Effects of IRAK-4 phosphomimetics on the Myddosome assembly

By

David Difateng Thipa

10667092

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

13/12/2018

TABLE OF CONTENTS

PLAGIARISM DECLARATION

Full name: David Difateng Thipa   Student number: 10667092

Title: Effects of IRAK4 phosphomimetics on the Myddosome complex

Declaration:

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

2. I declare that this proposal is my own work. Where someone else work was used (either from printed source, internet or any other source) due acknowledgement was given and references were given according to departmental requirements.

3. I did not make use of another student's previous work and submit it as my own.

4. I did not allow and will not allow anyone to copy my work with the intention of presenting it his or her own work.

Signature: _____   Date: __13/12/2018__

ACKNOWLEDGEMENTS

SUMMARY

MyD88 and IRAK-4 are two intracellular proteins that are involved in innate immunity. They form an oligomeric complex known as the Myddosome when a Toll-like receptor is activated upon ligand binding. This complex forms part of post-receptor signalling and is assembled by the interaction of death domains of both proteins. The regulatory mechanisms of assembly and dissociation of the Myddosome are not yet fully understood. One possible mechanism is that autophosphorylation regulates the assembly of the Myddosome.

IRAK-4 is a serine/threonine kinase that phosphorylates another kinase, IRAK-1 during TLR signaling. In addition to phosphorylating IRAK-1, it also phosphorylates itself, a process known as autophosphorylation. Its autophosphorylation is an intramolecular event that starts in the kinase loop and extends to its death domain. It is reported that two autophosphorylation sites namely serine 8 and threonine 62 were mapped to its death domain. We substituted these two amino acids with aspartic acid to create phosphomimetics. The mutant phosphomimetics were expressed in *E. coli* cells and thereafter purified with the aim of determining if they will form the Myddosome when mixed with MyD88 death domain.

Wild type and phosphomimetic complexes were formed and characterized using size exclusion chromatography, dynamic scattering and analytical ultracentrifugation. Our results showed that phosphomimetic mutants were able to form an oligomeric complex similar in size to the Myddosome. This result was contrary to what has been reported in literature. We therefore concluded that phosphomimetics were not ideal replacements for phosphorylation.

ABBREVIATIONS

ATP	Adenosine triphosphate
APS	Ammonium persulphate
AUC	Analytical ultracentrifugation
bp	base pair
DD	Death domain
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ECD	Ectodomain
EDTA	Ethylenediaminetetra-acetic acid
FL	Full length
GS	Glutathione sepharose
GST	Glutathione S-transferase
hrs	Hours
ID	Intermediate domain
IFNs	Interferons
IPTG	Isopropyl-β-D-thiogalactopyranoside
IRAK	Interleukin-1 receptor associated kinase
IRF	Interferon regulatory factor
KD	Kinase domain
kDa	Kilo Dalton
L	Litre
LB	Lysogeny broth Luria Bertani
LBP	Lipid binding protein
LPS	Lipopolysaccharide
LRR	Leucine-rich repeat

M	Molar
Min	Minutes
MyD88	Myeloid differentiation factor 88
mg	Milligram
mL	Millilitre
mM	Millimolar
MW	Molecular weight
MWCO	Molecular weight cut-off
NF-κB	Nuclear factor-KappaB
nM	Nanomolar
OD	Optical density
PAMPs	Pathogen associated molecular patterns
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PRRs	Pathogen recognition receptors
RNA	Ribonucleic acid
rpm	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SMOCs	Supra molecular organising centers
TAE	Tris-acetate/EDTA
TEMED	N,N,N,N-Tetramethylethylene
TIR	Toll/IL-1 receptor
TLR	Toll-like receptor
Tris	Tris (hydroxylmethyl) aminomethane
WT	Wild type
°C	Degrees Celsius
μg	Microgram
μM	Micromolar

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Innate immunity

Innate immunity is the first line of host defense against infections and is present from birth (Hato and Dagher, 2015). An innate immune response is critical in inducing immediate responses against invading pathogens. This is in contrast with the adaptive immune response which is delayed because it needs to produce antibodies which are used to ward off infections (Medzhitov, 2001; Uematsu and Akira, 2008). The innate immune system depends on a limited number of germ line encoded receptors known as pattern recognition receptors (PRRs) to recognise pathogens. The adaptive immune system on the other hand is mediated by T and B lymphocytes bearing antigen-specific receptors generated by somatic recombination (Vivier et al., 2011).

PRRs recognize conserved molecules that are common in microbes (Takeuchi and Akira, 2010). These molecules are often essential for the survival and pathogenicity of microbes and are known as pathogen associated molecular patterns (PAMPs). However, these molecules can also be found in non-pathogenic microbes. In that case, they are referred to as microbial associated molecular patterns (MAMPs). PRRs also play a role in the recognition of endogenous molecules released from damaged or necrotic cells (Takeuchi and Akira, 2010). These are known as damage associated molecular patterns (DAMPs).

There are five different classes of PRRs. These include Toll-like receptors (TLRs), C-type lectin receptors (CLRs), Nucleotide-binding oligomerization domain (NOD)-like receptor (NLRs), Retinoic acid-inducible gene (RIG)-I-like receptors (RLRs) and Absent in melanoma 2 (AIM2)-like receptors (ALRs) (Kumar et al., 2011). Although, there are five classes of PRRs, I will only describe TLRs in the rest of this thesis because our research is much more focused on TLR signalling.

1.2 Toll-like receptors

The protein Toll was first discovered in *Drosophila* fruitfly by Christiane Nusslein-Vollard (Anderson et al., 1985). She found it to be involved in dorsoventral patterning during embryogenesis. However, it was later found to be a key component of the fruitfly immunity against fungal infections (Lemaitre et al., 1996). The discovery of similarity between the cytoplasmic domain of toll and mammalian interleukin-1 receptors (IL-1Rs) led to the search for similar receptors in mammals from expressed sequence tags (EST) database. Mammalian

orthologs were found and named toll-like receptors because they resemble the *Drosophila* Toll protein. So far, there are thirteen TLRs that have been identified in mammals (Shan et al., 2018).

TLRs are type 1 integral membrane glycoproteins (Akira and Takeda, 2004). They have an ectodomain and a cytoplasmic domain (Uematsu and Akira, 2008). Their cytoplasmic domain is similar to that of the IL-1R family and is known as the Toll/IL-1R (TIR) domain (Bowie and O'Neill, 2000). It has 125-200 amino acid residues which have regions known as Box 1, Box 2 and Box 3 (Ve et al., 2015). TLR ECDs on the other hand have leucine-rich repeat (LRR) motifs. These motifs are responsible for the recognition of PAMPs (Uematsu and Akira, 2006).

TLRs are very specific in the ligands or PAMPs they bind as shown in table 1.1. For example, TLR4 binds LPS or endotoxin and will not bind any other ligand. Other TLRs work in pairs by forming heterodimer complexes to bind certain ligands. Examples of these include TLR1-TLR2, TLR2-TLR6. These bind lipoproteins and zymosan as shown in table 1.1.

Table 1.1: Mammalian TLRs and their ligands.

Toll-like receptor	Ligand
TLR2-TLR1 heterodimer TLR2-TLR6 heterodimer	Lipoproteins (diacyll and triacyl lipoproteins) Lipoteichoic acids, Lipomanans, Zymosan
TLR3	Double-stranded RNA from viruses
TLR4 (plus MD2 and CD14)	LPS from Gram-negative bacteria
TLR5	Bacterial flagellin
TLR7	Single-stranded RNA from viruses
TLR8	Responds to RNA from bacteria and viruses
TLR9	Bacterial and viral DNA rich in CpG-DNA motifs and hemozoin
TLR10-TLR2 heterodimer	Ligands from *Listeria*
TLR11	Flagellin
TLR12	Profillin from *Toxoplasma gondii*
TLR13	23S RNA

1.3 Toll-like receptor signalling pathways

Each toll-like receptor signals in its own unique way. The signalling pathways are many and to a certain extent complicated. I will only discuss the TLR4 signalling pathway because research in our lab is focused on it. TLR4 is the known receptor for endotoxin to which humans are hypersensitive to (Angus et al., 2001). Its binding triggers a signalling cascade through the innate immune system to produce an inflammatory response. This response sometimes has dire or fatal consequences when the innate immune system is overwhelmed by the endotoxin.

TLR4 uses two sets of adaptors for signalling (Kawasaki and Kawai, 2014). The first set uses MyD88 (Myeloid differentiation primary response protein 88) and MAL (MyD88 adaptor like), while the second set uses TRIF (TIR domain-containing adaptor protein inducing interferon-β) and TRAM (TRIF-related adaptor molecule). The pathway that uses the first set of adaptors is known as the MyD88 dependant pathway whereas the one that uses the second set is the MyD88 independent pathway.

1.3.1 MyD88 dependent pathway.

LPS does not bind directly to TLR4. It binds to LBP (LPS binding protein) which is in a complex with CD14 as shown in figure 1.2. Thereafter, LPS will be transferred to a dimeric complex consisting of TLR4 and MD2. Then the TIR domain of TLR4 will interact with MAL and MyD88 through homotypic TIR-TIR interactions. Interleukin-1 receptor associated kinase-4 (IRAK-4) will interact and form a complex with MyD88 and interleukin-1 receptor associated kinase-1 (IRAK-1) through homotypic death domain (DD) interactions. Another kinase namely interleukin-1 receptor associated kinase-2 (IRAK-2) is also recruited through homotypic DD interactions. A complex formed between the DDs of MyD88, IRAK-4 and IRAK-2 is known as the Myddosome (Lin et al., 2010). The Myddosome will be discussed in detail later on.

During the signalling cascade, IRAK-4 phosphorylates IRAK-1. The phosphorylated IRAK-1 will dissociate from IRAK-4 and form a complex with tumour necrosis factor associated factor 6 (TRAF6) (Akira and Takeda, 2004). Signals will then be transmitted to several other proteins leading to the translocation of NF-κB to the nucleus. Once in the nucleus, NF-κB will induce the transcription of proinflammatory cytokines.
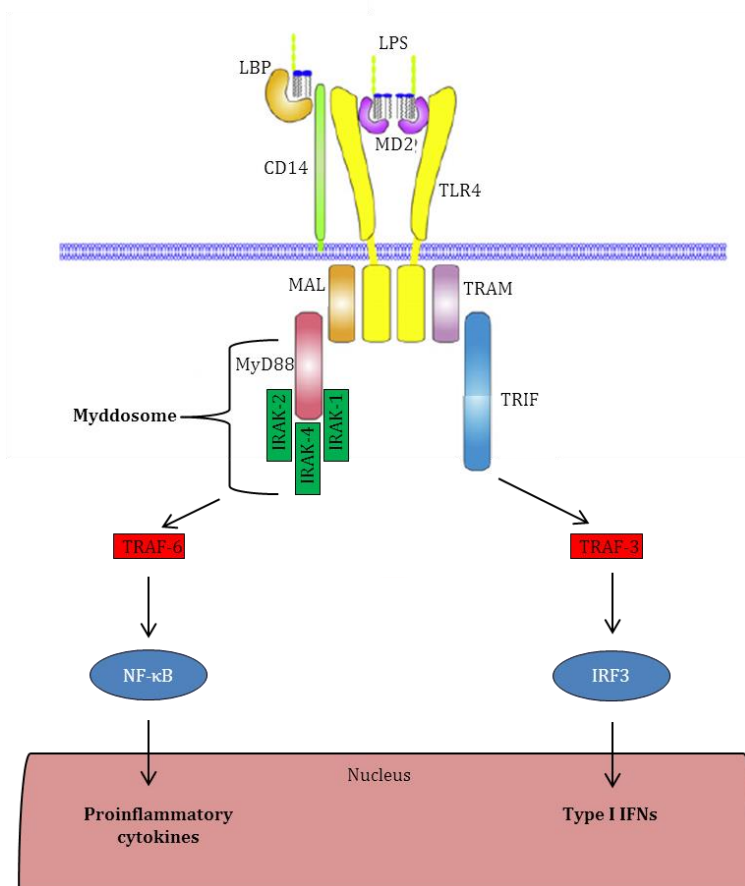
Figure 1.1: Overview of TLR4 signalling pathway (adapted from Brubaker et al., 2015). TLR4 in complex with coreceptor CD14 and accessory protein MD2 binds LPS leading to the activation of MyD88 dependent and MyD88 independent signalling pathways. These pathways culminate in activation of NF-κB and IRFs and the production of proinflammatory cytokines and type I IFNs, respectively.

1.3.2 MyD88 independent pathway.

LPS is transferred to TLR4 in the same way as in the MyD88 dependent pathway during signalling. The difference is that the adaptors used are TRAM and TRIF as shown in figure 1.1. TRAM interacts with TLR4 through homotypic TIR-TIR interactions. Thereafter, TRIF is recruited to the TLR4/TRAM complex, also through homotypic TIR-TIR interactions. Thereafter TRIF interacts with another protein TRAF3. From TRAF3, the signalling cascade will proceed all the way to IRF3. During the process, IRF3 is phosphorylated and then dimerizes (Li et al., 2009). Dimerised IRF3 will then translocate to the nucleus to stimulate the transcription of type I interferons.

1.4 TLR Structure and signalling at a molecular level

As previously mentioned, TLRs have two domains namely the ectodomain and the TIR domain. The ectodomain (ECD) has conserved leucine rich motifs (LRRs) which consist of approximately 20-30 amino acids that are arranged into a β-turn-α structure (Kobe and Kajava, 2001; Yin et al., 2015). These LRRs are glycosylated and causes the ECD to adopt a horseshoe structure as shown in figure 1.2.
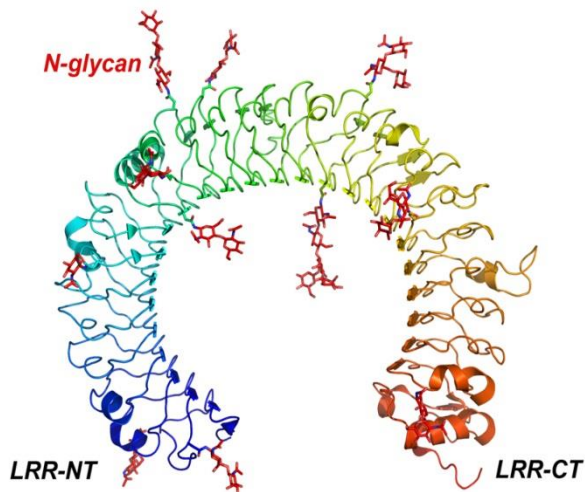
Figure 1.2: Crystal structure of human TLR3 ectodomain (Bell et al., 2005, adapted by Botos et al., 2011). As shown, the LRR motifs are assembled into a horse-shoe like structure and his highly glycosylated. The ECD is composed of 23 LRRs from the N-terminus to the C-terminus. Each LRR consists of a β-strand and an α-helix connected by loops.

Interestingly, ECDs transmit signals when they are in a dimerised state. A crystal structure of TLR4 dimer in complex with MD2 and LPS has been determined as shown in figure 1.3 (Park et al., 2009). There are more crystal structures of TLR ECDs that have been solved. However, we chose to show that of TLR4 for reasons that were mentioned earlier. The process of TLR dimerization has been a subject of considerable debate. Previously, there were conflicting views as to whether TLR dimers pre-exist or only form upon binding. It now seems their pre-existence is a widely accepted view (Berglund et al., 2015).
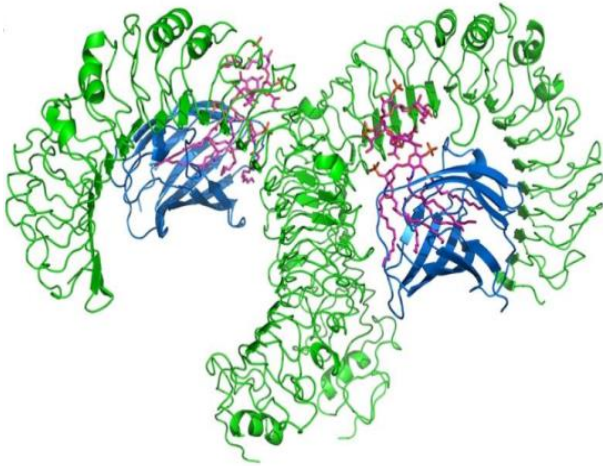
Figure 1.3: Structure of TLR4-ECD/MD2-LPS complex (Park et al., 2009, adapted by Botos eta l., 2011). The ribbon diagram shows TLR4-ECD dimer in green and two molecules of accessory protein MD2 (blue) bound to two LPS molecules in magenta.

TIR domains are made of five β-sheets that are surrounded by five α-helices (Jin and Lee, 2008). Of great importance, they have a "BB" loop, which connects βB strand and αB helix (Ve et al., 2015). This loop is essential for dimerization of the TIR domain (Jin and Lee, 2008). The BB loop also has a conserved proline residue that is critical for signalling (O'Neil and Bowie 2007; Ohnishi et al., 2009; Yin et al., 2015). For example, a P712H mutation on TLR4 in mice completely abolished signalling in response to LPS (Poltorak et al., 1998).

Isolated TIR domains of TLR1 and TLR2 were found to be monomeric in solution (Xu et al., 2000). This is in not surprising considering that the experiments were carried out *in vitro* in the absence of binding ligands and the ECD. In a different study, the structure of the isolated TLR10 TIR domain was solved as a dimer (Nyman et al., 2008). The structure of the TLR10 dimer is shown in figure 1.4.
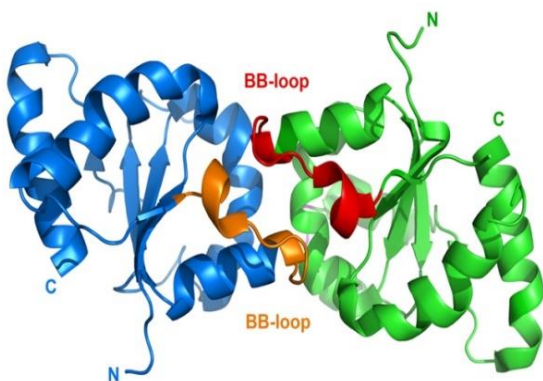
Figure 1.4: Structure of TLR10 TIR domain dimer (Nyman et al., 2008, adapted by Botos et al., 2011). The ribbon diagram shows TLR10 homodimer with two interacting BB-loops highlighted in red and gold.

1.5 Domains involved in TLR post-receptor signalling

There are two domains that are involved TLR post-receptor signalling. These are the TIR domain and the DD. The TIR domain has been described in detail in the previous section and this section will put more emphasis on the DD.

There are four adaptor proteins that link the TLRs to downstream kinases. These are MyD88, Mal, TRAM and TRIF. They all have a death domain and a TIR domain. The TIR domain is used to interact with the receptor through homotypic TIR-TIR interactions, whereas the DD is used to link the receptor to downstream kinases. So far, the TIR domain structures of all the adaptor proteins have been solved. They also have five β-sheets that are surrounded by five α-helices. However, there are structural differences in the loop and helical regions, between the TIR domains of adaptor proteins and receptors (Ve et al., 2015). For example, MyD88 has a longer and more flexible BB loop as compared to TLRs.

DDs are protein-protein interaction motifs belonging to a large superfamily. These motifs consist of ~ 120 amino acids (Gay and Gangloff, 2007). The name death domain was given to the novel TNF receptor 1 (TNFR1), domain, which was found to be involved in programmed cell death (apoptosis) through deletion mutagenesis (Tartaglia et al., 1993). The DD superfamily is considered one of the largest superfamily of protein domains (Reed et al., 2004). It comprises four subfamilies, namely death domain (DD), caspase recruitment domains (CARD), death effector domain (DED), and pyrin domain (PYD) (Park et al., 2007). Members of the DD superfamily are critical role players for apoptotic and inflammatory signalling (Park et al., 2007). Over 33 mammalian DD encoding genes have been identified, including TNF receptors (e.g. TNFR1) and numerous intracellular signalling proteins, particularly those involved in apoptosis, inflammation and necrosis such as MyD88, IRAKs, RIP1, Fas-associated protein with a DD (FADD), p53-induced protein with a DD (PIDD), RIP-associated ICH-1 homologous protein with a DD (RAIDD) and tumour necrosis factor protein with a DD (TRADD) (Park, 2011).

The unifying feature of the DD superfamily is a conserved antiparallel six-helical bundle fold arranged in the Greek key topology as shown in figure 1.5. Structurally, subfamilies that differ most are the DDs and CARDs, based on the orientation of the six α-helices (Weber and Vincenz, 2001). Members of the DD superfamily are able to self-associate and form hetero-complexes with high binding affinities (Weber and Vincenz, 2001).
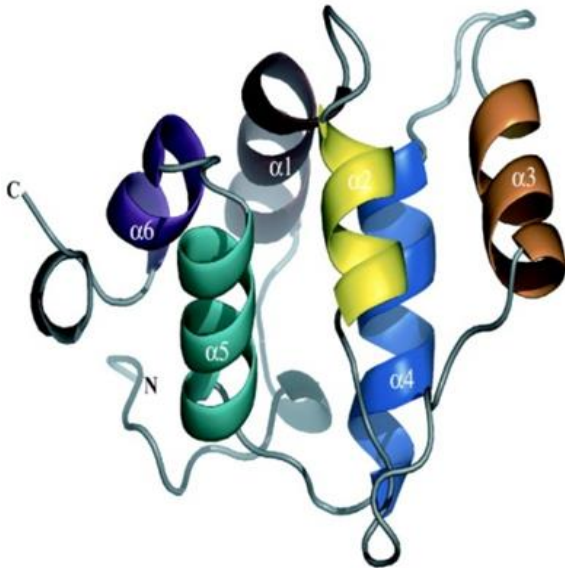
Figure 1.5: Crystal structure of murine IRAK-4 death domain (Lasker et al., 2005). The death domain is composed of six α-helices (α1-α6) connected together by loops as indicated.

1.6 Supramolecular organising centres in signalling

DD interactions facilitate the assembly of oligomeric signalling complexes known as supramolecular organising centres (SMOCs) (Kagan et al., 2014). An example of such an oligomeric assembly includes the Myddosome, a complex formed between the DDs of MyD88 and IRAK-4 with stoichiometries 7:4 and 8:4 (Motshwene et al., 2009). Another study published the crystal structure of the Myddosome variant which included IRAK-2 (Lin et al., 2010). This variant was found to have a stoichiometry of 6 MyD88 to 4 IRAK-4 and 4 IRAK-2 DD molecules as shown in figure 1.6.

To date, there is no evidence that suggests the Myddosome is physiological. The structure of the Myddosome variant is only made up of individual DDs and not full length proteins. The TIR domain of MyD88 and the kinase domains of IRAK-4 and IRAK-2 are the missing link in the structure. Questions have been raised if full length proteins will also form oligomeric structures with unusual stoichiometries. Another important question which we hope to address is: how is the assembly of the Myddosome regulated?
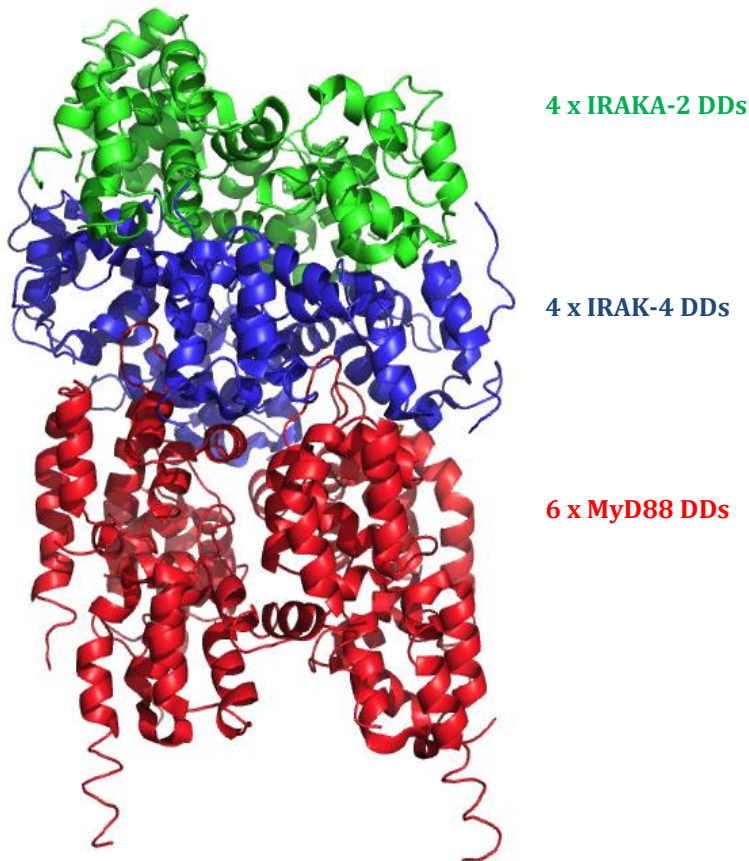
Figure 1.6: Crystal structure of the ternary Myddosome complex (Lin et al., 2010). The complex is formed through homotypic death domain interactions by 6 MyD88 molecules in green, 4 IRAK-4 molecules in blue and 4 IRAK-2 molecules in red.

1.7 Importance of MyD88 and IRAK-4 in TLR post-receptor signalling

We are interested in two proteins namely MyD88 and IRAK-4. These two proteins play an important role in the assembly of the Myddosome. The importance of each will be briefly discussed.

1.7.1 MyD88

As previously mentioned, MyD88 is an adaptor protein that plays a role in TLR signalling. It has a DD on its N-terminus and a TIR domain on its C-terminus. The importance of MyD88 in TLR signalling was first demonstrated using MyD88-deficient mice (Adachi et al., 1998 and Kawai et al., 1999). They were injected with high doses of LPS and they all survived. In contrast, the entire cohort of MyD88 wild type mice succumbed and died within 96 hours of receiving LPS injections. Studies on MyD88 deficient mice showed that lack of MyD88 leads to a poor inflammatory response.

1.7.2 IRAK-4

IRAK-4 is a kinase that phosphorylates IRAK-1 in addition to autophosphorylating itself during TLR signalling. It has an N-terminal DD, followed by a linker and a kinase domain (Kuglstatter et al., 2007). Structures of both its domains have been solved. The crystal structure of its DD was solved first and it showed a five helical bundle as expected (Lasker et al., 2005). Crystal structures of its kinase domain were solved roughly at the same time (Wang et al., 2006; Kuglstatter et al., 2007). These structures show its ATP binding site (K213) and its catalytic site (D311). They also showed autophosphorylation sites which we will discuss in the next section.

Functional studies on IRAK-4 knockout mice showed that they were resistant to lethal doses of LPS (Suzuki et al., 2002). They lacked a cytokine response when compared to IRAK-4 WT mice. Another study identified three children who do not express IRAK-4 (Picard et al., 2003). These children had a poor inflammatory response despite having recurrent bacterial infections due to *Streptococcus pneumoniae* and *Staphylococcus aureus*. Both studies show that the absence of IRAK-4 leads to a poor inflammatory response. These findings make IRAK-4 an ideal drug target for inflammatory diseases.

1.8 IRAK-4 autophosphorylation

Phosphorylation is a common post-translational modification of proteins in eukaryotic cells (Olsen and Mann, 2013). Enzymatic phosphorylation was first reported in 1954 on the protein casein (Burnett and Kennedy, 1954). Phosphorylation is a reversible event that is catalyzed by protein kinases which covalently attach a dianionic γ-phosphate group from MgATP to protein side chains (Chen and Cole, 2015). Eukaryotic protein kinases mainly phosphorylate the hydroxyl side chains of serine, threonine and tyrosine residues whereas prokaryotic protein kinases frequently phosphorylate histidine residues. However, some kinases do not only phosphorylate substrate but also phosphorylate themselves, an event known as autophosphorylation (Beenstock et al., 2016).

The first paper to be published on IRAK-4 gave hints that it was autophosphorylating (Li et al, 2002). IRAK-4 was incubated with a peptide substrate in the presence of γ-P^{32}ATP. Autoradiographs showed that there was a signal at the expected molecular weight of the substrate and that of IRAK-4. This was the first indication that IRAK-4 was autophosphorylating. A follow up study mapped IRAK-4 autophosphorylation sites using LC-MS/MS and went further to show that its autophosphorylation was an intramolecular event (Cheng et al., 2007). Some of the mapped sites, especially those in the activation loop had been previously seen on the crystal

structure of the IRAK-4 kinase domain shown in figure 1.7 (Wang et al., 2006). The structure was solved in complex with an inhibitor known as compound 1. It shows that the N-terminal lobe is mainly composed of a five-stranded antiparallel β-sheet (β1-5) and a single α-helix (αC), and a unique loop present between the N-terminal helix and the first β-strand. The C-terminal lobe is larger, primarily α-helical and has an activation loop on which several autophosphorylation sites were identified. These are Thr-342, Thr-345 and Ser-346.
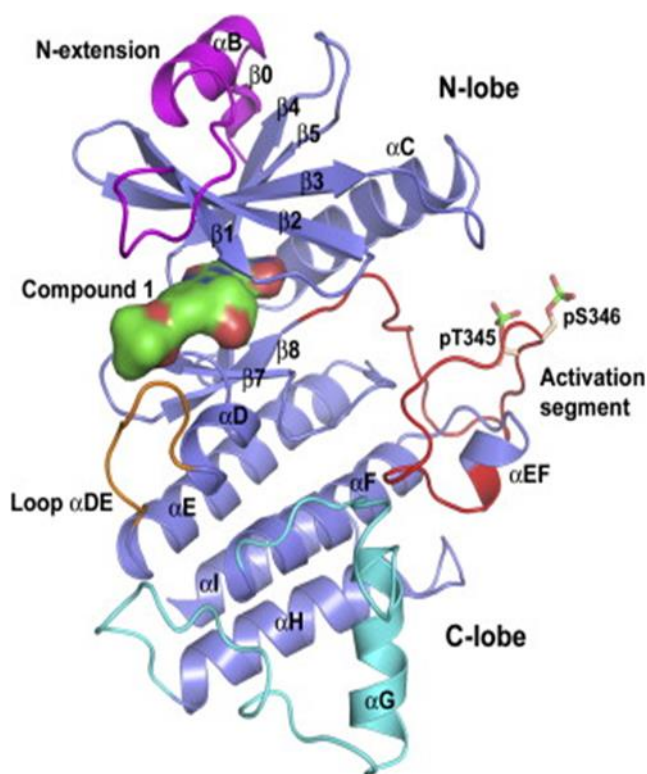
Figure 1.7: Crystal structure of IRAK-4 kinase domain attached to inhibitor compound 1 (Wang et al., 2006). The N-lobe of the KD is composed of five-stranded antiparallel β-sheet and a single α-helix (αC), followed by a C-terminal lobe that is primarily α-helical. The activation loop is part of the C-lobe shown in red. Autophosphorylation sites seen in the activation loop are T345 and S346. The part shown in magenta represents the N-terminal extension of the KD.

Autophosphorylation within IRAK-4 is not only limited to the kinase domain but also occurs within the DD. So far, two autophosphorylation sites have been identified in the DD (Motshwene, 2008). These are Serine 8 and Threonine 62 which are shown on figure 1.8. We generated this figure by loading the protein data bank (PDB) code 2A9I (Lasker et al., 2005) on PyMOL and went on to show the two amino acids as a red and green circle. It is evident from the structure that Serine 8 and Threonine 62 are located in the flexible regions of the IRAK-4 DD.
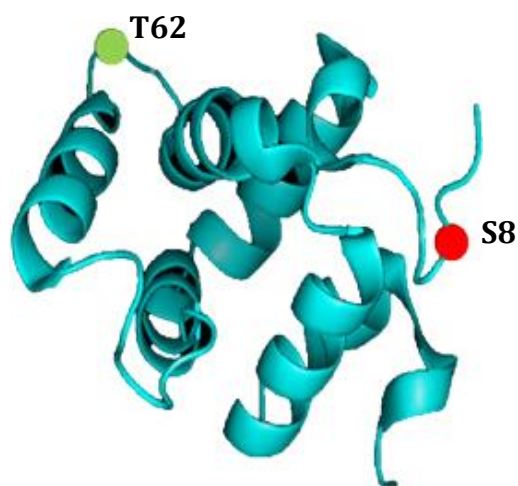
Figure 1.8: IRAK-4 death domain phosphorylation sites (adapted from Lasker et al., 2005). Two phosphorylation sites, S8 and T62, are indicated on the structure of murine IRAK-4 DD. S8 is located on the N-terminal loop whereas T62D is located on the loop that links H3 and H4.

1.9 Phosphomimetics in studying autophosphorylation and phosphorylation

We are interested in determining the functional roles of the two autophosphorylation sites within the DD of IRAK-4. IRAK-4 is a serine/threonine kinase. It adds phosphate groups on serine and threonine resulting in a modification as shown in figure 1.9. A double negative charge is added on these two amino acids thus making them larger.
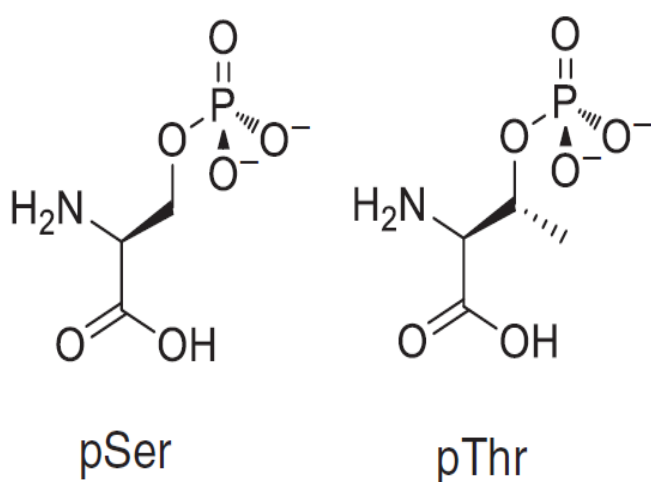
Figure 1.9: Structures of phosphoserine and phosphothreonine (adapted from Chen and Cole, 2015). A phosphate group which carries a double negative charge is attached onto the amino acids, making them larger.

It has not been possible for us to create phosphorylated IRAK-4 DDs. We therefore came up with an idea of mimicking autophosphorylation by creating phosphomimetics. Phosphomimetics are amino acids substitutions where amino acids of interest are replaced by negatively charged amino acids. These negatively charged amino acids are usually aspartic and glutamic acid as shown in figure 1.10. By creating phosphomimetics, we are hoping to mimic autophosphorylation on the DDs of IRAK-4 *in vitro*.
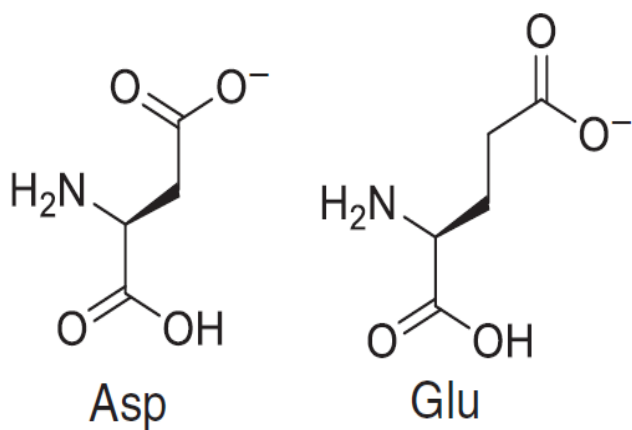
Figure 1.10: The structures of phosphomimetic aspartate and glutamate (adapted from Chen and Cole, 2015). The acidic Asp and Glu residues have a single negative charge.

1.10 Aim and problem identification

DDs of MyD88 and IRAK-4 form an oligomeric complex known as the Myddosome. The Myddosome assembles into DDs of MyD88 and IRAK-4 with stoichiometries 7:4 and 8:4. This complex has a large molecular weight considering that the individual DDs are less than 20 KDa. So far, the regulatory mechanism of assembly and dissociation of the Myddosome is not known. This project is aimed at finding a mechanism that regulates the assembly of this complex. Our hypothesis is that autophosphorylation is a switch that regulates this assembly. Two amino acids namely serine 8 and threonine 62 within the DD of IRAK-4 were found to autophosphorylate. We intend to substitute these two amino acids with aspartic acid, a negatively charged amino acid that will mimic autophosphorylation. Thereafter, we will assemble complexes by mixing the DD of MyD88 with either wild type or phosphomimetic mutants of IRAK-4 DDs. By so doing, we hope to see if the phosphomimetic mutants will assemble into the Myddosome. Failure to assemble will tell us that autophosphorylation is important for complex formation and may play a regulatory role.

1.11 Hypothesis

IRAK-4 phosphomimetics disrupt the assembly of the Myddosome complex.

1.12 Objectives

(a) Express and purify MyD88 and its death domain

(b) Express and purify IRAK-4 death domain

(c) Express and purify S8D and T62D phosphomimetic mutants of IRAK-4 death domain

(d) Reconstitute wild type and mutant Myddosome complexes

(e) Characterize the wild type and mutant Myddosome complexes using biophysical techniques

CHAPTER 2: MATERIALS AND METHODS

2.1 Introduction

This chapter provides the details of how the study was conducted. It is divided into three sections, namely molecular biology methods, general biochemical methods and biophysical techniques. Maps of plasmids which were used to clone the genes of interest by my predecessors are shown in the appendix.

Table 2.1: Buffers used in the study.

Buffer/Solution	Composition
Lysis buffers: MyD88^{FL-WT}	100 mM Tris pH 8.0, 200 mM NaCl, and 0.1% Triton X-100
MyD88^{DD-WT}/IRAK-4^{DD-WT}	50 mM Tris pH 8.0 and 30 mM NaCl
IRAK-4^{DD-S8D}/ IRAK-4^{DD-S8D}	100 mM Tris pH 8, 200 mM NaCl, 0.5% Triton X-100 and 5 mM DTT
PBS buffer, pH 7.4	137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO$_4$, 1.8 mM KH2PO$_4$, pH 7.4
TE buffer, pH 8.0	10 mM Tris pH 8.0 and 1 mM EDTA
Elution buffer for GST-tagged proteins	50 mM Tris pH 8.0, 50 mM NaCl and 40 mM reduced glutathione
HiTrap Chelating column equilibration buffer	50 mM Tris pH 8.0 and 30 mM NaCl
Elution buffer for His-tagged proteins	50 mM Tris pH 8.0 30 mM NaCl and 1 M imidazole
Size exclusion chromatography (SEC) buffer	30 mM Tris pH 8.0 and 50 NaCl
SDS-PAGE running buffer	25 mM Tris-HCl, 192 mM glycine, 0.1% (w/v) SDS
4X SDS-PAGE sample buffer	8% SDS (w/v), 40% Glycerol (v/v), 5% beta-mercaptoethanol, 240 mM Tris pH 6.8 and 0.04% bromophenol blue
TAE buffer	40 mM Tris-HCl pH 7.5, 20 mM sodium acetate, 1 mM EDTA, adjusted to pH 8.2 with acetic acid

2.2 Molecular biology methods

Table 2.2: Expression constructs used in this study.

Plasmid DNA (tag)	Insert	Expressed as:
pDEST-544 (His$_6$-NusA)	MyD88^{WT-FL}	N-His$_6$-NusA-3CP-MyD88^{WT-FL}
pMCSG7 (His$_6$)	MyD88^{WT-DD}	N-His$_6$-MyD88^{WT-DD}-Strep
pMCSG7 (His$_6$)	IRAK-4^{WT-DD}	N-His$_6$-IRAK-4^{WT-DD}
pETG30A (His$_6$-GST)	IRAK-4^{DD-S8D}	N-His$_6$-GST-3CP-IRAK4^{DD-S8D}
pETG30A (His$_6$-GST)	IRAK-4$^{DD-T62D}$	N-His$_6$-GST-3CP-IRAK4$^{DD-T62D}$

NB: 3CP – 3C protease cleavage site

2.2.1 Bacterial growth media, agar plates and antibiotics

Table 2.3: Bacterial growth media.

Media	Composition (1 L)
LB	10 g tryptone, 5 g yeast extract and 5 g NaCl
2X YT	16 g of tryptone, 10 g of yeast extract and 5 g of NaCl
Autoinduction	20 g tryptone, 5 g yeast extract, 5 g NaCl, 3 g KH$_2$PO$_4$ and 6 g Na$_2$HPO$_4$ supplemented with 10 mL of 60% glycerol, 5 mL of 10% glucose and 25 mL of 8% lactose, pH adjusted to 7.2.

Bacterial growth media was prepared in deionized water as shown in table 2.3. Media was then autoclaved and then cooled and stored at room temperature before use. LB agar was prepared by making LB media as in table 2.3 and adding 15g/L of bacterial agar to it before autoclaving. The LB agar was allowed to cool down to 42°C before adding the appropriate antibiotic. Two antibiotics were used throughout the study, namely ampicillin and chloramphenicol. The stock and working concentration for the former were 30 mg/mL and 0.03 mg/mL respectively and 100 mg/mL and 0.1 mg/mL for the latter.

2.2.2 Competent cells

E. coli competent cells were prepared using the Hanahan's calcium chloride method described by Chang et al., 2017.

2.2.3 Transformation of competent *E. coli* cells

Chemically competent *E. coli* cells (either Top10 or DH5α, BL21-CodonPlus (DE3)-RIL or Rosetta 2) were thawed on ice and thereafter mixed with 2 μL plasmid. The mixture was incubated on ice for 20 min, followed by heat shocking at 42°C for 50 sec and then back on ice for 2 min. Thereafter, the transformed cells were mixed with 200 μL LB media and incubated at 37°C with constant shaking at 170 rpm 1.5 hrs. About 50-100 μL of cells were plated on agar plates containing appropriate antibiotic/s and thereafter incubated at 37°C overnight. Single colonies were picked and tested for the presence of plasmid DNA with insert of interest by PCR.

2.2.4 Polymerase chain reactions

Table 2.4: Components of PCR reaction mixture.

Template DNA	1 μL (20 ng)
Forward primer	0.5 μL
Reverse primer	0.5 μL
ThermoPol Buffer (10X)	5 μL
$MgSO_4$	3 μL
dNTPs	5 μL
Vent polymerase (NEB)	1 μL
Deionized water	34 μL
Total	50 μL

Table 2.5: Steps for PCR amplification.

Step	Time (min)	Temperature (°C)
(1) Initial denaturation	1	95
(2) Denaturation	1	95
(3) Annealing	1	65
(4) Extension	2	72
(5) Final extension	10	72

PCR was prepared and run for 29 cycles in a PCR thermocycler (BioRad, USA). Amplification was checked by running 20 μL samples on agarose gel electrophoresis. Shown below are primers for PCR amplification of WT MyD88 full length, together with corresponding amino acids (underlined). The forward and reverse complement primers were used for PCR amplification. Amino acids highlighted in red represent the recognition sequence for 3C protease.

Primers for amplification of MyD88 DD

(a) MyD88 full-length forward

<div style="text-align:center">L E V L F Q G P S M A A G G P G A G</div>

5' C ACC CTG GAA GTT CTG TTC CAG GGG CCC TCT ATG GCT GCA GGA GGT CCC GGC GCG GGG TC 3'

(b) MyD88 full-length reverse

5' CGC CTT GCC AAG GCC TTG TCC CTG CCC TGA 3'

2.2.5 Gateway cloning: LR recombination of WT MyD88 full length

Table 2.6: Components of LR Recombination reaction.

Entry clone: pENTR/D-TOPO-MyD88^{FL-WT} (143 ng/μL)	1 μL
Destination vector: pDEST-544 (129 ng/μL)	1.2 μL
TE buffer	6.8 μL
Gateway® LR Clonase™ II (Invitrogen)	1 μL
Total	10 μL

WT MyD88 full length (MyD88^{FL-WT}) expression clone (pDEST-544-His$_6$-NusA-MyD88^{FL-WT}) was generated through LR recombination reaction following manufacturer's instructions. In LR recombination, a gene cloned in the *attL1* and *attL2* sites of an entry vector (entry clone) is transferred into a destination or expression vector, betweenn *attR1* and *attR2* sites. In this case, the MyD88^{FL-WT} gene was transferred from pENTR/D-TOPO entry vector into pDEST-544 destination vector shown in figure 2.1 (Plasmid maps of vectors used for cloning are shown in figure 9.1, appendix). The LR recombination reaction is catalysed by an enzyme called LR clonase

II (Invitrogen). The reaction mixture (shown in table 2.6) was set up, mixed gently and then incubated at room temperature for 1 hr. After incubation, 1 µL of proteinase K was added to the reaction mixture and incubated at 37ºC for 10 min in a water bath. Thereafter, 2 µL of the reaction mixture was used to transform chemically competent Top10 or DH5α cells which were then plated on ampicillin plates as explained in section 2.2.3. Single colonies were picked and cultured as described previously. Plasmid DNA isolated from these cells was used to transform chemically competent BL21-CodonPlus (DE3)-RIL or Rosetta 2 cells, which were also plated on ampicillin plates.
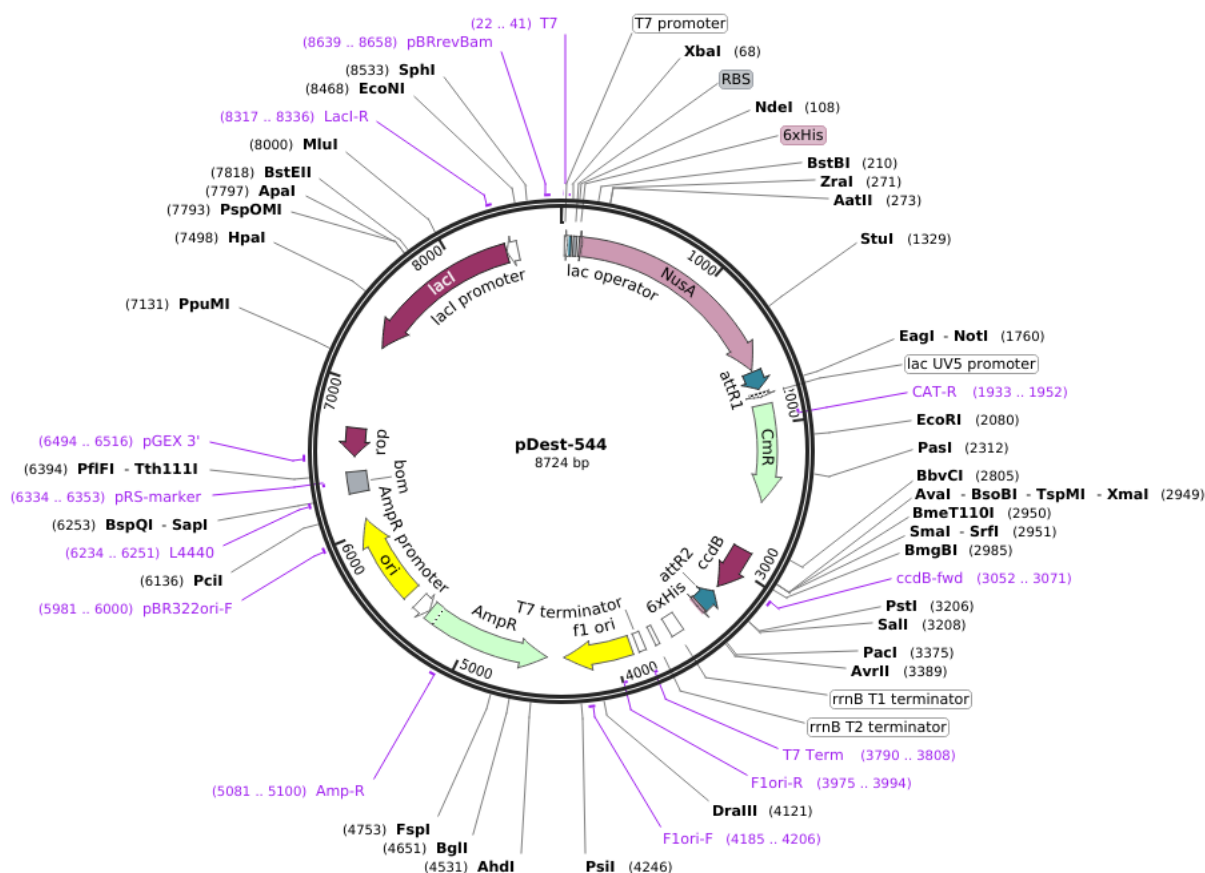
Figure 2.1: Plasmid map of pDest-544 expression vector. MyD8[-FL-WT] gene cloned in an entry vector was transferred into pDest-544 expression vector by LR recombination reaction as explained in section 2.2.5.

2.2.6 DNA sequencing

Expression constructs used in this study were sent for sequencing at central analytical facilities, Stellenbosch University and Department of Biochemistry, Cambridge University. Forward and reverse primers for the T7 promoter were used for sequencing.

2.2.7 Plasmid DNA isolation and purification

A single colony of transformed DH5α or Top10 cells was picked from an agar plate and inoculated into 5 mL LB media with ampicillin (100 µg/mL) and allowed to grow overnight at 37°C with constant shaking at 170 rpm. Plasmid DNA was isolated using the GeneJET plasmid mini prep kit (Thermo Fisher Scientific, USA) as per manufacturer's instructions and eluted with 50 µL Milli-Q water. Concentration of eluted plasmid DNA was measured at 260 nm (A_{260}) using a Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies).

2.2.8 Agarose gel electrophoresis

A 1% (w/v) agarose gel was prepared by dissolving 1 g of agarose in 100 mL TAE buffer. PCR products were mixed with GelRed (Thermo Fisher Scientific, USA) supplemented with loading dye using a 3:1 ratio (18 µL DNA sample + 6 µL GelRed) and then electrophoresed at 100 V in 1X TAE buffer for 45 to 60 min depending on the DNA marker used. DNA bands were visualized with a Gel Doc XR+ imager (BioRad, USA).

2.3 General biochemical methods

2.3.1 Protein production, purification and analysis

(a) Protein expression trials

A single colony of BL21 Codon Plus transformed cells carrying a plasmid DNA of interest was inoculated in 5 mL LB media with ampicillin (100 µg/mL) and chloramphenicol (30 µg/mL) and incubated with shaking until an OD_{600} between 0.6 and 1 was reached. The culture was divided into 1 mL (uninduced) and 4 mL (induced). Thereafter, the 4 mL culture was then induced to express protein by adding 1 mM IPTG and then incubated together with the 1 mL (uninduced) culture with shaking at 170 rpm and 37°C overnight.

The following morning, 1 mL was taken from the induced culture and analysed for protein expression together with the uninduced culture. From the expression trial cell cultures, 1 mL the induced and uninduced culture were used for analysis. Cells were harvested by centrifugation at 7000 rpm (Eppendorf Centrifuge 5417C) for 5 min. Thereafter, the cell pellets were resuspended in 50 µL of BugBuster Protein Extraction Reagent (Novagen) and incubated at room temperature

for 10 to 15 min. The lysates were then centrifuged at 14000 rpm (Eppendorf Centrifuge 5417C) for 3 min and the supernatants (soluble fractions) were transferred to new Eppendorf tubes. The cell pellets were resuspended in the same volume of 1/5 diluted BugBuster and spun down at the 14000 rpm for 3 min. Thereafter, the pellet was resuspended in the same volume of 1/5 diluted BugBuster and this was the insoluble fraction. Soluble and insoluble fractions were analysed by SDS-PAGE.

(b) Protein production by IPTG induction: Starter culture and main culture

Starter culture was prepared by inoculating BL21 Codon-Plus cells or Rosetta 2 cells expressing protein of interest into 100 mL of LB media containing ampicillin and chloramphenicol. The starter culture was grown overnight at 37°C at 170 rpm. The cells were harvested by centrifugation the following morning at 7000 rpm (Sorval Instrument RC5C) and 4°C for 10 min and thereafter resuspended in 5 to 10 mL fresh LB media. The fresh cell suspension was used to inoculate the main culture. Each litre of LB media containing ampicillin and chloramphenicol was inoculated with 1 mL of cells and incubated with shaking at 170 rpm at 37°C until an OD_{600} between 0.6 and 1 was reached. The culture was then induced to express protein by adding 1 mM IPTG and then incubated with shaking at 170 rpm and 15°C overnight.

(c) Protein expression by autoinduction: Starter culture and main culture

A starter culture for autoinduction was prepared as described in 2.3.1(a). Each litre of autoinduction media was supplemented with heat sterilised 60% glycerol (0.6% v/v), 10% filtered glucose (0.05% v/v) and 8% filtered lactose (0.2% v/v), ampicillin (100 μg/ mL) and chloramphenicol (30 μg/mL) and then inoculated with 1 mL of cells (starter culture). The inocula were cultured for 3 hrs with shaking at 170 rpm and 37°C. Thereafter, the temperature was lowered to 15°C and the culture was left to grow for 48 hrs while shaking at 170 rpm.

(d) Cell harvesting and cell lysis

Cell cultures were harvested by centrifugation at 7000 rpm for 15 min at 4°C. Cell pellets resulting from 4 L and 8 L cultures were resuspended in 80 mL and 160 mL of ice cold lysis buffer (20 mL lysis buffer per 1 L cell pellet) respectively. Two methods, sonication and homogenisation using an Emulsiflex were either used to lyse cells. The equipment was rinsed thoroughly with deionized water prior to use. Using sonication, the ultrasonic probe was adjusted into the cell suspension, which was placed on ice. The cells were then lysed at a pulse of 30 and 40% amplitude for 10 cycles and 20 cycles (30 sec sonication and 30 sec break) per 80 and 160 mL cell suspension respectively. Cell lysis by homogenisation was done by passing the cell suspension through an Emulsiflex two and four times per 80 and 160 mL cell suspension respectively. The resulting cell lysates were spun down at 28000 rpm for 1 hr in a Beckman Coulter Optima LE-80K Ultracentrifuge using an SW28 rotor (Beckman Coulter). The supernatant was collected and loaded either on a column with GST resin or a HiTrap column (GE Healthcare) depending on the protein.

(e) GST affinity chromatography

Approximately 80 mL of supernatant containing either His_6-GST tagged IRAK-4^{DD-S8D} or IRAK-$4^{DD-T62D}$ mutants was loaded onto a column containing 5 mL Glutathione Sepharose 4B resin (GE Healthcare) that had been pre-equilibrated with PBS. The sample was loaded by gravity flow in the cold room (4°C). Thereafter, unbound proteins were washed off with 100 mL PBS. A buffer containing 40 mM reduced glutathione, 50 mM NaCl and 50 mM Tris (pH 8.0) was used to elute the IRAK-4 mutants. About 10 mL fractions were collected and the presence of protein was determined using the Bradford reagent. Thereafter, the double His_6-GST tag was cleaved off at room temperature for 2 hrs by adding 3C protease (200 μl) to each 10 mL protein sample or fraction. Protein samples were then collected and analysed by SDS-PAGE.

(f) Nickel affinity chromatography

A pre-packed 5 mL HiTrap chelating column (GE Healthcare) was charged with 100 mM nickel sulphate and equilibrated with buffer A (50 mM Tris pH 8, 30 mM NaCl). 10 mM imidazole was added to the supernatants containing either MyD88^{FL-WT}, MyD88^{DD-WT}, IRAK-4^{DD-WT}, IRAK-4^{DD-S8D} or IRAK-$4^{DD-T62D}$. It was then loaded on the column and eluted with buffer B (50 mM Tris pH 8, 30 mM NaCl, 1 M imidazole) on the AKTA purifier by linear gradient. Thereafter, protein samples were collected and analysed by SDS-PAGE.

(g) Size exclusion chromatography

Following purification by nickel affinity chromatography, MyD88 and IRAK-4 DDs were concentrated to 2.1 mg/mL (MyD88^{DD-WT}), 2.4 mg/mL (IRAK-4^{DD-WT}), 2.6 mg mL (IRAK-4^{DD-S8D}) and 4.1 mg/mL (IRAK-4$^{DD-T62D}$) at 5000 x g using Amicon Ultra-15 (30000 MWCO) concentrators (Merck). Concentrated MyD88 and IRAK-4 DDs were then further purified by size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare) that had been equilibrated with 1 column volume of SEC buffer (30 mM Tris, pH 8.0, 50 mM NaCl). The flow rate was set at 0.5 mL/min and 500 μL fractions were collected and analysed for purity by SDS-PAGE.

(h) Reconstitution of the MyD88/IRAK-4 Myddosome complexes

The WT and mutant Myddosome complexes were reconstituted by mixing 2 mL of MyD88^{DD-WT} (2.1 mg/mL) with either 2 mL of IRAK-4^{DD-WT} (2.4 mg/mL), 2 mL of IRAK-4^{DD-S8D} (2.5 mg/mL) or 1 mL of IRAK-4$^{DD-T62D}$ (4 mg/mL) sample. The samples were then concentrated in a Vivaspin (6 5000 MWCO) concentrator to 500 μL at 5000 x g. Thereafter, 250 μL from each of the concentrated samples were loaded on a Superdex 200 10/300 GL column that had been pre-equilibrated with 1 column volume of SEC buffer. The flow rate was set at 0.5 mL/min and 500 μL fractions were collected and analysed on SDS-PAGE.

(I) Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS acrylamide gels (12% and 15%) were prepared using the Novex 1 mm and 1.5 mm gel cassettes (Life technologies) as shown in table 2.7 and 2.8. Protein samples were mixed with 4X sample buffer (20 μL protein sample + 5 μL sample buffer) and heat denatured at 95°C for 5 min. Thereafter, they were electrophoresed in a Mini Gel Tank (Invitrogen) and run at 200V for 45 min to 1h depending on the gel percentage. Upon completion of the run, gels were washed with deionized water, stained with Simply Blue SafeStain (Invitrogen) as per manufacturer's instructions. Deionized water was used to destain the gels for protein band visualisation.

Table 2.7: Solutions for preparing 12% and 15% separating gels for SDS-PAGE.

	12%	**15%**
30% Acrylamide	16 mL	20 mL
1.5 M Tris, pH 8.8	10 mL	10 mL
10% SDS	0.4 µL	0.4 mL
10% APS	0.4 µL	0.4 mL
TEMED	0.01 µL	0.016 mL
Deionized water	13.2 mL	9.2 mL
Total	40 mL	40 mL

Table 2.8: Solutions for preparing 5% stacking gels for SDS-PAGE.

30% Acrylamide	1.7 mL
1 M Tris, pH 6.8	1.25 mL
10% SDS	0.1 mL
10% APS	0.1 µL
TEMED	0.01 µL
Deionized water	6.8 mL
Total	10 mL

2.3.2 Quantification of DNA and protein concentration

Plasmid DNA concentration was measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies) at 260 nm. Protein concentration on the other hand was measured using the Bradford assay. 10 µL of BSA standards (0.2, 0.4, 0.6, 0.8 and 1 mg/mL) were mixed with 200 µL Bradford reagent (Bio-Rad) and the total volume was made up to 1 mL with deionised water. The absorbance of the mixture was measured at 595 nm using a spectrophotometer. A standard curve was then plotted from which concentrations of unknown protein samples could be determined.

2.3.3 Mass spectrometry

SDS-PAGE gel with two bands corresponding to the size of MyD88[FL-WT] was rinsed with deionized water. The bands were excised ($\sim$ 1 mm x 1mm pieces), placed in a 1.5 mL Eppendorf tube and then suspended and/or dehydrated in 1 mL of 100% acetic acid with vortexing. Dehydrated bands were then sent to the Central Analytical Facility, Stellenbosch University for identification by liquid chromatography-mass spectrometry (LC-MS).

2.3.4 Biophysical analysis

(a) Dynamic light scattering (DLS)

A ZETASIZER NANO-S (Malvern) was used to carry out dynamic light scattering experiments. Samples containing the Myddosome complexes in SEC buffer were loaded in 12.5 x 12.5 x 45 mm cuvettes. For each sample of complex, three independent DLS measurements (16 scans or cycles) were performed at 20°C.

(b) Analytical ultracentrifugation (AUC)

Analytical ultracentrifugation of the three Myddosome complexes was performed using the Beckman Optima XL-I ultracentrifuge. The eppon double-sector compartment of the ultracentrifuge was loaded with 400 μl of the complex (in SEC buffer) in the sample compartment and 400 μl of size exclusion chromatography buffer in the reference compartment. Sample compartment for cell 1 (WT Myddosome complex, 1.8 mg/mL), cell 2 (IRAK-4$^{DD\text{-}S8D}$ Myddosome complex, 1.8 mg/mL) and cell 3 (IRAK-4$^{DD\text{-}T62D}$ Myddosome complex, 1.2 mg/mL). The AUC machine was run at 60 000 rpm and data was recorded at a wavelength of 280 nm per minute intervals.

CHAPTER 3: Expression and purification of MyD88 full length, MyD88 and IRAK-4 DDs

3.1 INTRODUCTION

E. coli is the preferable organism of choice for the expression of recombinant proteins. It grows rapidly and can attain high densities within short periods of time (Rosano and Ceccarelli, 2014). Its growth media is relatively cheap when compared to other expression hosts like Sf9 cells and mammalian cell lines (Gechelle et al., 2015). Just like any other host, there is no guarantee that that recombinant protein will express at high levels or if it will have functional activity. Also, there remains a possibility that the recombinant protein will be insoluble, leading to the formation of inclusion bodies (Rosano and Ceccarelli, 2014). In this study, we expressed human MyD88 and IRAK-4 in *E. coli*. Our aim was to isolate and purify them so that they can be used in the reconstitution of the Myddosome complex in later experiments.

3.2 RESULTS

3.2.1 Cloning, expression of and purification of WT MyD88 full length

WT MyD88 full length (MyD88^{FL-WT}) was initially cloned in previous work by my supervisor, Dr Motshwene into pENTR/D vector. This construct was passed onto me for use in my project. I firstly performed an LR recombination to transfer the MyD88 gene onto an expression vector namely pDEST 544. This vector expresses proteins that have a His$_6$-NusA tag attached onto the N-terminus of the protein of interest, in this case MyD88^{FL-WT}. After the LR recombination reaction was performed, we wanted to verify the presence of MyD88^{FL-WT}. We opted to do this in three ways. Firstly, by a PCR reaction to check if we can amplify a gene fragment corresponding to its size. Secondly, by doing expression trials to determine if we can express the fusion protein (His$_6$-NusA-MyD88^{FL-WT}) of corresponding molecular weight. Thirdly, by doing peptide mapping using a mass spectrometer on the expressed protein.

A PCR reaction was set up using the pDEST 544 vector that was supposed to contain full length MyD88 from the LR recombination reaction as template. Bands that have a molecular weight of 783 bp were observed on the agarose gel as shown in figure 3.1, lane 1. This then led us to believe that full length MyD88 was successfully transferred from the pENTR/D vector into the pDEST 544 vector.
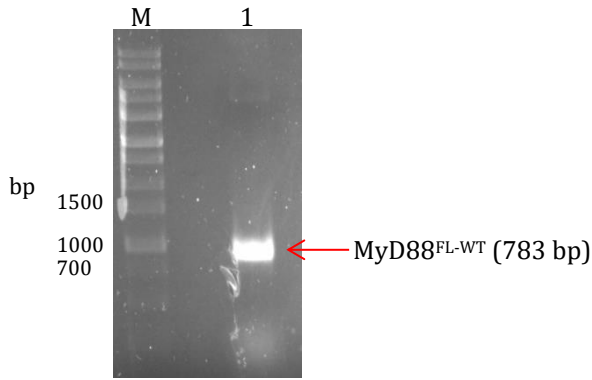
Figure 3.1: Agarose gel electrophoresis (1% agarose) of MyD88^{FL-WT}. MyD88^{FL-WT} (783 bp) was amplified by PCR using MyD88^{FL-WT} specific primers. Lane 1 is examined MyD88^{FL-WT} products. Lane M contains 1kb size marker.

We then went on to do expression trials in order to verify our earlier result. BL21-CodonPlus (DE3)-RIL cells that were transformed with pDEST 544 that contained full length MyD88 were grown in 5 ml autoinduction media. Cells resulting from the expression trial were harvested and lysed. The cell lysate was centrifuged and the resulting supernatant and pellet were analysed by SDS-PAGE as soluble and insoluble fraction, respectively as shown in figure 3.2A. Most of the fusion was found in the insoluble fractions (indicated with red arrows in lane 2 and 4, figure 3.2A).

Since most of the expected MyD88 fusion was in the insoluble fraction, we then decided to lower the growth temperature to 15°C. It has been shown that low temperatures enhance solubility due to the slow folding dynamics involved (San-Miguel et al., 2013). A 1 L culture was grown in autoinduction media and cells were harvested, lysed and spun down. The resulting supernatant was loaded onto a 5 ml HiTrap Chelating column that had been charged with nickel. The flowthrough was collected and ran on an SDS gel as shown in figure 3.2B, lane 1. Thereafter, the column was washed to remove unbound proteins. Collected samples of unbound proteins were also run on an SDS gel, in sequence of their collection (figure 3.2B, lane 2, 3). After unbound proteins were washed, the fusion protein was eluted from the column in a linear gradient of imidazole up to 500 mM. The eluted fusion protein was mixed and incubated with 3C protease in order to cleave MyD88^{FL-WT} from its binding partner, His$_6$-NusA. After the incubation step, the sample was analysed by an SDS gel as shown in figure 3.2B, lane 4. Two bands migrating close together and corresponding to the expected size of MyD88^{FL-WT} (33 kDa) in lane 4 (shown in a red box) were observed. These bands, T1 (top band) and T2 (bottom band) were excised from the gel and sent for peptide mapping in order to confirm if they were indeed that of MyD88^{FL-WT}. Peptide mapping was carried out by liquid chromatography-mass spectrometry (LC-MS). Briefly, the excised gel bands were treated with trypsin which cleaves peptide bonds on the C-terminus

of amino acids lysine or arginine. The resulting peptides were separated by HPLC and loaded onto a nanocapillary in the LC-MS. A voltage of 2kV was applied on the nanocapillary and this resulted in the emission of nanodrops. The mass to charge ratio of the nanodrops were acquired as shown in figure 3.3. These were then submitted onto a database to determine the identity of our protein. The results confirmed that our protein of interest was MyD88[FL-WT] as shown on table 3.1. Interestingly, the two bands, T1 and T2 were both of MyD88[FL-WT]. This result suggested that MyD88[FL-WT] was degrading.
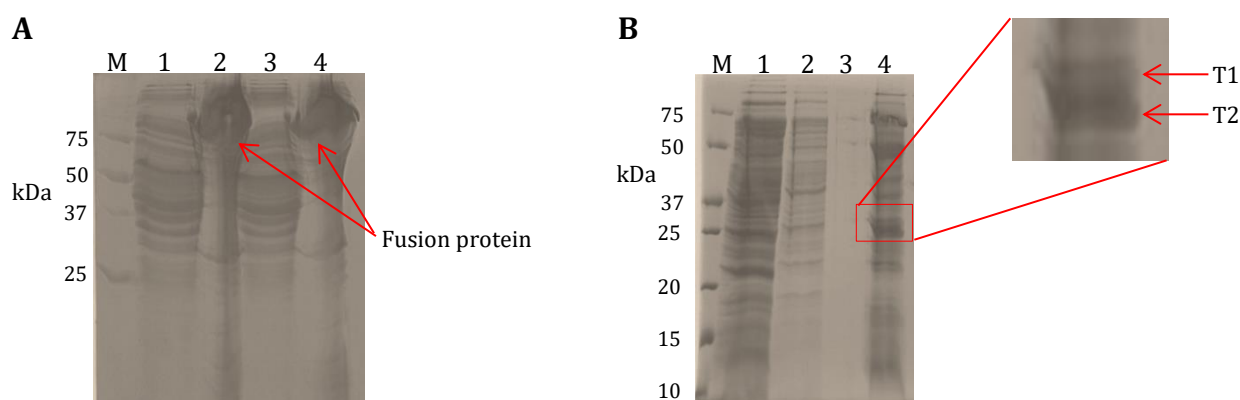
Figure 3.2: SDS-PAGE analysis of MyD88[FL-WT] expression and purification. (A) MyD88[FL-WT] expression trial following successful LR recombination. MyD88[FL-WT] was expressed as His_6-NusA fusion (89 kDa) by autoinduction in *E. coli* BL21 (DE3)-RIL at 37ºC. Lane 1 and 3 (soluble fractions); lane 2 and 4 (Insoluble fractions); the protein was expressed (indicated by the red arrow), albeit in the insoluble fractions. (B) MyD88[FL-WT] large scale expression from 1 L cell culture at 15ºC. Lane 1 (flowthrough); lane 2 and 3 (unbound proteins); lane 4 (beads after cleaving). MyD88[FL-WT] was not detected in the flowthrough but two close bands corresponding to its size (33 kDa) were observed in lane 4 and 6. The bands (T1 and T2) that were sent for peptide mapping are shown in a red box and then enlarged.
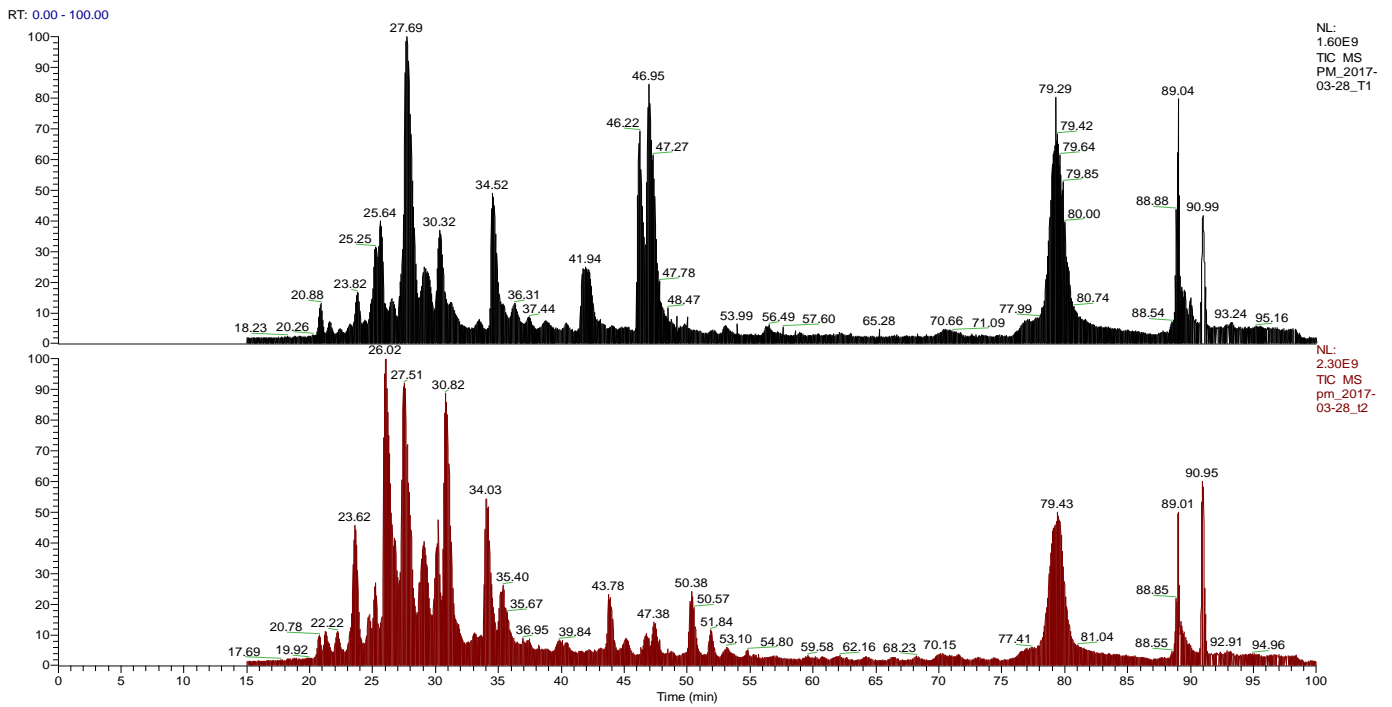
Figure 3.3: Liquid chromatography mass spectra of MyD88^FL-WT. The MyD88^FL-WT protein samples were analysed on a Thermo Scientific Fusion mass spectrometer supplied with a Nanospray Flex ionization source. Data were collected in positive mode with spray voltage and ion transfer capillary set to 2kV and 275°C respectively. An orbitrap detector set at 120 000 resolution was used to perform MS1 scans, while MS2 acquisitions were performed in an orbitrap mass analyzer set to 15 000 resolution and data collected in centroid mode.

Table 3.1: Proteins detected by LC-MS in excised gel bands T1 and T2

	Accession	Description	Score	Coverage	MW [kDa]
T1	H0Y4G9	Myeloid differentiation primary response protein MyD88 (Fragment) OS=Homo sapiens GN=MYD88 PE=2 SV=1 - [H0Y4G9_HUMAN]	300.21	65.51%	35.3
T2	H0Y4G9	Myeloid differentiation primary response protein MyD88 (Fragment) OS=Homo sapiens GN=MYD88 PE=2 SV=1 - [H0Y4G9_HUMAN]	95.01	56.96%	35.3

3.2.2 Expression and purification of WT MyD88 death domain

Our results in the previous section (3.2.1) showed that MyD88^{FL-WT} was degrading. We now decided to work on its DD instead, as it was previously expressed and purified successfully in *E. coli* (Motshwene et al., 2009).

A construct of WT MyD88 DD (MyD88^{DD-WT}) was provided by Miss Jenny Whitby, University of Cambridge for me to use. MyD88^{DD-WT} had been cloned in pMCSG7 expression vector (appendix, figure 9.1C). It expressed as a fusion protein with N-terminal His$_6$ tag and C-terminal Strep tag. We therefore expressed MyD88^{DD-WT} in *E. coli* Rosetta 2 cells and purified by nickel affinity chromatography. A broad peak was observed on the chromatograph shown in figure 3.4A. Samples from this peak were analysed by SDS-PAGE and a protein band that migrated at 20 kDa (figure 3.4B, lanes 2-9) was observed. It was likely to be MyD88^{DD-WT} as it migrated at the expected molecular weight. Other protein bands below 20 kDa were also observed. These could have been either degradation products of MyD88^{DD-WT} or other unknown proteins.

Figure 3.4: Purification of MyD88^{DD-WT} by nickel affinity chromatography. (A) Nickel affinity chromatography of MyD88^{DD-WT} using a 5 mL HiTrap chelating column. The chromatogram shows a broad peak of a sample containing MyD88^{DD-WT}. (B) SDS-PAGE analysis of affinity purified MyD88^{DD-WT}. Fractions collected from the broad peak in figure 3.4A were run on an SDS-PAGE for evaluation of purity. MyD88^{DD-WT} migrated at the expected molecular weight of 22 kDa (lane 1-9) but it was not pure enough due to apparent degradation. Lane M, protein molecular marker.

The affinity purified MyD88^{DD-WT} was then further purified by size exclusion chromatography (SEC) in a Superdex 200 10/300 GL column to remove impurities. Several peaks were observed from the size exclusion chromatograph as can been seen in figure 3.5A. Peak A had a retention volume of 7.5 ml which corresponded to the void volume of the column. Samples from the void volume are usually aggregates or proteins that are too big to be separated and resolved by the column. These were analysed by SDS-PAGE as shown in figure 3.5B, lane 1 and 2. Peak B on the other hand had a retention volume of 15.8 ml and we expected it to contain MyD88^{DD-WT} based on our protein calibrations of the column (data not shown). Samples of this peak were also analysed by SDS-PAGE as shown in figure 3.5B, lane 3-5. The protein was not pure as expected. There were still other protein bands below 20 kDa. We were not deterred by this and decided that we will use the protein anyway for further experiments. Peak C and D were run on a separate SDS gel and there was nothing significant to see. As a result, the data was not shown.
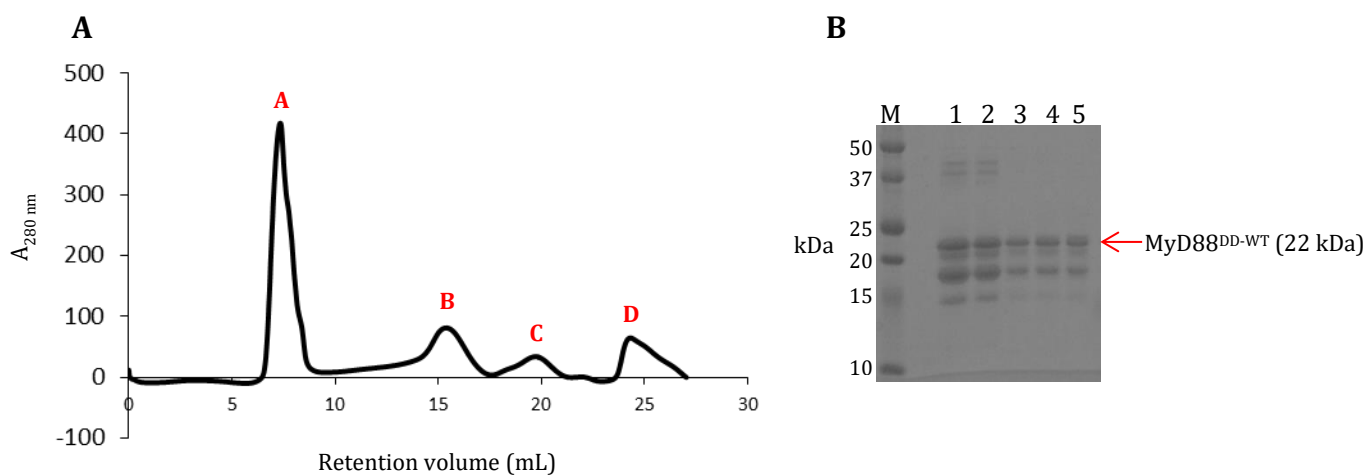
Figure 3.5: Purification MyD88^{DD-WT} by size exclusion chromatography. (A) Size exclusion chromatography of MyD88^{DD-WT} using Superdex 200 10/300 GL column. Four peaks with different retention volumes were observed on the chromatogram. (B) SDS-PAGE analysis of size exclusion chromatography purified MyD88^{DD-WT}. Peak A and B with respective retention volumes of 7.5 and 15.8 mL were analysed on SDS-PAGE. Peak A corresponded to the void volume and contained MyD88^{DD-WT} aggregates (lane 1 and 2) whereas peak B contained monomeric MyD88^{DD-WT} (lane 3-5). Lane M, protein molecular marker.

3.2.3 Expression and purification of WT IRAK-4 death domain

Having successfully expressed and purified MyD88^{DD-WT}, we now wanted to prepare WT IRAK-4 DD (IRAK-4^{DD-WT}). We were going to mix it with MyD88^{DD-WT} to reconstitute the Myddosome complex. A construct of IRAK-4^{DD-WT} was also provided to us by Miss Jenny Whitby, University of Cambridge. IRAK-4^{DD-WT} had been cloned into pMCSG7 expression vector (appendix, figure 9.1C). This vector expresses as a protein of interest with N-terminal His$_6$ tag. IRAK-4^{DD-WT} was also expressed in *E. coli* Rosetta 2 cells and purified by nickel affinity chromatography. Three peaks eluted on the nickel column as shown on the chromatograph in figure 3.6A. SDS-PAGE analysis of the samples from these peaks showed that all of the peaks contained a protein corresponding to the expected size of IRAK-4^{DD-WT} (12.5 kDa) as shown in figure 3.6B, lane 1-14. An unexpected protein with a molecular weight above 25 kDa was also observed.
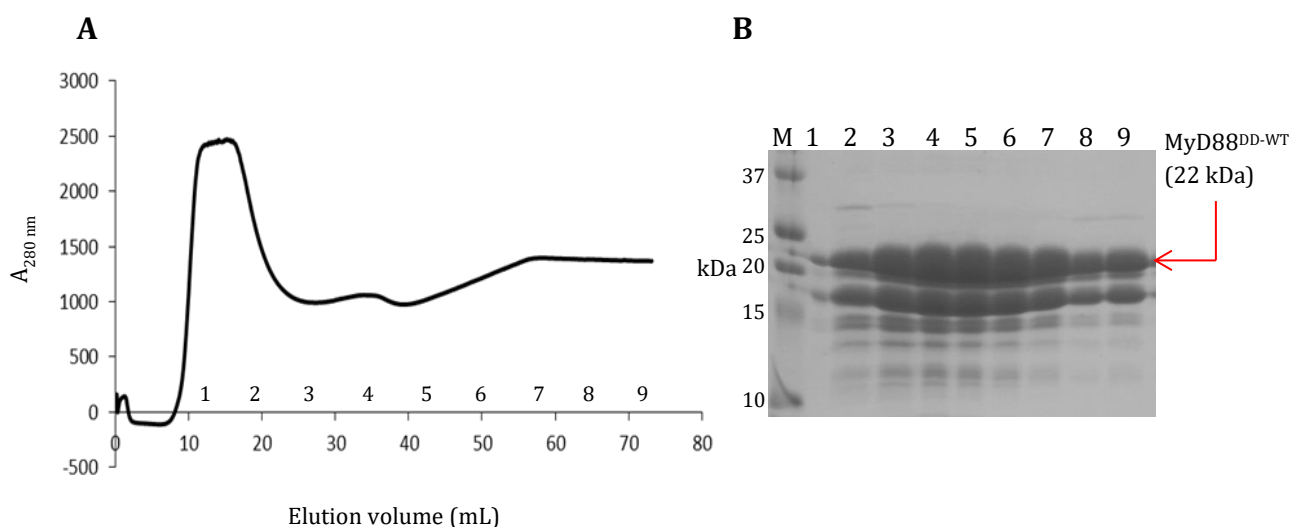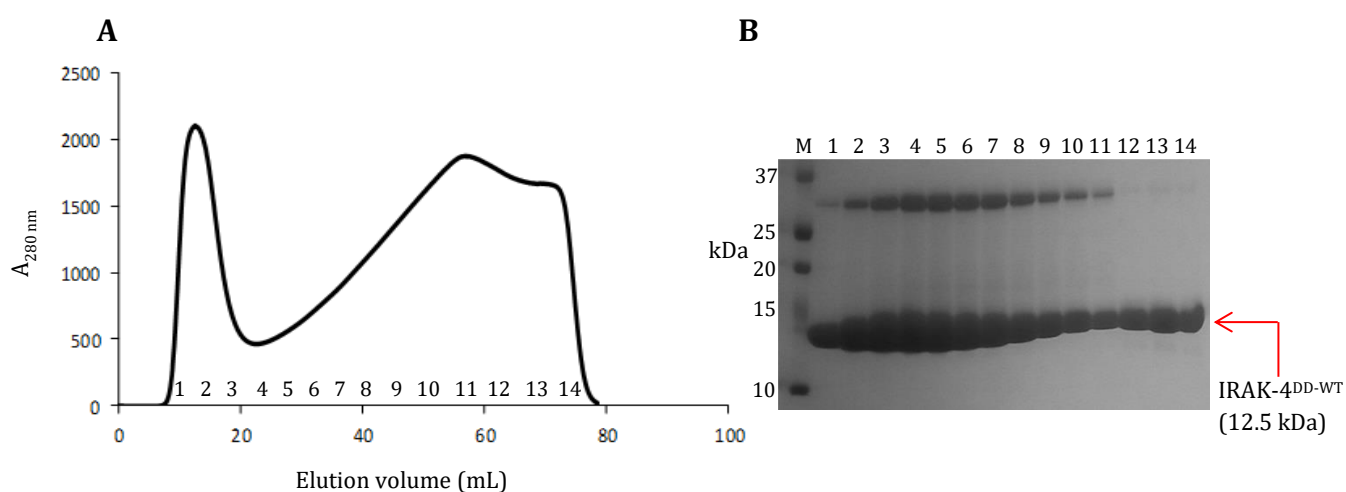

Figure 3.6: Purification of IRAK-4^{DD-WT} by nickel affinity chromatography. (A) Nickel affinity chromatography of IRAK-4^{DD-WT} using 5 mL HiTrap chelating column. Two unresolved peaks were observed on the chromatogram. (B) SDS-PAGE analysis of affinity purified IRAK-4^{DD-WT}. Fractions 1-14 collected from the peaks in figure 3.6A were run on SDS-PAGE. IRAK-4^{DD-WT} migrated at the expected molecular weight of 12.5 kDa (lane 1-14). A protein contaminant was also observed at molecular weight between 25 and 37 kDa (lane 1-11). Lane M, protein molecular marker.

IRAK-4^{DD-WT} that eluted from the affinity column was further purified by SEC in a Superdex 200 10/300 GL column to try to remove the protein with a molecular weight above 25 kDa. Two peaks (A and B) were observed from the size exclusion chromatograph shown in figure 3.7A. They had retention volumes of 17.5 and 24.3 mL, respectively. Samples from peak A were analysed by SDS-PAGE as shown in figure 3.7B, lane 1-4. A protein band with an expected molecular weight of IRAK-4^{DD-WT} was observed at 12.5 kDa. This band was pure and the protein contaminant above 25 kDa was no longer present. Samples from peak B were also analysed by SDS-PAGE. The data were not shown since there was nothing significant to see.

A

B
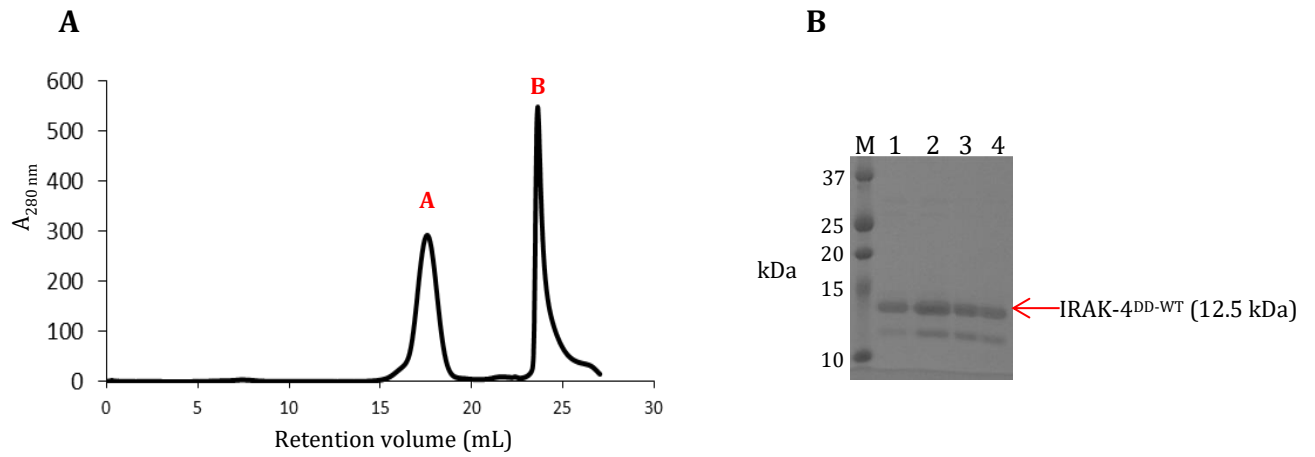
Figure 3.7: Purification of IRAK-4$^{\text{DD-WT}}$ by size exclusion chromatography. (A) Size-exclusion chromatography of IRAK-4$^{\text{DD-WT}}$ using Superdex 200 10/300 GL column. The chromatogram shows two peaks with different retention volumes. SDS-PAGE analysis of size exclusion chromatography purified IRAK-4$^{\text{DD-WT}}$. Peak A had a retention volume of 17.5 mL and contained IRAK-4$^{\text{DD-WT}}$ which migrated at the expected molecular weight of 12.5 kDa (lane 1-4). A band below 12.5 kDa resulting from possible degradation of IRAK-4$^{\text{DD-WT}}$ below 12.5 kDa was observed (lane 1-4). Lane M, protein molecular marker.

3.3 DISCUSSION

In this chapter we discuss the expression and purification of full-length MyD88, its DD and IRAK-4 DD. These proteins were expressed with the aim of reconstituting the Myddosome complex. We needed to obtain enough yields of protein to reconstitute the Myddosome complex.

3.3.1 Expression of WT MyD88 full length

To date, MyD88 full length has never been expressed in any host, whether prokaryotic or eukaryotic. We tried expressing it in *E. coli* and we experienced some challenges like cleaving it from its fusion partner, precipitation and degradation. One of the possible reasons why full length MyD88 failed to express in *E. coli* may be due to a high number of rare codons on its DD. There are 20 amino acids which are found in proteins, all specified by a sequence of three bases known as a codon. Out of 64 possible three-base codons, 61 code for amino acids including N-terminal formyl-methionine while the remaining three serve as stop codons. For most amino acids there is more than one codon and depending on the organism, there is a bias in codon usage. Some codons for the same amino acid are used more often than others known as rare codons (Clark and Clark, 2008). The bias in codon usage correlates with the availability of corresponding tRNAs. It is also a result of non-random choices between pyrimidine-ending codons. Thus heterologous gene expression can be hindered in cases were previously major codons are rare in the expression host. Possible outcomes include inhibition of protein synthesis, suboptimal cell growth, reduced mRNA stability, frame shift mutations and deletions (Berg and Silva, 1997).

In order to bypass this problem, we used BL21-CodonPlus cells. Codon plus cells supply additional copies of specific tRNA genes that are rare in *E. coli* in order to improve protein expression. We also used autoinduction media to improve protein yields. Autoinduction produces higher protein yields when compared to IPTG induction (Studier, 2005).

Full length MyD88 (MyD88[FL-WT]) was expressed as a His-NusA fusion protein. The N-terminal double His-NusA tag was fused to MyD88[FL-WT] through a cleavable linker, in the form of a 3C protease site. The *E. coli* NusA protein was used as a fusion partner to MyD88[FL-WT] because it yields fusion proteins that are highly soluble and stable and when compared to other tags (Ario de Marco, 2006). However, large protein tags like NusA (54 kDa) utilise more metabolic energy during overproduction (Waugh, 2005). This may affect the properties of the fusion protein. MyD88[FL-WT] is smaller than NusA in terms of size and during purification we found that NusA was significantly more than MyD88[FL-WT] as seen on SDS-PAGE. It would have been better to use GST as an alternative fusion partner because of its smaller size. GST has a molecular weight of 27

kDa whereas NusA is about 54 kDa. We tried using GST but the fusion failed to express. Since NusA does not have affinity property, we used it in conjunction with a His tag that binds to immobilized transition metals. While NusA stabilizes the MyD88^{FL-WT} during translation, the His tag enables affinity purification of the fusion via immobilized metal ion affinity chromatography.

Although we obtained soluble MyD88^{FL-WT}, the yield was low and not enough for the reconstitution of the Myddosome complex. In addition, there were two protein bands corresponding to the size of MyD88^{FL-WT} as seen on SDS-PAGE following affinity purification. We did MS to identify these bands and found that they were both MyD88^{FL-WT}. Having identified these bands we wanted to use the MS data to determine the boundaries for cloning a better construct that would be stable and free of degradation. However, MS could not pick up all the trypsin digested peptides as the highest sequence coverage was 65%.

3.3.2 Expression of WT MyD88 death domain

Since we could not determine boundaries for creating a stable MyD88^{FL-WT} construct, we abandoned the plan and rather opted to express its DD. MyD88 DD has been previously expressed with success (Lin et al., 2010; Motshwene et al., 2009). Unlike the MyD88^{FL-WT}, WT MyD88 DD (MyD88^{DD-WT}) had been cloned in pMCSG7 and expressed as a His$_6$-fusion; which could theoretically be purified in one affinity step. MyD88^{DD-WT} was expressed by IPTG induction in Rosetta 2 cells without any major difficulties. The affinity purification step gave a broad peak on the chromatograph, which was an indication that the sample was not pure enough. MyD88^{DD-WT} migrated at the expected size of 22 kDa when analysed SDS-gel. However, there were two additional proteins that purified with it. These bands were present after both affinity purification and size exclusion chromatography. It is highly likely that these bands resulted from the degradation of MyD88^{DD-WT}.

Since the MyD88^{DD-WT} sample was not pure enough, we performed size exclusion chromatography to further purify the sample. This technique was going to give us information regarding the oligomeric status, size and aggregation status of the protein. There was a sample that eluted in the void volume and it contained aggregates of MyD88^{DD-WT}. However, the peak with a retention volume of 15.8 mL also contained MyD88^{DD-WT} as determined on SDS-PAGE. This retention volume showed that it was highly likely to be a monomer and that it was unlikely to contain aggregates.

3.3.3 Expression of WT IRAK-4 death domain

IRAK-4^{DD-WT} was expressed by IPTG induction in Rosetta 2 cells without any problems. During affinity purification, IRAK-4^{DD-WT} co-eluted with a prominent protein of unknown identity which had a molecular weight above 25 kDa. This band was removed following purification by size exclusion chromatography. Similar to MyD88^{DD-WT}, IRAK-4^{DD-WT} was found to be monomeric based on its retention volume. Taken together, our observations are consistent with what has been reported in literature on the oligomeric status of both MyD88 and IRAK-4 DDs (Lin et al., 2010; Motshwene et al., 2009).

3.4 CONCLUSION

We managed to express full length MyD88 even though it was not stable in solution. Its lack of stability shifted our focus onto expressing its DD. The MyD88 DD expressed and purified well for further experiments. In addition we also expressed and purified IRAK-4 DD, which put us in a good position to proceed to the next stage of the project which was to reconstitute the Myddosome.

CHAPTER 4: Characterization of WT Myddosome complex

4.1 INTRODUCTION

DD interactions between MyD88 and IRAKs are essential for TLR signalling (Park, 2011). MyD88 adaptor protein links TLRs to IRAKs and facilitates the formation of a large signalling complex known as the Myddosome (Motshwene et al., 2009). The stoichiometry and size of the Myddosome complex has been determined by two independent research groups (Lin et al., 2010; Motshwene et al., 2009). Here we reconstitute the Myddosome complex using WT DDs of MyD88 and IRAK-4 and further characterize it using three biophysical techniques namely size exclusion chromatography, dynamic light scattering and analytical ultracentrifugation. Our aim is to see whether we can form the Myddosome complex similar to that reported in literature. After we have formed this complex, we will later do other experiments where we test if phosphomimetic mutants of IRAK-4 DDs also assemble into a Myddosome.

4.2 RESULTS

4.2.1 Reconstitution of wild type (WT) Myddosome complex

In the previous section, we purified MyD88^{DD-WT} and IRAK-4^{DD-WT} with the aim of using them to reconstitute the Myddosome, an oligomeric complex formed between the DDs of MyD88 and IRAK-4. Purified DDs of MyD88 and IRAK-4 were concentrated and thereafter mixed to reconstitute the Myddosome. The mixture was then loaded on a Superdex 200 10/300 GL column to determine the formation of the complex. Five peaks (A-E) were observed on the size exclusion chromatograph shown in figure 4.1A. Peak A had a retention volume of 7.5 ml which corresponded to the void volume of the column as previously mentioned. Peak B had a retention volume 11.5 mL and was more likely to contain constituents of the Myddosome complex based on previous work reported in literature. An analysis of samples from peak A on SDS-PAGE showed that there was no IRAK-4^{DD-WT} (figure 4.1B, lane 1). Samples from peak B on the other hand showed the presence of the Myddosome complex (figure 4.1B, lane 2-5). Protein bands corresponding to MyD88^{DD-WT} and IRAK-4^{DD-WT} were seen at the expected molecular weights of 22 and 12.5 kDa respectively. There was also another band below 20 kDa which was observed. This band was present in our earlier purification of MyD88^{DD-WT}. Samples from peak C and D were also analysed by SDS-PAGE as shown in figure 4.1B, lane 6-11. IRAK-4^{DD-WT} was mainly found in those two peaks. Peak E on the other hand contained no protein as determined by Bradford assay and was not loaded on SDS-PAGE.
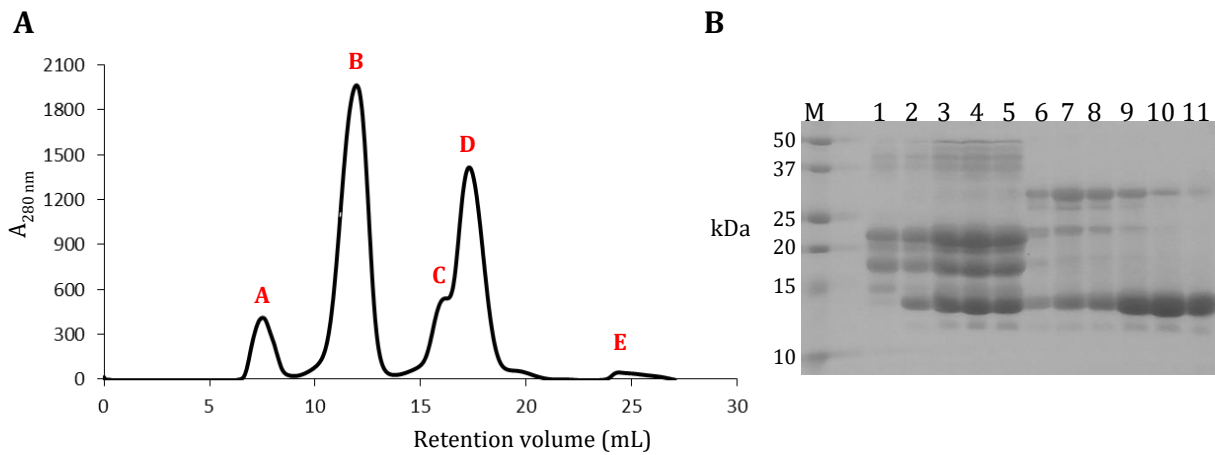
Figure 4.1: Size exclusion chromatography and SDS-PAGE analysis of WT Myddosome complex. (A) Size exclusion chromatography purified WT Myddosome complex using Superdex 200 10/300 GL column. Five peaks (A-E) with respective retention volumes of 7.5, 11.5, 15.8 and 17.5 and 24 mL were observed on the chromatogram. (B) SDS-PAGE of size exclusion chromatography purified WT Myddosome complex. All sample samples were analysed except sample from E. Peak A (lane 1) corresponded to the void volume and contained aggregates of MyD88[DD-WT]; peak B (2-5) contained the WT Myddosome complex; peak C and D (lane 6-11) contained mainly IRAK-4[DD-WT]. Lane M, protein molecular marker.

4.2.2 Characterization of WT Myddosome complex by Dynamic Light Scattering

Now that we had formed the Myddosome complex, we wanted to determine if it was a *bona fide* complex or an aggregate. We used dynamic light scattering (DLS), a technique that measures the size of particles by monitoring the speed at which these particles (protein particles in this case) are diffusing in Brownian motion. Small particles move faster and larger ones slower in Brownian motion. Diffusion rates are used to calculate the size distribution of protein molecules by light intensity or mass using the Stokes Einstein equation, $d(H) = kT/3\pi\eta D$. Particle size is calculated and given in terms of hydrodynamic diameter $d(H)$, where k is the Boltzmann's constant, T is the absolute temperature, η is the viscosity, and D is the diffusion coefficient.

Three independent DLS measurements were performed on the size exclusion chromatography purified Myddosome complex. DLS data on the complex was collected as shown in figure 4.2. A single peak with a $d(H)$ of 6.772 (peak 1, distribution results) was observed for the complex. In addition, the sample was found to have a polydispersity index (PDI) of 0.189. These results showed that the Myddosome complex was monodisperse and free of aggregates.
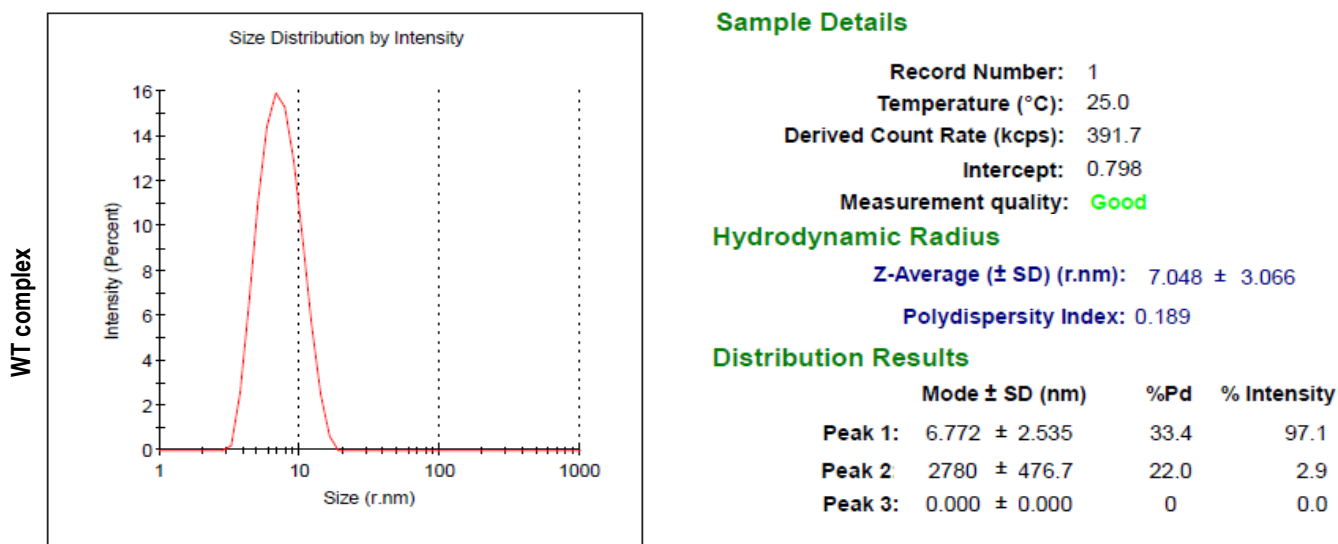
Figure 4.2: Dynamic light scattering (DLS) profile for WT Myddosome complex. The DLS hydrodynamic diameter distribution plot showing a peak of the WT Myddosome complex with a d(H) of 6.772 and a PDI of 0.189.

4.2.3 Characterization of WT Myddosome complex by AUC

We had ascertained by DLS that the WT Myddosome complex we reconstituted was free of aggregates. The next experiment that we did was analytical ultracentrifugation (AUC) to determine the size of the complex. AUC is a technique that uses very high centrifugal forces to effect the separation of macromolecules based on their size and shape (Cole et al., 2009; Howlett et al., 2006). The centrifugation process is monitored by an optical system. AUC also allows for the determination of binding affinities and stoichiometries depending on the nature of the experiment that is carried out. There are two experiments that can be carried out with AUC, namely sedimentation velocity (SV) and sedimentation equilibrium (SE). The former uses relatively high centrifugal forces to measure the rate at which molecules sediments in a sample cell, in response to centrifugal force, yielding information on the sedimentation coefficient (related to particle mass), molecular mass and shape of molecules (Cole et al., 2009). This process normally takes a few hours to complete. SE on the other hand uses lower centrifugal forces until a balance is attained between diffusion and sedimentation, and takes few days to complete.

We used SV to determine the size of the WT Myddosome complex. A series of sequential scans of the protein sample were performed at minute intervals as shown in figure 4.3A. The scan shows how the protein complexes move through solution in the sample chamber in real-time. SEDNTERP software was used to calculate the sedimentation coefficient from the scan results. A sedimentation velocity profile as shown in figure 4.3B was generated, from which sedimentation

coefficients could be determined. From the SV profile, a prominent peak with a sedimentation coefficient of 7.5 S was observed. This sedimentation coefficient was similar to the one previously reported on the Myddosome complex (Motshwene et al., 2009). We were therefore confident that we had assembled a Myddosome complex of a similar size.

A

B
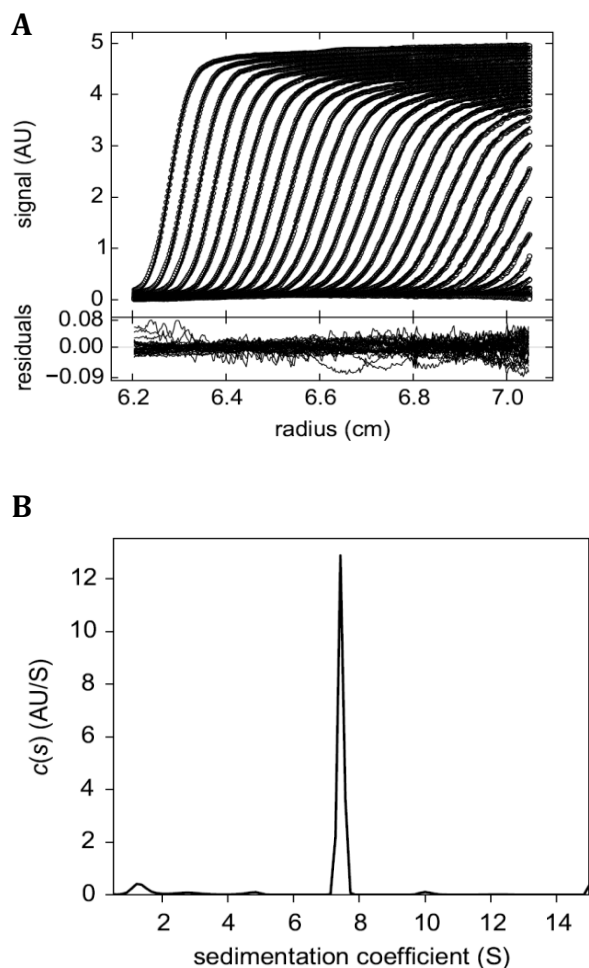
Figure 4.3: Sedimentation velocity (SV) profile of WT Myddosome complex. (A) SV graph showing scans of the WT Myddosome complex recorded at a wavelength of 280 nm. The SV scans show the behaviour of the complex in the sample chamber in response to centrifugal force over time. (B) Sedimentation coefficient distribution plot of the WT Myddosome complex derived from the SV scans. The outstanding peak centred at a sedimentation coefficient of 7.5 S corresponds to the WT Myddosome complex.

4.3 DISCUSSION

We successfully reconstituted the Myddosome complex and used three biophysical techniques to characterise it. These techniques are size exclusion chromatography, dynamic light scattering and analytical ultra-centrifugation. Our findings are discussed in relation to these.

The Myddosome complex was reconstituted by mixing WT MyD88 and IRAK-4 DDs. Thereafter, size exclusion chromatography was used to determine if the complex had formed. Our results showed that both proteins were present in the peak that had a retention volume of 11.5 mL. This volume did not differ from the one initially reported on the formation of the Myddosome (Motshwene et al., 2009), even though our constructs were different. For example, our MyD88 DD had a molecular weight of 22 kDa and the one used in the initial published study on the Myddosome was 16.5 kDa. However, the IRAK-4 DDs used were similar in molecular weight as they were expressed from the same construct.

We then analysed the Myddosome complex that we formed by dynamic light scattering as we were partly concerned about the slight degradation of the MyD88^{DD-WT}. By using DLS, we would be in a position to determine if the complex was aggregated or free of aggregates. DLS data showed the presence of one sharp peak. This was an indication that the complex existed as one population, meaning it was monodisperse. Also, a low polydispersity index was obtained. This clearly meant that the reconstituted Myddosome complex was stable in the buffer conditions that were used. We could have used the sample for crystallisation but it was not necessary since the structure of the Myddosome complex had already been published (Lin et al., 2010).

The size of the Myddosome had been previously determined by AUC. We also used this technique to compare if our Myddosome complex had a similar size to the reported one. Our data showed that it had a sedimentation coefficient of 7.5 S whereas that of the reported one was 6.3 S. This is not surprising because our MyD88 DD protein had a molecular weight of 22 kDa. It was 6 kDa bigger and based on the published stoichiometry (Motshwene et al., 2009), we were expecting our Myddosome complex to be 45 kDa bigger on average.

From the AUC data, we could not determine the stoichiometry of the Myddosome complex because we performed a SV and not a SE experiment. Had we done a sedimentation equilibrium experiment, we might not have been able to determine the stoichiometry of the Myddosome. A sedimentation equilibrium experiment is used to determine simple stoichiometries (Cole et al., 2009). Its stoichiometry is unusual and two independent research groups obtained different stoichiometries. The first group that published a paper on the Myddosome, found it to have a

stoichiometry of 7 MyD88 to 4 IRAK-4 DD molecules (Motshwene et al., 2009). On the contrary, the second group found a stoichiometry of 6 MyD88 to 4 IRAK-4 DD molecules but they also had IRAK-2 DD as part of the complex (Lin et al., 2010). It is not surprising that there are variations in stoichiometry between the two groups because they used two different techniques to determine it. The one group used non-dissociative mass spectrometry and the other determined it from the solved crystal structure.

4.4 CONCLUSSION

From the data presented, we cannot conclusively say that our Myddosome has exactly the same stoichiometry to that reported by Motshwene et al., 2009. We can only make inferences that it is highly likely to be the same based on the size obtained by AUC. Non-dissociative mass spectrometry would have given us a definitive answer.

CHAPTER 5: Expression and purification of IRAK-4 death domain mutants

5.1 INTRODUCTION

Many autophosphorylation sites have been mapped in full length IRAK-4. Among them is Serine 8 and Threonine 62 which are found in the DD. These two are of particular interest to us as they might play a role in regulating the assembly of the Myddosome. There has been a study that showed that serine 8 phosphorylated IRAK-4 is unable to form the Myddosome assembly (Dossang et al, 2016). We were therefore interested in studying the role of autophosphorylation in the assembly of the Myddosome. It was not going to be possible for us to conduct such studies since we were only working on individual DDs. We therefore came up with an alternative of using phosphomimetics. Our plan was to mimic autophosphorylation by introducing negative charges to the amino acids that get autophosphrylated.

5.2 RESULTS

5.2.1 Expression of IRAK-4 death domain mutants

We previously used IRAK-4^{DD-WT} to assemble the Myddosome complex. In order for us to study the effects phosphomimetics on the Myddosome assembly, we needed to mutate serine 8 and threonine 62 to an amino acid that carries a negative charge like aspartic acid or glutamic acid. Mutants of IRAK-4^{DD-S8D} and IRAK-4$^{DD-T62D}$ already existed in our lab. These were generated by site-directed mutagenesis by my supervisor, Dr Motshwene. Unlike the WT previously used, these were in a pETG30A expression vector (appendix, figure 9.1B). This vector expresses a protein of interest as a fusion protein with N-terminal His$_6$-GST tag.

The phosphomimetic constructs that were provided, were sequenced before use. This was done to verify the presence of the S8D and T62D mutations. DNA sequence results confirming these mutations are shown in appendix, figure 9.2.

5.2.2 Expression and purification of IRAK-4^{DD-S8D}

IRAK-4^{DD-S8D} was expressed in *E. coli* Rosetta 2 cells by autoinduction and processed as previously described. The proteins were purified by GST affinity chromatography. Proteins that eluted from the affinity column were mixed and incubated with 3C protease to cleave IRAK-4^{DD-S8D} from its binding partner, GST. A sample of IRAK-4^{DD-S8D} that eluted from the GST column was analysed by SDS-PAGE as shown in figure 5.1A, lane 1. This sample contained several protein

bands. A protein band with an approximate molecular weight of 37 kDa was observed. This band was more likely to be the IRAK-4^{DD-S8D} fusion protein. There was also another band with a molecular weight of approximately 27 kDa. This band ran at the expected molecular weight of GST. The cleaved IRAK-4^{DD-S8D} sample was also analysed by SDS-PAGE as shown in figure 5.1A, lane 2-6. As expected, there was very little or no fusion protein observed as GST could be clearly seen at 27 kDa and our protein of interest, IRAK-4^{DD-S8D} at 12.5 kDa. The presence of the 3C protease was clearly visible as a band that migrated at 39 kDa.

In order to get a much purer IRAK-4^{DD-S8D}, we pulled the cleaved material together and passed it onto a 5 mL HiTrap chelating column that had been charged with nickel. This was done to eliminate the GST contaminant. GST from our construct has a His tag on its N-terminus and it would be captured on the nickel column. The pooled cleaved material is shown in figure 5.1B, lane 1 and the flow through from the sample that was passed onto the nickel column in lane 2. It seemed like there was a reduction in the amount of His$_6$-GST. However, the concentrated sample of the flow through showed that His$_6$-GST was still present in significant quantities (figure 5.1B, lane 3). This therefore meant that the His$_6$-GST was not binding to the column.

Figure 5.1: SDS-PAGE analysis of IRAK4^{DD-S8D} purified by affinity chromatography (A) SDS-PAGE of GST affinity purified IRAK4^{DD-S8D}. Lane 1 contains sample of IRAK4^{DD-S8D} sample that eluted from the GST column. A band corresponding to the size of the fusion protein (37 kDa, red box) was observed. Lane 2-6 contains IRAK4^{DD-S8D} sample that was cleaved with 3C protease. Band corresponding to the size of 3C protease (39 kDa), His$_6$-GST double tag (27 kDa) and IRAK4^{DD-S8D} (12.5 kDa) were seen. (B) SDS-PAGE analysis nickel affinity purified IRAK4^{DD-S8D} sample. Lane 1 contains the cleaved material that was pooled together. Lane 2 contains the flowthrough of the samples that passed onto a 5 mL HiTrap chelating column. This sample shows a decrease in the amount of GST. Lane 3 shows the concentrated sample of the flowthrough. Lane M, protein molecular marker.

Concentrated samples of IRAK-4^{DD-S8D} were further purified by size exclusion chromatography in a Superdex 200 10/300 GL column. Six peaks (A-F) were observed on the size exclusion chromatograph shown in figure 5.2A. From our previous results (chapter 3 & 4) we expected IRAK-4^{DD-S8D} to have a retention volume somewhere between 15.0 and 19.0 mL. We therefore analysed samples within this range, which were that of peak B and D. SDS-PAGE analysis showed that peak B contained a protein band of 27 kDa (figure 5.2B, lane 1), which could be GST. Peak D on the other hand contained a pure protein band at 12.5 kDa (figure 5.2B, lane 2 & 3) corresponding to the size IRAK-4^{DD-S8D}.

Figure 5.2: Size exclusion chromatography and SDS-PAGE analysis of IRAK4^{DD-S8D}. (A) Size exclusion chromatograph of IRAK4^{DD-S8D} obtained by using a Superdex 200 10/300 column. Six peaks (A-F) were observed on the chromatograph that had retention volumes of 7.5, 15.0, 17.5, 19.0, 22.0 and 24.0 mL respectively. (B) SDS-PAGE of size exclusion chromatography purified IRAK4^{DD-S8D}. Only the samples from peak B and D were analysed. Peak A (lane 1) contained a protein corresponding to the size of His$_6$-GST double tag at 27 kDa. Peak D contained a band migrating at the expected molecular weight of IRAK4^{DD-S8D} at 12.5 kDa (red box). Lane M, protein molecular marker.

5.2.3 Expression and purification of IRAK-4$^{DD-T62D}$

IRAK-4$^{DD-T62D}$ was expressed and purified exactly like IRAK-4^{DD-S8D}. An SDS-PAGE analysis of the GST affinity purified IRAK-4$^{DD-T62D}$ showed the presence of the fusion protein at 37 kDa (figure 5.3A, lane 1). The cleaved fusion protein showed the presence of three bands that migrated at 39, 27 and 12.5 kDa as shown in figure 5.3A, lane 2-6. These were bands of the 3C protease, His$_6$-GST and IRAK-4$^{DD-T62D}$. Just like we did with IRAK-4^{DD-S8D}, we pooled the cleaved fusion proteins samples from figure 5.3A, lane 2-6. These were loaded on SDS-PAGE (figure 5.3B, lane 1) to compare them with a sample that was passed onto a nickel column. The sample that was passed onto a nickel column contained significantly less amounts of His$_6$-GST as shown in figure 5.3B, lane 2. This sample was concentrated in preparation for size exclusion chromatography but it was first analysed by SDS-PAGE to check for the presence of the His$_6$-GST contaminant as shown

in figure 5.3B, lane 3. The presence of His$_6$-GST was less and we were confident that we would eliminate the traces of the 3C protease from IRAK-4^{DD-S8D} by size exclusion chromatography.
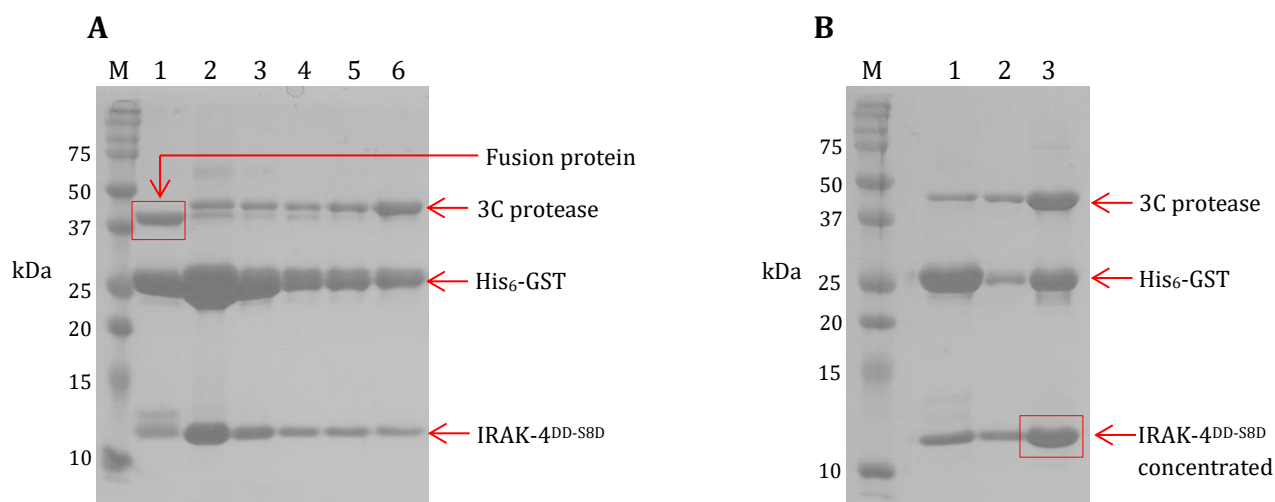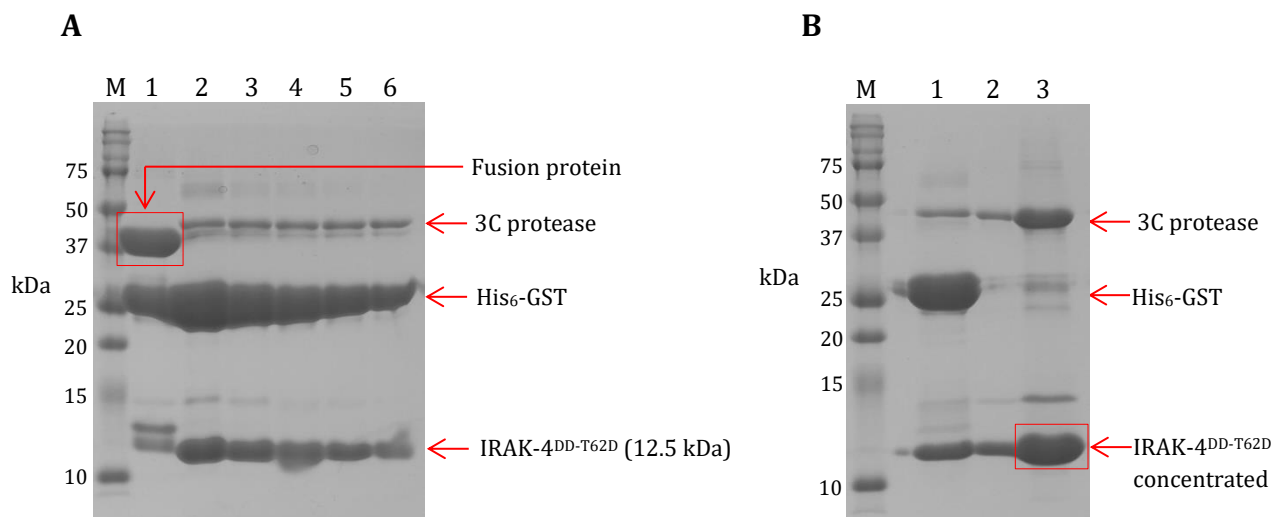

Figure 5.3: SDS-PAGE analysis of IRAK-4$^{DD-T62D}$ purified by affinity chromatography (A) SDS-PAGE of GST affinity purified IRAK4^{DD-S8D}. Lane 1 contains sample of IRAK-4$^{DD-T62D}$ sample that eluted from the GST column. A band corresponding to the size of the fusion protein (37 kDa, red box) was observed. Lane 2-6 contains IRAK-4$^{DD-T62D}$ sample that was cleaved with 3C protease. Band corresponding to the size of 3C protease (39 kDa), His$_6$-GST double tag (27 kDa) and IRAK-4$^{DD-T62D}$ (12.5 kDa) were seen. (B) SDS-PAGE analysis nickel affinity purified IRAK4^{DD-S8D} sample. Lane 1 contains the cleaved material that was pooled together. Lane 2 contains the flowthrough of the samples that were passed onto a 5 mL HiTrap chelating column. This sample shows a significant reduction in the amount of His$_6$-GST. Lane 3 shows the concentrated sample of the flowthrough. Lane M, protein molecular marker.

Size exclusion chromatography was performed on the concentrated nickel affinity purified sample (figure 5.2B, lane 2) mentioned earlier. Five peaks (A-E) were observed on the size exclusion chromatograph shown in figure 5.4 A. We only analysed samples from peak A and C on SDS-PAGE for reasons previously mentioned. Samples from peak A (15.0 mL) only contained GST (figure 5.4B, lane 1-2), whereas samples from peak C (19 mL) contained pure IRAK-4$^{DD-T62D}$ (figure 5.4B, lane 5-6).
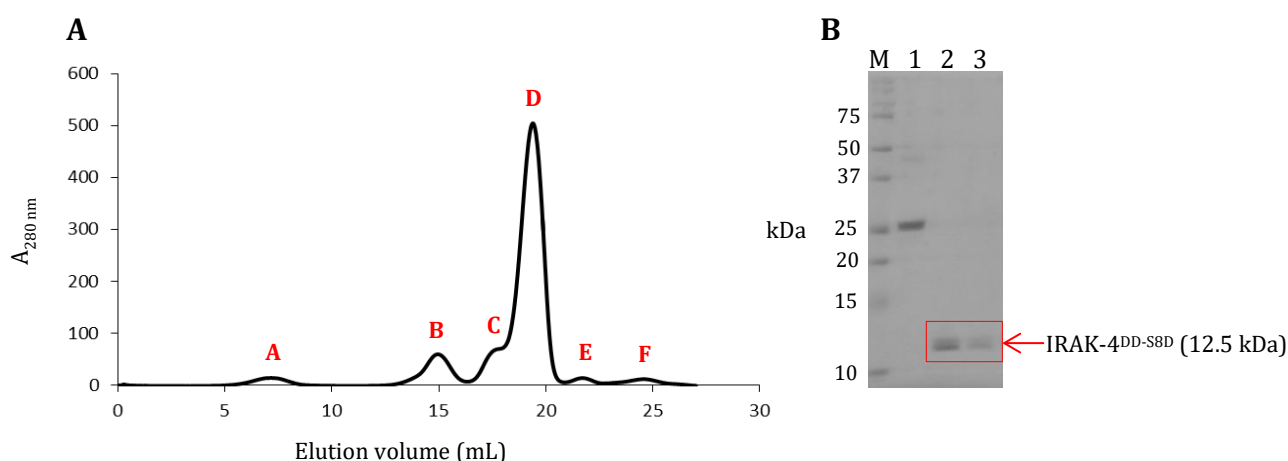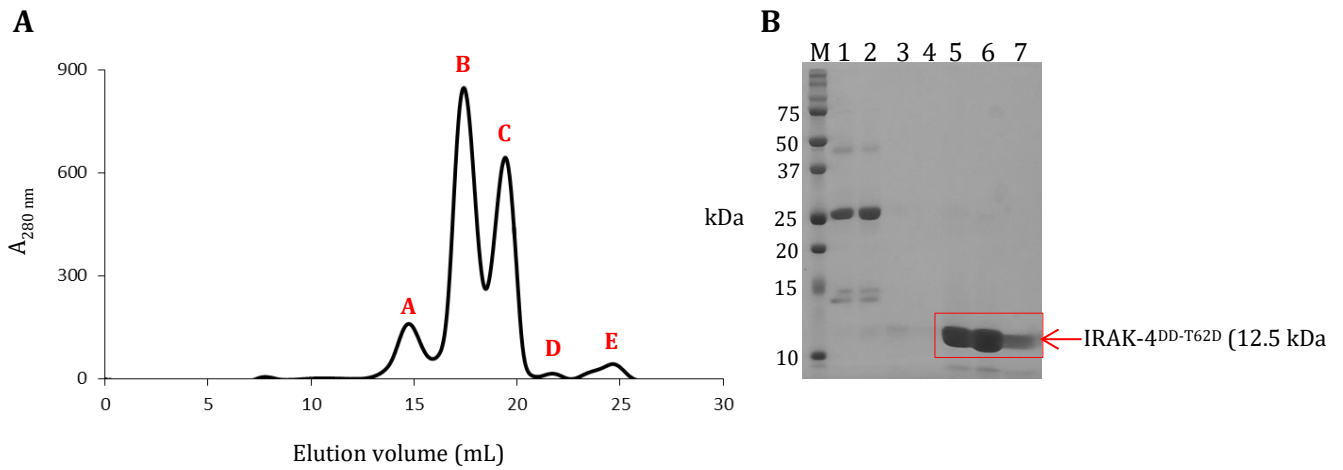
Figure 5.4: Size exclusion chromatography and SDS-PAGE analysis of IRAK-4$^{DD-T62D}$. (A) Size exclusion chromatograph of IRAK-4$^{DD-T62D}$ obtained by using a Superdex 200 10/300 column. The chromatograph showed five peaks (A-E) which had retention volumes of 15.0, 17.5, 19.0, 22.0 and 24.0 mL respectively. (B) SDS-PAGE of size exclusion chromatography purified IRAK-4$^{DD-T62D}$. Only the samples from peak A and C were analysed. Peak A contained a protein corresponding to the size of His$_6$-GST double tag at 27 kDa (lane 1 and 2). Peak C contained a band corresponding to the size of IRAK-4$^{DD-T62D}$ at 12.5 kDa (lane 5-7, red box). Lane M, protein molecular marker.

5.3 DISCUSSION

Phosphorylation is a post-translational modification that is found in certain proteins. It normally occurs when a kinase phosphorylates another protein that it interacts with. Some kinases do not only phosphorylate their substrate but can also phosphorylate themselves, a process known as autophosphorylation. Autophosphorylation alters the electrostatics and may even affect or lead to a conformational change on the structure of a protein (Beenstock et al., 2016; Chen and Cole, 2015).

We replaced Serine 8 and Threonine 62 with aspartate in order to mimic autophosphorylation. These two amino acids are located in the flexible regions of the IRAK-4 death domain, namely in the loops. It was therefore our expectation that the mutations on these amino acids were not going to drastically affect the structure of the IRAK-4 death domain. Our aim in creating these phosphomimetics was to mix them with MyD88 death domain and to determine if they would form the Myddosome complex.

The S8D phosphomimetic was expressed first in *E. coli* cells. It managed to express but did not express as well as wild type IRAK-4 death domain. Mutants sometimes do not express as well wild type proteins (Dailey and Dailey, 1997). For example, an S34Y mutation in the MyD88 death domain led to poor expression and aggregated protein (George et al, 2011). Although the MyD88 S8D expressed poorly, there were not much aggregates as determined by size exclusion chromatography. On the contrary, the T62D mutant expressed very well, even better than the wild type.

There is another approach that we could have taken to obtain phosphorylated S8 and T62 instead of using phosphomimetics. Isolated full length IRAK-4 can be autophosphorylated *in vitro* and thereafter subjected to limited proteolysis to separate the death domain from the kinase domain (Motshwene, 2008; Dossang et al, 2016). Thrombin cleaves full length IRAK-4 at the linker that joins the death domain to the kinase domain. The reported finding was verified by N-terminal sequencing (Motshwene, 2008). If we had full length IRAK-4, we could have also tried autophosphorylating it *in vitro* and to cleave it with thrombin. This could have resulted in a phosphorylated death domain as reported by Dossang et al., 2016. However, it would have required us to do phosphopeptide mapping to determine the exact phosphorylated amino acid. We would know if S8, T62 or both were autophosphorylated. A possibility of both being autophosphorylated would exist under these circumstances. This would have depended on the duration of IRAK-4 incubation with ATP. In order to avoid the possibility of a double phosphorylated IRAK-4 DD we could have created two mutants where either residue was mutated to a non-phosphomimetic amino acid such as alanine.

5.4 CONCLUSION

We were able to express and purify IRAK-4 phosphomimetics that were soluble and stable in solution. These were pure enough to be used in forming complexes with MyD88^{DD-WT} that we had purified earlier.

CHAPTER 6: Characterization of mutant Myddosome complexes

6.1 INTRODUCTION

In this chapter, we use phosphomimetic mutants of IRAK-4 death domain (IRAK-4^{DD-S8D} and IRAK-4$^{DD-T62D}$) to assemble the Myddosome complex. The reconstituted complexes are then further characterized by three biophysical methods mentioned in section 4.1 in order to compare them with the WT Myddosome complex. By so doing, we hope to have an answer to our hypothesis.

6.2 RESULTS

6.2.1 Reconstitution of IRAK-4^{DD-S8D} Myddosome complex

Mutants of IRAK-4 DD were purified with the aim of reconstituting the Myddosome complex with MyD88^{DD-WT} as shown in the previous section. Purified MyD88^{DD-WT} and IRAK-4^{DD-S8D} were concentrated and thereafter mixed to reconstitute the Myddosome. The mixture was then loaded on a Superdex 200 10/300 GL column to determine the formation of the complex. Six peaks (A-F) were observed on the size exclusion chromatograph shown in figure 6.1A. Peak A had a retention volume of 7.5 mL which corresponded to the void volume. Peak B had a retention volume 11.5 mL which is the same as that of WT Myddosome complex described in section 4.2.1. This peak most likely contained constituents of the mutant Myddosome complex. Samples from these peaks A and B were analysed by SDS-PAGE. As expected, the complex was detected in peak B as shown in figure 6.1B, lane 3-6. Protein bands corresponding to MyD88^{DD-WT} and IRAK-4^{DD-S8D} were seen at the expected molecular weights of 22 and 12.5 kDa respectively. A band below 22 kDa was also observed. This band was more likely to be a degradation product of MyD88^{DD-WT}. Samples from peak C were also analysed by SDS-PAGE as shown in figure 6.1B, lane 7-11. This peak contained three bands with molecular weights of 22, 27 and 46 kDa. The band at 22 kDa was undoubtedly that of MyD88^{DD-WT}. The remaining two were that of GST and excess 3C protease from the cleavage of the IRAK-4^{DD-S8D}. IRAK-4^{DD-WT} was not be detected in the sample from peak C. Samples from peaks D, E and F were also analysed by SDS-PAGE but there was nothing significant to see. As a result, the data was not shown.
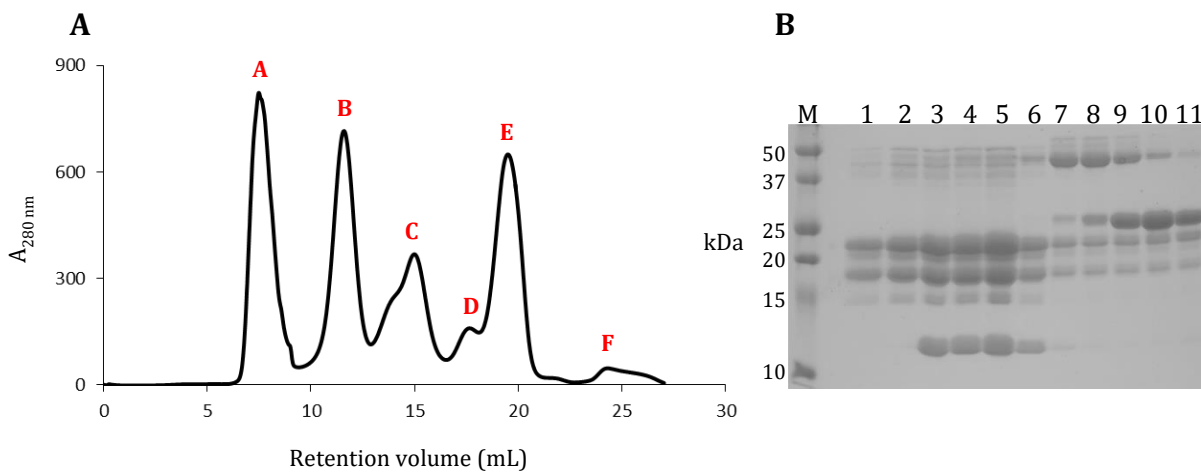
Figure 6.1: Size exclusion chromatography and SDS-PAGE analysis of IRAK-4^{DD-S8D} Myddosome complex. (A) Size exclusion chromatograph of purified IRAK-4^{DD-S8D} Myddosome complex obtained by using Superdex 200 10/300 GL column. Six peaks (A-F) with respective retention volumes of 7.5, 11.5, 15.0 and 17.5, 19.0 and 24 mL were observed on the chromatograph. (B) SDS-PAGE of size exclusion chromatography purified IRAK-4^{DD-S8D} Myddosome complex. Only the samples from peak A, B and C were analysed. Peak A contained aggregates of MyD88^{DD-WT} as expected (lane 1 and 2); peak B contained the IRAK-4^{DD-S8D} Myddosome complex (lane 3-6); peak C contained bands corresponding to MyD88^{DD-WT}, His$_6$-GST double tag and excess 3C protease at 22, 27 and 39 kDa respectively. Lane M, protein molecular marker.

6.2.2 Characterization of IRAK-4$^{DD\text{-}S8D}$ Myddosome complex by DLS

The mutant IRAK-4$^{DD\text{-}S8D}$ Myddosome complex was further characterised by DLS to determine if it was a *bona fide* complex or an aggregate. Three independent DLS measurements were performed. A single peak with a d(H) of 6.772 was observed for the complex as shown in figure 6.2. In addition, the complex was found to be monodisperse and having a PDI of 0.218. These results mean that the mutant IRAK-4$^{DD\text{-}S8D}$ Myddosome complex is not an aggregate.
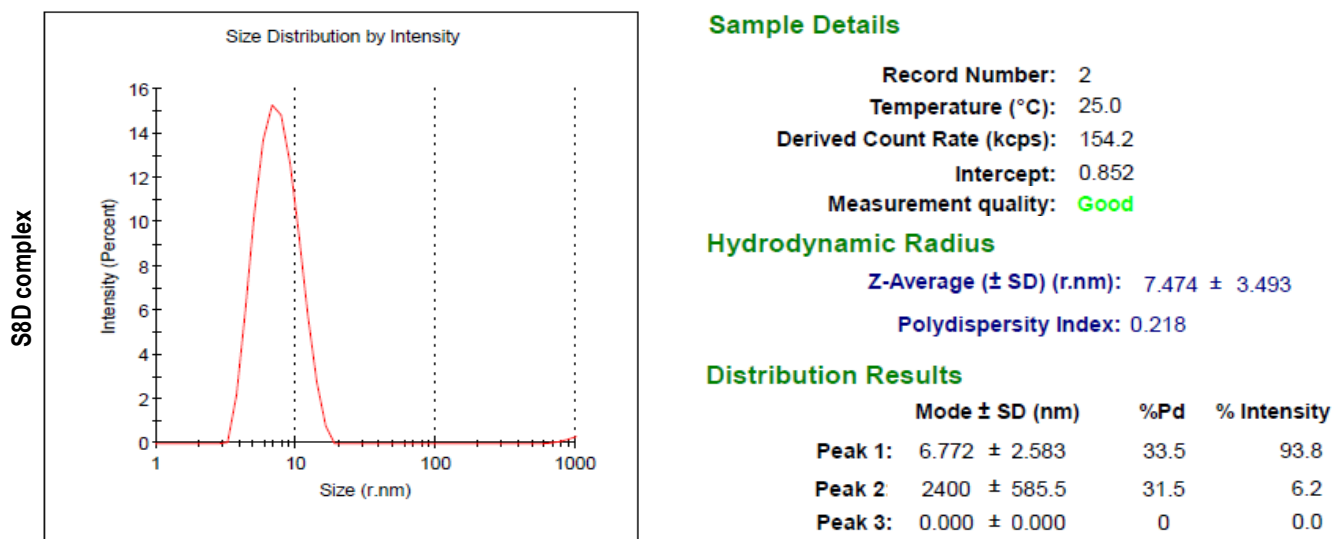
Figure 6.2: Dynamic light scattering (DLS) profile of IRAK-4$^{DD\text{-}S8D}$ Myddosome complex. DLS hydrodynamic distribution plot indicating a peak of the IRAK-4$^{DD\text{-}S8D}$ Myddosome complex with a d(H) of 6.772 and a PDI of 0.218.

6.2.3 Characterization of the complex by AUC

We determined by DLS that the IRAK-4$^{DD\text{-}S8D}$ Myddosome complex was free of aggregates. The next experiment we wanted to do was to determine its size. SV experiments were performed on the complex. A series of sequential scans of the protein sample were performed at one minute intervals as shown in figure 6.3A. As previously mentioned, the scan shows how the protein complexes move through solution in the sample chamber in real-time. A sedimentation coefficient was then calculated using the scan results to produce a profile shown in figure 6.3B. A prominent peak with a sedimentation coefficient of 7.5 S was observed from the graph. This sedimentation coefficient corresponded to the one previously reported on the Myddosome complex, meaning that the IRAK-4$^{DD\text{-}S8D}$ Myddosome complex had the same size as the WT complex.

A

B
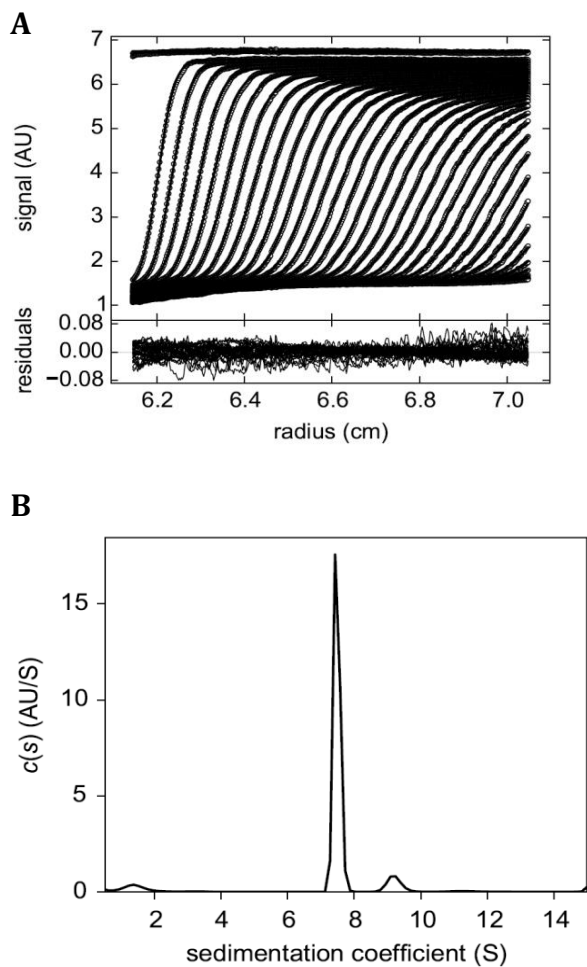
Figure 6.3: Sedimentation velocity (SV) profile of IRAK-4[DD-S8D] Myddosome complex. (A) SV graph showing scans of IRAK-4[DD-S8D] Myddosome complex recorded at a wavelength of 280 nm. The SV scans show the behaviour of the complex in the sample chamber in response to centrifugal force over time. (B) Sedimentation coefficient distribution plot of IRAK-4[DD-S8D] Myddosome complex derived from the SV scans. The outstanding peak centred at a sedimentation coefficient of 7.5 S corresponds to IRAK-4[DD-S8D] Myddosome complex.

6.2.4 Reconstitution of IRAK-4$^{DD-T62D}$ Myddosome complex

The next step was to reconstitute another mutant Myddosome complex, this time using IRAK-4^{DD-T62}. MyD88^{DD-WT} and IRAK-4$^{DD-T62D}$ were mixed together and loaded on SEC column. Complex formation was determined as previously described. Six peaks (A-F) were also observed on the size exclusion chromatograph shown in figure 6.1A. Based on previous results we only focused on Peak B since it has the same retention volume (11.5 mL) as WT and IRAK-4^{DD-S8D} Myddosome complex. Samples from peak B were analysed by SDS-PAGE to check for the presence of the IRAK-4$^{DD-T62D}$ Myddosome complex as shown in figure 6.1B, lane 2-5. The presence of the IRAK-4$^{DD-T62D}$ Myddosome was indeed detected. Protein bands corresponding to MyD88^{DD-WT} and IRAK-4$^{DD-T62D}$ were seen at the expected molecular weights of 22 and 12.5 kDa respectively.
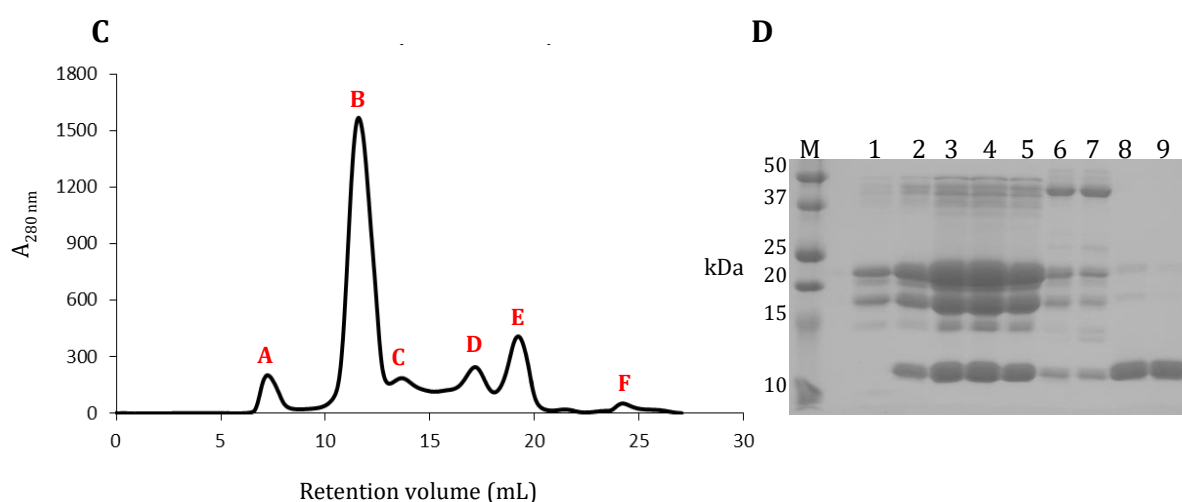
Figure 6.4: Size exclusion chromatography and SDS-PAGE analysis of IRAK-4$^{DD-T62D}$ Myddosome complex. (A) Size exclusion chromatograph purified IRAK-4$^{DD-T62D}$ Myddosome complex obtained by using Superdex 200 10/300 GL column. The chromatograph shows six peaks (A-F) which have retention volumes of 7.5, 11.5, 15.0 and 17.5, 19.0 and 24 mL respectively. (B) SDS-PAGE of size exclusion chromatography purified IRAK-4$^{DD-T62D}$ Myddosome complex. Only the samples from peak A, B and D were analysed. Peak A contained aggregates of MyD88^{DD-WT} (lane 1); both peak B (lane 2-5) and peak C (lane 6 and 7) contained the IRAK-4$^{DD-T62D}$ Myddosome complex (lane 2-5); peak D contained bands corresponding to IRAK-4$^{DD-T62D}$ at 12.5 kDa. Lane M, protein molecular marker.

6.2.5 Characterization of the IRAK-4$^{DD-T62D}$ Myddosome complex by DLS

Similarly, the IRAK-4$^{DD-T62D}$ Myddosome complex was further analysed by DLS as shown in figure 6.5. A single peak with a d(H) of 6.772 was observed and found to have a PDI of 0.165. These results showed that the sample contained only the complex and no other protein species or aggregates.
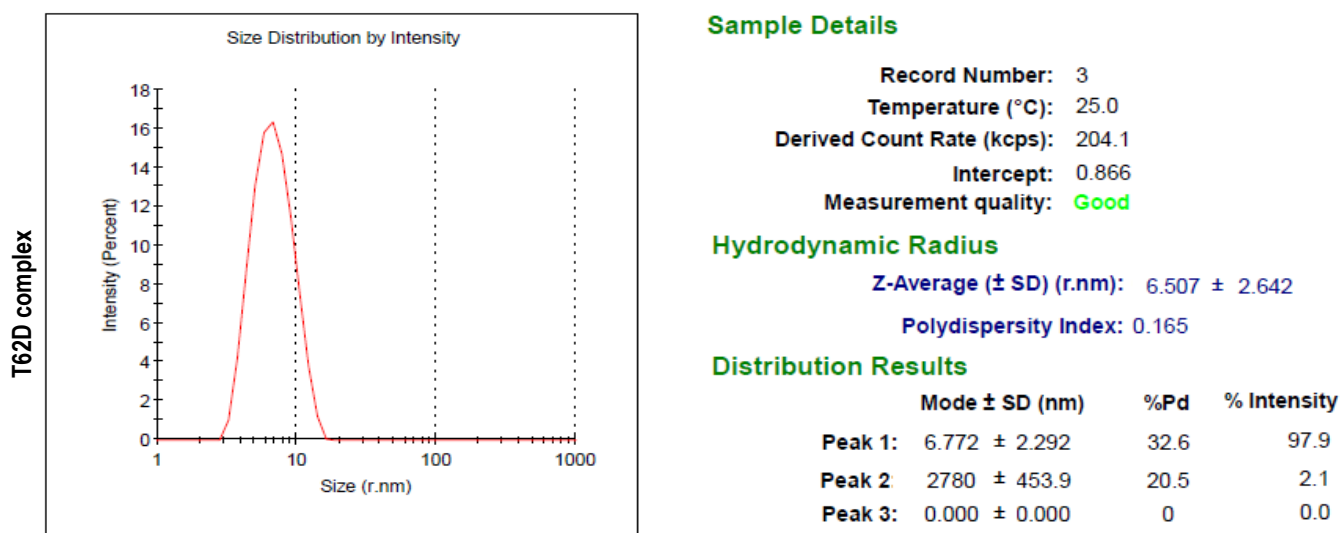
Figure 6.5: **Dynamic light scattering (DLS) profile of IRAK-4$^{DD-T62D}$ Myddosome complex**. DLS The hydrodynamic diameter distribution plot showing a peak of the IRAK-4$^{DD-T62D}$ Myddosome complex with a d(H) of 6.772 and a PDI of 0.165.

6.2.6 Characterization of IRAK-4$^{DD\text{-}T62D}$ Myddosome complex by AUC

The size of the IRAK-4$^{DD\text{-}T62D}$ Myddosome complex was also determined by AUC. A SV experiment was carried out and the resulting wavelength scans are shown in figure 6.6A. A sedimentation coefficient was then calculated and plotted as shown in figure 6.6B. From the plot, the sedimentation coefficient was found to be 7.5 S. This result, taken together with all the AUC results in this report, clearly shows that phosphomimetics had no effect on the Myddosome assembly.

A

B
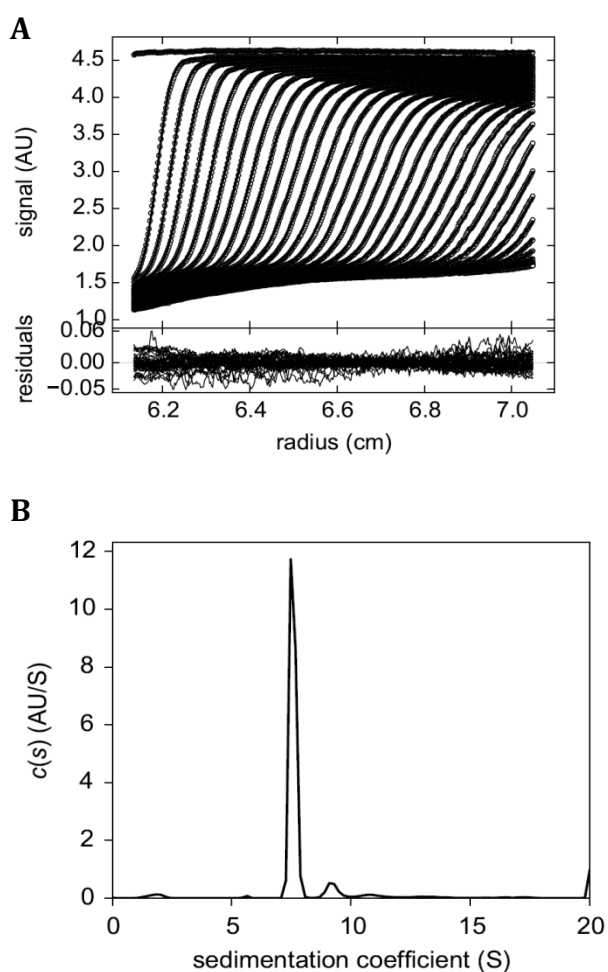
Figure 6.6: Sedimentation velocity (SV) profile of IRAK-4$^{DD\text{-}T62D}$ Myddosome complex. (A) SV graph showing scans of IRAK-4$^{DD\text{-}T62D}$ Myddosome complex recorded at a wavelength of 280 nm. The SV scans show the behaviour of the complex in the sample chamber in response to centrifugal force over time. (B) Sedimentation coefficient distribution plot of IRAK-4$^{DD\text{-}T62D}$ Myddosome complex derived from the SV scans. The outstanding peak centred at a sedimentation coefficient of 7.5 S corresponds to IRAK-4$^{DD\text{-}T62D}$ Myddosome complex.

6.3 DISCUSSION

Initial studies on IRAK-4 showed that its wild type failed to interact with MyD88, its binding partner in TLR signalling pathways (Li et al., 2002). Instead, the dead kinase IRAK-4 (K213A) interacted with MyD88 (Li et al., 2002). This was shown by immunoprecipitation. An important question arose: why does wild type IRAK-4 fail to interact with MyD88? So far, there had not been any answer but a hypothesis that autophosphorylation might be a switch that controls this interaction. This is not surprising considering the number of autophosphorylation sites that have been identified on IRAK-4. So far, there are thirteen autophosporylation sites that have been identified (Cheng et al., 2007). Of interest to us are those that are located on its DD because it interacts with MyD88 through homotypic DD interactions. Two sites namely S8 and T62 were found to be autophosphorylated in IRAK-4 (Motshwene, 2008). We were interested in studying if these sites act as a switch that regulates the interaction between IRAK-4 and MyD88.

The Myddosome is an oligomeric complex that was reconstituted by mixing the DDs of IRAK-4 and MyD88. This complex was found to have an unusual stoichiometry of 7 MyD88 to 4 IRAK-4 DD molecules. We used two phosphomimetic mutants of IRAK-4 in which S8 and T62 were replaced with aspartate. The aim of our experiment was to determine if these phosphomimetic mutants will form a complex with MyD88^{DD-WT}. If there was complex formation, we wanted to compare its size and stoichiometry with that of the wild type Myddosome complex.

Complexes were reconstituted using IRAK-4 DD phosphomimetic mutants and MyD88^{DD-WT}. The complexes formed and were characterised using three biophysical techniques namely, size exclusion chromatography, dynamic light scattering and analytical ultracentrifugation. Our results showed that the complexes that we reconstituted were similar in size to the Myddosome. This meant that the stoichiometry was more likely similar. Therefore our findings suggested that the phosphomimetic mutants did not have any effect on the Myddosome assembly.

A study by Dossang et al., 2016 showed that IRAK-4 DD that is phosphorylated at serine 8 failed to assemble into the Myddosome when mixed with the DD of MyD88. The phosphorylated DD was isolated in an unusual way. Full length IRAK-4 was expressed in Baculovirus and thereafter purified. It is known that IRAK-4 autophosphorylates when expressed in Baculovirus (Cheng et al., 2007). The purified autophosphorylated IRAK-4 was then treated with thrombin which cleaved it into the DD and KD. The cleaved DD was isolated and subjected to phosphopeptide mapping which showed phosphorylation at Serine 8. An anion exchange monoQ column was then used to separate phosphorylated from the unphosphorylated DD.

Our IRAK-4^{DD-S8D} phosphomimetic interacted with MyD88^{DD-WT} to form the Myddosome. This was unexpected and is in contrast with the study by Dossang et al., 2016. How do we explain this discrepancy? Phosphomimetics in the form of aspartate or glutamate have a single negative charge unlike phosphate monoesters that have two negative charges. Also, phosphomimetics are smaller and have a different geometry when compared to phosphoserine or phosphothreonine (Chen and Cole, 2015). More importantly, a phosphate group adds hydrogen bond acceptor hydrogen atoms that increase the size of the amino acid. This addition is capable of altering protein structure and affect intermolecular interaction in protein complexes (Chen and Cole, 2015). That could be the reason why phosphomimetics sometimes fail to work.

6.4 CONCLUSSION

We have successfully reconstituted the mutant Myddosome complexes with wild type MyD88 DD and mutant DDs of IRAK-4. The biophysical parameters of the mutant Myddosome were found to be the similar to that of the WT Myddosome complex. Altogether, our results show that IRAK-4 phosphomimetic mutations did not abolish the formation of the Myddosome complex with MyD88. This is in contrast with our hypothesis. We attribute our finding to the fact that phosphomimetics do not always mimic phosphorylation. Alternative methods to study IRAK-4 autophosphorylation such as in vitro kinase assays will be considered in future.

CHAPTER 7: CONCLUSION

Our aim was to try to understand the regulatory mechanism for the assembly of the Myddosome complex. This complex is important in TLR postreceptor signalling cascades that lead to the production of proinflammatory cytokines. It has a stoichiometry of DDs of MyD88 and IRAK-4 of 7:4 and 8:4. So far, it is not clear how such an oligomeric complex dissociates. The switch that controls its disassembly has not been determined but subject to postulation. Autophosphorylation has been suggested as a possible switch that controls the association and dissociation of the Myddosome complex. Key to the suggestion, are the two autophosphorylation sites that have been mapped in the death domain of IRAK-4. These are serine 8 and threonine 62. We created phosphomimetic mutants by substituting these two amino acid residues with aspartic acid. The created phosphomimetics were then used to study the effects of autophosphorylation on the Myddosome assembly.

We achieved the objectives of our study by expressing and purifying all the desired constructs. Thereafter, we reconstituted the Myddosome complex using wildtype and mutant phosphomimetics of IRAK-4 death domain. The reconstituted complexes were characterised using three biophysical techniques namely, size exclusion chromatography, dynamic light scattering and analytical ultracentrifugation.

We found that the Myddosome was able to reconstitute when phosphomimetics were used. This finding was contrary to what we had expected. Our data were very clear and convincing that phosphomimetics had no effect on the Myddosome assembly. This then led us to reject our hypothesis.

In conclusion, phosphomimetics are not ideal replacements for phosphorylation. They carry a single negative charge and therefore failed to weaken the interaction between the death domains of MyD88 and IRAK-4.

CHAPTER 8: REFERENCES

ADACHI, O., KAWAI, T., TAKEDA, K., MATSUMOTO, M., TSUTSUI, H., SAKAGAMI, M., NAKANISHI, K. & AKIRA, S. 1998. Targeted disruption of the MyD88 gene results in loss of IL-1-and IL-18-mediated function. *Immunity,* 9, 143-150.

AKIRA, S. & TAKEDA, K. 2004. Toll-like receptor signalling. *Nature Reviews Immunology,* 4, 499–511.

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, 779-789.

ANGUS, D. C., LINDE-ZWIRBLE, W. T., LIDICKER, J., CLERMONT, G., CARCILLO, J. & PINSKY, M. R. 2001. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Critical Care Medicine,* 29, 1303-1310.

BEENSTOCK, J., MOOSHAYEF, N. & ENGELBERG, D. 2016. How do protein kinases take a selfie (autophosphorylate)? *Trends in Biochemical Sciences,* 41, 938-953.

BELL, J. K., BOTOS, I., HALL, P. R., ASKINS, J., SHILOACH, J., SEGAL, D. M. & DAVIES, D. R. 2005. The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proceedings of the National Academy of Sciences,* 102, 10976-10980.

BERG, O. G. & SILVA, P. J. 1997. Codon Bias in Escherichia Coin: the Influence of Codon Context on Mutation and Selection. *Nucleic Acids Research,* 25, 1397-1404.

BERGLUND, N. A., KARGAS, V., ORTIZ-SUAREZ, M. L. & BOND, P. J. 2015. The role of protein–protein interactions in Toll-like receptor function. *Progress in Biophysics and Molecular Biology,* 119, 72-83.

BOTOS, I., SEGAL, D. M. & DAVIES, D. R. 2011. The structural biology of Toll-like receptors. *Structure,* 19, 447-459.

BOWIE, A. & O'NEILL, L. A. 2000. The interleukin-1 receptor/Toll-like receptor superfamily: signal generators for pro-inflammatory interleukins and microbial products. *Journal of Leukocyte Biology,* 67, 508-514.

BRUBAKER, S. W., BONHAM, K. S., ZANONI, I. & KAGAN, J. C. 2015. Innate immune pattern recognition: a cell biological perspective. *Annual Review of Immunology,* 33, 257-290.

BURNETT, G. & KENNEDY, E. P. 1954. The enzymatic phosphorylation of proteins. *Journal of Biological Chemistry,* 211, 969-980.

CHANG, A. Y., CHAU, V., LANDAS, J. A. & PANG, Y. 2017. Preparation of calcium competent Escherichia coli and heat-shock transformation. *JEMI Methods,* 1, 22-25.

CHEN, Z. & COLE, P. A. 2015. Synthetic approaches to protein phosphorylation. *Current Opinion in Chemical Biology,* 28, 115-122.

CHENG, H., ADDONA, T., KESHISHIAN, H., DAHLSTRAND, E., LU, C., DORSCH, M., LI, Z., WANG, A., OCAIN, T. D. & LI, P. 2007. Regulation of IRAK-4 kinase activity via autophosphorylation

within its activation loop. *Biochemical and Biophysical Research Communications,* 352, 609-616.

CLARKE IV, T. F. & CLARK, P. L. 2008. Rare codons cluster. *PloS One,* 3, e3412.

COLE, J. L., LARY, J. W., MOODY, T. P. & LAUE, T. M. 2008. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods in Cell Biology,* 84, 143-179.

DAILEY, T. A. & DAILEY, H. A. 1997. Expression, purification, and characteristics of mammalian protoporphyrinogen oxidase. *Methods in Enzymology,* 281, 340-349.

DE MARCO, A. 2006. Two-step metal affinity purification of double-tagged (NusA–His6) fusion proteins. *Nature Protocols,* 1, 1538–1543.

DOSSANG, A. C., MOTSHWENE, P. G., YANG, Y., SYMMONS, M. F., BRYANT, C. E., BORMAN, S., GEORGE, J., WEBER, A. N. & GAY, N. J. 2016. The N-terminal loop of IRAK-4 death domain regulates ordered assembly of the Myddosome signalling scaffold. *Scientific Reports,* 6, 37267.

GAY, N. J. & GANGLOFF, M. 2007. Structure and function of Toll receptors and their ligands. *Annual Review of Biochemistry,* 76, 141-165.

GECCHELE, E., MERLIN, M., BROZZETTI, A., FALORNI, A., PEZZOTTI, M. & AVESANI, L. 2015. A comparative analysis of recombinant protein expression in different biofactories: bacteria, insect cells and plant systems. *Journal of Visualized Experiments*, 97, 52459.

GEORGE, J., MOTSHWENE, P. G., WANG, H., KUBARENKO, A. V., RAUTANEN, A., MILLS, T. C., HILL, A. V., GAY, N. J. & WEBER, A. N. 2011. Two human MYD88 variants, S34Y and R98C, interfere with MyD88-IRAK4-myddosome assembly. *Journal of Biological Chemistry,* 286, 1341-1353.

HATO, T. & DAGHER, P. C. 2015. How the innate immune system senses trouble and causes trouble. *Clinical Journal of the American Society of Nephrology,* 10, 1459-1469.

HOWLETT, G. J., MINTON, A. P. & RIVAS, G. 2006. Analytical ultracentrifugation for the study of protein association and assembly. *Current Opinion in Chemical Biology,* 10, 430-436.

JIN, M. S. & LEE, J.-O. 2008. Structures of the toll-like receptor family and its ligand complexes. *Immunity,* 29, 182-191.

KAGAN, J. C., MAGUPALLI, V. G. & WU, H. 2014. SMOCs: supramolecular organizing centres that control innate immunity. *Nature Reviews Immunology,* 14, 821.

KAWAI, T., ADACHI, O., OGAWA, T., TAKEDA, K. & AKIRA, S. 1999. Unresponsiveness of MyD88-deficient mice to endotoxin. *Immunity,* 11, 115-122.

KAWASAKI, T. & KAWAI, T. 2014. Toll-like receptor signaling pathways. *Frontiers in Immunology,* 5, 461.

KOBE, B. & KAJAVA, A. V. 2001. The leucine-rich repeat as a protein recognition motif. *Current Opinion in Structural Biology,* 11, 725-732.

KUGLSTATTER, A., VILLASEÑOR, A. G., SHAW, D., LEE, S. W., TSING, S., NIU, L., SONG, K. W., BARNETT, J. W. & BROWNER, M. F. 2007. Cutting edge: IL-1 receptor-associated kinase 4 structures reveal novel features and multiple conformations. *The Journal of Immunology,* 178, 2641-2645.

KUMAR, H., KAWAI, T. & AKIRA, S. 2011. Pathogen recognition by the innate immune system. *International Reviews of Immunology,* 30, 16-34.

LASKER, M. V., GAJJAR, M. M. & NAIR, S. K. 2005. Cutting edge: molecular structure of the IL-1R-associated kinase-4 death domain and its implications for TLR signaling. *The Journal of Immunology,* 175, 4175-4179.

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, 973-983.

LI, M., ZHOU, Y., FENG, G. & SU, S. B. 2009. The critical role of Toll-like receptor signaling pathways in the induction and progression of autoimmune diseases. *Current Molecular Medicine,* 9, 365-374.

LI, S., STRELOW, A., FONTANA, E. J. & WESCHE, H. 2002. IRAK-4: a novel member of the IRAK family with the properties of an IRAK-kinase. *Proceedings of the National Academy of Sciences,* 99, 5567-5572.

LIN, S.-C., LO, Y.-C. & WU, H. 2010. Helical assembly in the MyD88–IRAK4–IRAK2 complex in TLR/IL-1R signalling. *Nature,* 465, 885-890.

MEDZHITOV, R. 2001. Toll-like receptors and innate immunity. *Nature Reviews Immunology,* 1, 135–145.

MOTSHWENE, P. G., MONCRIEFFE, M. C., GROSSMANN, J. G., KAO, C. C., AYALURU, M., SANDERCOCK, A. M., ROBINSON, C. V., LATZ, E. & GAY, N. J. 2009. An oligomeric signalling platform formed by the toll-like receptor signal transducers MyD88 and IRAK4. *Journal of Biological Chemistry*, 284, 25404–25411.

NYMAN, T., STENMARK, P., FLODIN, S., JOHANSSON, I., HAMMARSTRÖM, M. & NORDLUND, P. 2008. The crystal structure of the human toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer. *Journal of Biological Chemistry,* 283, 11861-11865.

OHNISHI, H., TOCHIO, H., KATO, Z., ORII, K. E., LI, A., KIMURA, T., HIROAKI, H., KONDO, N. & SHIRAKAWA, M. 2009. Structural basis for the multiple interactions of the MyD88 TIR domain in TLR4 signaling. *Proceedings of the National Academy of Sciences,* 106, 10260-10265.

OLSEN, J. V. & MANN, M. 2013. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Molecular & Cellular Proteomics*, 12, 3444–3452.

O'NEILL, L. A. & BOWIE, A. G. 2007. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nature Reviews Immunology,* 7, 353–364.

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, 1191–1195

PARK, H. H. 2011. Structural analyses of death domains and their interactions. *Apoptosis,* 16, 209-220.

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., SOUDAIS, C., DUPUIS, S., FEINBERG, J. & FIESCHI, C. 2003. Pyogenic bacterial infections in humans with IRAK-4 deficiency. *Science,* 299, 2076-2079.

POLTORAK, A., HE, X., SMIRNOVA, I., LIU, M.-Y., VAN HUFFEL, C., DU, X., BIRDWELL, D., ALEJOS, E., SILVA, M. & GALANOS, C. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science,* 282, 2085-2088.

REED, J. C., DOCTOR, K. S. & GODZIK, A. 2004. The domains of apoptosis: a genomics perspective. *Science Signaling,* 2004, re9-re9.

ROSANO, G. L. & CECCARELLI, E. A. 2014. Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in Microbiology,* 5, 172.

SAN-MIGUEL, T., PÉREZ-BERMÚDEZ, P. & GAVIDIA, I. 2013. Production of soluble eukaryotic recombinant proteins in E. coli is favoured in early log-phase cultures induced at low temperature. *SpringerPlus,* 2, 89.

SHAN, S., LIU, D., LIU, R., ZHU, Y., LI, T., ZHANG, F., AN, L., YANG, G. & LI, H. 2018. Non-mammalian Toll-like receptor 18 (Tlr18) recognizes bacterial pathogens in common carp (Cyprinus carpio L.): indications for a role of participation in the NF-κB signaling pathway. *Fish & Shellfish Immunology,* 72, 187-198.

STUDIER, F. W. 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification,* 41, 207-234.

SUZUKI, N., SUZUKI, S., DUNCAN, G. S., MILLAR, D. G., WADA, T., MIRTSOS, C., TAKADA, H., WAKEHAM, A., ITIE, A. & LI, S. 2002. Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4. *Nature,* 416, 750-756.

TAKEUCHI, O. & AKIRA, S. 2010. Pattern recognition receptors and inflammation. *Cell,* 140, 805-820.

TARTAGLIA, L. A., AYRES, T. M., WONG, G. H. & GOEDDEL, D. V. 1993. A novel domain within the 55 kd TNF receptor signals cell death. *Cell,* 74, 845-853.

UEMATSU, S. & AKIRA, S. 2006. Toll-like receptors and innate immunity. *Journal of Molecular Medicine,* 84, 712-725.

UEMATSU, S. & AKIRA, S. 2008. Toll-like Receptors (TLRs) and Innate Immunity: Toll-Like receptors (TLRs) and their ligands. *Handbook of Experimental Pharmacology*, 183, 1-20.

VE, T., WILLIAMS, S. J. & KOBE, B. 2015. Structure and function of Toll/interleukin-1 receptor/resistance protein (TIR) domains. *Apoptosis,* 20, 250-261.

VIVIER, E., RAULET, D. H., MORETTA, A., CALIGIURI, M. A., ZITVOGEL, L., LANIER, L. L.,

YOKOYAMA, W. M. & UGOLINI, S. 2011. Innate or adaptive immunity? The example of natural killer cells. *Science,* 331, 44-49.

WANG, Z., LIU, J., SUDOM, A., AYRES, M., LI, S., WESCHE, H., POWERS, J. P. & WALKER, N. P. 2006. Crystal structures of IRAK-4 kinase in complex with inhibitors: a serine/threonine kinase with tyrosine as a gatekeeper. *Structure,* 14, 1835-1844.

WAUGH, D. S. 2005. Making the most of affinity tags. *Trends in Biotechnology,* 23, 316-320.

WEBER, C. H. & VINCENZ, C. 2001. The death domain superfamily: a tale of two interfaces? *Trends in Biochemical Sciences,* 26, 475-481.

XU, Y., TAO, X., SHEN, B., HORNG, T., MEDZHITOV, R., MANLEY, J. L. & TONG, L. 2000. Structural basis for signal transduction by the Toll/interleukin-1 receptor domains. *Nature,* 408, 111–115.

YIN, Q., FU, T.-M., LI, J. & WU, H. 2015. Structural biology of innate immunity. *Annual Review of Immunology,* 33, 393-416.
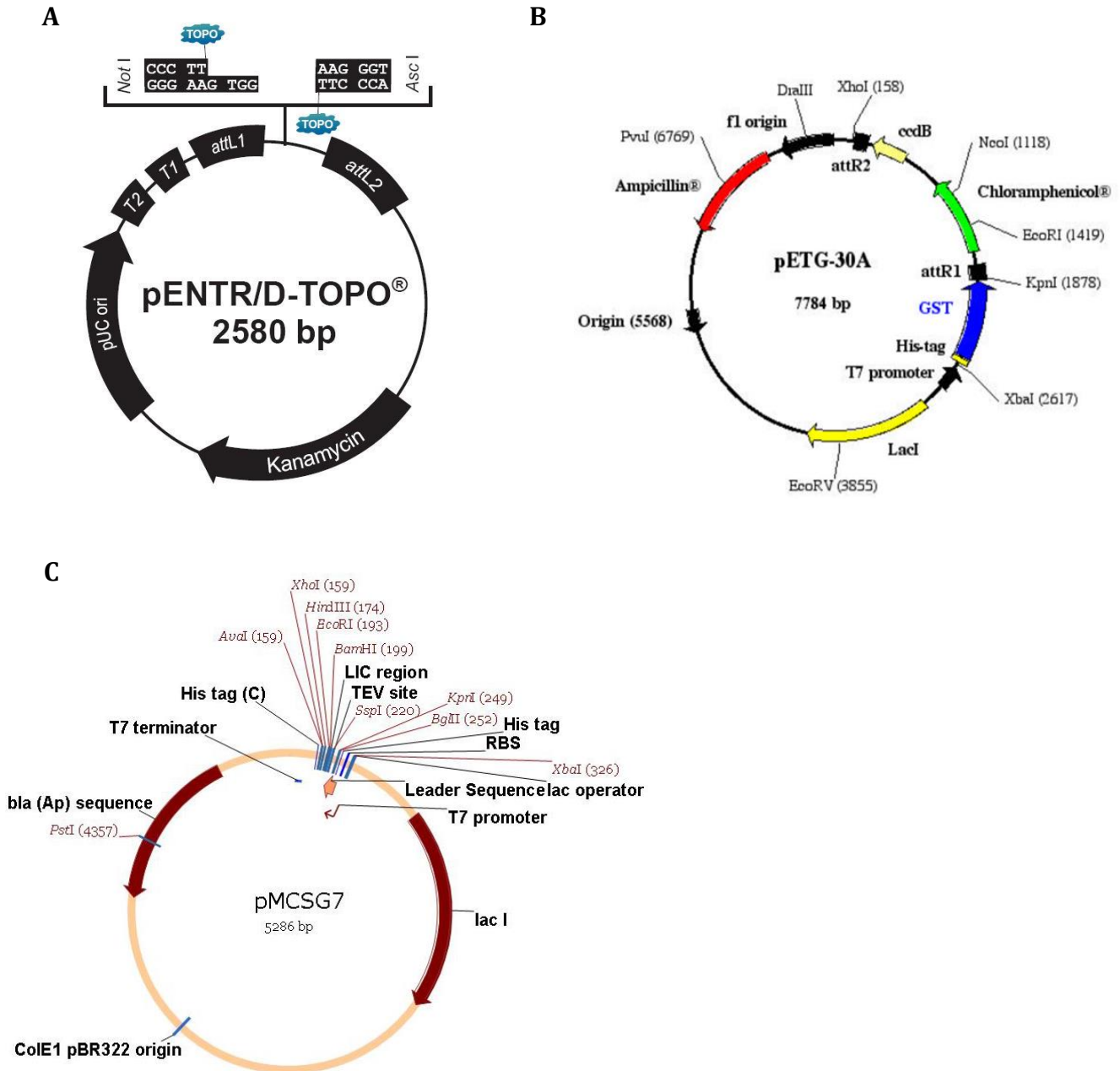
CHAPTER 9: APPENDIX

A

B

C

Figure 9.1: **Plasmid maps of Gateway cloning vectors used by our predecessors**. (A) pENTR/D-TOPO entry vector; Genes coding for MyD88^{WT-FL}, MyD88^{WT-DD} and IRAK-4$^{DD-WT/S8D/T62D}$ were cloned in the *att*L1 and *att*L2 sites of the entry vector. Ligation was catalyzed by topoisomerases indicated in blue (B) pETG30A expression vector. IRAK-4$^{DD-S8D/T62D}$ gene cloned in an entry vector was transferred into pETG30A expression vector by LR recombination reaction. IRAK-4$^{DD-S8D/T62D}$ was expressed as His$_6$-GST fusion in this study. (C) pMCSG7 expression vector. MyD88^{DD-WT}/IRAK-4^{DD-WT} gene cloned in an entry vector was transferred into pMCSG7 expression vector by LR recombination reaction.
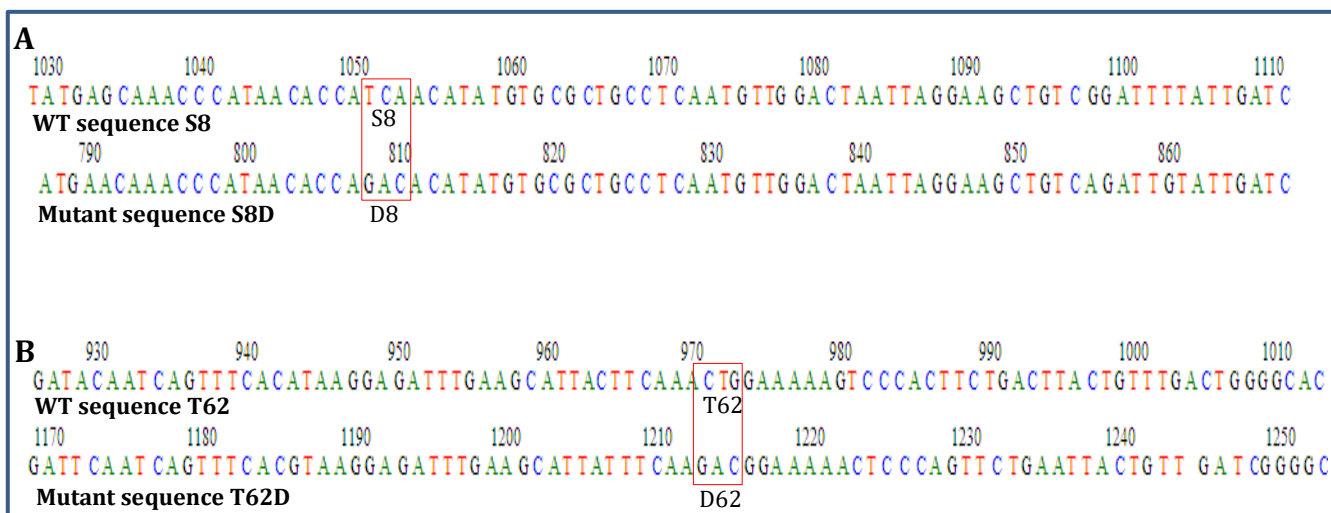
Figure 9.2: Comparison of DNA sequences of WT and mutant IRAK-4 DDs. (A) DNA Sequence comparison showing the phosphomimetic mutation Ser to Asp at codon 8, (B) DNA sequence comparison indicating phosphomimetic mutation Thr to Asp at codon 62, courtesy of our predecessors. The mutated sites are shown in red boxes.