gene) and 21 bases downstream (which spanned the first 21 residues of the MyD88 intermediate domain gene.) Therefore, the forward primer was composed of 45 nucleotide bases as shown in Table 2. The reverse primer was generated using the reverse compliment sequence of the forward primer. Both sets of primers were synthesized by Inqaba Biotec. A PCR reaction for the mutagenesis was set up as shown in Table 3 and the corresponding cycling conditions in table 4.

Table 3. PCR composition for SDM (50 µL reaction)

	Volume (µL)
Phusion™ High-Fidelity DNA polymerase (2 Units/µL)	0.5
5X Phusion™ HF Buffer (1.5 mM $MgCl_2$)	10
dNTPs (10 mM)	1
SDM forward primer (10 µM)	2.5
SDM reverse primer (10 µM)	2.5
MyD88-DD-ID-pETG-30A plasmid (50 ng/µL)	1
PCR grade ddH_2O	32.5
Total	**50**

Table 4. PCR cycling conditions for SDM (50 µL reaction)

	Temperature (°C)	Duration	Cycles
Initial denaturation	98	30 sec	1
Denaturation	98	10 sec	25
Annealing	55	30 sec	
Extension	72	4 min	
Final extension	72	10 min	1
Hold	4	∞	1

2.3.3 DpnI restriction enzyme digest

To isolate the SDM mutant from the wild type template MyD88-DD-ID in solution, the PCR product from the SDM was digested with DpnI to degrade the methylated template plasmid. The DpnI reaction was set-up as shown in table 5. The DpnI digestion was incubated at 37° C for 1 hour. Thereafter, the sample was prepared for DNA sequencing to verify the presence of the mutation.

Table 5. DpnI restriction enzyme digest composition (50 µL reaction)

	Volume (µL)
PCR product	44
DpnI (20 Units/µL)	1
*1X CutSmart™ Buffer**	5
Total	**50**

** 50 mM Potassium acetate, 20 mM Tris-acetate, 10 mM Magnesium acetate*

2.3.4 DNA sequencing

2.3.4.1 Materials for DNA sequencing preparation

Table 6. Kit for DNA sequencing preparation

Kit name	Kit supplier
Big Dye v3.1 sequencing kit	Thermo Fisher Scientific

2.3.4.2 Method – DNA sequencing

Plasmids were cleaned up using the Big Dye V3 sequencing kit in preparation for DNA sequencing. A PCR reaction was setup following the kit manufacturers protocol. The PCR composition and cycling conditions are shown in Table 7 and 8. Afterwards, the PCR product was mixed with 10 µL of PCR grade ddH_2O and 80 µL of the Big Dye solution A (ethanol (62.5 %) (v/v), sodium acetate (3M) pH 4.6). The mixture was centrifuged at 14 000 x *g* for 30 minutes. Next, the supernatant was discarded, before a wash step was conducted by resuspending the pellet in 250 µL of ice-cold ethanol (70 %). The suspension was centrifuged 14 000 x *g* for 5 minutes and this wash step was repeated two more times. After the final wash step, the supernatant was discarded. The pellet was air dried in a lamina flow cabinet for 30 minutes to allow evaporation of residual ethanol before sending it for Sanger sequencing. The raw sequencing data output was analysed using the SnapGene® viewer program (version 5.3.1) and the confirmation of the mutagenesis was further verified using the Expasy translation software.

Table 7. Big Dye PCR composition (10 µL reaction)

	Volume (µL)
Big Dye™ terminator 3.1 ready reaction mix	2
5X sequencing buffer	2
**MyD88-DD forward primer (10 mM)*	1
SDM plasmid DNA (55 ng/µL)	5
Total	**10**

** Sequence for the MyD88 forward primer is shown in table 9*

Table 8. Big Dye PCR cycling conditions (10 µL reaction)

	Temperature (°C)	Duration	Cycles
Initial denaturation	96	1 min	1
Denaturation	96	10 sec	25
Annealing	50	5 sec	
Extension	60	4 min	
Final extension	72	4 sec	1
Hold	4	∞	1

2.4 Background: Gateway system

The gateway system is an innovative cloning method that by-passes the need for restriction enzymes and ligases for typical cloning in biological experiments (Esposito et al., 2009). This unique system generates constructs using recombination technology. There is a high level of specificity in gateway cloning which makes this technique very favorable for experimental consideration (Reece-Hoyes & Walhout, 2018). The gateway system utilizes conserved directional attachment sites (*att* sites) to transfer gene segments, that are flanked by these *att* sites, between different vectors (Reece-Hoyes & Walhout, 2018). In essence, there are two main steps when using the gateway system. Step one is to clone the gene of interest into a vector to produce an entry clone as shown in Figure 8. This cloning process requires the critical inclusion of a "CACC" sequence in the beginning of the forward primer during amplification of the gene of interest. This sequence is important because it serves as the recognition and critical point of attachment for the "GTGG" overhang sequence that is present on the entry vector as shown in Figure 8. Subsequently, the amplified gene of interest (which contains the CACC sequence) is ligated into the entry vector by topoisomerases.

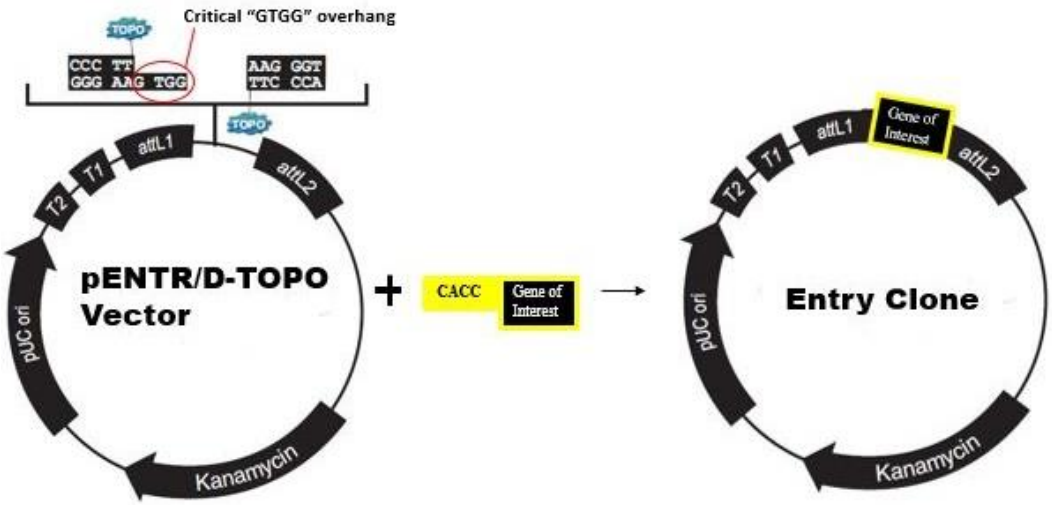


Figure 8. Basic outline for generating an entry clone. pENTR/D-TOPO plasmid forms the vector backbone for the entry clone. The gene of interest is amplified with the inclusion of the critical "CACC" sequence to anneal to the "GTGG" overhang sequence. Ligation of the gene of interest into the vector is carried out by topoisomerase enzymes to produce the entry clone.

The second step in gateway cloning is to generate an expression clone as shown in Figure 9. This requires a verified entry clone with the gene of interest flanked by *att*L1/*att*L2 sites and a destination vector that contains *att*R1/*att*R2 sites. These sites allow a recombination reaction whereby DNA segments for these *att* sites are swapped between the entry clone and the destination vector. The reaction is known as LR recombination and is carried out by the LR Clonase enzyme to produce an expression clone.

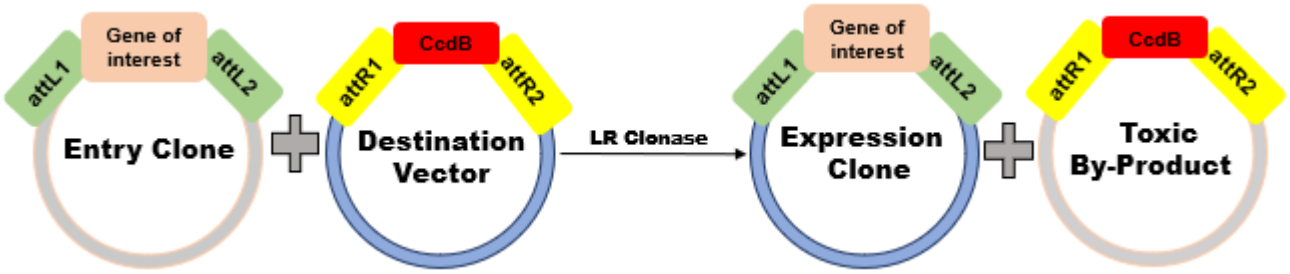


Figure 9. LR recombination reaction. The LR recombination reaction swaps the attL and attR flanked DNA fragments between the entry clone and the destination vector using the LR Clonase enzyme. The two end products are an expression clone with the gene of interest and a toxic by-product.

One of the key features of the gateway system, is that the entry clone can be used for expression in a variety of host expression systems as shown in Figure 10. This is due to the highly specific *att*L sites on an entry clone and the *att*R sites available on all gateway destination vectors. Therefore, LR recombination reactions can be conducted with a single entry clone (containing the gene of interest) to generate different expression clones utilizing any destination vector. As such, a diverse variety of expression clones can be produced quickly without having to individually re-clone, verify and sequence a new entry clone for each expression clone. Additionally, researchers can easily pivot their studies to different expression host systems using the same entry clone. For example, studies that generate expression clones for optimal expression in bacteria, can directly use the initial entry clone and conduct LR recombination reactions to produce constructs for optimal expression in plants.

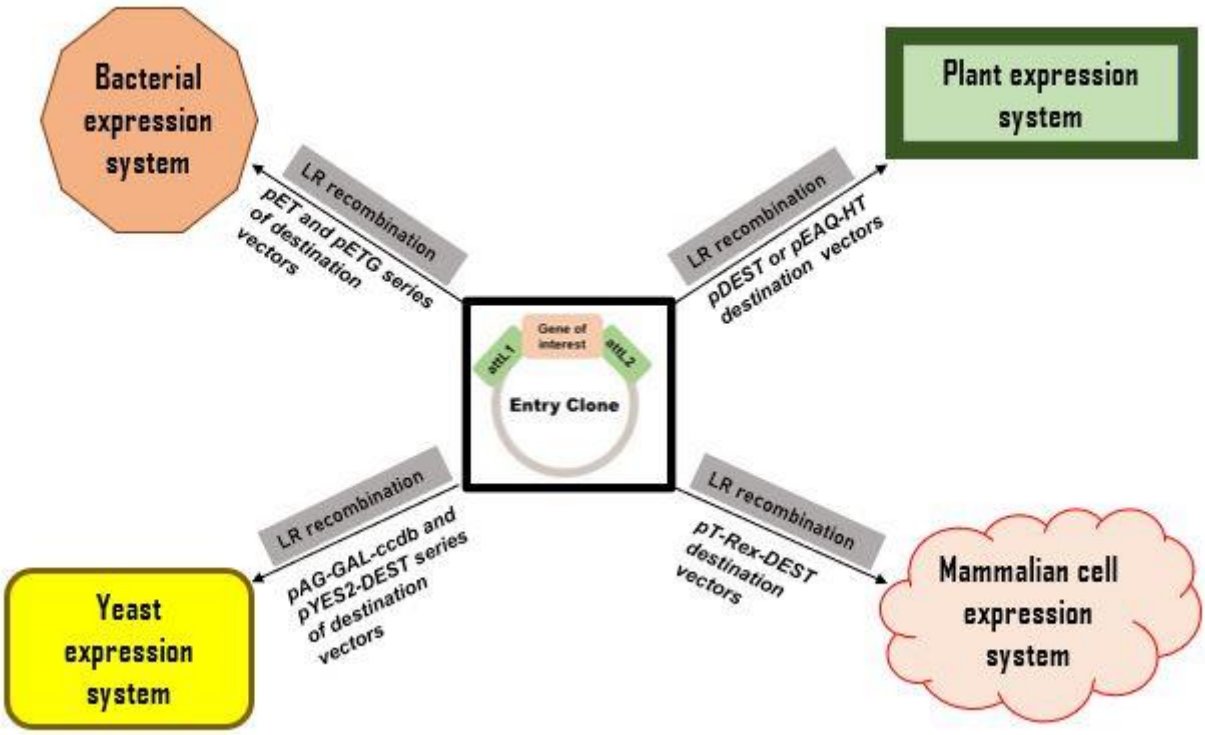


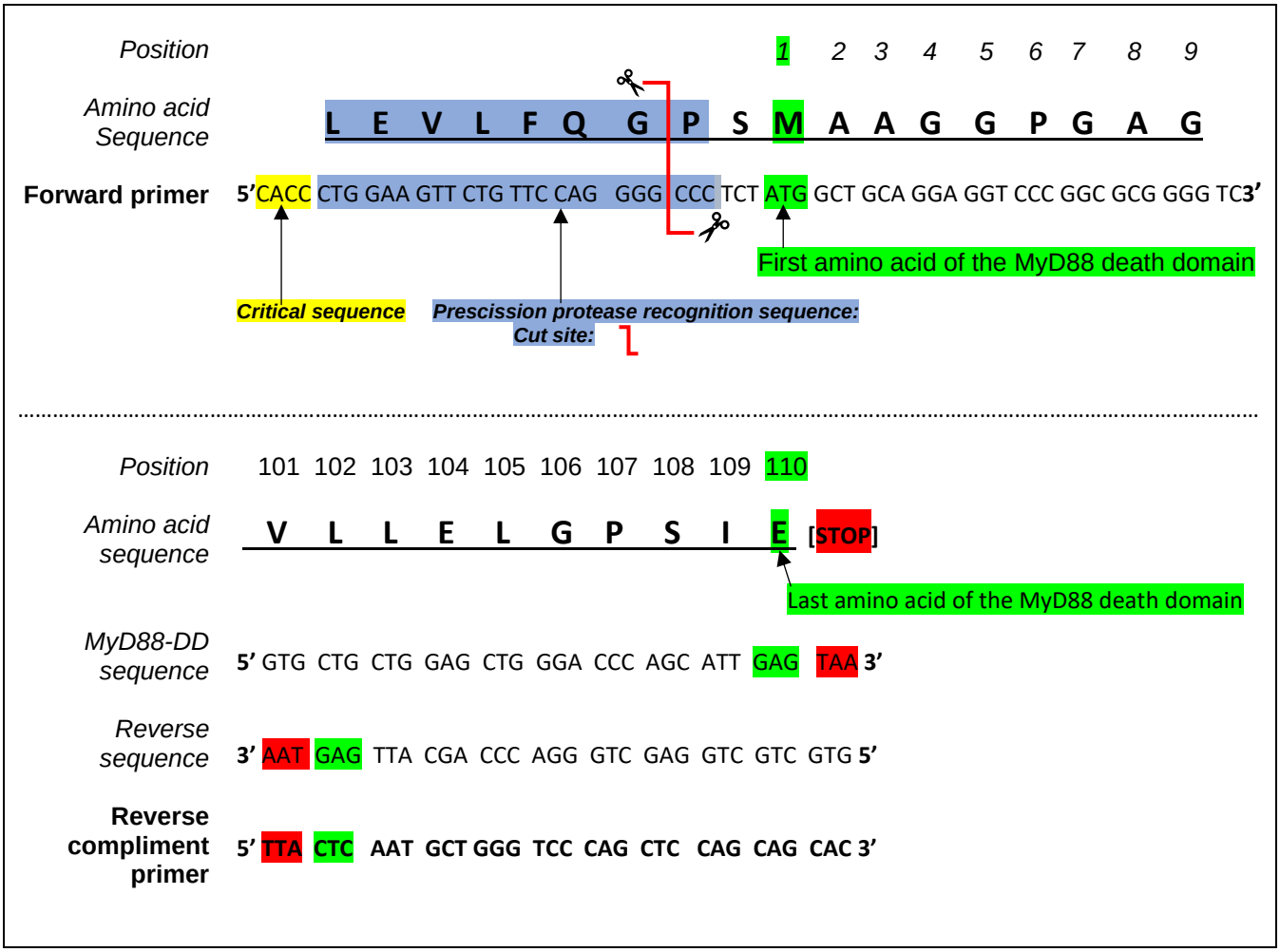
Figure 10. Versatile applications of an entry clone across different host expression systems. The same entry clone containing the gene of interest can be used in LR recombination reactions to produce constructs for optimal expression in different hosts like bacteria, plant, yeast and mammalian cells. A few examples of optimal destination vectors include pETG vectors for a bacterial expression system, pDEST vectors for a plant expression system, pYES2-DEST vectors for a yeast expression system and pT-Rex-DEST vectors for a mammalian cell expression system.

2.5 Cloning of MyD88 death domain (MyD88-DD)

2.5.1 Primer design

The first step in gateway cloning is to generate an entry clone that contains the gene of interest. We wanted to create an entry clone that contained the ID-deficient MyD88-DD gene. Therefore, we designed forward and reverse primers for amplifying the MyD88 death domain gene (amino acids 1-110) only as shown in Table 9. We included the critical "CACC" sequence at the beginning of the forward primer, followed by a Prescission protease recognition sequence for downstream removal of fusion tags. After the Prescission protease recognition sequence, we added a serine to allow for flexibility between the protease site and MyD88-DD. This was done to ensure that the Prescission protease has easy access to the cleavage site. The primers were synthesized by Inqaba Biotec.

Table 9. MyD88-DD primers



2.5.2. MyD88-DD gene amplification

To amplify the MyD88 death domain gene fragment (amino acids 1-110), we set up a 50 µL PCR reaction using our MyD88-DD primers as shown in table 10. The MyD88-DD-ID-pETG-30A plasmid was used as the template DNA. The cycling conditions of the PCR reaction are shown in Table 11. Upon completion of the PCR, the reaction mix was analysed by agarose gel electrophoresis.

Table 10. Kapa Taq PCR reaction composition (50 µL reaction)

	Volume (µL)
*Kapa Taq ready mix**	25
MyD88 forward primer (10 µM)	1
MyD88 reverse primer (10 µM)	1
MyD88-DD-ID-pETG-30A plasmid (50 ng/µL)	2
PCR grade ddH_2O	21
Total	**50**

**Kapa Taq ready mix*: dNTPs (0.2 mM each), Taq polymerase (0.02 Units/µL), *Kapa Taq* buffer, $MgCl_2$ (1.5 mM)

Table 11. PCR cycling conditions (50 µL reaction)

	Temperature (°C)	**Duration (Sec)**	**Cycles**
Initial denaturation	95	60	1
Denaturation	95	30	30
Annealing	54.1	30	
Extension	72	60	
Final extension	72	60	1
Hold	4	∞	1

2.5.3 Agarose gel electrophoresis

2.5.3.1 Materials

Table 12. Buffer for agarose gel electrophoresis

Application	Name	Composition
Agarose gel electrophoresis	1 x TAE buffer	40 mM Tris (pH 8.0), 20 mM Acetic acid and 1 mM EDTA

2.5.3.2 Methods

A 1 % agarose gel was prepared by dissolving 0.5 grams of agarose powder in 50 mL of a 1 X TAE buffer (Tris-base (40 mM), acetic acid (20 mM) and EDTA (1 mM)) in a microwave oven for 1 minute. Thereafter, 3.0 µL of SYBR Safe DNA gel stain was added to the dissolved hot agar. A 10-well comb was positioned in a casting tray (10 cm x 15 cm) before the hot agar was poured into the casting tray. The hot agar was left to cool and solidify at room temperature before removing the comb. A volume of 15 µL of PCR product was mixed with 5 µL of DNA gel loading dye before loading the resulting 20 µL mixture into the set agarose gel wells. The first well was loaded with 5 µL of the DNA ladder (NEB 10 kb). The gel was electrophoresed at 100 volts for 1 hour using 1 X TAE as a buffer. Thereafter, the gel was visualized using the Molecular Imager Gel Doc XR+ Imaging system.

2.5.4 Gel extraction

2.5.4.1 Materials

Table 13. Kit for gel extraction

Kit name	Kit supplier
GeneJET gel extraction kit	Thermo Fisher Scientific

2.5.4.2 Method - Gel extraction of amplified MyD88-DD gene

A band corresponding to the molecular weight (0.3 kb) of MyD88-DD was manually excised from the DNA gel using a sterile blade. The mass of the excised gel slice was recorded for use in the GeneJET gel extraction kit (Table 13) protocol as follows:

A ratio of 1:1 (w/v) of binding buffer was added to the excised gel slice and the sample was heated for 10 minutes at 60° C. During the heating process, the sample was mixed by inversion until the gel was completely dissolved, yielding a homogenous yellow solution. Up to 800 µL of the yellow gel mix solution was loaded onto a GeneJET purification column and centrifuged at 12 000 x *g* for 1 minute. The supernatant was discarded, and a wash step proceeded by adding 700 µL of wash buffer to the purification column before centrifuging at 12 000 x *g* for 1 minute. The supernatant was discarded once again and the purification column was centrifuged once again to remove residual wash buffer. The DNA fragment was eluted from the purification column with 30 µL of PCR grade ddH_2O into a new clean 1.5 mL Eppendorf tube by centrifuging at 12 000 x *g* for 1 minute.

2.5.5 TOPO cloning

The pENTR/D-TOPO plasmid was mixed with the gel purified MyD88-DD gene as shown in Table 14 to generate an entry clone. The mixture was incubated at 25° C for 30 minutes and thereafter transformed into DH5α (*E. coli*) cells.

Table 14. TOPO cloning reaction

	Volume (µL)
Fresh gel extracted PCR product (37 ng/µL of DNA)	4
Salt solution (1.2 M NaCl, 0.06 M $MgCl_2$)	1
*TOPO® vector**	1
Total	**6**

* 15-20 ng/µL linearized plasmid DNA in: 50 mM Tris-HCl, pH 7.4, 50% glycerol, 1 mM EDTA, 2 mM DTT, 0.1% Triton X-100, 100 µg/mL BSA

2.5.6 Cell transformation for MyD88-DD-pENTR/D entry clone

2.5.6.1 Materials

Table 15. Bacterial growth media

Application	Name	Composition
***Growth media**	Luria-Bertani (LB) broth	Tryptone 1 % (w/v), yeast extract 0.5 % (w/v), sodium chloride 0.5 % (w/v)
	LB agar (Solid media)	Tryptone 1 % (w/v), yeast extract 0.5 % (w/v), sodium chloride 0.5 % (w/v), agarose 1.5 % (w/v)

**All media was sterilized by autoclaving prior to use*

2.5.6.2 Method - Cell transformation for MyD88-DD-pENTR/D entry clone

The solid growth media was prepared by sterilizing LB agar (tryptone 1 % (w/v), yeast extract 0.5 % (w/v), sodium chloride 0.5 % (w/v), agarose 1.5 % (w/v)) in an autoclave. Afterwards, the hot LB agar broth was allowed to cool down (but not solidify) for 15 minutes before adding the kanamycin antibiotic at a working concentration of 50 µg/mL. Approximately 20 mL of the lukewarm LB agar was subsequently poured into a petri dish (90 mm x 15 mm) and left to completely solidify.

For cell transformation, 3 µL of plasmid DNA was added to 50 µL of ice cold DH5α chemi-competent cells. The transformation mixture was incubated on ice for 15 minutes before heat shocking at 42° C for 1 minute followed by brief incubation on ice for a further 2 minutes. A 200 µL volume of LB media was added to the chilled cell transformation mixture and subsequently incubated for 80 minutes at 37° C on a shaking incubator. A minimum of 50 µL of the transformation culture was spread using glass beads on the solid LB agar media. The LB agar plates were then incubated at 37° C overnight for colonies to grow.

2.5.7 Colony PCR for MyD88-DD-pENTR/D entry clone

Colonies observed on the overnight LB agar plates were analysed by colony PCR to verify the presence of the MyD88-DD gene. A master mix for the PCR reaction was prepared as shown in Table 16. A colony was scraped off the LB agar plate using a pipette tip. The pipette tip containing the scraped off colony was then stirred into a 25 µL aliquot of the PCR master mix. Thereafter, the same pipette tip was used to inoculate 15 mL of sterilized LB broth that contained kanamycin at a final concentration of 50 µg/mL. The 25 µL aliquots (that were stirred with a scraped colony) were placed in a thermal cycler to conduct the PCR reaction using the cycling conditions shown in Table 17. Once the PCR was completed, the PCR products were analysed by agarose gel electrophoresis on a 1 % gel as described in Section 2.5.3. Upon verification of the MyD88-DD gene (0.3 kb) on the agarose gel, we then proceeded to grow the 15 mL LB broth samples by incubating them overnight at 37° C on a shaking incubator.

Table 16. Kapa Taq colony PCR master mix (12 x 25 µL reactions)

	Volume (µL)
Kapa Taq ready mix*	150
MyD88 forward primer (10 µM)	6
MyD88 reverse primer (10 µM)	6
PCR grade ddH_2O	138
Total	**300**

**Kapa Taq ready* mix: dNTPs (0.2 mM each), Taq polymerase (0.02 Units/µL), Kapa Taq buffer, $MgCl_2$ (1.5 mM)

Table 17. Colony PCR cycling conditions (25 µL reaction)

	Temperature (°C)	Duration (sec)	Cycles
Initial denaturation	95	60	1
Denaturation	95	30	30
Annealing	54.1	30	
Extension	72	30	
Final extension	72	60	1
Hold	4	∞	1

2.5.8 Plasmid extraction of MyD88-DD-pENTR/D entry clone

Plasmid extraction was conducted using the GeneJET plasmid miniprep kit. The protocol used was as follows:

15 mL overnight cell cultures were centrifuged at 3 000 x *g* for 10 minutes at 4° C. The supernatant was discarded and the pellet was resuspended in 250 µL of chilled resuspension solution (containing RNase A) and transferred to a clean 1.5 mL eppendorf tube. Then 250 µL of lysis solution was added to the cell suspension with gentle inversion of no more than 5 times. Afterwards, 350 µL of neutralization solution was added and mixed by inversion. The cell suspension was centrifuged at 13 000 x *g* for 5 minutes to pellet the cell debris. The supernatant was gently transferred (without disturbing the white precipitate) to the GeneJET spin column and centrifuged at 13 000 x *g* for 1 minute. The flowthrough was discarded and thereafter 500 µL of wash solution was added to the column. The column was centrifuged again at 13 000 x *g* for 1 minute. After the centrifugation step, the flowthrough was discarded again. The wash step was repeated once more and thereafter the empty column was centrifuged to remove residual wash solution. The column was then transferred to a new eppendorf tube. 30 µL of PCR grade ddH_2O was added to the column followed by a 2-minute incubation period at 25° C before centrifugation at 13 000 x *g* for 2 minutes to elute the MyD88-DD-pENTR/D entry clone plasmid DNA.

2.5.9 M13 primer analysis for orientation of the insert in MyD88-DD-pENTR/D entry clone

A PCR reaction was conducted using the extracted MyD88-DD-pENTR/D entry clone (as template) to verify if the MyD88-DD gene (0.3 kb) was inserted in the correct orientation in the plasmid. The PCR composition and cycling conditions that were used are described in section 2.5.7 using our MyD88 forward primer and an M13 reverse primer (provided in the TOPO cloning kit). The M13 reverse primer is expected to extend the original length of the amplified gene of interest by 0.2 kb if the insert is in the correct orientation. The PCR product was analysed by agarose gel electrophoresis on a 1 % gel. Subsequently, plasmids which showed an amplified band of 0.5 kb were cleaned and sent for DNA sequencing as described in (Section 2.3.4).

2.5.10 LR recombination for MyD88-DD-pENTR/D entry clone

Once we verified the presence of the MyD88-DD gene in the entry clone through DNA sequencing, the next step was to generate expression clones using gateway destination vectors. Expression clones are generated through LR recombination reactions. Therefore, were performed LR recombination of the MyD88-DD entry clone into the four destination vectors shown in figure 11. These are pETG-10A which expresses a His-tagged fusion protein, pETG-20A which expresses a His-Trx fusion protein, pETG-30A which expresses a His-GST fusion protein and lastly pETG-41A which expresses a His-MBP fusion protein. The pETG-10A, pETG-20A, pETG-30a and pETG-41A vectors were propagated in DB3.1 cells prior to setting up LR recombination reactions. Only DB3.1 cells are used to propagate these vectors

because they have a resistance gene against the toxin produced by the ccdb gene (McCardell, Pandey, & Murray, 2019).

LR recombination reactions were set up as shown in Table 18, using the four destination vectors already mentioned. The reactions were incubated at 25° C for 2 hours. Thereafter, 1 µL of Proteinase K solution was added and the whole reaction mix was incubated at 37° C for 15 minutes. The addition of Proteinase K was to terminate the LR recombination reaction.

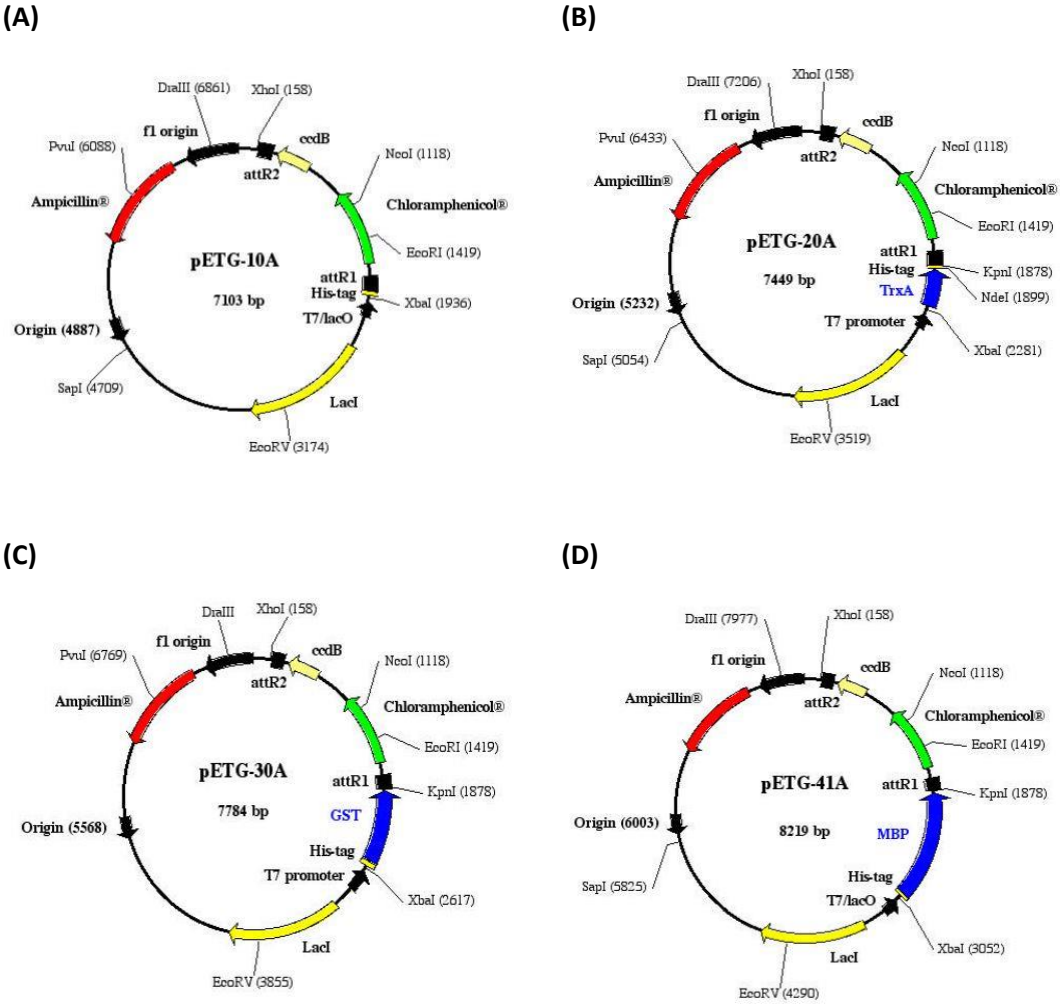


Figure 11. Gateway destination vector maps. The destination vector maps sourced from the European Molecular Biology Laboratory (EMBL) Heidelberg repository **(A)** pETG-10A plasmid with an N-terminal Histidine (His) tag. **(B)** pETG-20A plasmid with an N-terminal Histidine-thioredoxin tag (His-Trx) tag. **(C)** pETG-30A plasmid with an N-terminal Histidine-Glutathione S-transferase (His-GST) tag. **(D)** pETG-41A plasmid with an N-terminal Histidine-Maltose binding protein (His-MBP) tag.

Table 18. LR recombination reaction

	Volume (µL)
MyD88-DD-pENTR/D entry clone (50 ng/µL)	3
**Destination vector (50 ng/µL)*	3
TE buffer (Tris 1mM, EDTA 10mM)	2
LR Clonase enzyme	2
Total	**10**

**Available destination vectors: pETG-10A, pETG-20A, pETG-30A, pETG-41A*

2.6 Protein expression, purification and analysis of death domains

2.6.1 Small scale culturing – Expression trials

To conduct expression trials, verified plasmids were first transformed into BL21 (DE3) cells and streaked on LB agar plates that contained 100 µg/mL of ampicillin and incubated at 37° C overnight. Colonies were picked and inoculated in 5 mL of LB broth that contained ampicillin. The inocula were then placed in a shaking incubator (140 rpm) at 37° C. The optical density (OD_{600}) of these shaking cultures was monitored regularly until it was between 0.8 to 1.0. Thereafter, IPTG was added to a final concentration of 1 mM. The cultures were then left to grow overnight at 37° C while shaking.

2.6.2 Protein extraction of small scale expression trial cultures.

Protein extraction from the expression trial cultures was conducted using BugBuster® protein extraction reagent (B-PER) using the following protocol:

1 mL of culture was transferred to a 1.5 mL eppendorf tube and centrifuged at 13 400 x *g* for 1 minute. The supernatant was discarded and the pellet was resuspended in 50 µL of B-PER, then mixed well by vortexing before incubation at 25° C for 10 minutes to lyse the cells. The lysate was centrifuged at 13 400 x *g* for 3 minutes. The supernatant (soluble fraction) was transferred to a clean eppendorf tube, and the pellet was resuspended in 50 µL of chilled 20 % (v/v) B-PER. The pellet suspension was vortexed before centrifugation at 13 400 x *g* for 3 minutes. The supernatant was discarded and the pellet (insoluble fraction) was resuspended in 50 µL of chilled 20 % (v/v) B-PER. Protein fractions were analysed by SDS-PAGE.

2.6.3 SDS-PAGE analysis of death domain fusion proteins

2.6.3.1 Materials

Table 19. Buffers for SDS-PAGE

Name	Composition
Destaining buffer	0.1 % (v/v) acetic acid, 0.4 % (v/v) ethanol
1 X SDS running buffer	192 mM glycine, 25 mM Tris, 0.1 % (w/v) SDS
Sample buffer	100 mM Tris (pH 6.8), 4 % (w/v) SDS, 200 mM β mercaptoethanol, 0.2 % (w/v) bromophenol blue, 20 % (v/v) glycerol
Staining buffer	0.1 % (v/v) acetic acid, 0.4 % (v/v) ethanol, 0.3 % (w/v) Coomassie brilliant blue G 250

2.6.3.2 Method – SDS-PAGE

SDS-PAGE was used to analyse protein samples. Protein samples with an expected molecular weight that was greater than 50 kDa were analysed on a 10 % resolving gel while those with an expected molecular weight of less than 50 kDa were analysed on a 12 % resolving gel. The protocol for making a 10 and 12 % separating gel is shown in Table 20 and Table 21. TEMED was added last before pouring the mixture into casting plates, leaving a 2.5 cm gap at the top which was filled with water to remove air bubbles. The separating gel was allowed to set for 15 minutes before discarding the water. The stacking gel was prepared as shown in Table 20 and Table 21. The mixture was poured into the 2.5 cm gap - on top of the resolving gel. Thereafter, a 10 well comb was inserted and the gel was allowed to solidify for 15 minutes. The comb was removed and protein samples were loaded onto the gel in a 4:1 ratio of protein to sample buffer. The loaded samples were then electrophoresed at 40 mA (constant current) for 1 hour in 1 X SDS running buffer (Table 19). After electrophoresis, the gel was stained with coomassie dye for 20 minutes on a shaker (Stovall Belly Dancer Shaker). The stain was discarded and a destaining buffer was used to destain the gel for 1 hour. Protein bands on the gel were imaged using the Molecular Imager Gel Doc XR+ Imaging system.

Table 20. Composition for a 10 % SDS polyacrylamide gel

	Resolving gel (mL)	Stacking gel (mL)
40 % (w/v) Bis-acrylamide	2.00	0.75
1.5 M Tris (pH 8.8)	2.00	n/a
0.5 M Tris (pH 6.8)	n/a	1.25
10 % (w/v) SDS	0.08	0.05
10 % (w/v) APS*	0.08	0.05
TEMED	0.008	0.005
ddH_2O	3.80	2.90
Total	**8.00**	**5.00**

Table 21. Composition for a 12 % SDS polyacrylamide gel

	Resolving gel (mL)	Stacking gel (mL)
40 % (w/v) Bis-acrylamide	2.40	0.75
1.5 M Tris (pH 8.8)	2.00	n/a
0.5 M Tris (pH 6.8)	n/a	1.25
10 % (w/v) SDS	0.08	0.05
10 % (w/v) APS	0.08	0.05
TEMED	0.008	0.005
ddH_2O	3.40	2.90
Total	**8.00**	**5.00**

2.6.4 Western blot of His-tagged death domain proteins

2.6.4.1 Materials

Table 22. Buffers/reagents for western blot

Category	Name	Supporting information
Buffer	Blocking buffer	5 % (w/v) skimmed milk, 1 x PBST buffer: 140 mM sodium chloride, 3 mM potassium chloride, 10 mM disodium phosphate, 2 mM dipotassium phosphate, 0.05 % (w/v) Tween-20
Buffer	Transfer buffer	50 mM Tris, 40 mM glycine, 20 % (v/v) methanol
Buffer	Wash buffer	140 mM sodium chloride, 3 mM potassium chloride, 10 mM disodium phosphate, 2 mM dipotassium phosphate, 0.05 % (w/v) Tween-20
Reagent	Clarity western ECL	Commercially purchased. Supplier: Bio Rad
Reagent	TMB enhanced one component HRP membrane	Commercially purchased. Supplier: Sigma-Aldrich

2.6.4.2 Methods – Western blot

Protein samples were first separated by SDS-PAGE (as described in Section 2.6.3); however, the gel was kept unstained after electrophoresis. Next, a PVDF membrane was incubated in 5 mL of 100 % methanol for activation for 3 minutes. Afterwards, two pieces of blotting paper were soaked in 30 mL of transfer buffer. The transfer stack was then assembled as shown in Figure 12. Subsequently, the assembled stack was compressed using a roller to remove air bubbles and excess buffer before the stack was placed in the Trans Blot cassette. The transfer was done at a current of 1.3 amps and at a voltage of 25 for 20 minutes. Afterwards, the PVDF membrane was incubated in blocking buffer for 1 hour at 4° C on a shaker. The blocking buffer was discarded and then the PVDF membrane was incubated in an anti-histidine antibody buffer (20 µg/µL HRP conjugated anti-His antibody, 1 x blocking buffer) for 1 hour at 4° C on a shaker. Then the PVDF membrane was washed 3 times in a wash buffer (1 x PBS, 0.05 % (w/v) Tween-20) at 25° C for 10 minutes before discarding the wash buffer. After the wash steps, 3 mL of Bio Rad Clarity western ECL substrate was poured on the PVDF membrane to initiate the chemiluminescence reaction. The PVDF membrane was then immediately visualized using a Molecular Imager Gel Doc XR+ Imaging system. Thereafter, another visualization technique that uses colorimetric analysis was used on the same PVDF membrane to confirm the chemiluminescence result. The PVDF membrane was then washed 3 times with distilled water followed by the addition of 2 mL of TMB

enhanced one component HRP membrane substrate to initiate the colorimetric reaction. The PVDF membrane was visualized again using the Molecular Imager Gel Doc XR+ Imaging system.

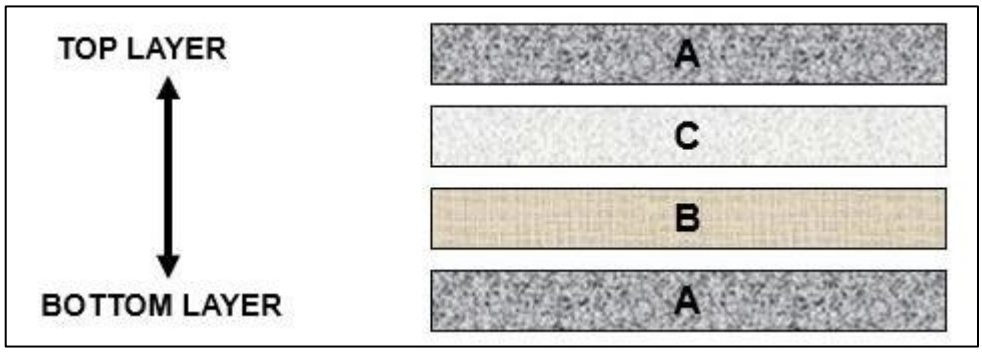


Figure 12. Western blot transfer stack. Assembly of the western blot stack for transference of proteins from the SDS polyacrylamide gel to the PVDF membrane. **A** – Blotting paper. **B** – PVDF membrane. **C** - Unstained SDS polyacrylamide gel containing separated proteins.

2.6.5 Large scale expression of death domain proteins

Colonies were first grown using BL21 (DE3) cells transformed with expression constructs using the protocol described in Section 2.5.6. Thereafter, the colonies were picked and inoculated in 20 mL of LB broth that contained 100 µg/mL of ampicillin. The inoculated growth media was incubated overnight at 37° C on a shaking incubator. Afterwards, the overnight culture was centrifuged at 4° C for 10 minutes at 2000 x g. The supernatant was then discarded and the pellet was resuspended in 4 mL of fresh LB broth before inoculating it in 4 L of LB broth. The inocula was then incubated at 37° C on a shaking incubator and the optical density (OD_{600}) was regularly checked until it reached 0.5. Once, the OD_{600} reached 0.5, the cultures were placed on ice until the OD_{600} gradually reached a range of 0.8 – 1.0 at which point IPTG was added to a final concentration of 1 mM. The cultures were then left to grow overnight at 15° C while shaking.

2.6.6 Protein extraction for large scale cultures

The overnight 4 L cultures were centrifuged at 5 213 x *g* for 15 minutes at 4° C. Afterwards, the supernatant was discarded and the pellet was resuspended in 80 mL of lysis buffer (100 mM Tris (pH 8.0), 200 mM NaCl, 0.5 % (v/v) Triton X-100). The cell suspension was sonicated while on ice for 10 minutes at 40 % amplitude with 30 second pulses alternating between on/off settings. Thereafter, the lysate was centrifuged at 37 500 x *g* for 45 minutes at 4° C and the pellet was discarded. The supernatant was collected (soluble fraction) and stored on ice for nickel affinity chromatography purification.

2.6.7 Nickel affinity chromatography purification of fusion proteins

2.6.7.1 Materials

Table 23. Column and buffers for nickel affinity chromatography

Category	Name	Supporting information
Column	HiTrap chelating HP column (5 mL) 	Commercially purchased. Supplier: Cytiva life sciences
Buffer	Buffer A	Composition: 20 mM Tris (pH 8.0), 30 mM NaCl, 10 mM Imidazole
Buffer	Buffer B	Composition:20 mM Tris (pH 8.0), 30 mM NaCl, 50 mM Imidazole
Buffer	Buffer C	Composition: 20 mM Tris (pH 8.0), 30 mM NaCl, 250 mM Imidazole

2.6.7.2 Method – Nickel affinity chromatography

A 5 mL HiTrap chelating HP column was cleaned by passing 15 mL of ddH_2O through it to remove the storage ethanol before using the buffers shown above in Table 23. It was then charged with 6 mL of 0.1 M nickel sulphate. A volume of 10 mL of buffer C was passed through the column to remove excess nickel sulphate. Thereafter, 10 mL of buffer A was passed through the column to equilibrate it. In order to prevent non-specific binding, buffer C was added to the unpurified protein sample to give a 10 mM concentration of imidazole. The sample was mixed by inversion and thereafter loaded onto the column using a 10 mL syringe. Afterwards, the first wash step was conducted by passing 10 mL of buffer A in a syringe through the column. The second wash step was conducted by passing 10 mL of buffer B in a syringe through the column. Once the second wash step was completed, 1 mL fractions were eluted from the column into clean 1.5 mL eppendorf tubes using buffer C that was in a syringe. The eluted fractions were then analysed by SDS-PAGE.

2.6.8 Bradford assay for MyD88 death domain fusion proteins

2.6.8.1 Materials

Table 24. Buffers/reagents for Bradford assay

Category	Name	Supporting information
Reagent	Bio Rad protein assay dye reagent concentrate (5X Bradford reagent)	Commercially purchased. Supplier: Bio Rad
Buffer	Diluent buffer	Composition: 20 mM Tris (pH 8.0), 30 mM NaCl

2.6.8.2 Method – Bradford assay

Preliminary detection of protein fractions after affinity purification was conducted in a clear 96 well clear Greiner plate by mixing 5 µL of protein fraction with 95 µL of 1X dilution of Bio Rad protein assay dye reagent concentrate (Table 24). The presence of protein was detected by the colour change of the reagent from brown to blue.

To determine protein concentration of nickel affinity chromatography purified fractions, bovine serum albumin (BSA) standards were prepared. A 2-fold dilution series from a 2 mg/mL BSA stock was set up to make 5 standards having concentrations of 2.0 mg/mL, 1.0 mg/mL, 0.5 mg/mL, 0.25 mg/mL, and 0.125 mg/mL. The purified protein fractions were diluted in the ratio of 1 to 10 using the diluent buffer (20 mM Tris (pH 8.0), 30 mM NaCl). The Bradford assay was set up in a 96 well clear Greiner microplate and each well was loaded as shown in Table 25. The assay was done in triplicate with the diluent buffer used as the blank. The loaded microplate was analysed using the SpectraMax® Paradigm® multi-mode detection platform. The samples in the microplate were analysed using the SoftMax Pro 6.4 software package with the detection wavelength set at 595 nm. Absorbance readings were exported to Microsoft Excel to generate the line of best fit from the mean averages of the raw data. Protein concentrations were then calculated, taking into account the dilution factor.

Table 25. Bradford assay composition

	Volume (µL)
BSA standard/protein fraction	1
5X Bradford reagent	20
ddH_2O	79
Total	**100**

.

2.6.9 Prescission protease cleavage of death domain fusion proteins

The Prescission protease cleaves off all the tags in the death domain fusion proteins used in this project. Therefore, Prescission protease was expressed from a 3C gene in the pGEX-6P-2 plasmid and purified with assistance of Mr. Clifford Ntui, a fellow student. To set up the cleavage reaction, 50 µg of Prescission protease was added to each purified protein fraction and mixed gently 5 times by inversion. The reactions were then incubated at 25° C for 2 hours and then analysed by SDS-PAGE.

2.6.10 Size exclusion chromatography (SEC) of death domain proteins

2.6.10.1 Materials

Table 26. Column and buffer for size exclusion chromatography

Category	Name	Supporting information
Column	ENrich SEC 650-10 x 300 mm column (24 mL) 	Commercially purchased. Supplier: Bio Rad
Buffer	SEC buffer	Composition: 20 mM Tris (pH 7.5), 150 mM NaCl, 2 % (v/v) glycerol

2.6.10.2 Method – SEC

Protein samples were purified by size exclusion chromatography (SEC) using the ÄKTA Purifier w/UPC-900 system. An ENrich SEC 650-10 x 300 mm column (24 mL) was connected to the ÄKTA system and subsequently equilibrated with 2 column volumes of SEC buffer. A 500 µL sample loading loop was cleaned with 3 loop volumes of SEC buffer, before injecting 500 µL of the protein sample into it. The SEC purification protocol was loaded on the UNICORN^{TM} FPLC interface using the parameters shown in Table 27. Afterwards, eluted fractions were analysed by SDS-PAGE.

Table 27. Size exclusion chromatography FPLC parameters

Block	Variable	Value
Main	Column	SEC_650_24_ml
Flow rate	Flow rate (mL/min)	0.30 – 0.50
Column pressure limit	Column pressure limit (MPa)	10.00
Start instructions	Wavelength (nm)	280
	Averaging time UV (sec)	5.12
	Empty loop with (mL)	1.00
Fractionation	Eluate fraction size (mL)	1.00
Length_of_elution	Length of elution (CV)	1.00

2.6.11 Reconstituting the death domain Myddosome complexes

The Myddosome complexes were reconstituted by mixing 3 mL of IRAK-4 death domain (8 mg/mL) with 3 mL of MyD88-DD (10 mg/mL) for the intermediate domain (ID) deficient sample. A 3 mL volume of IRAK-4 death domain (8 mg/mL) was mixed with 3 mL of MyD88-DD-ID (8 mg/mL) to reconstitute the ID-inclusive Myddosome complexes. The samples were concentrated using 10 kDa MWCO centrifugal filter columns to a final volume of 500 µL. Thereafter, samples were purified by size exclusion chromatography and the eluted fractions were analysed by SDS-PAGE.

2.6.12 Dynamic light scattering (DLS) analysis for MyD88-DD

The hydrodynamic diameter of the MyD88 death domain protein was measured by dynamic light scattering (DLS). The protein sample (1 mL) was loaded into a clear polystyrene semi-micro cuvette and analysed using the Malvern Zetasizer Nano ZS (with the Zetasizer software version 7.11.) The SOP parameters for sample analysis were loaded as shown in Table 28. The hydrodynamic diameter was measured and reported along with a corresponding graph.

Table 28. SOP parameters for Zetasizer Nano ZS DLS analysis

Tab	Value
Sample material	Proteins
Measurement index	Size
Dispersant RI	1.33
Material RI	1.45
Viscosity (McPherson)	1.57
Cuvette class	ZEN 0040
Record number	4
Plot	Average only

2.7 Instruments

Table 29. Instruments

Category	Equipment	Manufacturer	Headquarters
Autoclaves	Hiclave™ HVE-50 sterilizer	Hirayama manufacturing corporation	Saitama, Japan
	HL-341 vertical type steam sterilizer	GMP-Echung machinery corporation	Taoyuan City, Taiwan
Balances	Mettler Toledo AE 163	Mettler Toledo® solutions	Ohio, USA
	Precisa weighing balance XB 6200D	Precisa instruments AG	Zürich, Switzerland
	Sartorius 1712MP8 analytical balance	Sartorius laboratory instruments	Göttingen, Germany
Cell density meter	CO8000 CD meter	WPA-Biowave	Connecticut, USA
Centrifuges	Eppendorf centrifuge 5415	Eppendorf	Hamburg, Germany
	Heraeus Megafuge 8R	Thermo Fisher Scientific	Massachusetts, USA
	Sorvall Lynx 6000	Thermo Fisher Scientific	Massachusetts, USA
Electrophoresis apparatus	PowerPac™ basic	Bio Rad	California, USA
	Mini-Protean tetra cell	Bio Rad	California, USA
Fast protein liquid chromatography	ÄKTA Purifier FPLC P-900 System	Amersham biosciences corporation	Buckinghamshire, England
Imaging systems	Molecular Imager Gel Doc XR+ Imaging system	Bio Rad	California, USA
	Transilluminator Safeview	Cleaver Scientific	Warwickshire, UK
Incubators	Orbital incubator SI500	Cole-Parmer Stuart equip	Illinois, USA
	Refrigerated shaker laboratory incubator	MRC laboratory instruments	Essex, England
PCR	T100 Thermal cycler	Bio Rad	California, USA
pH Meter	MT 320 pH meter	Mettler Toledo	Ohio, USA
Shaker	Thermomixer, model 5436	Sigma-Aldrich	Missouri, USA
Size analyser	Zetasizer Nano ZS	Malvern Panalytical	Worcestershire, UK
Sonicator	Q500 Sonicator	Qsonica	Connecticut, USA
Spectrophotometers	NanoDrop ND-1000	Peqlab Biotechnologie	Erlangen, Germany
	SpectraMax® Paradigm®	Molecular devices LLC	California, USA
Transfer system	Trans-Blot® Turbo™	Bio Rad	California, USA
Vortex	Wizard vortex mixer	VELP Scientifica	Lombardia, Italy

Chapter 3: Expression, purification and analysis of MyD88-DD-ID protein

3.1 Introduction

In this chapter, we wanted to produce and purify an ID-inclusive MyD88-DD protein. This protein would offer the necessary contrast with the ID-deficient MyD88-DD protein in our efforts to investigate the role of the ID in the reconstitution of the Myddosome complex. Furthermore, there is evidence in literature on the known interactions with IRAK-4 death domain using this ID-inclusive protein (Motshwene et al., 2009). As such, this ID-inclusive protein would also serve as a positive control for interaction studies later on. The workflow for this chapter includes DNA sequencing, followed by production, purification and cleaving of the ID-inclusive MyD88-DD fusion protein.

3.2 Results

3.2.1 DNA sequencing of MyD88-DD-ID-pETG-30A expression clone.

We had an MyD88-DD-ID-pETG-30A construct that had been cloned by my supervisor, Dr. Motshwene. It had expressed very well in previous studies, and it was our intention to use it for the current study. However, we firstly needed to confirm the presence of the MyD88-DD-ID gene. Therefore, the plasmid was cleaned using the Big Dye V3 sequencing kit followed by Sanger sequencing using the MyD88 forward primer. The raw sequencing data was then analysed using the SnapGene® viewer program (version 5.3.1). Thereafter, once we were satisfied with quality of the data based on the clearly resolved peaks, we exported the data to the Expasy translation software. The result of the gene translation is shown in Figure 13.

The first amino acid of the translated gene sequence is a methionine as shown in Figure 13. This methionine corresponds to the 54th amino acid of the MyD88 gene and is part of the death domain region. Amino acids downstream of this methionine, were consistent with those of the MyD88 sequence. The first amino acid of the MyD88 intermediate domain was also identified in our translated sequence. It is the glutamic acid highlighted with a red circle. This glutamic acid corresponds to the 111th amino acid of MyD88.

There are at least three codons that the sequencing machine failed to resolve. These are shown by a non-existing codon ‘nnn’ in figure 13. As expected, the Expasy translation software failed to translate the ‘nnn’ codons. It assigned them a single letter code X. A single letter code X is used to denote any amino acid. The last residue of the translated sequence is an aspartic acid which we identified as the 151st amino acid of MyD88. Therefore, we were certain that we had the correct construct in our possession. Our gene sequencing results showed the presence of both the MyD88 death domain and the majority of the intermediate domain.

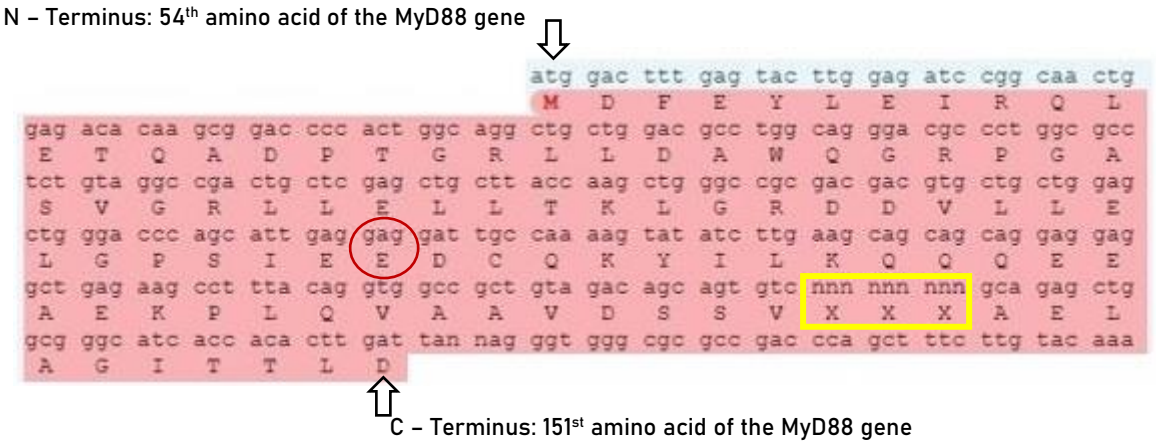


Figure 13. Expasy translation of raw DNA sequencing data for MyD88-DD-ID-pETG-30A plasmid. Expasy software translation of sequencing data showing the amino acid sequence for the MyD88-DD-ID gene. The first amino acid of the translation was a methionine residue which corresponded to the 54th amino acid of the MyD88 gene. The red circle indicates the first amino acid of the intermediate domain which is a glutamic acid. The yellow box indicates codons that were not resolved by the sequencer. The last amino acid of the translation was an aspartic acid which corresponded to the 151st amino acid of the MyD88 gene.

3.2.2 Nickel affinity chromatography purification and Prescission protease cleavage of His-GST-MyD88-DD-ID fusion protein

Since we verified that we had the correct MyD88-DD-ID-pETG-30A plasmid, we now wanted to produce soluble protein. Therefore, we prepared a starter culture by inoculating BL21 cell colonies (transformed with the MyD88-DD-ID-pETG-30A plasmid) into LB broth and incubated them overnight at 37° C while shaking. Thereafter, four litres of LB broth were inoculated with this starter culture followed by 1 mM IPTG induction and then incubating overnight at 15° C while shaking. The induced cells were harvested and subsequently lysed by sonication. The lysate was centrifuged, after which the supernatant was collected as the soluble fraction. Nickel affinity chromatography was then used to purify the soluble fraction. From the purified sample, 20 µL was aliquoted and kept aside as the uncleaved control. The remainder of the purified sample was mixed with 50 µg of Prescission protease and incubated for 2 hours at 25° C to cleave MyD88-DD-ID from its fusion partner His-GST. Afterwards, both the cleaved and uncleaved samples were analysed by SDS-PAGE as shown in Figure 14.

The uncleaved sample contained several bands that include the expected 42 kDa His-GST-MyD88-DD-ID fusion protein as shown in lane 1 of figure 14. It had the highest intensity when compared to other bands in that lane. Therefore, this result suggested we were successful in producing soluble His GST-MyD88-DD-ID fusion protein. The cleaved sample in lane 2 of figure 14 also had several protein bands. There was a 42 kDa band (of significantly smaller intensity compared to the 42 kDa band in lane 1), a 25 kDa band and a 17 kDa protein band. Based on the specificity of the Prescission protease, we therefore deduced that the 42 kDa band was the uncleaved His GST MyD88 DD ID protein, while the 25 kDa band was the His GST tag and the 17 kDa band was the MyD88 DD ID protein. The obtained yields as determined by visual inspection of the SDS gel, were enough for subsequent experiments.

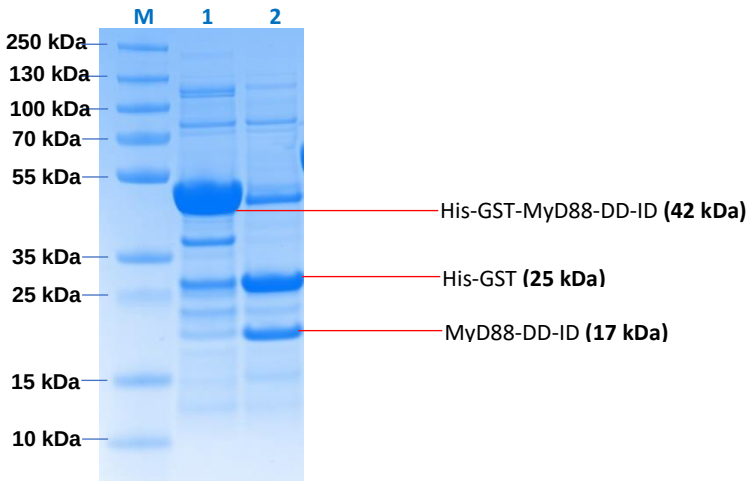


Figure 14. SDS gel of nickel affinity chromatography purified and Prescission protease cleaved His-GST-MyD88-DD-ID protein. Soluble protein fraction expressed from the ID-inclusive MyD88-DD-ID-pETG-30A plasmid was purified by nickel affinity chromatography using a HiTrap chelating HP column. A 20 µL aliquot of this purified sample was kept aside (as the uncleaved fraction) while the rest of the purified soluble protein was cleaved by Prescission protease. Both the cleaved and uncleaved fractions were analysed by SDS-PAGE on a 12 % SDS polyacrylamide gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa molecular weight reference ladder. Lane 1 - uncleaved fraction. Lane 2 - cleaved fraction.

3.2.3 Size exclusion chromatography of cleaved His-GST-MyD88-DD-ID protein

The successful cleavage of the His-GST-MyD88-DD-ID in the previous section allowed us to obtain the MyD88-DD-ID protein. Therefore, we now wanted to further purify the cleaved sample in order to separate the MyD88-DD-ID protein from the His-GST tag. We opted to purify the cleaved protein sample by size exclusion chromatography (SEC). This technique would purify the sample and in addition allow us to determine the retention volume of the ID-inclusive MyD88-DD protein. The retention volume would give us an indication of the size and oligomeric status of the ID-inclusive MyD88-DD.

The cleaved His-GST-MyD88-DD-ID protein sample was concentrated to a final volume of 500 µL prior to loading on an ENrich SEC 650 (10 x 300 mm) column. This SEC purification resulted in a chromatogram that has six peaks as shown in Figure 15. The corresponding fractions for these peaks were collected and labelled as peak 1 up to 6. Each peak had its own retention volume and absorbance as shown in Table 30. Peak 1 had a retention volume of 11.45 mL and an absorbance of 103 mAU. Peak 2 had a retention volume of 12.95 mL and an absorbance of 126 mAU. Peak 3 had a retention volume of 14.39 mL and an absorbance of 150 mAU. Peak 4 had a retention volume of 15.95 mL and an absorbance of 620 mAU. Peak 5 had a retention volume of 20.42 mL and an absorbance of 650 mAU. Peak 6 had a retention volume of 22.71 mL and an absorbance of 350 mAU. The results suggested that peak 4, 5 and 6 had the highest concentration of protein based on the high absorbance values.

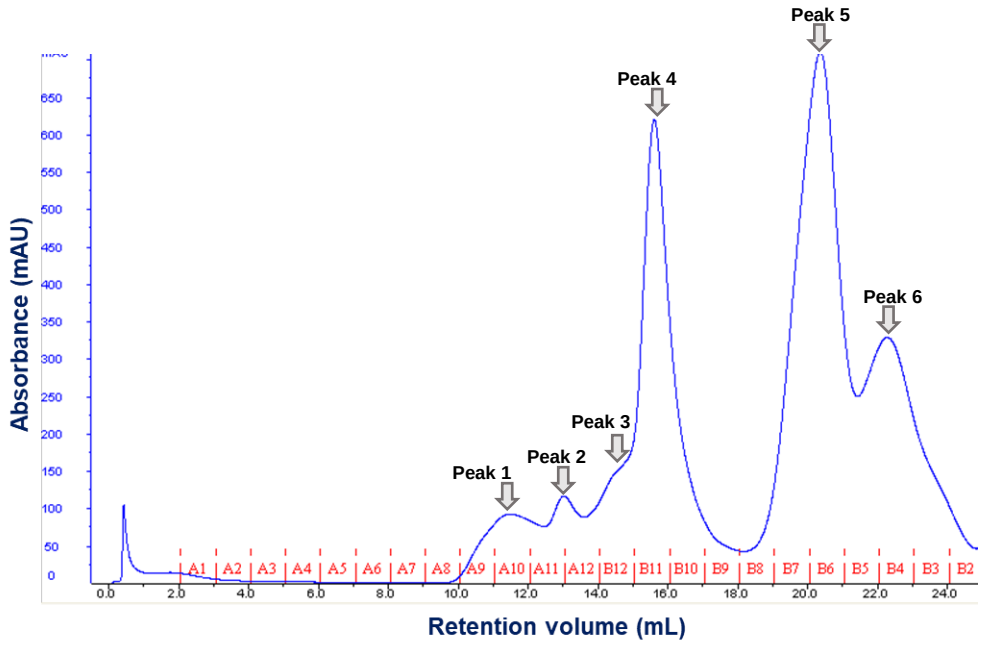


Figure 15. Chromatogram for SEC purification of cleaved His-GST-MyD88-DD-ID. The cleaved His-GST-MyD88-DD-ID protein sample was purified by size exclusion chromatography (SEC) and a chromatogram was generated from this purification by an ÄKTA FPLC machine. UV absorbance readings at 280 nm were captured for prominent peaks and collected as follows: Peak 1 – [fraction: A9 and A10]. Peak 2 – [A11 and A12]. Peak 3 – [B12]. Peak 4 – [B10 and B11]. Peak 5 – [B7 and B6]. Peak 6 – [B4].

Table 30. SEC data output for cleaved His-GST-MyD88-DD-ID protein

	Retention volume (mL)	Absorbance (mAU)
Peak 1	11.45	103
Peak 2	12.95	126
Peak 3	14.39	150
Peak 4	15.95	620
Peak 5	20.42	650
Peak 6	22.71	350

After our SEC purification, we now wanted to verify if the ID-inclusive MyD88-DD-ID protein had been separated from the His-GST tag. Samples of fractions that eluted from the six peaks were loaded and analysed by SDS-PAGE in the order of their elution. A sample from peak 1 was loaded in lane 1 as shown in Figure 16. It showed no significant protein bands. Similarly, a sample from peak 2, loaded in lane 2 also showed no significant protein bands. A sample from peak 3, loaded in lane 3 showed very faint protein bands. One such band was a 17 kDa band which led us to believe that this protein was the MyD88-DD-ID protein based on the observed molecular weight. A sample from peak 4, loaded in lane 4 showed several bands of which the most notable one was a large 25 kDa band. This band was highly likely to be the His-GST tag based on its observed molecular weight on SDS-PAGE. Peak 5 and 6 were very peculiar because they showed high absorbance values of 650 mAU and 350 mAU respectively and yet there were no protein bands detected when their fractions were analysed on SDS-PAGE.

The result obtained in figure 16 was not satisfactory. It showed that we were losing the MyD88-DD-ID protein during SEC. The intensity of the MyD88-DD-ID protein band in figure 16, lane 3 was very low when compared the intensity obtained prior to SEC in figure 14, lane 2. Therefore, this meant that we would be better off using the cleaved sample prior to SEC for reconstituting the Myddosome complex. However, we did manage to obtain a retention volume of MyD88-DD-ID from the SEC purification. The retention volume that was obtained was 14.39 mL, which is comparable to what has been found in literature.

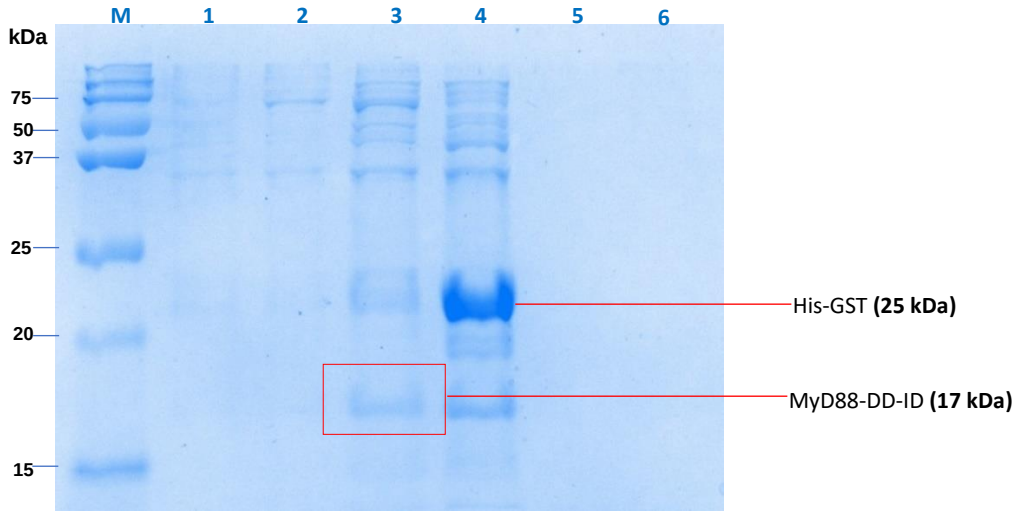


Figure 16. SDS gel of SEC purification fractions for cleaved His-GST-MyD88-DD-ID. Fractions for the peaks that were identified in SEC purification of cleaved His-GST-MyD88-DD-ID protein were collected and analysed by SDS-PAGE on a 12 % SDS polyacrylamide gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa protein marker. Lane 1 – peak 1. Lane 2 - peak 2. Lane 3 – peak 3. Lane 4 – peak 4. Lane 5 – peak 5. Lane 6 – peak 6.

3.3 Discussion

Before we could start with our experiments, we had to verify if we had the correct MyD88-DD-ID-pETG-30A plasmid. The construct was therefore sent for DNA sequencing and the raw data (data not shown) showed resolved peaks. However, there were regions where the sequencing machine could not resolve the peaks properly. These areas only covered a small proportion of the raw data, that is three codons and were thus insignificant. It is not uncommmon to have unresolved peaks during sequencing (Crossley et al., 2020). Such an occurrence happens in almost every sequencing reaction (Van Dijk, Jaszczyszyn, Naquin, & Thermes, 2018). The quality of our sequencing data was therefore not compromised as we identified the MyD88 sequence from the plasmid that we sent for sequencing. As expected, the sequence did not cover the first amino acid of MyD88 because of the forward primer we used. It annealed at the beginning of the MyD88 gene. It is therefore quite common during DNA sequencing for certain bases upstream of the primer not to be detected. Otherwise, we were able to positively identify a sequence of at least 94 amino acids that belonged to MyD88. It covered both the death and intermediate domains.

Now that we were sure we had the correct construct, our aim was to express and purify the MyD88-DD-ID fusion protein. Optimal conditions for the expression of the MyD88-DD-ID fusion protein from the pETG-30A construct were already established from previous studies (Motshwene et al., 2009; George et al., 2011). We therefore used the same conditions for our large scale production. The His-GST-MyD88-DD-ID fusion protein of the expected size was expressed and subsequently cleaved with Prescission protease. Cleavage was successful as determined by the presence of the MyD88-DD-ID protein and His-GST. However, there was some uncleaved protein sample. This could possibly be attributed to sample heterogeneity, meaning that there could be have been some aggregated fusion protein that failed to cleave (Young, Britton, & Robinson, 2012).

In order to isolate the MyD88-DD-ID protein, the cleaved fusion protein sample was purified using size exclusion chromatography. Size exclusion chromatography separates protein molecules on the basis of size and shape (Held & Kilz, 2021). By using it, we had hoped that it would separate the MyD88-DD-ID protein from the His-GST tag and the minute His-GST-MyD88-DD-ID fusion that failed to cleave. Our results showed that the 17 kDa MyD88-DD-ID protein had a retention volume of 14.39 mL. It eluted earlier than the 25 kDa His-GST that has a retention volume of 15.95 mL. Therefore, we needed to come up with an explanation on how this anomaly could have occurred. In size exclusion chromatography, larger molecules elute first, followed by intermediate molecules and the smaller molecules last. However, we have to bear in mind that shape also plays a role in the retention volumes of molecules in size exclusion chromatography (Held & Kilz, 2021). The observed retention volume of MyD88-DD-ID protein may lead one into thinking that it has oligomerised, and has possibly formed dimers. There are some studies that have reported MyD88 as a dimer (Padova, Quesniaux, & Ryffel, 2018). However,

these have relied on the use of pull down assays. Motshwene et. al, 2009, has conclusively shown using analytical ultracentrifugation that MyD88 death domain is a monomer. Therefore, its observed retention in our study can mainly be attributed to its shape. Moncrieffe et. al., 2020 has shown using cryo-electron microscopy that MyD88 forms helical filaments *in vitro*. It is therefore our belief that it elutes earlier than the His-GST protein due to the formation of these filaments.

The yields of MyD88-DD-ID protein after size exclusion chromatography were low and not ideal for future experiments. Low yields were an indication that we were losing protein during size exclusion chromatography. Motshwene et. al, 2009 and George et al., 2011 obtained good yields of MyD88-DD-ID protein in their studies. Unlike us, they started by purifying the His-GST-MyD88-DD-ID fusion protein on a glutathione sepharose 4B resin. We on the other hand started by purifying it by nickel affinity chromatography because we were not able to procure the glutathione sepharose 4B resin due to our limited budget. A study that compared protein purification via His-tag and GST-tag, also observed differences in yield between using these two tags (Scheich, Sievert, & Büssow, 2003).

Since the yields of MyD88-DD-ID protein after size exclusion chromatography were low, we decided that we will use the cleaved sample prior to size exclusion chromatography as shown in Figure 14, lane 2 for future experiments. Those experiments will entail mixing the cleaved sample with purified IRAK-4 death domain in order to reconstitute the Myddosome complex.

3.4 Conclusion

We expressed the His-GST-MyD88-DD-ID fusion protein from our construct and obtained soluble protein. This fusion protein was cleaved with Prescission protease, and a satisfactory yield of the MyD88-DD-ID protein was obtained. Further purification of our cleaved sample using size exclusion chromatography resulted in low recovery of the MyD88-DD-ID protein. Therefore, we decided that we were going to use the cleaved sample prior to size exclusion chromatography in the reconstitution of the Myddosome complex in future experiments.

Chapter 4: Generating ID-deficient MyD88-DD construct by site directed mutagenesis and protein analysis

4.1 Introduction

In order to conduct our study on the role of the ID in the Myddosome assembly, we needed to have an MyD88-DD protein that was ID-deficient to compare against an ID-inclusive MyD88-DD protein. Therefore, in this chapter, we now wanted to produce the ID-deficient MyD88-DD by introducing a stop codon on the first amino acid of the intermediate domain using site directed mutagenesis (SDM). The first amino acid of the intermediate domain is glutamic acid 111 (E111). Primers were designed that mutated the GAG codon for glutamic acid 111 for a TAA stop codon. By so doing, the MyD88-DD-ID-pETG-30A plasmid in our possession had been rendered incapable of expressing the intermediate domain. This plasmid is the same one that was used in the previous chapter. Therefore, our aim in this chapter was to create an E111 stop mutant, followed by expression, purification and cleaving of the fusion protein to assess if the final product can be used to reconstitute the Myddosome.

4.2 Results

4.2.1 Site directed mutagenesis (SDM) sequencing

Site directed mutagenesis was performed on the MyD88-DD-ID-pETG-30A plasmid to introduce an E111 stop mutant. The mutant plasmids that were generated by site directed mutagenesis were sequenced using the MyD88 forward primer and the raw sequencing data was analysed using the SnapGene® viewer program (version 5.3.1). Thereafter, once we were satisfied with quality of the data based on the clearly resolved peaks, we exported the data to the Expasy translation software. The result of the gene translation is shown in Figure 17.

The first amino acid of the translated gene sequence is a methionine as shown in figure 17. This methionine corresponds to the 27th amino acid of the MyD88 gene, and is part of the death domain region. Amino acids downstream of this methionine, covered the entire MyD88 death domain sequence till the last glutamic acid (E110), which is annotated with an arrow and green square in figure 17. The next amino acid after E110, was found to be a stop codon; and it is annotated with a red circle. It would have represented the first amino acid of the intermediate domain had it not been mutated to a stop codon. Therefore, this sequencing result showed that our site directed mutagenesis was successful.

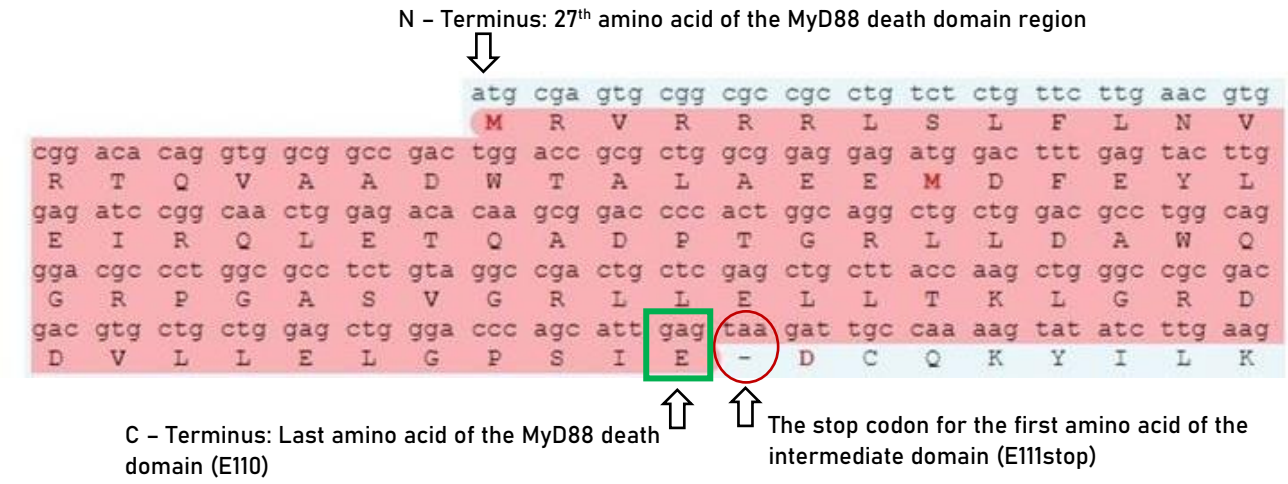


Figure 17. Expasy translation of raw DNA sequencing data for SDM construct. Expasy software translation of sequencing data showing the amino acid sequence for the SDM mutated MyD88-DD-ID gene. The first amino acid of the translation was a methionine residue which corresponded to the 27th amino acid of the MyD88 gene. The green square indicates the last amino acid of the MyD88 death domain (E110). The red circle indicates the stop codon on the first amino acid of the intermediate domain (E111stop).

4.2.2 Expression trial for MyD88-DD-pETG-30A construct (SDM mutant)

Based on our success in producing the MyD88 E111stop mutant, the next step was to do an expression trial to determine if the plasmid would express the intermediate domain deficient MyD88-DD fusion protein whose expected molecular weight is 37 kDa. We decided to also include the unmutated MyD88-DD-ID plasmid in the expression trial as a positive control. Therefore, both plasmids were transformed into BL21 (DE3) cells which were later cultured in 5 mL of LB broth containing ampicillin. Thereafter, they were induced with IPTG to a final concentration of 1 mM and left to grow overnight at 37° C in a shaking incubator as previously described. The cells were then harvested by centrifugation and the resulting cell pellets were lysed using BugBuster® protein extraction reagent. The lysate was centrifuged to obtain the soluble and insoluble protein fractions. Both soluble and insoluble fractions were analysed by SDS-PAGE as shown in Figure 18.

Lane 1 in figure 18 was loaded with the soluble fraction that resulted from the lysate of cells that expressed the unmutated MyD88-DD-ID protein. Several bands were observed. However, there was an intense band with a molecular weight of approximately 41 kDa that stood out. This band was highly likely to be the MyD88-DD-ID protein. Lane 2 on the other hand was loaded with the insoluble fraction that resulted from the lysis of cells that expressed the unmutated MyD88-DD-ID protein. There were also several bands that were observed. However, just like in lane 1, an intense 41 kDa band, likely to be the MyD88-DD-ID protein was observed. From the results seen in lane 1 and 2, it was clear that our positive control was expressing protein as expected. This result was also consistent with our observations in Chapter 3.

Lane 3 was loaded with the soluble fraction that resulted from the lysate of cells that expressed the MyD88 E111stop mutant. A banding pattern similar to lane 1 was also observed. An intense 41 kDa stood out. The insoluble fraction of the expressed MyD88 E111stop mutant was loaded in lane 4. There were many protein bands that were observed. However, there was an intense 35 kDa band with an expected molecular weight of the His-GST-MyD88-DD fusion protein. This band was highly likely to be the His-GST-MyD88-DD fusion protein.

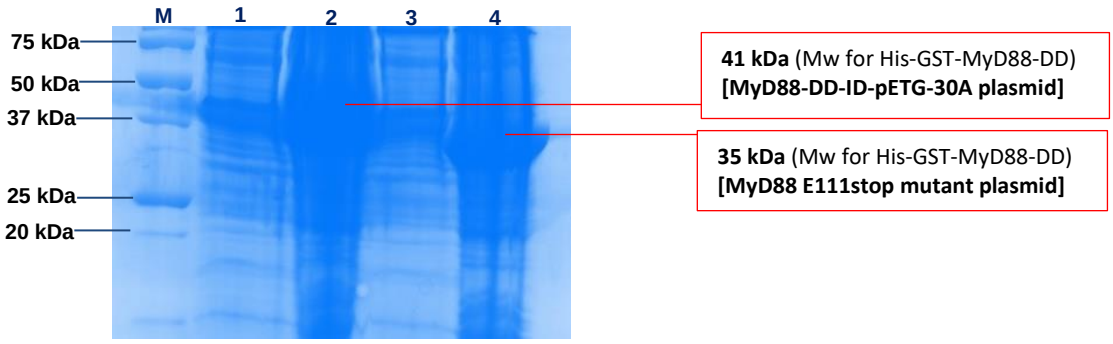


Figure 18. SDS gel of MyD88-DD protein expression trials. Small scale protein expression was conducted from the MyD88 E111stop mutant plasmid and the unmutated MyD88-DD-ID-pETG-30A plasmid. The extracted soluble [sol] and insoluble [ins] protein fractions from the expressions were analysed by SDS-PAGE on a 12 % SDS polyacrylamide gel. 18 µL samples from these fractions were loaded as follows: Lane M – 250 kDa protein ladder. Lane 1 – His-GST-MyD88-DD-ID [sol]. Lane 2 – His-GST-MyD88-DD-ID [ins]. Lane 3 – His-GST-MyD88-DD [sol - MyD88 E111stop mutant]. Lane 4 – His-GST-MyD88-DD [ins - MyD88 E111stop mutant].

4.2.3 Upscale and nickel affinity purification of SDM mutant MyD88-DD protein.

The presence of our expected His-GST-MyD88-DD fusion protein in the insoluble fraction did not deter us from doing a large-scale expression. There was a possibility that its expression levels were low, hence we could not see it in the soluble fraction. Therefore, doing a large-scale expression was going to eliminate whatever uncertainties we had.

In an attempt to get more soluble His-GST-MyD88-DD protein, we set up 4 litres of culture. These were induced with IPTG when they reached an optical density (OD_{600}) of 0.8. The temperature was then lowered to 15° C and the cultures were left to grow overnight. Cells were then harvested, lysed by sonication and then the lysate was centrifuged at 37 500 x *g*. The resulting soluble fraction was purified by nickel affinity chromatography. Fractions that were eluted from the nickel column were collected and analysed by SDS-PAGE as shown in figure 19.

Lane 1 to 3 of figure 19, showed the presence of three protein bands. These had molecular weights of 37, 27 and 24 kDa. The 37 kDa band was highly likely the target His-GST-MyD88-DD protein. This result suggested it was purified in a soluble form. The 27 kDa band was highly likely to be the His-GST tag and the 24 kDa was possibly its degradation product.

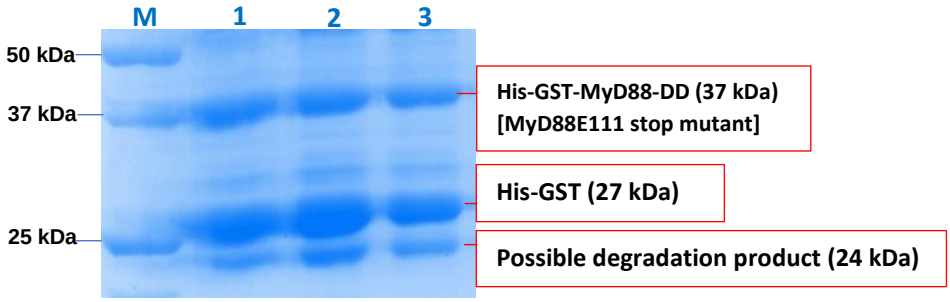


Figure 19. SDS gel of MyD88E111 stop mutant purified by nickel affinity chromatography. Soluble protein expressed from the MyD88E111 stop mutant plasmid (SDM construct) was purified by nickel affinity chromatography purification. Purified samples were analysed by SDS-PAGE on a 12 % SDS polyacrylamide gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa molecular marker. Lanes 1, 2, 3 – The first three 1 mL fractions eluted from the HiTrap chelating HP column during the nickel affinity chromatography purification.

4.2.4 Cleaving the His-GST-MyD88-DD fusion protein with Prescission protease

Having produced the soluble His-GST-MyD88-DD fusion protein, we next wanted to cleave MyD88-DD from its His-GST fusion partner. Therefore, 50 µg of Prescission protease enzyme was added to the nickel affinity chromatography purified protein from the previous section (4.2.3) and then incubated at 25° C for 2 hours. The cleaved protein sample was subsequently analysed by SDS-PAGE alongside an uncleaved sample as a negative control as shown in Figure 20.

Lane 1 in figure 20 was loaded with the uncleaved protein sample and it showed several bands of varying intensity, which were identical to those observed in Figure 19 of the previous section (4.2.3). The cleaved sample in lane 2, was also found to have an identical banding profile as compared to lane 1. However, there was a very small 12 kDa protein band. This band is annotated in a red circle. Based on its molecular weight, it is highly likely to be MyD88-DD. However, its yield was very low as determined by the intensity of the band observed. Such a very low yield could not be used further downstream applications.

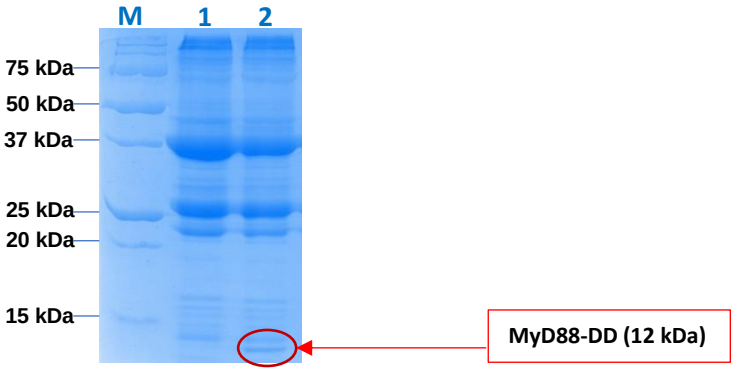


Figure 20. SDS gel of His-GST-MyD88-DD (SDM) fusion protein cleaved by Prescission protease. Purified His-GST-MyD88-DD protein was cleaved using Prescission protease. The cleaved sample as well as an uncleaved control sample were analysed by SDS-PAGE on a 12 % SDS polyacrylamide gel. 18 µL samples from these fractions were loaded as follows: Lane M – 250 kDa protein reference ladder. Lane 1 – Uncleaved His-GST-MyD88-DD protein (negative control). Lane 2 - Cleaved His-GST-MyD88-DD protein.

4.3 Discussion

We performed site directed mutagenesis on a construct known to express both the MyD88 death domain and the intermediate domain. The codon for the first amino acid of the intermediate domain was mutated to a stop codon, thus creating a new construct that would only express the death domain. Sequencing of the construct showed the presence of the stop codon. This result therefore meant we were successful in creating an intermediate domain deficient construct.

Expression trials were conducted on the intermediate domain deficient construct. The results showed that the expected His-GST-MyD88-DD protein was mainly in the insoluble fraction, meaning it had formed inclusion bodies. The formation of inclusion bodies stems from difficulties in protein synthesis, protein folding, cell metabolism, environmental conditions and target protein properties as the host cells are forced to express a foreign protein (Bhatwa et al., 2021). Furthermore, variations in codon usage that occur between different species, is a well documented factor that affects recombinant protein production (Menzella, 2011). Thus, unfavorable codon bias in *E. coli* during translation, could have negatively affected the protein synthesis of our human MyD88-DD protein.

In order to overcome the formation of inclusion bodies, we implemented a few changes in our culturing conditions given the observations from the expression trial. As mentioned earlier, environmental conditions are one of the key factors that contribute to inclusion body formation. Therefore, we shifted from culturing at 37° C to a lower temperature of 15° C for the production of MyD88-DD. Lower culturing temperatures are known to reduce stress on bacterial cell processes such as translation which are slowed down resulting in improved folding dynamics that promote solubility (Gopal & Kumar, 2013). We also upscaled our cultures to enhance the chances of obtaining higher protein yields. The adjustments that were made in the culturing conditions, yielded soluble protein when compared to the expression trial.

Although we were successful in expressing soluble His-GST-MyD88-DD protein, our attempts at cleaving this fusion protein were ineffective. There was very little MyD88-DD visible on SDS-PAGE. We doubt that increasing the concentration of the Prescission protease would have made any difference. The same batch of Prescission protease had been used with success at the same concentration to cleave the His-GST-MyD88-DD-ID fusion protein in the previous chapter (section 3.2.2). Another possibility why the His-GST-MyD88-DD protein failed to cleave is that it might have expressed as a soluble aggregate. It is known that aggregation does make protease cleavage sites inaccessible (Mazzini et al., 2022). Since our Prescission protease had previously worked, there was no point in changing its cleavage site from the His-GST-MyD88-DD for another protease like TEV or thrombin.

Unfused His-GST was also observed on SDS-PAGE. This is not something new during protein expression. It has been observed in other studies due to plasmids being leaky (Seniya, Yadav, & Jain,

2020). There was also some degradation which we believed to be likely His-GST. The degradation of unfused His-GST is of least importance to this study. However, it could have possibly been prevented by the inclusion of protease inhibitors during cell lysis.

4.4 Conclusion

We were successful in producing the MyD88 E111stop mutant using site directed mutagenesis. This mutant strictly expresses the MyD88 death domain. However, we could not use it for further experiments as it yielded extremely low concentrations of MyD88 protein. More losses of protein would have been incurred had it been purified further. Therefore, we decided to clone a new set of expression constructs to produce better yields of the MyD88-DD protein. Our plan was to use the Gateway cloning system, a technology available in our research group. It enables one to explore the use of various tags in an attempt to obtain potentially higher yields of the target protein.

Chapter 5: Generating ID-deficient MyD88-DD by gateway cloning system

5.1 Introduction

The MyD88 E111stop mutant that was generated using site directed mutagenesis in chapter 4 produced low yields of protein. As a result, we could not proceed to purify it further because there were going to be further losses of protein. We therefore decided to clone the intermediate domain deficient MyD88-DD from scratch. To do this, we opted to use the gateway cloning technology. This technology enables you to insert your gene of interest into many destination vectors, which are used for protein expression. The advantage with the gateway technology is that cloning is only done once.

5.2 Results

5.2.1 Amplifying the ID-deficient MyD88-DD gene insert

The first step in our gateway cloning process was to generate an entry clone containing only the MyD88 death domain (MyD88-DD) gene. Therefore, we designed MyD88 forward and reverse primers to only amplify the MyD88-DD gene. These primers were then used to perform a PCR reaction using the MyD88 template DNA. After the PCR reaction, the product was analysed by agarose gel electrophoresis as shown in Figure 21(A). Lanes 1 and 2 in Figure 21(A) served as the negative controls and they showed no bands. A 0.3 kb band with an expected molecular weight of the MyD88-DD gene was observed in lane 3. The observed band in lane 3 suggested that our PCR reaction was successful.

In order to clone the MyD88-DD gene, we needed more amplified PCR product. We therefore conducted another round of PCR and loaded the entire product into a bigger agarose gel well as shown in Figure 21(B). Lane 1 of this figure contained the expected 0.3 kb band. This band was excised from the gel and purified in preparation for gateway cloning.

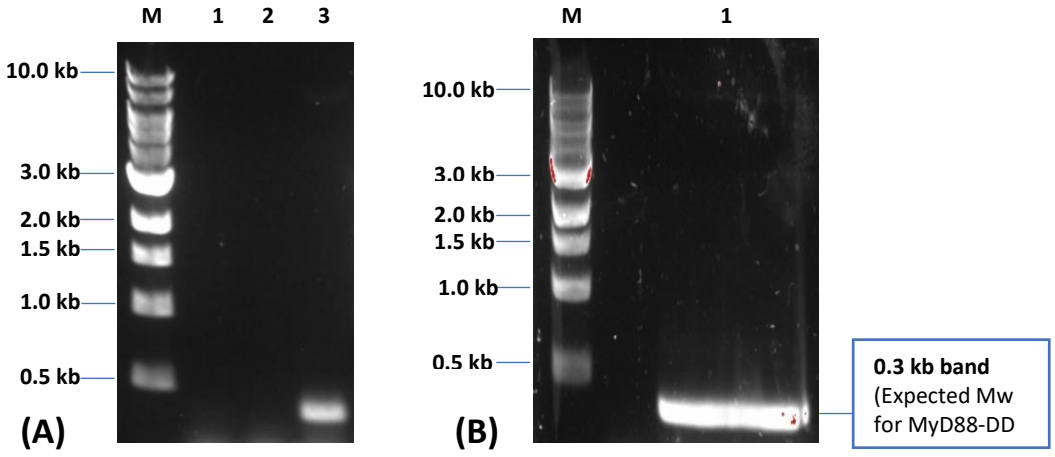


Figure 21. Agarose gel showing amplification of MyD88-DD gene by PCR. MyD88 forward and MyD88 reverse primers were used in a PCR reaction using MyD88 template DNA. The PCR product was analysed by agarose gel electrophoresis on a 1 % agarose gel. **(A)** Lane M – 1 kb DNA ladder. Lane 1 – negative control (no primers). Lane 2 – negative control (no MyD88 template DNA). Lane 3 - MyD88 template DNA. **(B)** Lane M - 1 kb DNA ladder. Lane 1 (expanded well) - MyD88 template DNA.

5.2.2 TOPO cloning and verification of MyD88-DD-pENTR/D entry clone.

The purified 0.3 kb band from the previous section was mixed with the pENTR/D vector to initiate a gateway cloning reaction. After 30 minutes of incubation at room temperature, the reaction mixture was transformed into DH5α cells. The DH5α cells were then streaked onto LB agar plates that contained 50 µg/mL of kanamycin and incubated overnight. Colonies were observed and subsequently analysed by colony PCR using the MyD88 forward and reverse primers. The PCR products were then electrophoresed on a 1 % agarose gel as shown in Figure 22. Lane 1 was loaded with the positive control which consisted of MyD88 template DNA and it showed the 0.3 kb band. Lanes 2-9 contained template DNA from the 8 colonies that were picked from the LB agar plates. They also showed the expected 0.3 kb band. However, the intensity of the band in lane 2 was higher when compared to the other lanes. Since 0.3 kb was the expected molecular weight of the MyD88-DD gene, we were convinced that the TOPO reaction was successful in producing the MyD88-DD-pENTR/D entry clone in all 8 colonies that we picked.

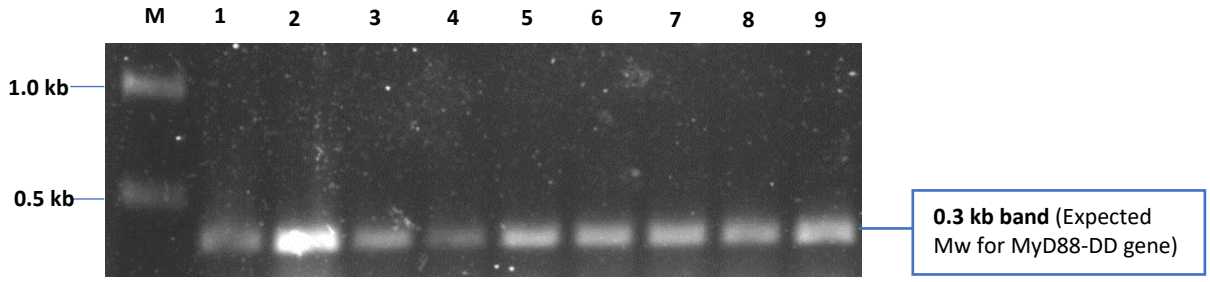


Figure 22. Agarose gel of colony PCR for TOPO cloning. Colony PCR was conducted using colonies that grew overnight from the TOPO cloning reaction DH5α cell transformation. The colony PCR product was analysed by agarose gel electrophoresis on a 1 % gel. Lane M – 1 kb DNA ladder. Lane 1 – (positive control) MyD88-DD template DNA. Lane 2-9 – Colonies picked from the LB agar plate which contained the kanamycin antibiotic.

Even though we believed thus far, that the 0.3 kb DNA fragment was that of the MyD88-DD gene, we still wanted to ensure that it was inserted in the correct orientation. The same eight pipette tips that were used to pick each colony, were also used to inoculate 15 mL aliquots of sterilized LB broth that contained kanamycin. The inoculated aliquots were incubated at 37° C and left on a shaking incubator overnight. Thereafter, a GeneJET plasmid miniprep kit was used to extract plasmids from the overnight cultures for all eight colonies.

The TOPO cloning kit comes with M13 forward and M13 reverse primers which are used to check if the gene of interest has been cloned in the correct orientation. There is also template DNA that can be used as a control to verify if the provided M13 primers are functional. We used these primers in lane 1 of Figure 23 to test their functionality by amplifying the provided DNA template. A PCR product with a molecular weight of 0.75 kb was observed. This was the expected molecular weight of the amplified DNA as per the kit manufacturers guide. The result therefore showed that the M13 primers were functioning as intended. Plasmids extracted from all eight colonies were used as template to verify the

correct orientation of the MyD88 death domain gene. The MyD88 forward primer and the M13 reverse primer were used for the amplification. A 0.5 kb band was observed in lanes 2 – 7 and there was no amplification in lanes 8 and 9. The M13 reverse primer adds a 0.2 kb to the amplified gene of interest, hence we see a 0.5 kb band. Taken together, the results from lanes 2 – 7 show that the MyD88 gene was cloned in the correct orientation.

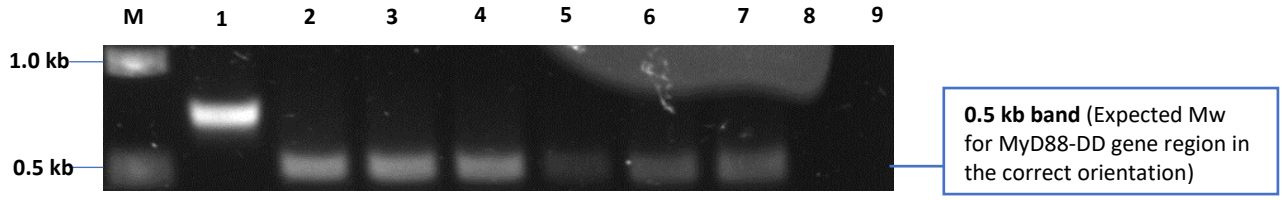


Figure 23. Agarose gel showing amplification of the MyD88-DD gene from the pENTR/D vector. PCR was conducted on the extracted TOPO reaction plasmids using the MyD88 forward primer and the M13 reverse primer. The PCR product was analysed by agarose gel electrophoresis on a 1 % agarose gel. Lane M – 1 kb DNA ladder. Lane 1 – Positive Control DNA supplied with TOPO cloning kit. Lane 2-9 - extracted plasmids from colony PCR verified samples.

We needed to further ascertain if the gene that was correctly inserted into the pENTR/D vector was indeed that of MyD88 death domain. The plasmids were therefore prepared for sequencing using the Big Dye V3 sequencing kit followed by Sanger sequencing using the MyD88 forward primer. The raw sequencing data was then analysed using the SnapGene® viewer program (version 5.3.1). Thereafter, once we were satisfied with quality of the data based on the clearly resolved peaks, we exported the data to the Expasy translation software. The result of the gene translation is shown in Figure 24.

The gene translation sequence shown in figure 24 is from a colony that had the highest DNA intensity in lane 2 of Figure 22. It shows the first amino acid of the translated gene sequence as a methionine. That methionine corresponds to the 54th amino acid of the MyD88 gene and is part of the death domain region. Amino acids downstream of that methionine, covered the entire MyD88 death domain sequence till the last glutamic acid (E110), which is annotated with an arrow. The next amino acid after E110, was found to be a stop codon; and it is annotated with a red circle. Therefore, the sequencing result shows that our cloning was successful. We had indeed cloned the MyD88 death domain gene into the pENTR/D vector.

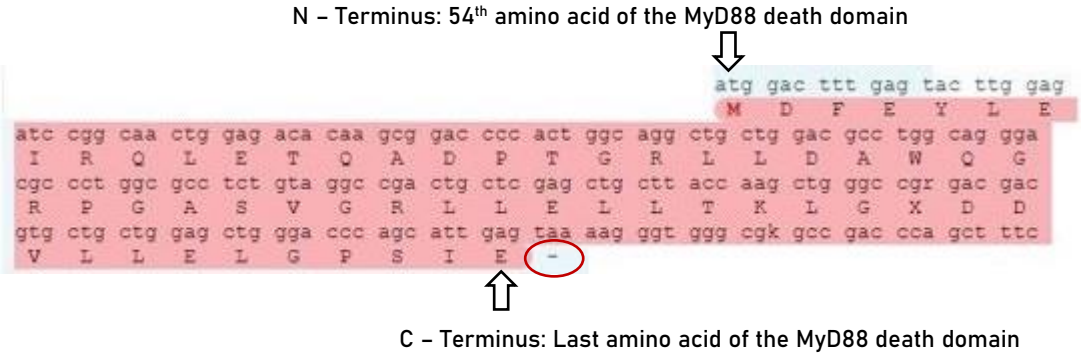


Figure 24. Expasy translation of raw DNA sequencing data for MyD88-DD-pENTR/D entry clone extracted from colony 1. Expasy software translation of sequencing data showing the amino acid sequence for the MyD88-DD. The first amino acid of the translation was a methionine residue which corresponded to the 54th amino acid of the MyD88 gene. The red circle indicates a stop codon on the first amino acid of the intermediate domain.

5.2.3 LR recombination of MyD88-DD-pENTR/D entry clone with pETG-destination vectors

Since we had cloned the MyD88 death domain gene into the pENTR/D vector, the next step was to transfer the gene to destination vectors using LR recombination. Gateway destination vectors are used for protein expression. We had four destination vectors at our disposal which were propagated in DB3.1 cells due to their toxicity in DH5α and DH10 cells. These were the pETG-10A vector (which confers a His-tag), the pETG-20A vector (which confers a His-Trx tag), pETG-30A vector (which confers a His-GST tag) and the pETG-41A vector (which confers a His-MBP tag).

Each individual destination vector was mixed with the pENTR/D vector and the LR Clonase enzyme was added to initiate the LR recombination. The LR recombination reaction was stopped after 2 hours by the addition of Proteinase K. The individual reaction mixtures were transformed into DH5α cells and grown overnight on LB agar plates that contained 100 µg/mL of ampicillin. Resulting colonies were picked and screened by PCR for the presence of MyD88 death domain. The PCR products were analysed by agarose gel electrophoresis as shown in Figure 25.

Lane 1 in Figures 25(A to D), was the positive control. The positive control showed the expected 0.3 kB band. All the colonies that we screened shown in figures 25(A to D) also contained the 0.3 kb band of varying but distinguishable intensity. The only exception was lane 6 of Figure 25(A) where the band intensity was very low, so we excluded this sample from further investigation. Nonetheless, the observed 0.3 kb band verified the presence of the MyD88-DD gene and confirmed to us that the LR recombination reaction was successful.

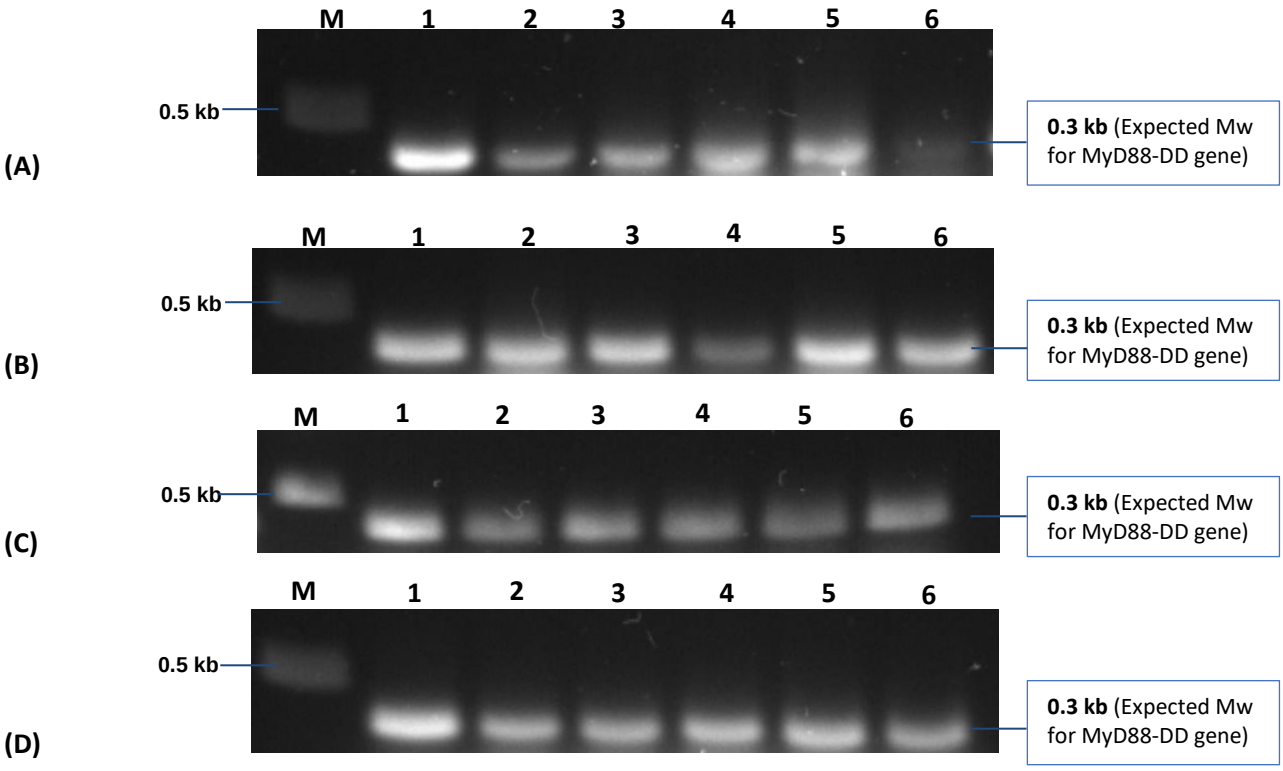


Figure 25. Agarose gel of colony PCR for LR recombination reactions. Colony PCR was conducted on DH5α cell colonies transformed with the LR recombination reaction of the MyD88-DD-pENTR/D entry clone with pETG destination vectors. The PCR products were analysed by agarose gel electrophoresis on a 1 % agarose gel. Sample loading was as follows: Lane M – 1 kb DNA ladder. Lane 1 –MyD88-DD-pENTR/D entry clone (positive control). Lanes 2-6 – screening of picked colonies. **(A)** LR recombination using the pETG-10A vector. **(B)** LR recombination using the pETG-20A vector. **(C)** LR recombination using the pETG-30A vector. **(D)** LR recombination using the pETG-41A vector.

5.2.4 Expression trials of MyD88-DD fusion proteins

Since we had successfully transferred the ID-deficient MyD88-DD gene into the four destination vectors namely, pETG-10A, pETG-20A, pETG-30A and pETG-41A, we now wanted to test for expression. These constructs were transformed into BL21 (DE3) cells and grown in 5 mL volumes of LB broth. The cultures were inoculated with IPTG to induce protein expression. Soluble and insoluble protein fractions were extracted from the induced cells using the BugBuster® protein extraction reagent and thereafter analysed by SDS-PAGE as shown in Figure 26.

Lane 1 in figure 26 was loaded with the soluble fraction expressed from the MyD88-DD-pETG-41A construct. It showed an assortment of protein bands and there was no prominent band detected. However, a very large distinct 54 kDa band was observed in lane 2 which was loaded with the insoluble fraction expressed from the MyD88-DD-pETG-41A construct. The 54 kDa protein band in the insoluble fraction was highly likely to be the His-MBP-MyD88-DD fusion protein. Lane 3 was loaded with the soluble fraction expressed from the MyD88-DD-pETG-30A construct and there were several faint protein bands of little significance that were observed. Lane 4 was loaded with the insoluble fraction expressed from the MyD88-DD-pETG-30A construct and there was a distinct 35 kDa protein band that was observed.

This band was highly likely to be the His-GST-MyD88-DD fusion protein based on its molecular weight. Lane 5 was loaded with the soluble fraction expressed from the MyD88-DD-pETG-20A construct and there were several faint protein bands that were observed. However, a 24 kDa protein band was observed in lane 6 which was loaded with the insoluble fraction expressed from the MyD88-DD-pETG-20A construct. This 24 kDa protein band, migrated at the expected molecular weight of the His-Trx-MyD88-DD fusion protein. Lane 7 was loaded with the soluble fraction expressed from the MyD88-DD-pETG-10A construct and there were also several faint protein bands of little significance that were observed. Lane 8 was loaded with the insoluble fraction expressed from the MyD88-DD-pETG-10A construct and there was a large 12 kDa band which migrated at the expected molecular weight of the His-MyD88-DD fusion protein.

The results from all the soluble fractions in Figure 26 did not show the expected proteins. Therefore, we could not determine if MyD88-DD fusion proteins were expressed in the soluble form. However, there was a possibility that they were expressed at levels that were beyond the detection of coomassie staining. Techniques like western blotting or silver staining would enable us to detect low expression levels of the fusion proteins. If there is one thing that was certain, it was that our fusion proteins formed inclusion bodies based on the high expression levels in the insoluble fraction.

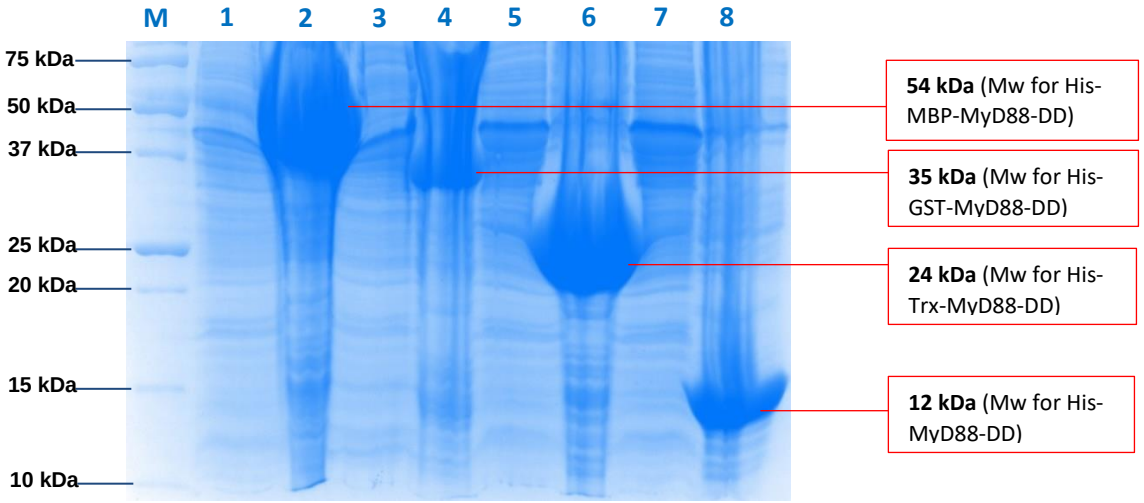


Figure 26. SDS gel of expression trials for LR recombination expression constructs. Small scale protein expression was conducted using the MyD88-DD expression clones from the LR recombination reactions. The extracted soluble [sol] and insoluble [ins] protein fractions from the expressions were analysed by SDS-PAGE on a 12 % SDS polyacrylamide gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa protein reference ladder. Lane 1 – His-MBP-MyD88-DD [sol]. Lane 2 – His-MBP-MyD88-DD [ins]. Lane 3 - His-GST-MyD88-DD [sol]. Lane 4 – His-GST-MyD88-DD [ins]. Lane 5 – His-Trx-MyD88-DD [sol]. Lane 6 – His-Trx-MyD88-DD [ins]. Lane 7 – His-MyD88-DD [sol]. Lane 8 – His-MyD88-DD [ins].

5.2.5 Western blot analysis of MyD88-DD fusion proteins.

We opted to do a western blot in order to determine if our four constructs expressed soluble protein since we failed to see it on a coomassie stained SDS gel. Another SDS gel that contained exactly the same fractions as Figure 26 was electrophoresed. The gel was not stained after it had electrophoresed. Instead, it was used to transfer proteins to a PVDF membrane. The blotted PVDF membrane was then blocked for non-specific binding using skimmed milk and thereafter incubated with an anti-His antibody. Two detection methods were used on this PVDF membrane which were chemiluminescence (shown in Figure 27(A)) and colorimetric detection (shown in Figure 27(B)).

We were able to observe chemiluminescent signals on the blot shown in Figure 27. These signals were also detected in soluble fractions. Our interest was only in the soluble fractions and hence the annotations shown in red circles were only done on them. A chemiluminescent signal was observed at the expected 54 kDa molecular weight of the His-MBP-MyD88-DD fusion protein in lane 1 of Figure 27(A) as annotated with a red circle. This lane had been loaded with the soluble fraction expressed from the MyD88-DD-pETG-41A construct. A very intense chemiluminescent signal was observed in lane 2 which had been loaded with the insoluble fraction of the MyD88-DD-pETG-41A construct. Such an intense signal had been expected because the stained gel loaded with the same fraction had a high concentration of protein that migrated as a smear. Lane 3 also had a chemiluminescent signal at the expected 37-kDa molecular weight of the His-GST-MyD88-DD fusion protein. This lane had been loaded with the soluble fraction from the MyD88-DD-pETG-30A construct. The insoluble fraction of the MyD88-DD-pETG-30A construct in lane 4 also had an intense chemiluminescent signal. However, its chemiluminescent signal did not show a smear when compared to lane 2. Lanes 5 and 6 contained the soluble and insoluble fractions from the MyD88-DD-pETG-20A construct. A chemiluminescent signal was observed in the soluble fraction at the expected 25 kDa molecular weight of the His-Trx-MyD88-DD fusion protein. However, the chemiluminescent signal was more intense in the insoluble fraction. Its level of intensity was comparable to that observed in lane 2. Lastly, chemiluminescent signals were also observed in lanes 7 and 8. These lanes contained fractions from the MyD88-DD-pETG-10A construct that expressed the His-MyD88-DD fusion protein. A signal was observed at 12 kDa in lane 7 which contained the soluble fraction. On the other hand, the insoluble fraction in lane 8 had a very intense signal that spread out as a smear.

The blotting pattern for the colorimetric detection was identical to that of the chemiluminescence as can be seen in Figure 27(B). Observations that were made in all the lanes of Figure 27(A) are exactly the same as in Figure 27(B). Since the observations are the same, we are not going to report them to avoid repetition.

Taken together, our western blot results show that all the four constructs do express soluble fusion protein. Therefore, our next aim was to upscale the production of the four fusion proteins to determine the most soluble one. Since all of them contain a His-tag, nickel affinity chromatography would be used to purify them.

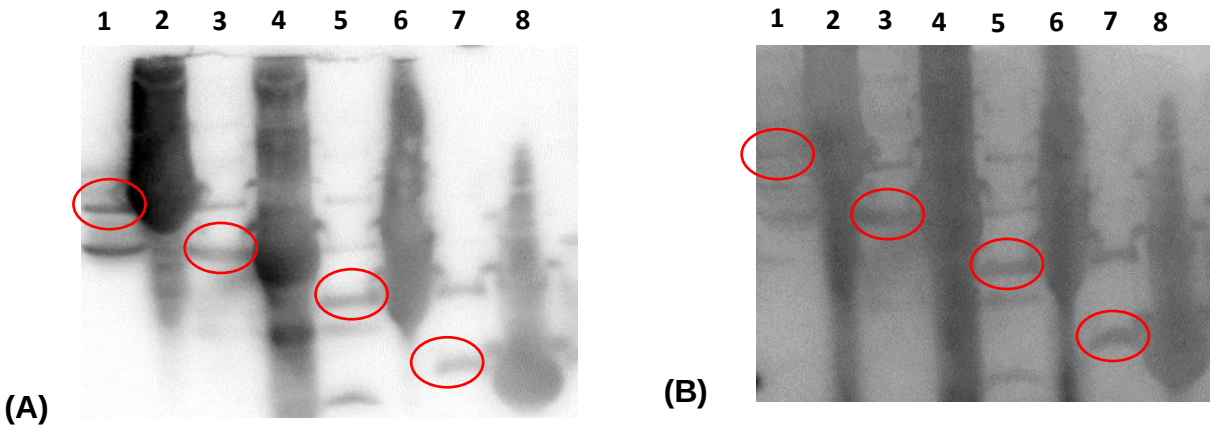


Figure 27. Western blot showing expression trial of MyD88-DD fusion proteins. Soluble and insoluble fractions of the expression trials for the ID-deficient MyD88-DD fusion proteins were transferred from an unstained SDS polyacrylamide gel to a PVDF membrane and analysed using two western bot detection methods. Target proteins in the soluble fractions were annotated by the red circles. Sample loading was consistent on all sub figures as follows: Lane 1 - His-MBP-MyD88-DD [sol]. Lane 2 – His-MBP-MyD88-DD [ins]. Lane 3 – His-GST-MyD88-DD [sol]. Lane 4 – His-GST-MyD88-DD [ins]. Lane 5 – His-Trx-MyD88-DD [sol]. Lane 6 – His-Trx-MyD88-DD [ins]. Lane 7 – His-MyD88-DD [sol]. Lane 8 – His-MyD88-DD [ins]. **(A)** Chemiluminescence detection using luminol substrate of the ECL kit. **(B)** Colorimetric detection using TMB enhanced one component HRP membrane kit.

5.2.6 Nickel affinity chromatography purification of MyD88-DD fusion proteins

Since we were aware that MyD88-DD proteins are more soluble when cultured at lower temperatures, we decided to implement this temperature change for an upscaled 4 litre production of each expression construct. Each expression construct was cultured in a total of 4 litres (4 x 1L) of LB broth. They were then induced with IPTG when they reached an optical density (OD_{600}) of 0.8. The temperature was then lowered to 15° C and the cultures were left to grow overnight. Cells were then harvested, lysed by sonication and the lysate was centrifuged at 37 500 x *g*. The resulting soluble fraction was purified by nickel affinity chromatography. Fractions that were eluted from the nickel column were collected and analysed by SDS-PAGE as shown in Figure 28.

Three fractions of the nickel affinity purified samples were collected from the expressed MyD88-DD-pETG-41A construct. A Bradford assay was done on the samples in order to have an idea if there was protein in the fractions prior to loading on SDS-PAGE as shown in Figure 28(A). All the three fractions that were assayed, showed a deep blue colour. A deep blue colour is an indication that there is a high concentration of protein. Since there was confirmation of the presence of protein in the samples, they were then analysed on SDS-PAGE as shown in Figure 28(A). All the three lanes showed

intense 54 kDa protein bands. These bands were highly likely to be the His-MBP-MyD88-DD fusion protein based on the observed molecular weight. Of great importance to us, is that the yields of soluble protein were high based on the high intensity of the bands observed on the SDS gel.

Similarly, three fractions of the nickel affinity purified samples were collected from the expressed MyD88-DD-pETG-30A construct. A Bradford assay was also done on the samples in order to have an idea if there was protein in the fractions prior to loading on SDS-PAGE as shown in Figure 28(B). Two of the three fractions that were assayed, showed a deep blue colour. This was an indication that there were high concentrations of protein in the two samples. The third sample that did not show a deep blue colour had less protein concentration. All the three samples were loaded on SDS-PAGE as shown in Figure 28(B). SDS-PAGE analysis of the three samples showed a 37 kDa protein band which decreased in intensity from lane 1 to lane 3. This band was highly likely to be the His-GST-MyD88-DD fusion protein based on the observed molecular weight. There was also a 20-25 kDa band observed in all three lanes which we assumed was GST tag. Protein yields observed from these fractions were satisfactory, but significantly lower than the yields for the presumed His-MBP-MyD88-DD protein in Figure 28(A).

Three fractions of the nickel affinity purified samples were collected from the expressed MyD88-DD-pETG-20A construct. Thereafter, the samples were assayed from the presence of protein using a Bradford assay as shown in Figure 28(C). The assayed samples showed a very pale light blue colour that was borderline grey for all the three fractions. This suggested very little protein was present. The SDS-PAGE confirmed our initial observation because there were very faint 24 kDa protein bands observed in all the three lanes as shown in Figure 28(C). The intensity of the bands was significantly less when compared to the protein bands we observed in Figures 28(A) and 28(B). However, the 24 kDa protein band corresponded to the expected molecular weight of the His-Trx-MyD88-DD fusion protein. These low yields ruled out the possibility of us continuing further experiments using the MyD88-DD-pETG-20A construct.

The three protein samples that were analysed in figure 28(D) were from the expression of the MyD88-DD-pETG-10A construct. A Bradford assay of the first and third sample showed a very pale light blue colour, which led us to believe there was very little protein. The second sample on the other hand, showed a deeper blue colour in comparison to the other two fractions, which suggested a high protein concentration. SDS-PAGE analysis of the sample in lane 1, showed a very faint 16 kDa protein band. The second sample in lane 2 showed the presence of two bands at 16 and 12 kDa. A molecular weight of 12 kDa corresponded to the expected size of the His-MyD88-DD fusion protein. However, the identity of the 16 kDa protein band could not be deduced from the gel. The third sample was loaded in lane 3 and it showed a very faint 12 kDa protein band. Therefore, the yields of our target fusion protein were too low to conduct further experiments with the MyD88-DD-pETG-10A construct.

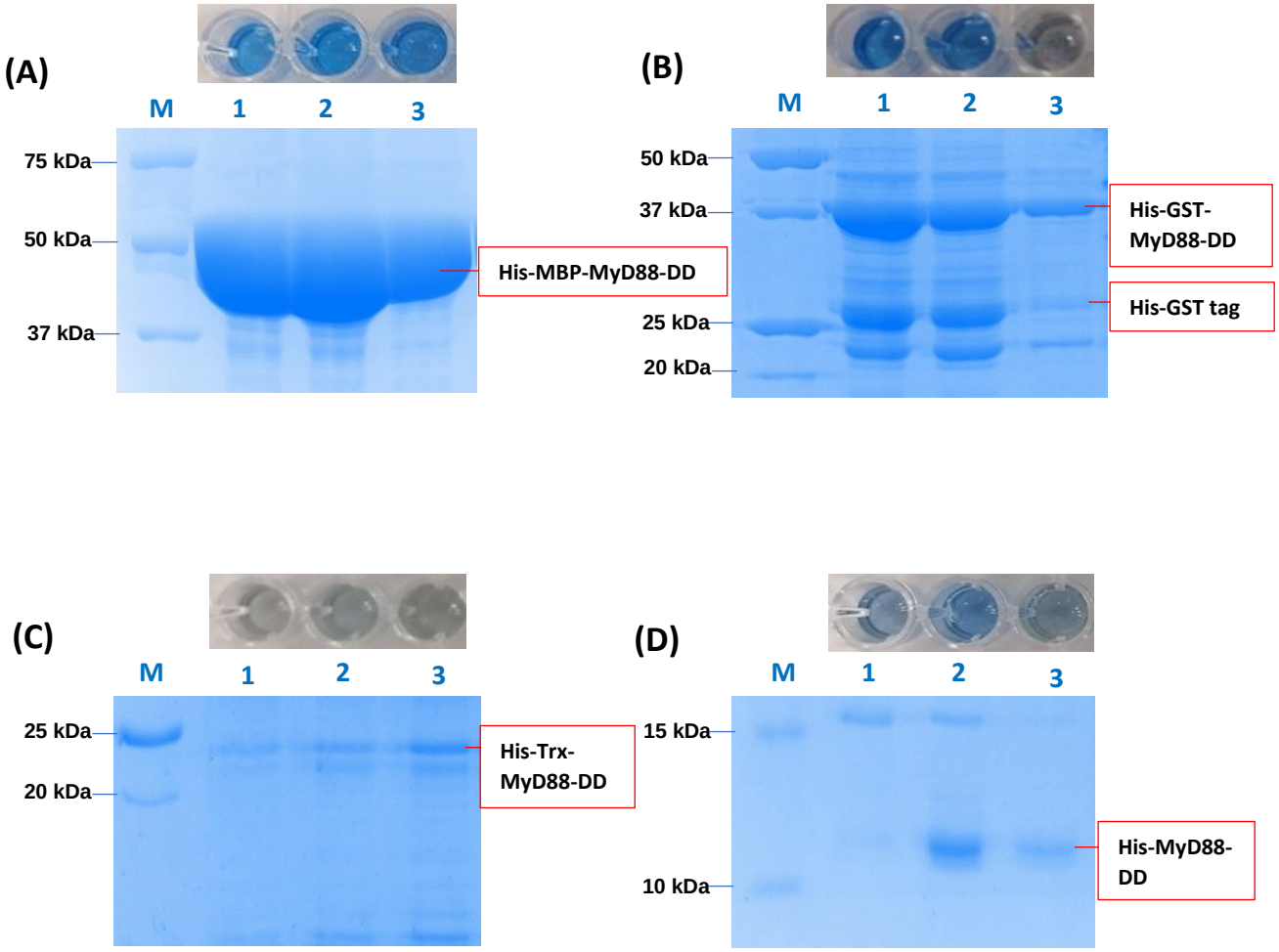


Figure 28. Bradford assay and SDS gels of nickel affinity chromatography purified MyD88-DD fusion proteins. Nickel affinity chromatography purification for large scale productions of the four MyD88-DD fusion proteins was conducted. 1 mL fractions were eluted during the purification and these fractions were analysed using a Bradford detection assay. The samples were further analysed by SDS-PAGE and 18 µL of these samples were loaded in lanes 1-3 in the order they were eluted. Lane M was loaded with a 250 kDa reference marker. **(A)** His-MBP-MyD88-DD protein (10 % resolving-gel). **(B)** His-GST-MyD88-DD protein (12 % resolving-gel). **(C)** His-Trx-MyD88-DD (12 % resolving-gel) **(D)** His-MyD88-DD (12 % resolving-gel)

5.2.7 Bradford assay to quantify the His-MBP-MyD88-DD protein.

The results from the previous section showed that the His-MBP-MyD88-DD fusion protein had a significantly higher yield than the rest of the fusion proteins (Figure 28) based on the high levels of intensity of the bands observed on the SDS gel. As such, we decided to focus only on this fusion protein for subsequent purification steps. However, before we could start with the purification, we wanted to get an estimate of the yields of our chosen fusion protein. Therefore, we decided to conduct a Bradford assay to quantify these yields. To conduct the Bradford assay, we first setup a 2-fold dilution series of BSA standards (starting with a 2 mg/mL concentration). These BSA standards were added to a 1X Bradford reagent in triplicate and then immediately analysed using a Spectra-paradigm. Absorbance readings for the samples were captured at a wavelength of 595 nm. The readings were then exported to a spreadsheet (Microsoft Excel) as shown in Table 31. The calculated mean averages of the absorbance readings were then plotted against their corresponding standards to generate a standard curve shown in Figure 29.

A 5-point standard curve with very small standard deviation error bars (+0.013) at each point is shown in Figure 29. The calculated R-squared regression value for these points across a line of best fit was 98.48 % which suggested that there was very little variance in our datapoints. We were satisfied with this R-squared value. Though, it was worth noting that the datapoint at 1 mg/mL was the only point that was not directly on the trendline, suggesting it may have been an outlier. Nonetheless, the straight-line equation for our trendline was calculated and shown to be $y = 0{,}3502x + 0{,}1013$. This was the equation we used to calculate the concentrations of unknown protein.

Table 31. Absorbance readings for BSA standards

Concentrations (mg/mL)	**BLANK**	**2**	**1**	**0.5**	**0.25**	**0.125**
1st Repeat	-0,0035	0,7700	0,4906	0,3038	0,1930	0,1194
2nd Repeat	0,0044	0,7861	0,5055	0,2976	0,1766	0,1209
3rd Repeat	-0,0009	0,7710	0,5101	0,2745	0,1594	0,1113
Mean average	0	0,7757	0,5020	0,2919	0,1763	0,1172
Standard deviation	0,0040	0,0090	0,0101	0,0154	0,0168	0,0051

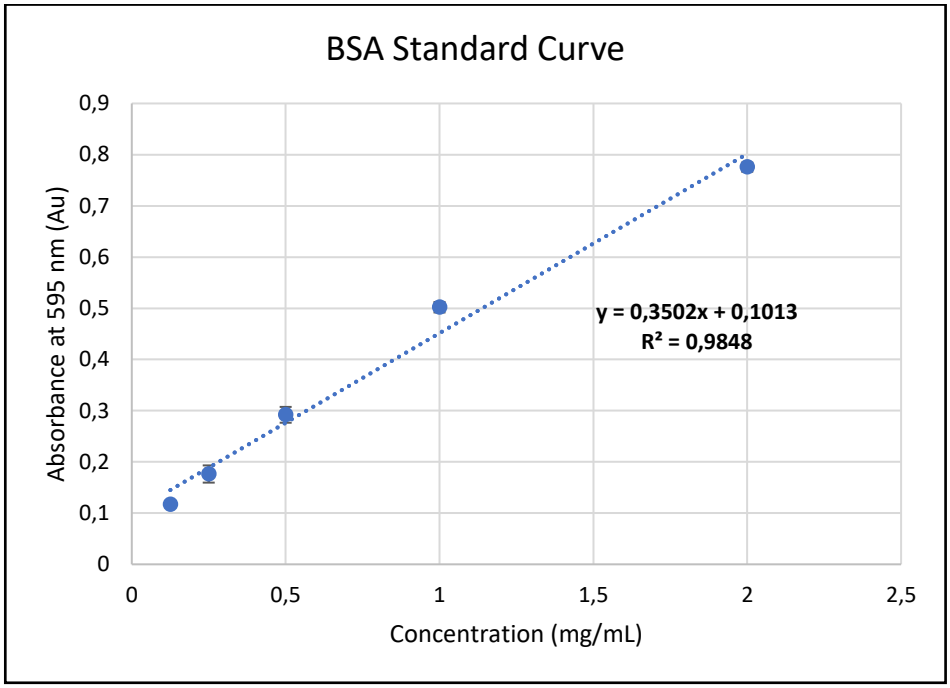
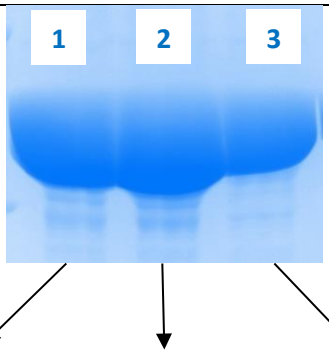


Figure 29. BSA standard curve for determining concentrations of unknown protein using the Bradford assay. Graph of BSA protein standard absorbance readings at 595 nm plotted against protein concentration. The R-squared value for the dataset was shown as R^2 and the trendline equation was shown in the y=mx + c format with calculated values for m and c calculated by Microsoft Excel.

Once we had our equation to quantify our unknown protein concentrations, we set up a Bradford assay for the three His-MBP-MyD88-DD fusion protein fractions we purified by nickel affinity chromatography in section 5.2.6. Since the protein bands in Figure 28(A) suggested there was a very high yield of protein. Therefore, we took 10 µL aliquots of each fraction and did serial dilutions with 20 mM Tris (pH 8.0), 30 mM NaCl till we got a 1:100 dilution. The Tris buffer is the same one that was used for the BSA standards to maintain consistency. The diluted His-MBP-MyD88-DD protein fractions were added to 1X Bradford reagent followed by analysis in a Spectra-paradigm. The absorbance values were read at 595 nm wavelength then exported to Microsoft Excel. The mean average of the absorbance readings was calculated as well as the protein concentrations using our BSA standard line equation as shown in Table 32. Therefore, the calculated concentrations of our three protein fractions (considering the 1:100 dilution factor) were: 10.66 mg/mL, 24,94 mg/mL and 7.12 mg/mL. We believed that these high yields of total fusion protein would be sufficient to carry out our study.

Table 32. Protein concentrations for His-MBP-MyD88-DD nickel affinity chromatography fractions

Eluted fractions			
	1st fraction	**2nd fraction**	**3rd fraction**
1st Replicate absorbance	0,1360	0,1957	0,1237
2nd Replicate absorbance	0,1417	0,1874	0,1258
3rd Replicate absorbance	0,1382	0,1828	0,1292
Mean average absorbance	0,1386	0,1886	0,1262
Mean protein concentration (mg/mL) calculated by substitution into BSA standard equation: *y = 0,3502x + 0,1013*	0,1066	0,2494	0,0712

**1:100 dilution of the analysed samples*

5.2.8 Prescission protease cleavage of His-MBP-MyD88-DD fusion protein

Protein quantification of His-MBP-MyD88-DD in the previous section, suggested we had an average yield of 14 mg/mL from the pooled three 1 mL fractions shown in Figure 28(A). This yield was good enough to conduct further studies. Therefore, the next step was to cleave this fusion protein to obtain the ID-deficient MyD88-DD protein. But before we could proceed with cleaving the protein, we needed to have a negative control in the form of an uncleaved sample. Therefore, 20 µL of uncleaved protein was taken from the pooled sample and set aside as a negative control, while the remaining pooled sample was cleaved with 50 µg of Prescission protease at 25° C for 2 hours. Thereafter, both the cleaved and uncleaved protein samples were analysed by SDS-PAGE as shown in Figure 30.

An intense 54 kDa band emanating from the uncleaved sample was observed in lane 1 of figure 30. This band was highly likely to be the His-MBP-MyD88-DD fusion protein as expected for the negative control. Lane 2 was loaded with the cleaved sample and three intense bands were observed with molecular weights of 54 kDa, 42 kDa and 12 kDa. This observation showed that protein cleavage did occur. We were therefore convinced that the 54 kDa protein band was some uncleaved His-MBP-MyD88-DD fusion protein because there was less of it when compared to the negative control. There was also an appearance of a 42 kDa protein which is absent in the uncleaved negative control.

This 42 kDa band is definitely the His-MBP tag based on its molecular weight. The 12 kDa band on the other hand is definitely the ID-deficient MyD88-DD protein. This result means we were successful in expressing and purifying soluble ID-deficient MyD88-DD protein.

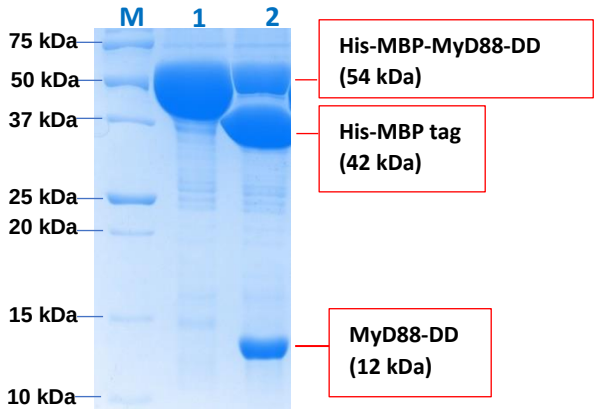


Figure 30. SDS gel of nickel affinity chromatography purified and Prescission protease cleaved His-MBP-MyD88-DD fusion protein. Purified His-MBP-MyD88-DD fusion protein was cleaved by Prescission protease. The cleaved sample as well as an uncleaved control sample were analysed by SDS PAGE on a 10 % SDS polyacrylamide gel. 18 µL samples from these fractions were loaded as follows: Lane M – 250 kDa molecular weight reference ladder. Lane 1 – Uncleaved fusion protein (negative control). Lane 2 – Cleaved protein.

5.2.9 Size exclusion chromatography purification of MyD88-DD protein.

The cleaved His-MBP-MyD88-DD sample from the previous section contained three proteins, with MyD88-DD among them. Based on the intensity of the visualised bands on SDS-PAGE (Figure 30), the yields of protein were satisfactory. Our next objective was to isolate MyD88-DD from the other two proteins in the sample. We hoped to accomplish this task by using size exclusion chromatography (SEC). Therefore, the cleaved His-MBP-MyD88-DD protein sample was concentrated to a volume of 500 µL and thereafter loaded into an ENrich SEC 650 (10 x 300 mm) column which was connected to an ÄKTA FPLC machine. The SEC purification resulted in a chromatogram in Figure 31 which shows the presence of five peaks and their corresponding fractions. The peaks were labelled as peak 1 up to 5. Each peak had its own retention volume and absorbance as shown in Table 33. Peak 1 had a retention volume of 10.51 mL and an absorbance of 380 mAU. Peak 2 had a retention volume of 13.04 mL and an absorbance of 41 mAU. Peak 3 had a retention volume of 15.41 mL and an absorbance of 76 mAU. Peak 4 had a retention volume of 16.73 mL and an absorbance of 1750 mAU. Peak 5 had a retention volume of 19.52 mL and an absorbance of 15 mAU. The results clearly show that peaks 1 and 4 had the highest concentration of protein based on the high absorbance values.

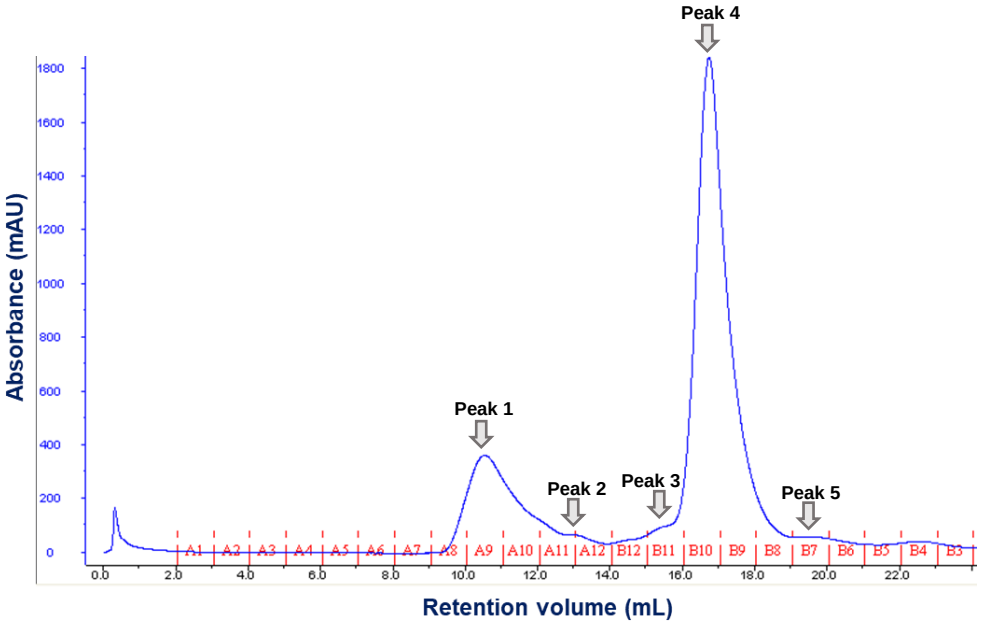


Figure 31. Chromatogram for SEC purification of cleaved His-MBP-MyD88-DD. The cleaved His-MBP-MyD88-DD protein sample was purified by SEC and a chromatogram was generated from this purification by an ÄKTA FPLC machine. UV absorbance readings at 280 nm were captured for prominent peaks and collected as follows: Peak 1 – [fraction: A9]. Peak 2 – [A11 and A12]. Peak 3 – [B11]. Peak 4 – [B10]. Peak 5 – [B7].

Table 33. SEC data output for cleaved His-MBP-MyD88-DD protein

	Retention volume (mL)	Absorbance (mAU)
Peak 1	10.51	380
Peak 2	13.04	41
Peak 3	15.41	76
Peak 4	16.73	1750
Peak 5	19.52	15

We next wanted to check if size exclusion chromatography had separated the ID-deficient MyD88-DD protein from the His-MBP tag. Samples of fractions that eluted from the five peaks were loaded and run on SDS-PAGE in the order of their elution. A sample from peak 1 was loaded in lane 1 as shown in Figure 32 and a 12 kDa protein band was observed. This 12 kDa band is the ID-deficient MyD88-DD protein. A sample from peak 2, loaded in lane 2 also showed the presence of the ID-deficient MyD88-DD protein, in addition to the His-MBP-MyD88-DD. A sample from peak 3, loaded in lane 3 showed very faint protein bands, none of which were noteworthy. A sample from peak 4, loaded in lane 4 showed a very intense 42 kDa His-MBP tag. A sample from peak 5, loaded in lane 5, showed very faint double bands with an approximate molecular weight of 42 kDa. Therefore, taken together the results suggested that SEC was successful in purifying our protein sample.

The identification of the 12 kDa ID-deficient MyD88-DD from peak 1 sample was totally unexpected. Peak 1 has a retention volume of 10.51 mL and we expect high molecular weight proteins to elute at that volume based on calibrations that were previously done on the column. This result suggested that the ID-deficient MyD88-DD protein was forming oligomers.

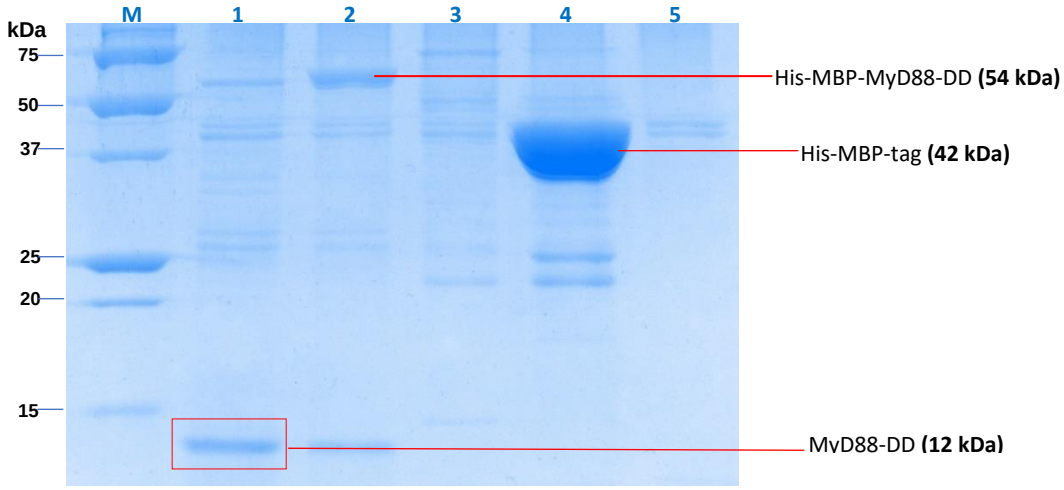


Figure 32. SDS gel of SEC purification fractions for cleaved His-MBP-MyD88-DD. Fractions for the peaks that were identified in SEC purification of cleaved His-MBP-MyD88-DD protein were collected and analysed by SDS-PAGE on a 10 % SDS gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa protein marker. Lane 1 – peak 1. Lane 2 - peak 2. Lane 3 – peak 3. Lane 4 – peak 4. Lane 5 – peak 5.

5.2.10 DLS analysis of ID-deficient MyD88-DD protein

While the SEC purification was a success in producing a purer sample of the ID-deficient MyD88-DD protein, we were very intrigued by the retention volume we observed for this protein. The earlier SEC elution of this 12 kDa MyD88-DD protein before the larger 42 kDa His-MBP tag suggested that our target protein was much larger in solution and possibly forming oligomers. Therefore, we wanted to determine if it was forming oligomers by using dynamic light scattering (DLS). DLS is able to distinguish an oligomer from an aggregate. The 12 kDa MyD88-DD protein was analysed by DLS using a Malvern Zetasizer Nano ZS as shown in Figure 33.

DLS analysis in Figure 33, showed the presence of three peaks. The first peak had a hydrodynamic diameter of 212.5 nm and constituted 97.2 % of the total molecules. The second peak had a hydrodynamic diameter of 5016 nm and represented 2.2 % of the total molecules of the sample. The third peak had a hydrodynamic diameter of 41 nm and represented 0.6 % of the sample. The results showed that the MyD88-DD protein sample had an average size of 189.1 nm and a low polydispersity index of 0.274. This result suggested that our sample was not an aggregate, but an oligomer.

System

Temperature (°C):	4.0	Duration Used (s):	60
Count Rate (kcps):	305.6	Measurement Position (mm):	4.65
Cell Description:	Disposable sizing cuvette	Attenuator:	8

Results

			Size (d.nm):	% Intensity:	St Dev (d.n...
Z-Average (d.nm):	189.1	Peak 1:	212.5	97.2	87.29
PdI:	0.274	Peak 2:	5061	2.2	610.5
Intercept:	0.952	Peak 3:	41.82	0.6	6.962
Result quality :	Good				

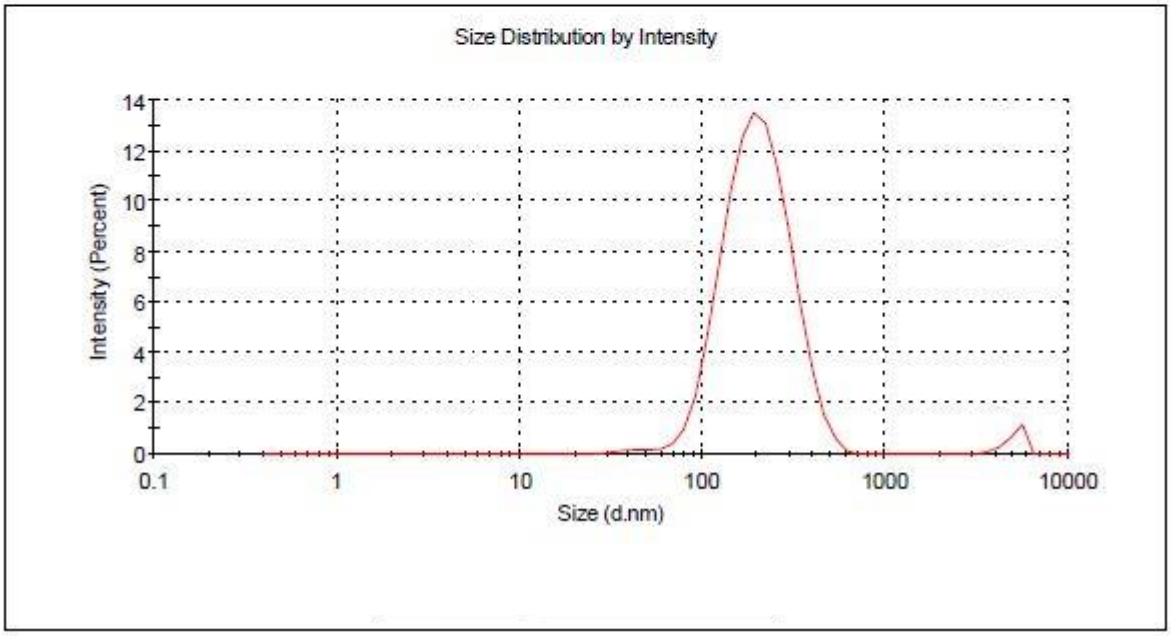


Figure 33. DLS profile of ID-deficient MyD88-DD protein. MyD88-DD protein fraction (from peak 1 of the SEC purification in Figure 27) was analysed through DLS using the Zetasizer Nano. Size determination was calculated (through the hydrodynamic diameter) and correlated to the relative intensity distribution to generate a plot denoted by the red trendline.

5.3 Discussion

In this chapter, our goal was to express and purify an ID-deficient MyD88-DD protein. Since we obtained very low yields of this protein from the site directed mutagenesis construct in the previous chapter, we opted to use the gateway cloning system to produce expression clones with different tags. This allowed us to assess the protein yields of these varying constructs efficiently given that the cloning would only be done once using this gateway cloning system (Reece-Hoyes & Walhout, 2018). Ideally, we were striving to obtain protein yields that are higher than the one obtained from the site directed mutagenesis construct.

We started our cloning strategy by designing MyD88 primers to amplify the ID-deficient MyD88-DD gene. The forward primer was designed to have the critical "CACC" sequence required for annealing with the pENTR/D vector. This "CACC" sequence was at the beginning of the forward primer, followed by a Prescission protease recognition sequence for downstream removal of fusion tags. Both the forward and reverse primers were used to amplify the MyD88-DD and an expected 0.3 kb gene fragment was obtained. The gene fragment was then mixed with the pENTR/D vector, resulting in its annealing and ligation by the two topoisomerases that are attached to the vector. Thereafter, the presence of the MyD88 gene was verified by DNA sequencing. This then made the pENTR/D vector that has the MyD88 gene our master clone. The master clone is transcriptionally silent (Esposito et al., 2009).

The verified MyD88 gene from the master clone, was transferred to four expression vectors namely pETG-10A, pETG-20A, pETG-30A and pETG-41A by LR recombination. These vectors express fusion proteins that have a His-tag (pETG-10A), His-Trx tag (pETG-20A), His-GST tag (pETG-30A) and His-MBP tag (pETG-41A). Furthermore, these vectors have attachment sites (*att* sites) which are used during recombination. There are at least four different types of attachment sites namely *attP, attB, attL, and attR* (Esposito et al., 2009). Our chosen expression vectors have the *attR* sites, while our pENTR/D vector has the *attL* site. These sites get swapped when the gene from the master clone is transferred to our chosen expression vectors. The process of transferring the MyD88 gene to our expression vectors is known as LR recombination. LR recombination means that the *attL* sites from the master clone are swapped with the *attR* sites in our chosen expression vectors. Once the expression clones are generated, they do not need to be sequenced because the *att* sites are site specific. More importantly, the process of recombination ensures that the gene of interest, MyD88 death domain in our case, is in the correct reading frame.

Since there was no need to sequence our four expression clones, we went ahead to test them for expression. We found that all of them were expressing the MyD88-DD fusion proteins of the expected molecular weights. However, expression was only detected in the insoluble fractions. This is a typical occurrence in heterologous protein expression whereby the recombinant protein is improperly folded and becomes trapped in inclusion bodies (Bhatwa et al., 2021). An absence of MyD88-DD in the soluble

fractions did not deter us from conducting western blots to determine with certainty if there was indeed no protein. Therefore, anti-His tag antibodies were used to detect the N-terminal His-tag that existed on all the expressed fusion proteins. Two visualisation techniques namely chemiluminescence and colorimetry were used to detect the presence of the MyD88-DD fusion from the blot. Both chemiluminescence and colorimetry showed that all the four expression clones produced soluble MyD88-DD fusion proteins. However, there was more background noise in colorimetry when compared to chemiluminescence. Studies have shown that this is usually the case due to the varying levels of sensitivity by the two techniques (Mishra et al., 2019).

The positive results obtained from the western blots implied that all the four expression clones expressed very little quantities of soluble MyD88-DD fusion protein. We therefore upscaled our protein production by growing larger cultures of cells. In addition, we lowered the culturing temperature after IPTG induction to 15° C because it has been shown that such low temperature improves the folding dynamics of heterologous protein production which promotes a more soluble end product (Gopal & Kumar, 2013). As expected, lowering the temperature and upscaling yielded more soluble MyD88-DD fusion protein on all the four clones based on the bands that were observed on SDS gels and the Bradford assay. However, the His-MBP-MyD88-DD fusion yielded the most amount of soluble protein when compared to the other three fusion proteins. Therefore, we opted to use it for subsequent purification steps.

Our initial attempt to quantify the His-MBP-MyD88-DD fusion resulted in high absorbance values. According to the standardised practice in spectrophotometry, samples should be adjusted to produce absorbance values that are less than one in accordance with the Beer-Lambert laws within a linear range (Khattab, Ghanem, & Brooks, 2017). Values greater than one, indicate high concentration of the sample being analysed. Therefore, we did a 1:10 dilution ratio of our protein fractions, but the absorbance values were still greater than 1. As such we diluted our samples further to a 1:100 dilution ratio. It was at this point that our absorbance values were finally less than 1. Therefore, we were now able to quantify our protein sample using the BSA standard curve equation. We obtained an average yield of 14 mg of protein which we were satisfied to proceed with.

The His-MBP-MyD88-DD protein was cleaved with Prescission protease to separate the MyD88-DD from its fusion partner His-MBP. Cleavage occurred at the Prescission protease recognition site which confers specificity. Therefore, we were confident that the 12 kDa band that was observed on the SDS gel was that of the MyD88-DD protein. However, the presence of the 54 kDa His-MBP-MyD88-DD fusion protein showed that there was incomplete cleavage. Factors that contribute to instances of incomplete cleavage include the length of incubation for cleaving as well as the incubation temperature (Zhu et al., 2017). We opted not to extend the cleaving time because it is associated with loss of protein stability and eventually degradation of the target protein (Wingfield, 2015).

The cleaved His-MBP-MyD88-DD fusion protein, was purified by size exclusion chromatography (SEC) in order to isolate the MyD88-DD. This technique separates biomolecules based on size and shape (Fekete, Beck, Veuthey, & Guillarme, 2014). Of great significance to us, was that the MyD88-DD with a molecular weight of 12 kDa eluted earlier than the larger 42 kDa His-MBP tag. This observation suggested that the MyD88-DD could either be an aggregate or an oligomer. The possibility of it being an aggregate was unlikely because it eluted within the column volume and not in the void volume. Therefore, it was highly likely to have oligomerised. Evidence of MyD88 oligomerising has been shown by cryo-electron microscopy (Moncrieffe et al., 2020). It was found to form larger oligomers. This finding is in support of the results that we obtained.

To further substantiate our finding that MyD88-DD was an oligomer, we conducted dynamic light scattering (DLS). This technique uses visible light that is focused on a suspension of particles that scatter this light when they are in Brownian motion (Falke & Betzel, 2019). The profile of the scattered light is then used to determine the hydrodynamic diameter of the particles and to determine if they are aggregates based on the polydispersity index. DLS data showed that the MyD88-DD sample had a low polydispersity index of 0.274, meaning that the sample was homogenous. It also showed that MyD88-DD had a hydrodynamic diameter 189.1 nm. This hydrodynamic diameter is larger than that of the monomeric intermediate domain inclusive MyD88-DD which was found to be 135 nm (George et al., 2011). Since the difference between these two proteins was the intermediate domain, we postulated that the absence of the ID likely led to the formation of the larger oligomer. Hence, we were very keen to see what effect (if any) this would have on reconstitution of the Myddosome complex given our observations in this chapter.

5.4 Conclusion

We cloned and expressed soluble MyD88-DD from four different vectors namely pETG-10A, pETG-20A, pETG-30A and pETG-41A. pETG-41A which expressed as a His-MBP fusion protein, produced the highest yield of protein when compared to the other three expression vectors. The MyD88-DD was purified and thereafter characterised using two biophysical techniques namely size exclusion chromatography and dynamic light scattering. Both techniques showed that it was an oligomer and not an aggregate.

Chapter 6: Expression, purification and analysis of IRAK-4-DD protein

6.1 Introduction

We produced MyD88 death domain with and without intermediate domain in chapter 3 and chapter 5 respectively. However, before we could conduct our protein-protein interaction studies to investigate our research question, we still needed the IRAK-4 death domain protein. This protein binds MyD88-DD proteins to form the Myddosome complex (Motshwene et al., 2009). Therefore, this chapter is about the expression of the IRAK-4-DD fusion protein from a construct that was previously cloned in our research group. Our workflow for this chapter includes, DNA sequencing, nickel affinity chromatography purification and size exclusion chromatography purification.

6.2. Results

6.2.1 DNA sequencing of the IRAK-4-DD-pETG-30A plasmid.

An IRAK-4-DD-pETG-30A expression clone that produces IRAK-4-DD protein had been cloned by my predecessors. This construct had worked very well in previous studies and was therefore made available to us to conduct our studies. However, we firstly needed to verify the presence of the IRAK-4-DD gene in this expression clone. Therefore, the construct was cleaned using the Big Dye V3 sequencing kit followed by Sanger sequencing using an IRAK-4-DD reverse primer that was available in our research group. The raw sequencing data was then analysed using the SnapGene® viewer program (version 5.3.1). Thereafter, once we were satisfied with quality of the data based on the clearly resolved peaks, we exported the data to the Expasy translation software as shown in Figure 34.

The translated gene sequence in Figure 34 shows a glycine residue as the first amino acid. Downstream of this glycine, we identified a Prescission protease recognition sequence as annotated in green on the figure. Thereafter, we observed a serine residue followed by a methionine. This methionine was identified as the 1st amino acid of the IRAK-4 death domain gene as the amino acids downstream of it matched the known amino acid sequence of the IRAK-4 gene. The last residue for the sequence was a threonine, which we identified as the 80th amino acid of the IRAK-4-DD gene. Therefore, it was evident that we sequenced the IRAK-4 gene.

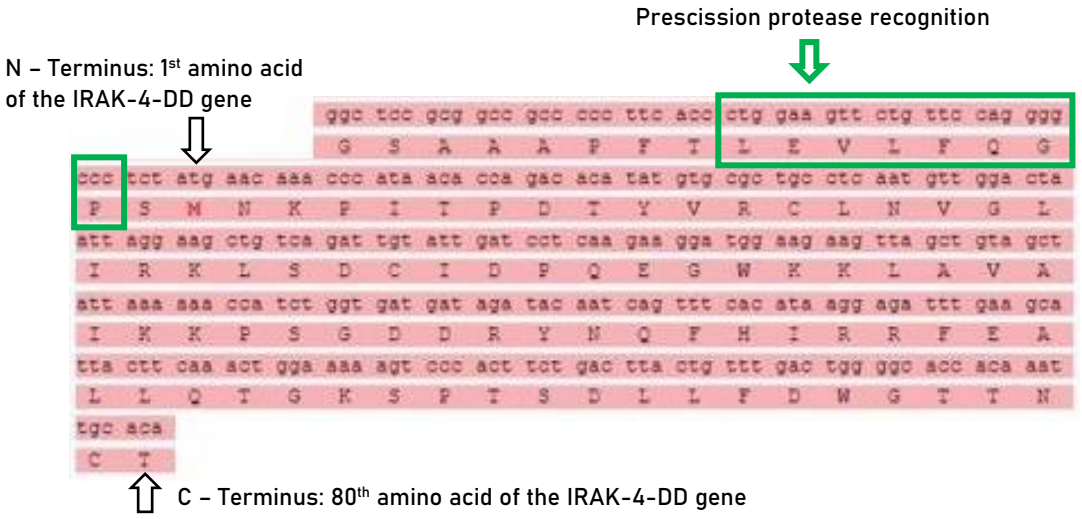


Figure 34. Expasy translation of raw DNA sequencing data for IRAK-4-DD-pETG-30A plasmid. Expasy software translation of sequencing data with pink highlighted region showing the amino acid sequence for the IRAK-4-DD gene. The annotated green box shows the Prescission protease recognition sequence. The methionine residue shown by the arrow in the figure corresponded to the first amino acid of the IRAK-4-DD gene. The last residue in the translation was a threonine which corresponded to the 80th amino acid of the IRAK-4-DD gene.

6.2.2 Nickel affinity chromatography purification and Prescission protease cleavage of His-GST-IRAK-4-DD protein

Having confirmed that we were working with the correct IRAK-4-DD-pETG-30A expression clone, our next objective was to produce the IRAK-4-DD fusion protein. This clone had previously expressed well (George et al., 2011) and as such, we decided to directly do a large scale production. Therefore, 4 litres of culture were grown at 37° C on a shaking incubator and induced with 1 mM IPTG at an OD_{600} of 0.8. The temperature was then lowered to 15° C and the cells were left to grow overnight. The cells were harvested the following day and lysed by sonication. The lysate was centrifuged after which the supernatant was collected as the soluble fraction. This soluble fraction was then purified by nickel affinity chromatography using a 5 mL HiTrap chelating HP column. From the purified sample, 20 µL was aliquoted and kept aside as the uncleaved control. The remainder of the purified sample was mixed with 50 µg of Prescission protease enzyme and incubated for 2 hours at 25° C to cleave the fusion protein. Afterwards, both the cleaved and uncleaved samples were analysed by SDS-PAGE as shown in Figure 35.

A 37 kDa band was observed in the uncleaved sample in lane 1. Its molecular weight corresponded to that of the expected His-GST-IRAK-4-DD fusion protein. Lane 2 on the other hand was loaded with the cleaved protein sample and it showed the presence of two prominent protein bands having molecular weights of 12 kDa and 25 kDa. When compared to lane 1, the 37 kDa His-GST-IRAK-4-DD fusion protein was absent in lane 2. This showed that there was complete cleavage of the His-GST-IRAK-4-DD fusion protein. Since the Prescission protease has a specific recognition sequence, it is highly unlikely

that another protease would have cleaved the His-GST-IRAK-4-DD fusion protein. Therefore, our analysis of the bands observed in lane 2 is that the 25 kDa band is the His-GST tag, while the intense 12 kDa band is the IRAK-4-DD protein. Furthermore, the observed yields in lane 2 based on the intensity of the two bands, were enough for use in reconstituting the Myddosome complex in later experiments.

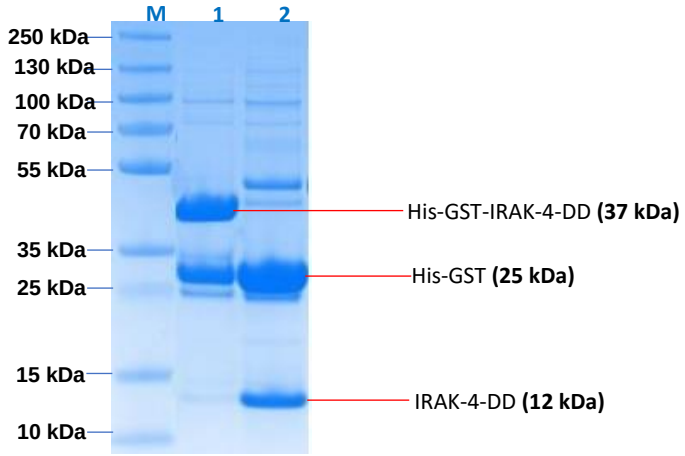


Figure 35. SDS gel of nickel affinity chromatography purified and Prescission protease cleaved His-GST-IRAK-4-DD fusion protein. Soluble protein fraction expressed from the IRAK-4-DD-pETG-30A plasmid was purified by nickel affinity chromatography using a HiTrap chelating HP column. A 20 µL aliquot of this purified sample was kept aside (as the uncleaved fraction) while the rest of the purified soluble protein was cleaved by Prescission protease. Both the cleaved and uncleaved fractions were analysed by SDS PAGE on a 12 % SDS polyacrylamide gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa molecular weight reference ladder. Lane 1 uncleaved fraction. Lane 2 cleaved fraction

6.2.3 Size exclusion chromatography of cleaved His-GST-IRAK-4-DD protein

Although we obtained good yields of the IRAK-4-DD in the previous section, we wanted to subject it to size exclusion chromatography to determine its retention volume. Knowing its retention volume will be helpful in the analysis of our data when we reconstitute the Myddosome in Chapter 7. Therefore, the cleaved His-GST-IRAK-4-DD protein sample was concentrated to a final volume of 500 µL and thereafter loaded onto an ENrich SEC 650 (10 x 300 mm) column. A chromatogram was generated during the size exclusion chromatography purification process as shown in Figure 36.

Six peaks were observed on the chromatogram in Figure 36. The corresponding fractions from these peaks were collected and labelled as peak 1 up to 6. Each peak had its own retention volume and absorbance as shown in Table 34. Peak 1 had a retention volume of 13.83 mL and an absorbance of 248 mAU. Peak 2 had a retention volume of 15.22 mL and an absorbance of 476 mAU. Peak 3 had a retention volume of 16.14 mL and an absorbance of 2995 mAU. Peak 4 had a retention volume of 18.75 mL and an absorbance of 1117 mAU. Peak 5 had a retention volume of 20.68 mL and an absorbance of 469 mAU. Peak 6 had a retention volume of 21.95 mL and an absorbance of 255 mAU. The results suggested that peak 3 and peak 4 had the highest concentration of protein based on the high absorbance values observed.

.

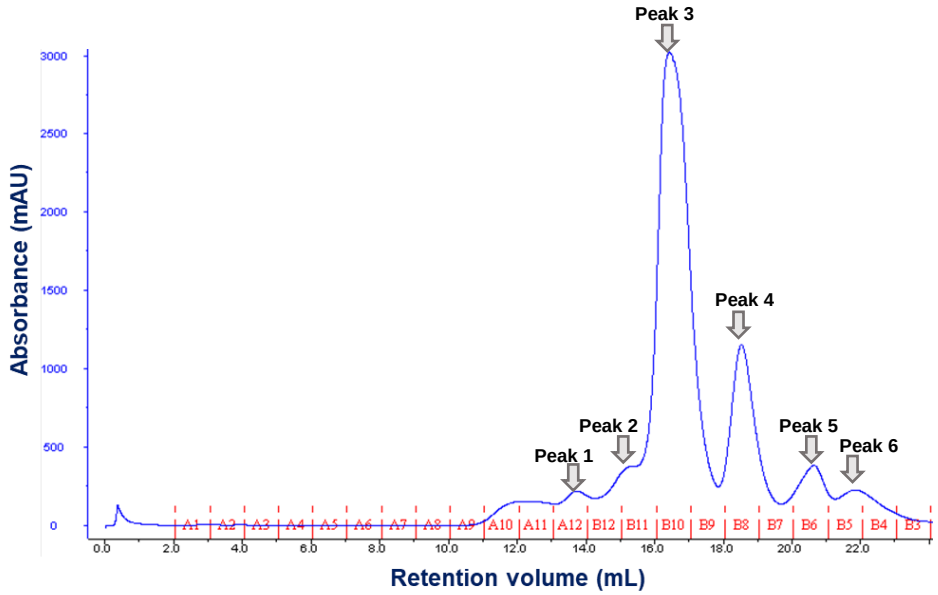


Figure 36. Chromatogram for SEC purification of cleaved His-GST-IRAK-4-DD. The cleaved His-GST-IRAK-4-DD protein sample was purified by SEC and a chromatogram was generated from this purification by an ÄKTA FPLC machine. UV absorbance readings at 280 nm were captured for prominent peaks and collected as follows: Peak 1 – [fraction: A12]. Peak 2 – [B11]. Peak 3 – [B10]. Peak 4 – [B8]. Peak 5 – [B6]. Peak 6 – [B5].

Table 34. SEC data output for cleaved His-GST-IRAK-4-DD protein

	Retention volume (mL)	Absorbance (mAU)
Peak 1	13.83	248
Peak 2	15.22	476
Peak 3	16.14	2995
Peak 4	18.75	1117
Peak 5	20.68	469
Peak 6	21.95	255

Samples of fractions that eluted from the six peaks were loaded and analysed by SDS-PAGE in the order of their elution as shown in Figure 37. A sample from peak 1 was loaded in lane 1 and it showed a few faint high molecular weight bands, none of which were particularly noteworthy. Similar high molecular weight bands were also observed in lane 2 which was loaded with a sample from peak 2. However, there was an intense 25 kDa band. This 25 kDa band was also observed in lane 3 which was loaded with a sample from peak 3. It was much more intense than in lane 2. The 25 kDa band was definitely the His-GST tag. There were also traces of the His-GST tag observed in lane 4. This lane had been loaded with a sample from peak 4. However, the prominent band in lane 4 was the 12 kDa protein which was definitely IRAK-4-DD. The observation of IRAK-4-DD was traced back to its elution on size exclusion chromatography where it was found to have a retention volume of 18.75 mL. Samples from peak 5 and 6 were loaded in lanes 5 and 6 respectively and there were no protein bands detected in both these lanes.

Therefore, based on Figure 36, Table 34 and Figure 37 we were able to determine the retention volume of IRAK-4-DD.

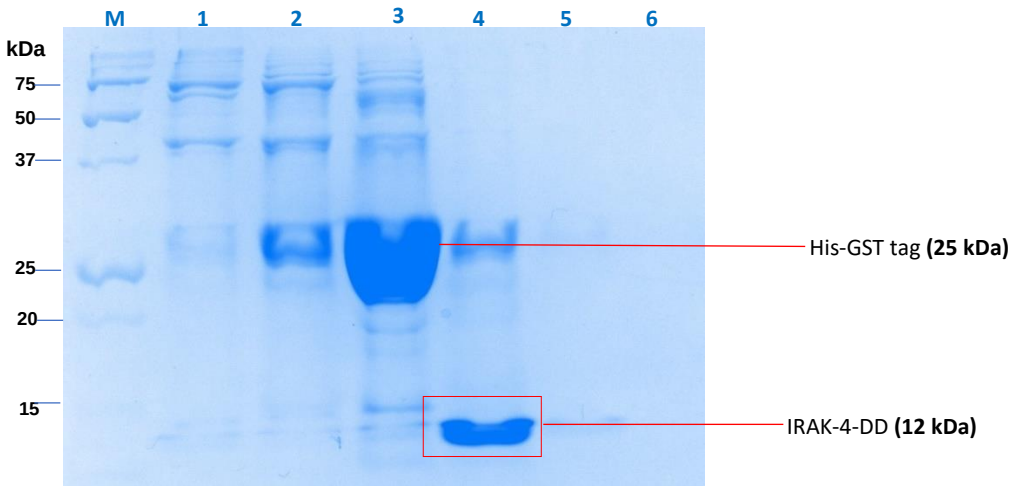


Figure 37. SDS gel of SEC purification fractions for cleaved His-GST-IRAK-4-DD. Fractions for the peaks that were identified in SEC purification of cleaved His-GST-IRAK-4-DD protein were collected and analysed by SDS-PAGE on a 12 % SDS gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa protein marker. Lane 1 – peak 1. Lane 2 - peak 2. Lane 3 – peak 3. Lane 4 – peak 4. Lane 5 – peak 5. Lane 6 – peak 6

6.3 Discussion

IRAK-4 is a binding partner of MyD88. It interacts with MyD88 through homotypic death domain interactions (Su, Xu, & Huang, 2020). We expressed the death domain of IRAK-4 using an expression clone that was readily available in our lab; and it had been shown to produce good yields (George et al., 2011). This expression clone was sequenced before we could start with our experiments. The DNA sequencing results confirmed that the clone contained the IRAK-4-DD gene and there was also a Prescission protease site to cleave IRAK-4-DD from its fusion partner His-GST.

The expression clone was grown on large scale to produce soluble His-GST-IRAK-4-DD fusion protein. Growth conditions that were optimal for the production of this fusion protein were used. They included lowering of the temperature to 15° C post IPTG induction. Lower temperatures are known to favour the production of soluble properly folded proteins (Gopal & Kumar, 2013).

Soluble His-GST-IRAK-4-DD fusion protein was obtained under low temperature growth conditions. This fusion protein was then cleaved with Prescission protease. Prescission protease was incorporated into the expression clone due to its specificity for its recognition sequence. It only cleaves between the glycine and proline in the following amino acid sequence LEVLFQGP (Abdelkader & Otting, 2021). Due to its specificity, we were therefore confident that it had cleaved the His-GST-IRAK-4-DD fusion protein into the His-GST tag and the IRAK-4-DD.

The cleaved His-GST-IRAK-4-DD fusion protein was purified by size exclusion chromatography in order to determine the retention volume of IRAK-4-DD. Our results showed that IRAK-4-DD had a retention volume of 18.75 mL. This retention volume was smaller than that of the 25 kDa His-GST tag which was at 16.15 mL. We also found a similar retention volume of His-GST in Chapter 3. Therefore, our analysis of the size exclusion chromatogram and SDS-PAGE results showed that the IRAK-4-DD protein was a monomer. This finding is similar to results reported elsewhere (Motshwene et al., 2009; George et al., 2011). Therefore, we would be in a position to reconstitute the Myddosome complex in the next chapter, knowing the retention volume of IRAK-4-DD. This known retention volume is going to assist us in identifying IRAK-4-DD among other proteins that have a similar molecular weight to it.

6.4 Conclusion

We expressed soluble IRAK-4-DD from the IRAK-4-DD-pETG-30A construct. It was expressed as a His-GST-IRAK-4-DD fusion protein, which was cleaved with Prescission protease to yield a His-GST tag and IRAK-4-DD. The cleaved sample was analysed by size exclusion chromatography and the result showed that IRAK-4-DD was a monomer.

Chapter 7: Protein-to-protein interaction studies in Myddosome assembly

7.1 Introduction

In this chapter, we wanted to investigate the role of the intermediate domain (ID) in the protein-to-protein interactions that form the Myddosome complex. As such, our plan was to reconstitute this complex in the presence and absence of the ID as a comparative study. The first section of the results are aimed at determining the effect of the presence of the ID domain in Myddosome assembly. Therefore, the cleaved sample that contains the MyD88-DD-ID protein was mixed with the cleaved sample that contains the IRAK-4-DD protein to reconstitute the Myddosome complex. Thereafter, size exclusion chromatography and SDS-PAGE were used to determine the presence of the complex. The second section of the results, are aimed at determining the effect of the absence of the intermediate domain in Myddosome assembly. As such, the cleaved sample that contains the intermediate domain deficient MyD88-DD protein was mixed with the cleaved sample that contains the IRAK-4-DD protein to reconstitute the Myddosome complex. The presence or absence thereof, was determined by both size exclusion chromatography and SDS-PAGE.

7.2 Results

7.2.1 Reconstitution of the Myddosome complex in the presence of the ID

Our aim in this section was to reconstitute the Myddosome complex. We had aliquots of the cleaved His-GST-MyD88-DD-ID protein and the cleaved His-GST-IRAK-4-DD that were produced in chapter 3 and 6 respectively. These were mixed and incubated at room temperature for 2 hours. Thereafter, the mixture was concentrated to a final volume of 500 µL prior to loading on a size exclusion chromatography column. A chromatogram was generated during the purification process as shown in Figure 38.

The chromatogram in Figure 38 showed six peaks. Corresponding fractions for these peaks were collected and labelled as peak 1 up to 6. Each peak had its own retention volume and absorbance as shown in Table 35. Peak 1 had a retention volume of 11.88 mL and an absorbance of 239 mAU. Peak 2 had a retention volume of 13.15 mL and an absorbance of 378 mAU. Peak 3 had a retention volume of 16.14 mL and an absorbance of 2759 mAU. Peak 4 had a retention volume of 18.17 mL and an absorbance of 832 mAU. Peak 5 had a retention volume of 19.92 mL and an absorbance of 1106 mAU. Peak 6 had a retention volume of 20.73 mL and an absorbance of 752 mAU. The results showed that peak 3 and peak 5 were highly likely to have the highest concentration of protein based on the high absorbance values observed.

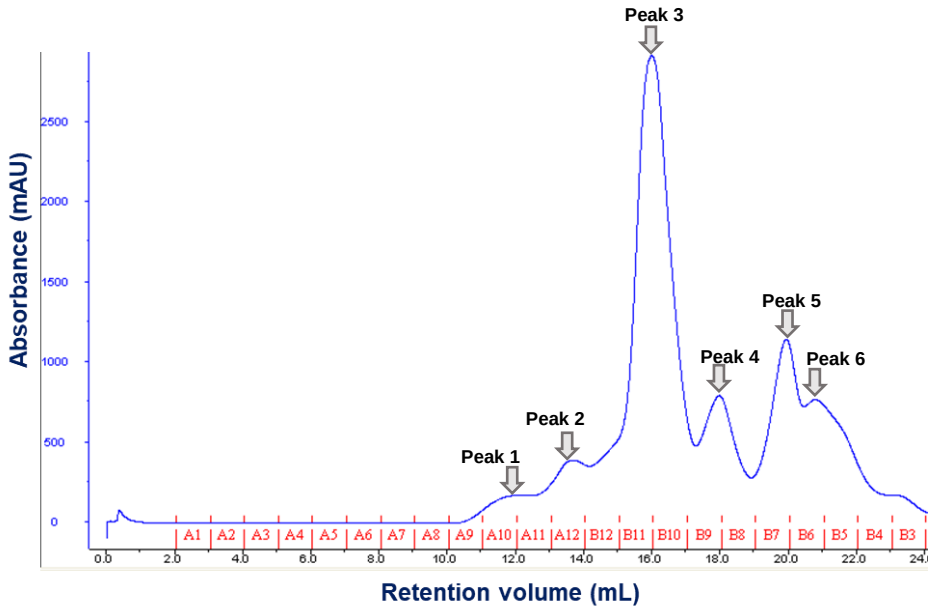


Figure 38. Chromatogram for SEC purification of ID-inclusive Myddosome complex reconstitution. The protein mixture of cleaved His-GST-IRAK-4-DD and cleaved His-GST-MyD88-DD-ID was purified by SEC and a chromatogram was generated from this purification by an ÄKTA FPLC machine. UV absorbance readings at 280 nm were captured for prominent peaks and collected as follows: Peak 1 – [fraction: A10 and A11]. Peak 2 – [A12]. Peak 3 – [B10]. Peak 4 – [B9]. Peak 5 – [B7]. Peak 6 – [B6].

Table 35. SEC purification data output for the ID-inclusive protein mixture

	Retention volume (mL)	Absorbance (mAU)
Peak 1	11.88	239
Peak 2	13.15	378
Peak 3	16.14	2759
Peak 4	18.17	832
Peak 5	19.92	1106
Peak 6	20.73	752

Our next objective was to verify if the Myddosome complex had been reconstituted. Therefore, samples of the SEC fractions that eluted from the six peaks were loaded and analysed by SDS-PAGE in the order of their elution as shown in Figure 39. A sample from peak 1 was loaded in lane 1 and it showed several faint protein bands which were not noteworthy. However, a sample from peak 2 which was loaded in lane 2 showed many protein bands of higher intensity. Amongst these proteins, we identified an intense 17 kDa protein band and a 12 kDa protein band as annotated with a red box in lane 2. These two protein bands were of MyD88-DD-ID and IRAK-4-DD protein. Therefore, co-elution of these proteins strongly suggested this was the Myddosome complex. A sample from peak 3, loaded in lane 3 showed a very large and intense 25 kDa protein band. This protein band was definitely the His-GST tag based on its molecular weight and retention volume of 16.14 mL. A sample from peak 4, loaded in lane 4 showed a very faint 25 kDa band and an intense 12 kDa band. The 12 kDa band is excess uncomplexed IRAK-4-DD protein based on its molecular weight and retention volume of 18.17 mL. A sample from

peak 5, loaded in lane 5 showed a faint 12 kDa band presumed to be an *E. coli* protein which we could not identify. A sample from peak 6, loaded in lane 6 showed no protein bands. Taken together, our results show that we managed to reconstitute the Myddosome complex. It eluted in peak 2 with a retention volume of 13.15 mL.

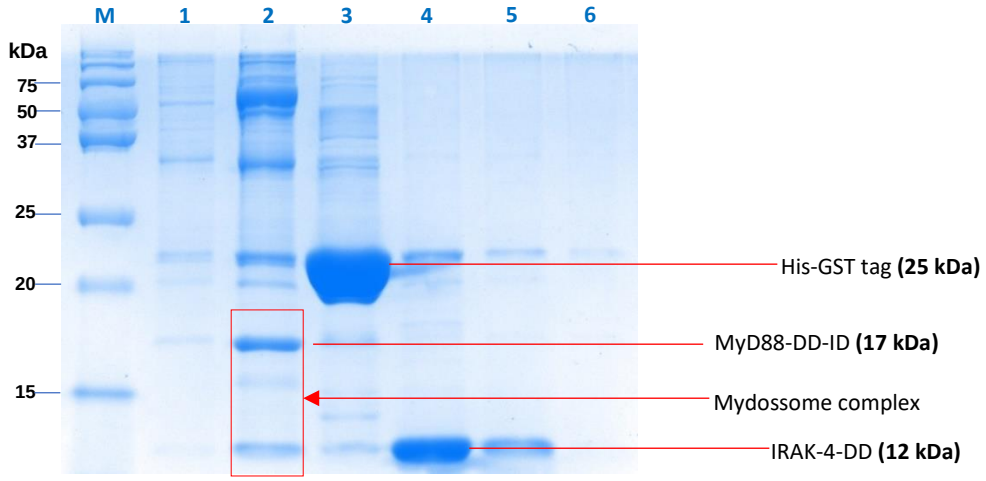


Figure 39. SDS gel of SEC purification fractions for ID-inclusive protein mixture. Fractions for the peaks that were identified in SEC purification of cleaved ID-inclusive protein mixture (cleaved His-GST-MyD88-DD-ID mixed with cleaved His-GST-IRAK-4-DD) protein were collected and analysed by SDS-PAGE on a 12 % SDS gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa protein marker. Lane 1 – peak 1. Lane 2 - peak 2. Lane 3 – peak 3. Lane 4 – peak 4. Lane 5 – peak 5. Lane 6 – peak 6

7.2.2 Reconstitution of the Myddosome complex in the absence of the ID

Our aim in this section was also to reconstitute the Myddosome complex using MyD88-DD that lacks the intermediate domain. We had aliquots of the cleaved His-GST-MyD88-DD protein and the cleaved His-GST-IRAK-4-DD that were produced in chapter 5 and 6 respectively. These were mixed and purified as described in section 7.2.1 and a chromatogram was generated during the purification process as shown in Figure 40.

Seven peaks were observed on the chromatogram shown in Figure 40. The corresponding fractions for these peaks were collected and labelled as peak 1 up to 7. Each peak had its own retention volume and absorbance as shown in Table 36. Peak 1 had a retention volume of 10.51 mL and an absorbance of 107 mAU. Peak 2 had a retention volume of 13.75 mL and an absorbance of 64 mAU. Peak 3 had a retention volume of 15.42 mL and an absorbance of 129 mAU. Peak 4 had a retention volume of 16.73 mL and an absorbance of 1528 mAU. Peak 5 had a retention volume of 17.91 mL and an absorbance of 1230 mAU. Peak 6 had a retention volume of 18.75 mL and an absorbance of 399 mAU. Peak 7 had a retention volume of 20.98 mL and an absorbance of 111 mAU. The results suggested that peak 4 and peak 5 had the highest concentration of protein based on the absorbance values.

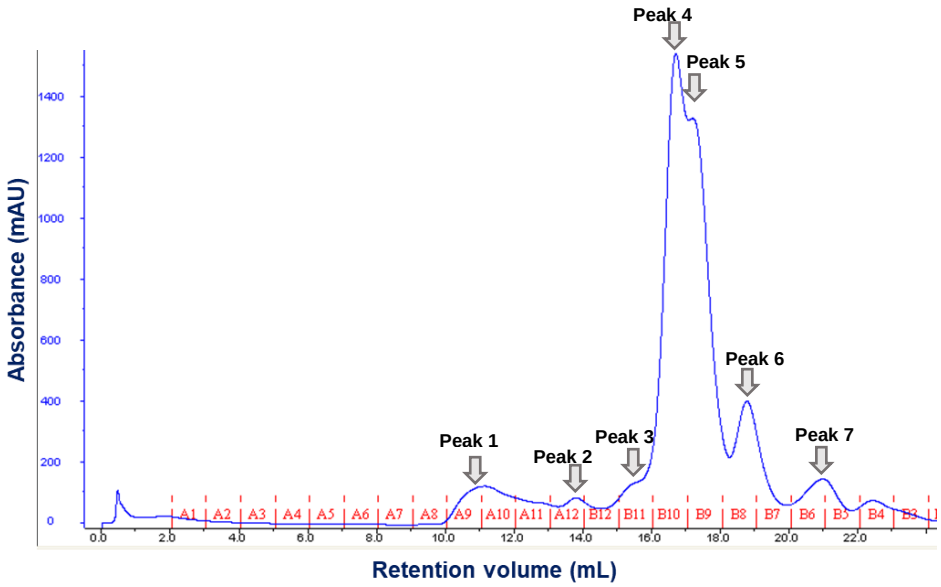


Figure 40. Chromatogram for SEC purification of ID-deficient Myddosome complex reconstitution. The protein mixture of cleaved His-GST-IRAK-4-DD and cleaved His-MBP-MyD88-DD was purified by SEC and a chromatogram was generated from this purification by an ÄKTA FPLC machine. UV absorbance readings at 280 nm were captured for prominent peaks and collected as follows: Peak 1 – [fraction: A9]. Peak 2 – [A12]. Peak 3 – [B11]. Peak 4 – [B10]. Peak 5 – [B9]. Peak 6 – [B8]. Peak 7 – [B6].

Table 36. SEC purification data output for the ID-deficient protein mixture

	Retention volume (mL)	Absorbance (mAU)
Peak 1	10.51	107
Peak 2	13.75	64
Peak 3	15.42	129
Peak 4	16.73	1528
Peak 5	17.91	1230
Peak 6	18.75	399
Peak 7	20.98	111

In order to verify if the ID-deficient Myddosome complex had been reconstituted, samples of the fractions that eluted from the seven peaks were loaded and analysed by SDS-PAGE in the order of their elution as shown in Figure 41. A sample from peak 1 was loaded in lane 1 and it showed a few high molecular weight bands and a distinct 12 kDa protein band. This 12 kDa protein band was the ID-deficient MyD88-DD based on its molecular weight and its retention volume of 10.51 mL. A retention volume of 10.51 mL for the ID-deficient MyD88-DD was also obtained in chapter 4. Lane 2 was loaded with a sample from peak 2 and it showed very few faint protein bands which were not noteworthy. A sample from peak 3, loaded in lane 3 showed some faint bands and a small 25 kDa protein band. Two large intense 25 kDa and 42 kDa bands were observed in a sample from peak 4, loaded in lane 4. Similarly, a sample from peak 5, loaded in lane 5 showed a large 42 kDa band and a relatively

smaller 25 kDa. It is very clear from the chromatogram that peak 3 and 5 are shoulders of peak 4. That is why we are seeing the same protein bands. The 25 kDa and 42 kDa bands that we are seeing in lanes 3 to 5 are the His-GST and His-MBP proteins. A sample from peak 6, loaded in lane 6 showed a faint 42 kDa band as well as the 12 kDa IRAK-4-DD protein. We are definitely sure that the 12 kDa is IRAK-4-DD based on its consistent retention volume which we previously determined in chapter 6. A sample from peak 7, loaded in lane 7 showed no protein bands. Taken together, our results show that the Myddosome complex failed to form in the absence of the intermediate domain. The intermediate domain deficient MyD88-DD protein failed to co-elute with IRAK-4-DD, hence we say there was no Myddosome complex formation.

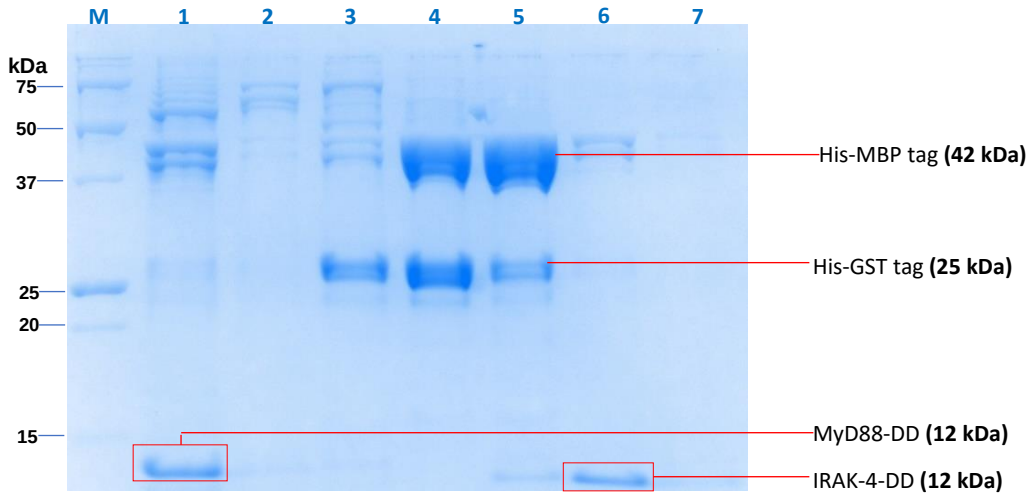


Figure 41. SDS gel of SEC purification fractions for ID-deficient protein mixture. Fractions for the peaks that were identified in SEC purification of cleaved ID-deficient protein mixture (cleaved His-MBP-MyD88-DD mixed with cleaved His-GST-IRAK-4-DD) protein were collected and analysed by SDS-PAGE on a 12 % SDS gel. 18 µL samples from these fractions were loaded as follows: Lane M - 250 kDa protein marker. Lane 1 – peak 1. Lane 2 - peak 2. Lane 3 – peak 3. Lane 4 – peak 4. Lane 5 – peak 5. Lane 6 - peak 6. Lane 7 – peak 7

7.3 Discussion

We used samples of His-GST-MyD88-DD-ID and His-GST-IRAK-4-DD proteins that were cleaved with Prescission protease to reconstitute the Myddosome complex. Our results showed that the Myddosome did assemble in the presence of the MyD88 intermediate domain. This finding is supported by previous studies that have also reconstituted the Myddosome complex (Motshwene et al., 2009). It is clear from our results that both the MyD88 containing the intermediate domain and IRAK-4 death domain eluted in the same peak in size exclusion chromatography. This finding was clear evidence that the two proteins were interacting with each other and that they had formed a Myddosome complex. Further evidence to support the formation of the Myddosome complex is that they eluted earlier on size exclusion chromatography; at a retention volume where we expected large molecular weight proteins to elute. Such a retention volume meant that there was oligomerisation. However, we did not perform some biophysical experiments to further characterise the Myddosome as this has been previously done using analytical ultra-centrifugation and non-dissociative mass spectrometry (Motshwene et al., 2009).

Our main objective in this study, was to determine if the Myddosome was going to form in the absence of the MyD88 intermediate domain. This was an important facet in the field of innate immunity that needed to be investigated. Earlier studies in TLR signalling had shown that the splice variant of MyD88, known as MyD88s, failed to interact with its physiological binding partner IRAK-4 due to its lack of the intermediate domain (Janssens et al., 2003). However, a study that published the crystal structure of the Myddosome contradicted earlier studies by showing that MyD88 death domain that lacked the intermediate was able to form a complex with IRAK-4 death domain (Lin et al., 2010). We therefore attempted to assemble the Myddosome complex using MyD88 death domain that lacked the intermediate domain. Our results showed that the Myddosome complex failed to form in the absence of the MyD88 intermediate domain. The 12 kDa MyD88 death domain that lacked the intermediate domain eluted on its own and at a retention volume where we expected larger molecular weight proteins. We had characterised it in chapter 5 using size exclusion chromatography and dynamic light scattering and found it to be an oligomer and not an aggregate. The IRAK-4 death domain that we used in attempting to reconstitute the Myddosome complex, also had 12 kDa molecular weight. It eluted much later on size exclusion chromatography, at an expected retention volume for its monomeric state.

Our results are in line with earlier studies that showed that MyD88 that lacks the intermediate fails to interact with IRAK-4 (Janssens et al., 2003). Granted, it is a known fact that MyD88 interacts with members of the IRAK family through homotypic death domain interactions (Gay, Gangloff, & O'Neill, 2011). However, our results have shown that its intermediate domain is essential for the formation of the Myddosome complex. We therefore cannot explain how the study that solved the crystal structure of the Myddosome managed to reconstitute it without the MyD88 intermediate domain. Therefore, we have to ask ourselves if the findings of that study are physiologically relevant.

7.4 Conclusion

It is our conclusion that the MyD88 intermediate domain is essential for the assembly of the Myddosome complex. We carried out two experiments that made us come to this conclusion. One of the experiments was a positive control where we used MyD88 death domain that also contained an intermediate domain. As expected, the Myddosome did assemble in our positive control. However, the Myddosome failed to assemble in our other experiment where we used the MyD88 death domain that lacked the intermediate domain.

Chapter 8: Project conclusion

The aim of this project was to investigate if the Myddosome complex would form or assemble in the absence of the MyD88 intermediate domain. Therefore, our objectives were to express MyD88 that had both the intermediate and death domain, MyD88 death domain that lacked the intermediate domain and IRAK-4 death domain. These three proteins were used to reconstitute the Myddosome complex.

Our data showed that the Myddosome complex was able to reconstitute in the presence of the MyD88 intermediate domain. However, the Myddosome complex failed to reconstitute in the absence of the MyD88 intermediate domain. This finding led us to accept our hypothesis that the intermediate domain deficient MyD88 death domain does not form the Myddosome complex. It is therefore our conclusion that the MyD88 intermediate domain is essential for the assembly of the Myddosome complex.

Chapter 9: References

Abdelkader, E. H., & Otting, G. (2021). NT*-HRV3CP: An optimized construct of human rhinovirus 14 3C protease for high-yield expression and fast affinity-tag cleavage. *Journal of biotechnology, 325*, 145-151.

Akar-Ghibril, N. (2022). Defects of the innate immune system and related immune deficiencies. *Clinical Reviews in Allergy & Immunology, 63*(1), 36-54.

Akira, S., Uematsu, S., & Takeuchi, O. (2006). Pathogen recognition and innate immunity. *Cell, 124*(4), 783-801.

Amarante-Mendes, G. P., Adjemian, S., Branco, L. M., Zanetti, L. C., Weinlich, R., & Bortoluci, K. R. (2018). Pattern recognition receptors and the host cell death molecular machinery. *Frontiers in immunology, 9*, 2379.

Anderson, K. V., Jürgens, G., & Nüsslein-Volhard, C. (1985). Establishment of dorsal-ventral polarity in the Drosophila embryo: genetic studies on the role of the Toll gene product. *Cell, 42*(3), 779-789.

Avbelj, M., Panter, G., & Jerala, R. (2018). The role of N-terminal segment and membrane association in MyD88-mediated signaling. *Biochemical and biophysical research communications, 495*(1), 878-883.

Bai, Z.-Q., Ma, X., Liu, B., Huang, T., & Hu, K. (2022). Solution structure of c-FLIP death effector domains. *Biochemical and biophysical research communications, 617*, 1-6.

Balka, K. R., & De Nardo, D. (2019). Understanding early TLR signaling through the Myddosome. *Journal of leukocyte biology, 105*(2), 339-351.

Behzadi, P., García-Perdomo, H. A., & Karpiński, T. M. (2021). Toll-like receptors: general molecular and structural biology. *Journal of Immunology Research, 2021*.

Bhatwa, A., Wang, W., Hassan, Y. I., Abraham, N., Li, X.-Z., & Zhou, T. (2021). Challenges associated with the formation of recombinant protein inclusion bodies in Escherichia coli and strategies to address them for industrial applications. *Frontiers in Bioengineering and Biotechnology, 9*, 630551.

Boraschi, D., Italiani, P., Weil, S., & Martin, M. U. (2018). The family of the interleukin-1 receptors. *Immunological reviews, 281*(1), 197-232.

Botos, I., Segal, D. M., & Davies, D. R. (2011). The structural biology of Toll-like receptors. *Structure, 19*(4), 447-459.

Brinkworth, J. F., & Valizadegan, N. (2021). Sepsis and the evolution of human increased sensitivity to lipopolysaccharide. *Evolutionary Anthropology: Issues, News, and Reviews, 30*(2), 141-157.

Bryant, C. E., Symmons, M., & Gay, N. J. (2015). Toll-like receptor signalling through macromolecular protein complexes. *Molecular immunology, 63*(2), 162-165.

Chan, J. Y. W., Tsui, J. C. C., Law, P. T. W., So, W. K. W., Leung, D. Y. P., Sham, M. M. K., . . . Chan, C. W. H. (2018). Regulation of TLR4 in silica-induced inflammation: An underlying mechanism of silicosis. *International Journal of Medical Sciences, 15*(10), 986.

Ciesielska, A., Matyjek, M., & Kwiatkowska, K. (2021). TLR4 and CD14 trafficking and its influence on LPS-induced pro-inflammatory signaling. *Cellular and Molecular Life Sciences, 78*, 1233-1261.

Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., . . . Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring. *Journal of Veterinary Diagnostic Investigation, 32*(6), 767-775.

Deguine, J., & Barton, G. M. (2014). MyD88: a central player in innate immune signaling. *F1000prime reports, 6*.

den Haan, J. M., Arens, R., & van Zelm, M. C. (2014). The activation of the adaptive immune system: cross-talk between antigen-presenting cells, T cells and B cells. *Immunology letters, 162*(2), 103-112.

Ding, J., Pan, X., Du, L., Yao, Q., Xue, J., Yao, H., . . . Shao, F. (2019). Structural and functional insights into host death domains inactivation by the bacterial arginine GlcNAcyltransferase effector. *Molecular cell, 74*(5), 922-935. e926.

Dossang, A. C., Motshwene, P. G., Yang, Y., Symmons, M. F., Bryant, C. E., Borman, S., . . . Gay, N. J. (2016). The N-terminal loop of IRAK-4 death domain regulates ordered assembly of the Myddosome signalling scaffold. *Scientific reports, 6*(1), 37267.

Esposito, D., Garvey, L. A., & Chakiath, C. S. (2009). Gateway cloning for protein expression. *Methods Mol Biol*, 31-54. doi:10.1007/978-1-59745-196-3_3

Falke, S., & Betzel, C. (2019). Dynamic Light Scattering (DLS) Principles, Perspectives, Applications to Biological Samples. *Radiation in Bioanalysis: Spectroscopic Techniques and Theoretical Methods*, 173-193.

Federico, S., Pozzetti, L., Papa, A., Carullo, G., Gemma, S., Butini, S., . . . Relitti, N. (2020). Modulation of the innate immune response by targeting toll-like receptors: a perspective on their agonists and antagonists. *Journal of medicinal chemistry, 63*(22), 13466-13513.

Fekete, S., Beck, A., Veuthey, J.-L., & Guillarme, D. (2014). Theory and practice of size exclusion chromatography for the analysis of protein aggregates. *Journal of pharmaceutical and biomedical analysis, 101*, 161-173.

Fitzgerald, K. A., & Kagan, J. C. (2020). Toll-like receptors and the control of immunity. *Cell, 180*(6), 1044-1066.

Gay, N. J., Gangloff, M., & O'Neill, L. A. (2011). What the Myddosome structure tells us about the initiation of innate immunity. *Trends in immunology, 32*(3), 104-109.

George, J., Motshwene, P. G., Wang, H., Kubarenko, A. V., Rautanen, A., Mills, T. C., . . . Weber, A. N. (2011). Two human MYD88 variants, S34Y and R98C, interfere with MyD88-IRAK4-myddosome assembly. *Journal of Biological Chemistry, 286*(2), 1341-1353.

Godfroy, J. I., Roostan, M., Moroz, Y. S., Korendovych, I. V., & Yin, H. (2012). Isolated Toll-like receptor transmembrane domains are capable of oligomerization. *PLoS One, 7*(11), e48875.

Gopal, G. J., & Kumar, A. (2013). Strategies for the production of recombinant protein in Escherichia coli. *The protein journal, 32*(6), 419-425.

Haley, R. M., & von Recum, H. A. (2019). Localized and targeted delivery of NSAIDs for treatment of inflammation: A review. *Experimental Biology and Medicine, 244*(6), 433-444.

Held, D., & Kilz, P. (2021). Size-exclusion chromatography as a useful tool for the assessment of polymer quality and determination of macromolecular properties. *Chemistry Teacher International, 3*(2), 77-103.

Huoh, Y. S., & Hur, S. (2022). Death domain fold proteins in immune signaling and transcriptional regulation. *The FEBS journal, 289*(14), 4082-4097.

Israel, L., Wang, Y., Bulek, K., Della Mina, E., Zhang, Z., Pedergnana, V., . . . Descatoire, M. (2017). Human adaptive immunity rescues an inborn error of innate immunity. *Cell, 168*(5), 789-800. e710.

Iwasaki, A., & Medzhitov, R. (2015). Control of adaptive immunity by the innate immune system. *Nature immunology, 16*(4), 343-353.

Jain, A., Kaczanowska, S., & Davila, E. (2014). IL-1 receptor-associated kinase signaling and its role in inflammation, cancer progression, and therapy resistance. *Frontiers in immunology, 5*, 553.

Janssens, S., & Beyaert, R. (2003). Role of Toll-like receptors in pathogen recognition. *Clinical microbiology reviews, 16*(4), 637-646.

Janssens, S., Burns, K., Vercammen, E., Tschopp, J., & Beyaert, R. (2003). MyD88S, a splice variant of MyD88, differentially modulates NF-κB-and AP-1-dependent gene expression. *FEBS letters, 548*(1-3), 103-107.

Jin, M. S., Kim, S. E., Heo, J. Y., Lee, M. E., Kim, H. M., Paik, S.-G., . . . Lee, J.-O. (2007). Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell, 130*(6), 1071-1082.

Kagan, J. C., Magupalli, V. G., & Wu, H. (2014). SMOCs: supramolecular organizing centres that control innate immunity. *Nature Reviews Immunology, 14*(12), 821-826.

Karananou, P., Alataki, A., & Papadopoulou-Alataki, E. (2020). Interleukin-1 Receptor-Associated Kinase 4 Deficiency in a Greek Teenager. *Case Reports in Immunology, 2020*.

Kawai, T., Adachi, O., Ogawa, T., Takeda, K., & Akira, S. (1999). Unresponsiveness of MyD88-deficient mice to endotoxin. *Immunity, 11*(1), 115-122.

Kawasaki, T., & Kawai, T. (2014). Toll-like receptor signaling pathways. *Frontiers in immunology, 5*, 461.

Khattab, R., Ghanem, A., & Brooks, M. S.-L. (2017). Quality of dried haskap berries (Lonicera caerulea L.) as affected by prior juice extraction, osmotic treatment, and drying conditions. *Drying Technology, 35*(3), 375-391.

Kiripolsky, J., Romano, R.-A., Kasperek, E. M., Yu, G., & Kramer, J. M. (2020). Activation of Myd88-dependent TLRs mediates local and systemic inflammation in a mouse model of primary Sjögren's syndrome. *Frontiers in immunology, 10*, 2963.

Konno, H., Chinn, I. K., Hong, D., Orange, J. S., Lupski, J. R., Mendoza, A., . . . Barber, G. N. (2018). Pro-inflammation associated with a gain-of-function mutation (R284S) in the innate immune sensor STING. *Cell reports, 23*(4), 1112-1123.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.-M., & Hoffmann, J. A. (1996). The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in Drosophila adults. *Cell, 86*(6), 973-983.

Lin, S.-C., Lo, Y.-C., & Wu, H. (2010). Helical assembly in the MyD88–IRAK4–IRAK2 complex in TLR/IL-1R signalling. *Nature, 465*(7300), 885-890.

Liu, S., Ge, D., Long, Z., Chi, C., Lv, Z., & Liu, H. (2020). Molecular features of interleukin-1 receptor-associated kinase-b and a in Mytilus coruscus, regulating their function by infection of aquatic pathogens and the expression of their serine/threonine protein kinase functional domains. *Fish & Shellfish Immunology, 102*, 469-479.

Lv, Z., Wang, Z., Luo, L., Chen, Y., Han, G., Wang, R., . . . Feng, J. (2019). Spliceosome protein Eftud2 promotes colitis-associated tumorigenesis by modulating inflammatory response of macrophage. *Mucosal immunology, 12*(5), 1164-1173.

Mazzini, G., Ricagno, S., Caminito, S., Rognoni, P., Milani, P., Nuvolone, M., . . . Merlini, G. (2022). Protease-sensitive regions in amyloid light chains: what a common pattern of fragmentation across organs suggests about aggregation. *The FEBS journal, 289*(2), 494-506.

McCardell, R. D., Pandey, A., & Murray, R. M. (2019). Control of density and composition in an engineered two-member bacterial community. *BioRxiv*, 632174.

McPherson, A. (2017). Protein crystallization. *Protein Crystallography*, 17-50.

Meng, H., Liu, Z., Li, X., Wang, H., Jin, T., Wu, G., . . . Yu, Q. (2018). Death-domain dimerization-mediated activation of RIPK1 controls necroptosis and RIPK1-dependent apoptosis. *Proceedings of the National Academy of Sciences, 115*(9), E2001-E2009.

Menzella, H. G. (2011). Comparison of two codon optimization strategies to enhance recombinant protein production in Escherichia coli. *Microbial cell factories, 10*(1), 1-8.

Mishra, M., Tiwari, S., Gunaseelan, A., Li, D., Hammock, B. D., & Gomes, A. V. (2019). Improving the sensitivity of traditional Western blotting via Streptavidin containing Poly-horseradish peroxidase (PolyHRP). *Electrophoresis, 40*(12-13), 1731-1739.

Moncrieffe, M. C., Bollschweiler, D., Li, B., Penczek, P. A., Hopkins, L., Bryant, C. E., . . . Gay, N. J. (2020). MyD88 death-domain oligomerization determines Myddosome architecture: implications for Toll-like receptor signaling. *Structure, 28*(3), 281-289. e283.

Motshwene, P. G., Moncrieffe, M. C., Grossmann, J. G., Kao, C., Ayaluru, M., Sandercock, A. M., . . . Gay, N. J. (2009). An oligomeric signaling platform formed by the Toll-like receptor signal transducers MyD88 and IRAK-4. *Journal of Biological Chemistry, 284*(37), 25404-25411.

Müller, U., Vogel, P., Alber, G., & Schaub, G. A. (2008). The innate immune system of mammals and insects. *Trends in innate immunity, 15*, 21-44.

Nanson, J. D., Kobe, B., & Ve, T. (2019). Death, TIR, and RHIM: Self-assembling domains involved in innate immunity and cell-death signaling. *Journal of leukocyte biology, 105*(2), 363-375.

Nimma, S., Ve, T., Williams, S. J., & Kobe, B. (2017). Towards the structure of the TIR-domain signalosome. *Current opinion in structural biology, 43*, 122-130.

Padova, F., Quesniaux, V., & Ryffel, B. (2018). MyD88 as a therapeutic target for inflammatory diseases. *Expert OpinTher Targets, 25*, 401-408.

Park, B. S., Song, D. H., Kim, H. M., Choi, B.-S., Lee, H., & Lee, J.-O. (2009). The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nature, 458*(7242), 1191-1195.

Park, H. H., Lo, Y.-C., Lin, S.-C., Wang, L., Yang, J. K., & Wu, H. (2007). The death domain superfamily in intracellular signaling of apoptosis and inflammation. *Annual review of immunology, 25*.

Picard, C., Puel, A., Bonnet, M., Ku, C.-L., Bustamante, J., Yang, K., . . . Fieschi, C. (2003). Pyogenic bacterial infections in humans with IRAK-4 deficiency. *Science, 299*(5615), 2076-2079.

Płóciennikowska, A., Hromada-Judycka, A., Borzęcka, K., & Kwiatkowska, K. (2015). Co-operation of TLR4 and raft proteins in LPS-induced pro-inflammatory signaling. *Cellular and Molecular Life Sciences, 72*, 557-581.

Reece-Hoyes, J. S., & Walhout, A. J. (2018). Gateway recombinational cloning. *Cold Spring Harbor Protocols, 2018*(1), pdb. top094912.

Riera Romo, M., Pérez-Martínez, D., & Castillo Ferrer, C. (2016). Innate immunity in vertebrates: an overview. *Immunology, 148*(2), 125-139.

Saikh, K. U. (2021). MyD88 and beyond: A perspective on MyD88-targeted therapeutic approach for modulation of host immunity. *Immunologic Research, 69*, 117-128.

Sameer, A. S., & Nissar, S. (2021). Toll-like receptors (TLRs): structure, functions, signaling, and role of their polymorphisms in colorectal cancer susceptibility. *BioMed Research International, 2021*.

Santoni, G., Cardinali, C., Morelli, M. B., Santoni, M., Nabissi, M., & Amantini, C. (2015). Danger-and pathogen-associated molecular patterns recognition by pattern-recognition receptors and ion channels of the

transient receptor potential family triggers the inflammasome activation in immune cells and sensory neurons. *Journal of neuroinflammation, 12*(1), 1-10.

Scheich, C., Sievert, V., & Büssow, K. (2003). An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinity chromatography. *BMC biotechnology, 3*, 1-8.

Seniya, S. P., Yadav, P., & Jain, V. (2020). Construction of E. coli—Mycobacterium shuttle vectors with a variety of expression systems and polypeptide tags for gene expression in mycobacteria. *PLoS One, 15*(3), e0230282.

Su, L.-C., Xu, W.-D., & Huang, A.-F. (2020). IRAK family in inflammatory autoimmune diseases. *Autoimmunity reviews, 19*(3), 102461.

Sušjan-Leite, P., Ramuta, T. Ž., Boršić, E., Orehek, S., & Hafner-Bratkovič, I. (2022). Supramolecular organizing centers at the interface of inflammation and neurodegeneration. *Frontiers in immunology, 13*.

Suzuki, N., Suzuki, S., Duncan, G. S., Millar, D. G., Wada, T., Mirtsos, C., . . . Li, S. (2002). Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4. *Nature, 416*(6882), 750-754.

Takeda, K., & Akira, S. (2015). Toll-like receptors. *Current protocols in immunology, 109*(1), 14.12. 11-14.12. 10.

Tartey, S., & Takeuchi, O. (2017). Pathogen recognition and Toll-like receptor targeted therapeutics in innate immune cells. *International reviews of immunology, 36*(2), 57-73.

Tran, N. T., Liu, H., Jakovlić, I., & Wang, W.-M. (2015). Blunt snout bream (Megalobrama amblycephala) MyD88 and TRAF6: characterisation, comparative homology modelling and expression. *International Journal of Molecular Sciences, 16*(4), 7077-7097.

Tsukamoto, H., Takeuchi, S., Kubota, K., Kobayashi, Y., Kozakai, S., Ukai, I., . . . Kanemitsu, Y. (2018). Lipopolysaccharide (LPS)-binding protein stimulates CD14-dependent Toll-like receptor 4 internalization and LPS-induced TBK1–IKKϵ–IRF3 axis activation. *Journal of Biological Chemistry, 293*(26), 10186-10201.

Van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The third revolution in sequencing technology. *Trends in Genetics, 34*(9), 666-681.

Vetreno, R. P., Qin, L., Coleman Jr, L. G., & Crews, F. T. (2021). Increased Toll-like Receptor-MyD88-NFκB-Proinflammatory neuroimmune signaling in the orbitofrontal cortex of humans with alcohol use disorder. *Alcoholism: Clinical and Experimental Research, 45*(9), 1747-1761.

Wang, J., Zhang, Z., Liu, J., Zhao, J., & Yin, D. (2016). Ectodomain architecture affects sequence and functional evolution of vertebrate Toll-like receptors. *Scientific reports, 6*(1), 26705.

Wang, Z., Wesche, H., Stevens, T., Walker, N., & Yeh, W.-C. (2009). IRAK-4 inhibitors for inflammation. *Current topics in medicinal chemistry, 9*(8), 724-737.

Wiese, M. D., Manning-Bennett, A. T., & Abuhelwa, A. Y. (2020). Investigational IRAK-4 inhibitors for the treatment of rheumatoid arthritis. *Expert Opinion on Investigational Drugs, 29*(5), 475-482.

Wingfield, P. T. (2015). Overview of the purification of recombinant proteins. *Current protocols in protein science, 80*(1), 6.1. 1-6.1. 35.

Xin, L., Wang, M., Zhang, H., Li, M., Wang, H., Wang, L., & Song, L. (2016). The categorization and mutual modulation of expanded MyD88s in Crassostrea gigas. *Fish & Shellfish Immunology, 54*, 118-127.

Yin, Q., Fu, T.-M., Li, J., & Wu, H. (2015). Structural biology of innate immunity. *Annual review of immunology, 33*, 393-416.

Yoon, S.-i., Kurnasov, O., Natarajan, V., Hong, M., Gudkov, A. V., Osterman, A. L., & Wilson, I. A. (2012). Structural basis of TLR5-flagellin recognition and signaling. *Science, 335*(6070), 859-864.

Young, C. L., Britton, Z. T., & Robinson, A. S. (2012). Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnology journal, 7*(5), 620-634.

Zhou, W., Kaneko, N., Nakagita, T., Takeda, H., & Masumoto, J. (2021). A comprehensive interaction study provides a potential domain interaction network of human death domain superfamily proteins. *Cell Death & Differentiation, 28*(11), 2991-3008.

Zhu, K., Zhou, X., Yan, Y., Mo, H., Xie, Y., Cheng, B., & Fan, J. (2017). Cleavage of fusion proteins on the affinity resins using the TEV protease variant. *Protein expression and purification, 131*, 27-33.

Zindel, J., & Kubes, P. (2020). DAMPs, PAMPs, and LAMPs in immunity and sterile inflammation. *Annual Review of Pathology: Mechanisms of Disease, 15*, 493-518.