UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Structural and Functional Analysis of *Listeria* Adhesion Protein

Clare Anne Boswell

15121187

Supervisor: Prof Wolf-Dieter Schubert

Submitted in partial fulfilment of the requirements for the degree

Masters in Biochemistry

In the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

1 February 2022

DECLARATION OF ORIGINALITY

UNIVERSITY OF PRETORIA

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

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

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

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

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

Full names of student: Clare Anne Boswell

Student number: 15121187

Topic of work: Structural and Functional Analysis of *Listeria* Adhesion Protein

Declaration

I, Clare Anne Boswell declare that the thesis/dissertation, which I hereby submit for the degree Masters in Biochemistry, at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: .....

DATE: .1 February 2022.....

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ActA	Actin assembly-inducing protein
AEC	Anion exchange chromatography
adh	Gene fragment encoding the alcohol dehydrogenase domain of *Listeria* adhesion protein
ADH	Alcohol dehydrogenase domain of *Listeria* adhesion protein
AdhE	Acetaldehyde-alcohol dehydrogenase
aldh	Gene fragment encoding the acetaldehyde domain of *Listeria* adhesion protein
ALDH	Acetaldehyde domain of *Listeria* adhesion protein
amp	Ampicillin
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
CoA	Coenzyme A
CoASH	Coenzyme A (reduced)
DLS	Dynamic light scattering
E. coli	*Escherichia coli*
EtOH	Ethanol
GST	Glutathione-S-transferase
Hsp	Heat shock protein
Hsp60	Heat shock protein 60
I_KB	Nuclear factor kappa B
IL-6	Interleukin 6
InlA	Internalin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kan	Kanamycin
k_{cat}	Catalytic efficiency
K_m	Michaelis constant
LAP	*Listeria* adhesion protein
lap	Gene encoding *Listeria* adhesion protein
LB	Lysogeny broth
L. monocytogenes	*Listeria monocytogenes*
Lm	*Listeria monocytogenes*
MLCK	Myosin light-chain kinase
NAD^+	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)

NaOAc	Sodium Acetate
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
o/n	Overnight
PCR	Polymerase chain reaction
PBS	Phosphate-buffered saline (buffer)
RE	Restriction enzyme
SEC	Size-exclusion chromatography
TAE	Tris-acetate-EDTA
TNFα	Tumour necrosis factor α
V_{max}	Maximum velocity

1. Abstract

Listera monocytogenes (*Lm*) is a gram-positive opportunistic foodborne pathogen. It is responsible for the disease listeriosis, which though rare, causes high morbidity and mortality. The pathogen targets the intestine for systemic entry. *Lm* uses several membrane proteins to breach the intestinal barrier and to translocate for systemic distribution. Additionally, the pathogen utilizes a normally cytosolic protein for translocation: bifunctional acetaldehyde alcohol dehydrogenase, which moonlights as *Listeria* adhesion protein (LAP). LAP has been reported to interact with mitochondrial heat shock protein 60 (Hsp60) presented on intestinal epithelial cells. The interaction allows for paracellular translocation, avoiding intracellular host immunity. LAP was cloned, produced and purified for downstream experimentation. The purified protein was characterized with enzyme activity assays and electron microscopy. The acetaldehyde dehydrogenase and alcohol dehydrogenase domains of LAP have V_{max} values of 0.56 $mM.min^{-1}$ and 1.17 $mM.min^{-1}$ respectively. LAP was also found to oligomerise into filaments potentially needed for activity.

2. Introduction

2.1. Foodborne pathogens

Consuming contaminated foods significantly impacts public health worldwide and is a major contributor of morbidity and mortality. The World Health Organization estimates 600 million cases of foodborne illnesses yearly with 420 000 associated deaths (Kirk *et al.*, 2015). Foodborne diseases occur globally though with significant skewing towards lower income countries. The main bacterial genera contributing to outbreaks are *Salmonella, Escherichia, Toxoplasma, Listeria* and *Campylobacter*. The pathogens are distributed via food products such as unpasteurized milk, cold-meats, raw vegetables, unwashed fruit, uncooked meat and various other foods such as ice cream, deserts, or similar contaminated prior to consumption. Foodborne pathogens mostly cause diarrhoea and diarrhoea related symptoms. Most cases are not lethal, but prolonged illness particularly in immuno-compromised patients may lead to death (Havelaar *et al.*, 2015, Lake *et al.*, 2010).

2.2. Listeriosis

Listeriosis is a foodborne disease caused by the opportunistic pathogen *Listeria monocytogenes* (*L. monocytogenes*, *Lm*). For most of the 20th century it was seen as a disease of ruminants such as cattle and sheep. The symptoms included anorexia, encephalitis and head tilt, causing the animals to be off balance and to walk in circles (Brugère-Picoux, 2008). In the 1980s listeriosis was discovered to affect humans and several large Listeriosis outbreaks were reported in Germany, Sweden, France, the United States of America, New Zealand among others (White *et al.*, 2002).

Listeriosis mainly affects immunosuppressed individuals such as neonates, pregnant women and the elderly (Lecuit *et al.*, 1997). Initially *Lm* infections cause symptoms such as nausea, diarrhoea, fever, headaches and muscle aches with later more severe symptoms being meningitis, encephalitis and spontaneous abortion. Outbreaks of listeriosis have been attributed to ready-to-eat products such as deli meats, fruit dairy products and uncooked meat (Swaminathan and Gerner-Smidt, 2007, Thomas *et al.*, 2020).

The mortality of infected individuals is high, with 20 to 30% confirmed cases leading to death (de Noordhout *et al.*, 2014). The largest listeriosis outbreak globally occurred in South Africa between January 2017 and September 2018. A total of 937 cases were linked to the outbreak, resulting in a mortality rate of 27%. Half the cases were associated with pregnancy including neonatal infection. The source of this outbreak was linked to the ready-to-eat meat product “polony”, which is not cooked or heated prior to consumption (Thomas *et al.*, 2020).

2.3. *Listeria monocytogenes*

2.3.1. Background

Lm is a gram-positive, rod-shaped bacterium first discovered in South Africa in 1924 during an outbreak affecting rabbits and guinea pigs (Murray *et al.*, 1926). The genus *Listeria* currently accommodates 17 known species. Prominent examples include: *L. monocytogenes, L. ivanovii, L. seeligeri, L. innocua, L. welshimeri, and L. grayi.* Only *L. monocytogenes* and *L. ivanovii* are pathogenic, infecting humans and ruminants or other animals respectively (Vazquez-Boland *et al.*, 2001). *Listeria monocytogenes* is a facultative anaerobe able to grow at low temperatures, low pH and high salt concentrations. This causes major concerns for the food industry as *Lm* can survive refrigeration temperatures and food preservation conditions. *Listeria* is ubiquitously found in water and soil indicating that it is adapted to many environments (Gandhi and Chikindas, 2007).

2.3.2. Pathogenesis

Lm has numerous regulatory factors which help signal transition between infection and non-infection conditions (de las Heras *et al.*, 2011). After ingestion of contaminated food *Lm* switches to an infectious lifestyle. It produces and presents some highly specialized proteins that allow it to cross the intestinal barrier. Similar to other enteropathogenic pathogens such as *Shigella* and *Salmonella* spp, *Lm* elicits phagocytosis in normally non-phagocytic mammalian cells inducing its own uptake. Upon crossing the intestinal barrier *Lm* disseminates into the lymph and blood where it targets the liver and occasionally other organs (Cossart, 2011). *Lm* is one of few bacteria able to overcome both the blood brain and placental barriers leading to encephalitis and meningitis in the first case and infection of the foetus, spontaneous abortions or neonatal fatality in the second (Rengarajan *et al.*, 2016).

2.3.3. Traditional route of infection

Once ingested, most *Lm* cells are killed in the stomach. Those that reach the intestine, use the surface protein internalin (InlA) to induce its uptake into the intestinal epithelium. InlA interacts with epithelial or E-cadherin (EC), a protein of epithelial adherence junctions. The extracellular interaction induces a complex intracellular signalling cascade and recruitment of endocytic proteins that cause the cell to engulf the bacterium despite intestinal epithelial cells normally being non-phagocytic (Cossart and Helenius, 2014). Within the resulting phagosome *Lm* secretes the toxin listeriolysin O which lyses the endocytic vacuole releasing the bacterium into the cytoplasm where it begins to proliferate. To evade the intracellular

immune mechanisms, *Lm* uses actin assembly-inducing protein (ActA) to polymerize host actin to propel itself away from autophagic vacuoles (Gouin *et al.*, 2005, Pizarro-Cerdá *et al.*, 2012). At the plasma membrane, *Lm* creates pseudopod-like protrusions into neighbouring cells. Some protrusions are severed by cytoskeletal machinery involving dynamin, creating a double-membraned vesicle in the neighbouring cell (Lambrechts *et al.*, 2008). Here *Lm* secretes listeriolysin O and phospholipases A and B to lyse the double membrane and once more release the bacterium into the cytoplasm. Individual *Lm* exit the basal membrane of host epithelial cells by an unknown mechanism to reach the lamina propria and complete the crossing of the intestinal barrier (Drolia *et al.*, 2018).

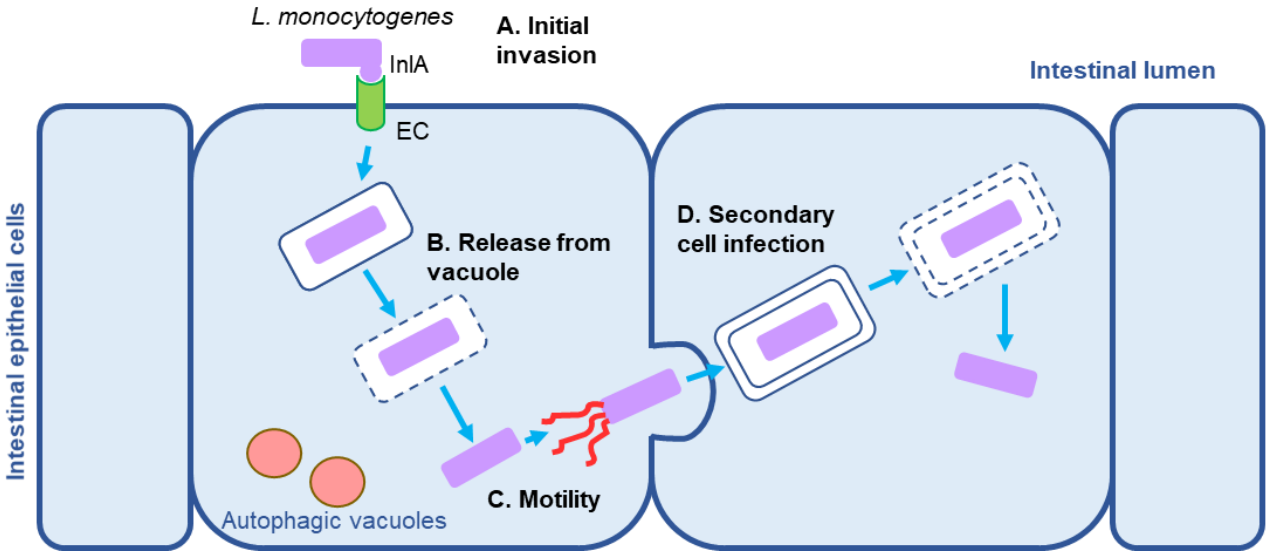

Figure 2.1 Internalin associated invasion route of *Listeria monocytogenes*.

A. Internalin (InlA) binds the N-terminal domains of E-cadherin to induce the internalization of *Lm.* **B.** *Lm* produces LLO that degrades the membrane of the vacuole to release *Lm* into the cytoplasm. **C.** The bacterium polymerizes actin for motility to avoid autophagic vacuoles and to move toward the plasma membrane. **D.** *Lm* infects the adjacent cell by deforming the plasma membrane to create protrusions into the neighbouring cell. Once the recipient cell pinches off the protrusion, the bacterium is concealed in a double membrane vacuole, which it lyses to continue the cycle of infection.

2.3.4. *Listeria* adhesion protein associated route of infection

Recently another listerial protein was found to be involved in the early invasion of the host. *Listeria* adhesion protein (LAP) was observed to interact with the eukaryotic heat shock protein 60 (Hsp60), normally found in the mitochondria of intestinal epithelial cells (Wampler *et al.*, 2004). Twelve to 48 hours after ingestion, the nutrient-limiting environment of the intestine induces LAP production by *Lm*. Most LAP molecules remains cytosolic, while a small fraction is secreted and associates with the bacterial cell wall. Surface-associated LAP interacts with Hsp60 molecules exposed on intestinal epithelial cell membrane. The interaction of LAP with Hsp60 initiates an intracellular signalling cascade involving IκB

kinase (IKK) and the activation of the “nuclear factor kappa-light-chain-enhancer of activated B cells” (NF-κB) pathway. The pathway induces proinflammatory cytokines, tumour necrosis factor α (TNFα) and interleukin-6 (IL-6) increasing epithelial cell membrane permeability. NF-κB activation also activates the myosin light-chain kinase (MLCK), which phosphorylates the myosin light-chain to actively reallocate junctional proteins. This resulting junctional opening provides *Lm* direct access to the lamina propria while also exposing EC for InlA-mediated uptake (Drolia *et al.*, 2018).

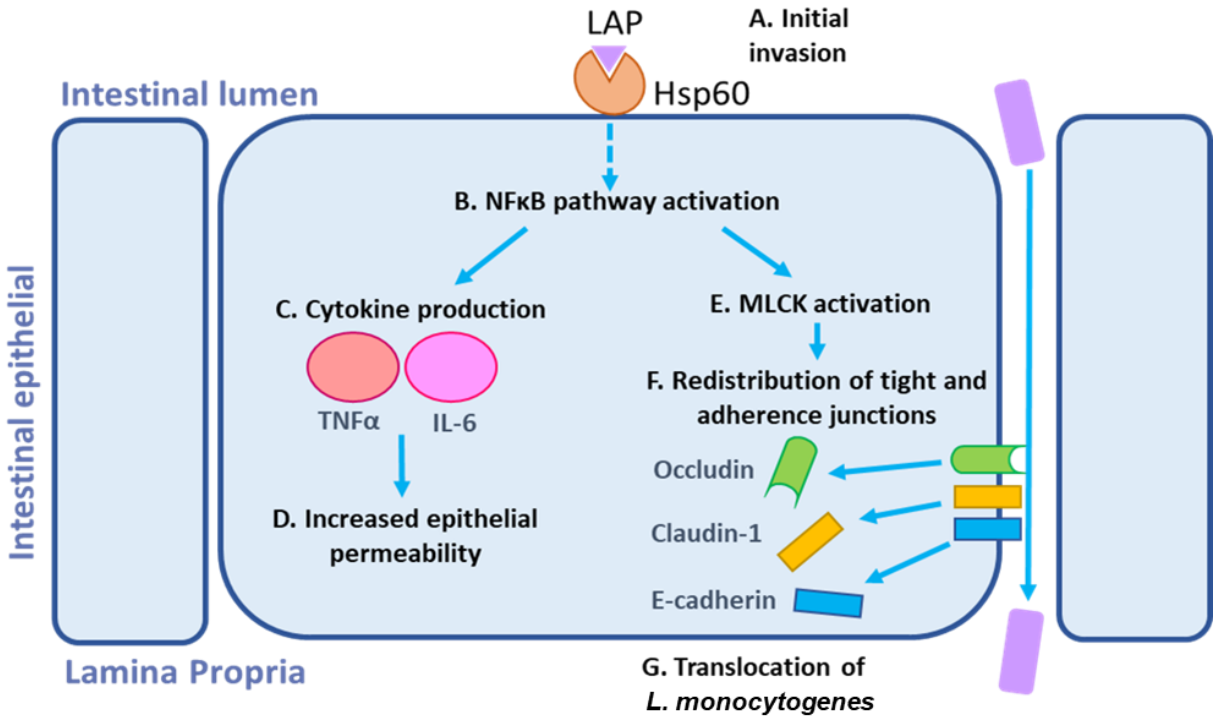

Figure 2.2 Crossing the intestinal barrier via the LAP pathway.

A. In the intestine, listerial LAP interacts with heat shock protein 60 (Hsp60), partly localized to the outer surface of the intestinal epithelial membrane. **B**. The interaction induces Hsp60 internalization activating the NFκB pathway to (**C**) ramp up cytokine production and (**E**) MLCK activation. **D**. The cytokines are proinflammatory and thus lead to increased epithelial permeability. **F**. MLCK activation causes the tight and adherence junctions to redistribute from the basolateral membrane adjoining neighbouring cells. **G**. The disruption of the tight and adherence junctions allows *Lm* to translocate between cells to the lamina propria.

2.4. Protein Chaperones

Chaperones are proteins that aid the folding, degradation and transport of other proteins *in vivo* (Ranson *et al.,* 1998). Although the amino acid sequence of any protein provides the blueprint for all secondary, tertiary and possibly quaternary structure, local minima on the folding path may create imperfect intermediates with incomplete partitioning of hydrophilic and hydrophobic residues to surface and core regions. Hydrophobic patches on the protein surface may lead to protein aggregation marking the protein for degradation. Chaperones

recognize proteins with hydrophobic surface patches and help to partially unfold such proteins to move them out of local minima to resume the path towards their native conformations. The *Escherichia coli* (*E. coli)* chaperones GroEL and GroES remain the most studied and best understood members of this family (Horwich *et al.*, 2007).

Mammalian chaperonins are named Heat shock proteins (Hsp) and were first discovered in 1962 (Ritossa, 1962). Heat shock proteins are involved in diverse processes such as secretion, protein degradation, transcription factor regulation and trafficking, but most notably as molecular chaperones. In this regard they maintain the integrity of proteins when exposed to environmental stress including pH and temperature fluxes as well as toxins (Ranford *et al.*, 2000). Hsps are generally named for their sizes such that Hsp110, Hsp90, Hsp70, Hsp60, Hsp40 have approximate masses of 110, 90, 70, 60 and 40 kDa while “small Hsps” are below 40 kDa in size (Johnson *et al.*, 1989).

Hsp60 is a 60 kDa oligomer organized into two stacked heptameric rings creating a central cavity (Cheng *et al.*, 1990). Both mitochondrial and cytoplasmic Hsp60 splice variants exist, where cytoplasmic Hsp60 has a degenerate 26-amino acid mitochondrial signal sequence. Cytoplasmic Hsp60 is involved in preventing apoptosis, and regulating immunity and cancer whereas mitochondrial Hsp60 is critical to mitochondrial protein folding and transport (Itoh *et al.*, 2002). In addition to cytoplasmic and mitochondrial localizations, Hsp60 is also found on the surface of epithelial cells. In this context it is considered a moonlighting protein as it may have additional functions from those listed above (Pockley and Henderson, 2017).

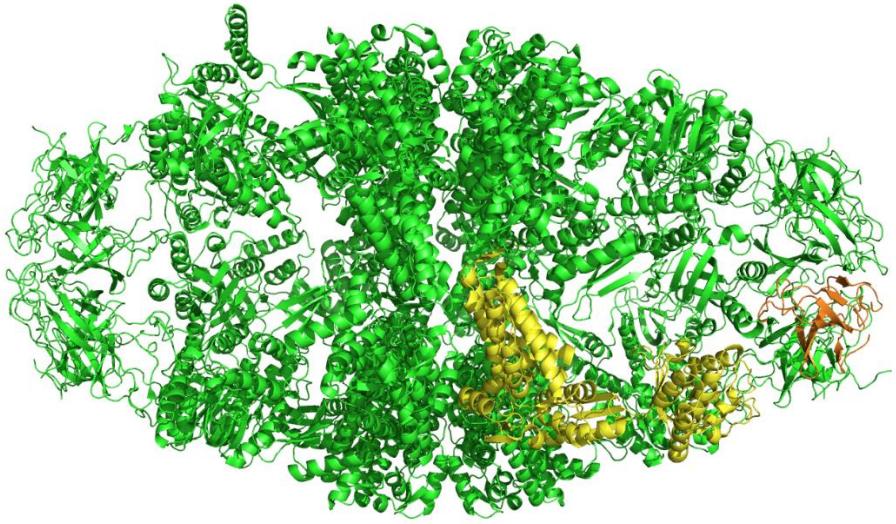

Figure 2.3 Structure of Human Heat Shock Protein 60.
The image displays the Hsp60 with the repeating monomers (orange and yellow). The two-fold symmetry can be seen from left to right. The image was generated in Pymol using the crystal structure of *Homo sapiens* Hsp60 (PDB: 4PJ1).

2.5. *Listeria* adhesion protein and Hsp60 interaction

Several attempts have been undertaken to determine the affinity of the LAP/Hsp60 interaction. In 2011, LAP subdomains were separately cloned, produced and purified. The first ALDH domain was separated into the constructs encoding Met1 to Pro223 and Gly224 to Gly411, while the second ADH domain was subdivided into Gly412 to Val648 and Pro649 to Val866. The subdomains as well as full-length LAP were added to Caco-2 cells with Hsp60 to allow for interaction. An immunofluorescent assay was used to determine a K_D of 7.2 ± 0.5 nM (Jagadeesan *et al.*, 2011).

A nanomolar interaction indicates a very tight interaction between LAP and Hsp60. With respect to LAP subdomains, Hsp60 was found to have the greatest affinity for Gly224 to Gly411 subdomain with a K_D of 9.5 ± 2.6 nM, indicating a slightly lower affinity than for the full-length protein interaction. The authors concluded that the Gly_{224}-Gly_{411} subdomain of LAP mediates the interaction between LAP and Hsp60 (Jagadeesan *et al.*, 2011). While the interaction of LAP with Hsp60 would be interesting and worthy of further study, these aspects form part of a project currently being undertaken by a PhD student. The project being reported on here exclusively concentrated on the protein LAP itself.

2.6. Development of multi-functional proteins

Many prokaryotic and eukaryotic proteins have been shown to adopt secondary roles in addition to their established main function. Especially, if the new function is linked to a distinct localization, the proteins are often said to be “moonlighting”. Multifunctional proteins were initially postulated to reflect limits to genome expansion though recently proteins evolving additional beneficial functions are more generally assumed to be evolutionarily advantageous (Huberts and van der Klei, 2010).

Bacterial proteins have variously been found to moonlight in molecular or cellular adhesion alongside their primary physiological roles. Due to their critical intracellular functions, the secretion and surface localisation of these moonlighting proteins is tightly regulated, sometimes involving stress stimuli (Yang *et al.*, 2011).

Acetaldehyde alcohol dehydrogenase (AdhE) has been found to have moonlighting properties in various bacteria. Thus AdhE in *Salmonella typhimurium* was observed to be upregulated during host invasion. Similarly AdhE in *Streptococcus pneumoniae* is a secreted virulence factor when ethanol concentrations are high (Luong *et al.*, 2015, Zetterström *et al.*, 2018). AdhE of *E. coli* contributes to colonization by regulating key virulence genes (Echave *et al.*, 2003). Similarly AdhE of *Lm* appears critical for listerial infection (see below).

2.7. Bifunctional Acetaldehyde-Alcohol Dehydrogenases

2.7.1. Structure

Acetaldehyde-alcohol dehydrogenases (AdhE) are bifunctional enzymes combining an N-terminal acetaldehyde dehydrogenase with an iron-dependent, C-terminal alcohol dehydrogenase domain. A critical NAD^+ cofactor binding site is shared by both domains (Jagadeesan *et al.*, 2010).

As early as the 1970s, bifunctional AdhEs were observed to form helical, macromolecular structures referred to as spirosomes composed of multiple AdhEs (Laurenceau *et al.*, 2015). In particular, the AdhE spirosome from *E. coli* was studied extensively revealing that the spirosome is critical for enzymatic activity of both domains. The spirosome switches between compact and extended conformations depending on cofactor availability. The ALDH active site is accessible from the outer surface of the spirosome and is hence not affected by a change in spirosome conformation. By contrast, the ADH active site is located within the spirosome such that it is largely inaccessible in the compact conformation of the spirosome with little to no activity. In the extended conformation, by contrast, the ADH active site is fully accessible such that catalytic activity can take place unhindered (Kim *et al.*, 2020, Pony *et al.*, 2020, Kim *et al.*, 2019).

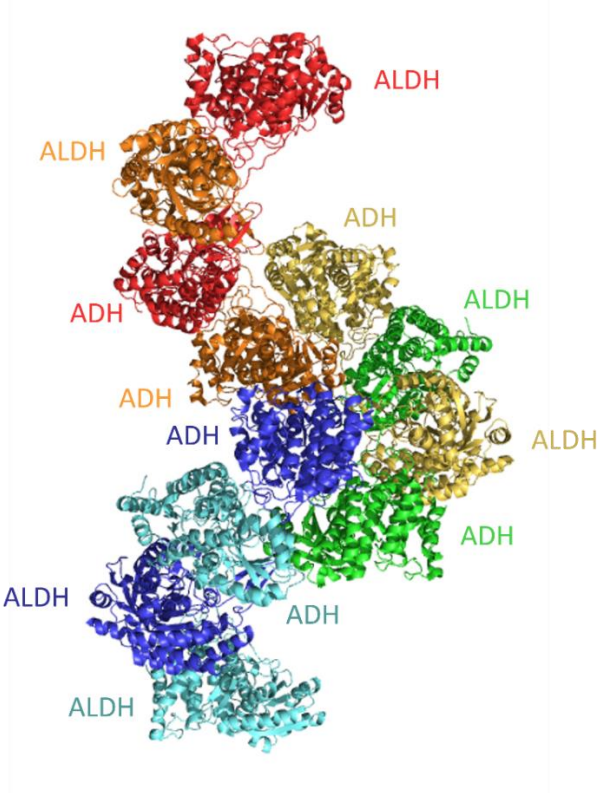

Figure 2.4 Spirosome structure of *E. coli* AdhE
Each AdhE monomer in the spirosome is assigned a unique colour. The image was generated in Pymol using the EM structure of *E. coli* AdhE (PDB: 7BVP).

2.7.2. Function

Acetaldehyde-alcohol dehydrogenases are housekeeping enzymes, catalysing reactions important for cell survival. Under anaerobic conditions bacteria use fermentation to produce adenosine triphosphate (ATP) for the cell. As ethanol is a frequent by-product of fermentation, the AdhE enzymes are of significant industrial importance (Peng *et al.*, 2008). AdhE catalyses two reactions in the formation of ethanol during fermentation: In a first step it reduces acetyl-CoA to acetaldehyde (1) and in a second step it reduces acetaldehyde to ethanol (2) (Tsuji *et al.*, 2016). Each domain achieves this reduction through a hydride transfer involving the simultaneous transfer of a proton and two electrons (Deno *et al.*, 1960).

$$CH_3COO\text{-}CoA + NADH + H^+ \rightleftharpoons CH_3CHO + NAD^+ + CoA \quad (1)$$

$$CH_3CHO + NADH + H^+ \rightleftharpoons CH_3CH_2OH + NAD^+ \quad (2)$$

The fermentation process also regenerates NAD^+ required for glycolysis, allowing the latter to generate high-energy ATP for bacterial metabolism (Pony *et al.*, 2020).

2.8. *Listeria* adhesion protein – an AdhE

Like other AdhE enzymes, the AdhE from *Listeria monocytogenes* is involved in ethanol synthesis. However, alongside this house-keeping function, it moonlights as an adhesion protein on the surface of the bacterial cell. In this secondary function, it was named *Listeria* adhesion protein or LAP. The latter function primarily appears to happen under stress conditions such as invasion of a higher eukaryote (Jagadeesan *et al.*, 2010).

In 1999, the 94 kDa protein LAP (Imo1634) was identified as a pathogenicity factor that aids in *Lm* adhesion to intestinal epithelial cells (Pandiripally *et al.*, 1999). Structurally and functionally, LAP or AdhE is a two-domain enzyme with an N-terminal acetaldehyde dehydrogenase (ALDH) domain and a C-terminal alcohol dehydrogenase (ADH) domain. It has both a single NAD^+ and a single Fe^{2+} binding site respectively involving amino acid residues between Gly_{427} and Gly_{432} in the N-terminal domain and Gly_{724} to Gly_{742} in the C-terminal domain (Figure 2.5) (Jagadeesan *et al.*, 2011).

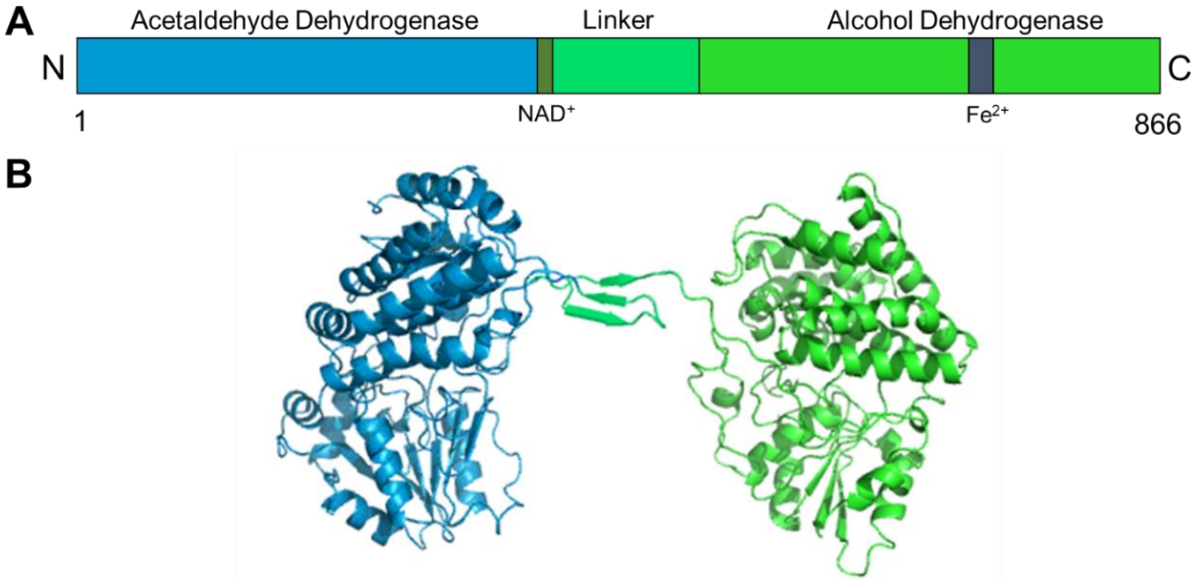

Figure 2.5 Structure of *Listeria* adhesion protein or AdhE
A. Domain structure of *Listeria* adhesion protein. B. A predicted three-dimensional structure of LAP derived by homology modelling based on *E. coli* AdhE using SwissProt. The two domains are connected by a linker structured as a small β-sheet.

LAP is found in both pathogenic and non-pathogenic *Listeria* spp sharing 98% nucleotide sequence identity. While LAP was initially postulated to be a general adhesion factor for all *Listeria* spp, this was not confirmed for non-pathogenic species (Jagadeesan *et al.*, 2010).

In *Lm,* LAP expression correlates inversely with the glucose concentration, such that it is upregulated in nutrient-limiting stress conditions but downregulated when glucose is available. LAP mostly remains in the cytosol with only small fraction being presented on the surface. During infection, secretion and surface presentation of LAP increase resulting in higher levels of cellular attachment (Jaradat and Bhunia, 2002). It is mainly secreted via the SecA2-dependent pathway despite LAP lacking an N-terminal secretion signal (Burkholder *et al.*, 2009, Jaradat *et al.*, 2003). The SecA2-dependent pathway has previously been implicated in the secretion of virulence factors in gram-positive bacteria and of house-keeping proteins that double as pathogenicity factors during infection in other bacteria such as *Streptococcus gordonii, Streptococcus parasanguinis, Mycobacterium smegmatis* and *Mycobacterium tuberculosis* (Rigel and Braunstein, 2008).

LAP has not been structurally characterised in detail. Due to its similarity to described AdhE structures, a structural model may be inferred by homology. However, it remains unclear how the atypical localization may affect the normal oligomerisation. Similarly, the mode of its interaction with Hsp60 remains unknown. This project aimed at characterising LAP to provide structural information and individual domain constructs to support the investigation of its interaction with human Hsp60.

2.9. Aim and Objectives

2.9.1. Aim

To characterise LAP (AdhE) both functionally and structurally.

2.9.2. Objectives

- Clone LAP construct into pGEX-6P-2 plasmid.
- Separate the individual domains of LAP at the genetic and protein level.
- Transform *E. coli* cells with the resulting plasmids.
- Produce and purify full-length LAP and LAP domains.
- Characterize full-length LAP and LAP domains in terms of solubility and stability.
- Determine enzyme kinetics of full-length LAP and LAP domains.
- Determine whether LAP forms spirosomes, like other AdhE proteins.
- Structurally characterise LAP and LAP domains by modelling or X-ray crystallography.

3. Materials and Methods

Chemical reagents were from SigmaAldrich (St. Louis, MI, USA), Merck (Darmstadt, Germany), ThermoFisher Scientific (Waltham, MA, USA) and Inqaba Biotec (Pretoria, South Africa).

A synthetic, codon-harmonized *lap* gene construct in a pUC57 plasmid was ordered from Gene Universal (Newark, DE, USA) that *inter alia* reduced the A-T richness of the *Lm* genome. Restriction enzyme (RE) sites were added to either side of the gene fragment, with BamHI and NotI sites at the 5' and the 3' ends respectively. The choice derived from the pGEX-6P-2 production vector into which the *lap* gene was to be transferred for the production of a GST-tagged fusion protein.

3.1. Cloning and DNA-based techniques

3.1.1. Competent cells

Pre-cultures of BL21 and DH10α *E. coli* cells were prepared from glycerol stocks by inoculating 5 mL lysogeny broth (LB) and incubating overnight (o/n) with shaking at 37°C. The pre-culture was transferred to a 100 mL of LB, incubated at 37°C with shaking at 200 rpm. OD_{600} readings were taken every 30 min until an OD of 0.4-0.6 was reached. The culture was cooled on ice and centrifuged (Heraeus megafuge 8R, ThermoFisher Scientific, Waltham, MA, USA) at 11 000 x *g* for 10 min at 4°C. The cell pellet was resuspended in ice-cold transformation buffer (10 mM Pipes pH 6.7, 55 mM $MgCl_2$, 15 mM $CaCl_2$, 250 mM KCl). The culture was incubated on ice for 15 min and centrifuged at 11 000 x *g* for 10 min at 4°C. The pellet was gently resuspended with ice-cold transformation buffer, then centrifuged at 4500 x *g* for 5 min at 4°C. The pellet was resuspended in 100 mM $CaCl_2$, 20% (v/v) glycerol. Cells were flash frozen in liquid nitrogen and stored at -80°C.

3.1.2. Transformation

The pUC57-*lap* plasmid (5µL) was added to 50 µL chemically competent *E. coli* DH10α cells and incubated on ice for 30 min. The cells were heat shocked at 42°C for 60 s and incubated on ice for 5 min. LB was added up to 1 mL without antibiotic and were incubated with shaking at 37°C for 1 h. After growth is observed using OD_{600}, 50 µL was plated on LB agar plates with 100 µg/mL Ampicillin (Amp) and incubated at 37°C o/n. Glycerol stocks were prepared from the freshly transformed cells. The transformed cells were added to a final concentration of 15% (v/v) glycerol and stored at -80°C.

3.1.3. Restriction enzyme digest

The *lap* gene as received in the synthetic pUC57 construct was flanked by BamHI and NotI restriction sites allowing the *lap* gene to be transferred to the pGEX-6P-2 plasmid with matching RE cut sites. The pUC57-*lap* plasmid was digested with BamHI, NotI and ScaI (ThermoFisher Scientific Waltham, MA, USA) to release the *lap* gene fragment from pUC57 and digest the latter into smaller fragments.

Table 3.1 Volumes of reagents for the double digest of pUC57-lap.

Reagent	Volume (µL)
10 × 3.1 buffer (NEB, MA, USA)	4
DNA (113 ng/µL)	10
BamHI (NEB, MA, USA)	1
Not1 (NEB, MA, USA)	1
$dddH_2O$	24
Final volume	40

The tubes containing the reaction mixture were incubated for 1 h at 37°C and then 65°C for 20 min to inactivate the restriction enzymes. The contents of the tubes after the reaction were analysed using a 1% (w/v) agarose gel with 1 x SYBR Safe (ThermoFisher Scientific MA, USA). Purple loading dye (New England Biolabs, MA, USA) was added and the 50 µL reaction mixture loaded into two wells. A 1 kb DNA ladder (New England Biolabs, MA, USA) was loaded into a separate lane. The samples were separated by electrophoresis at 90 V for 1 h and the resulting gel visualized in a Gel Doc XR+ System (Bio-Rad, CA, USA).

3.1.4. Agarose Gel Extraction

The agarose gel was inspected under a UV transilluminator (SigmaAldrich, MI, USA) and the band of interest was cut from the gel with a sterile scalpel. The excised agarose block was weighed in a pre-weighed tube such that 200 µL binding buffer could be added for every 100 mg of agarose (GeneJET Gel Extraction Kit, ThermoFisher Scientific MA, USA). Tubes were incubated at 50°C for 10 min or until the agarose gel had completely melted. The tube contents were decanted into GeneJET purification columns and the columns centrifuged at 10 000 x *g* for 2 min. The flow through discarded, 500 µL wash buffer added and columns centrifuged at 10 000 x *g* for 1 min. This procedure was repeated twice. After centrifuging the columns at 10 000 x *g* for 2 min to remove excess wash buffer, the column was transferred to a sterile tube, elution buffer added to the column, incubated for 2 min, then

centrifuged at 10 000 x *g* for 2 min. The flow through was collected and its concentration determined using ND-100 spectrophotometer (PEQLAB Biotechnologie, Germany) at an absorbance of 260 nm.

3.1.5. Ligation

The DNA extracted from the agarose gel was to be ligated into the plasmid pGEX-6P-2 to create pGEX-6P-2-*lap*. To achieve this, the purified linear *lap* gene with sticky ends was added to the purified linear pGEX-6P-2 with complimentary sticky ends in 3:1 and 7:1 ratios. T4 DNA ligase was added in T4 DNA ligase buffer (New England Biolabs, MA, USA) and dddH_2O added to a total volume of 20 µL. The tube containing the reaction was incubated at 4°C o/n. The contents of the reaction tube were used to transform chemically competent *E. coli* DH10α cells.

3.1.6. Plasmid isolation

The plasmids (see supplementary below) were extracted from cell cultures using the GeneJET plasmid miniprep kit (ThermoFisher Scientific, MA, USA) for sequencing and double digest – both to confirm that the inserts had correctly been incorporated above. Transformed colonies were picked from the agar plates, added to 10 mL LB and incubated overnight with shaking at 37°C. The cultures were centrifuged at 5 000 x *g* for 2 min and the supernatant discarded. The remaining cell pellets were resuspended in resuspension buffer. First, 200 µL lysis buffer were added to the 2 mL tubes (inverted several times) before 200 µL neutralization buffer was added and the tubes inverted several times. Tubes were centrifuged at 14 000 x *g* for 5 min. The supernatant solutions were decanted into GeneJET purification columns and centrifuged at 14 000 x *g* for 1 min. The flow through was discarded. Five hundred microlitres wash solution was added to the GeneJET purification columns and centrifuged at 14 000 x *g* for 1 min. This was repeated. The columns were placed into new microtubes and elution buffer was added to the columns. They were incubated at room temperature for 2 min and then centrifuged at 14 000 x *g* for 2 min. The plasmid concentrations were determined by absorbance at 260 nm using an ND-100 spectrophotometer (PEQLAB Biotechnologie, Germany). The plasmids were analysed by agarose gel electrophoresis. Purple loading dye (New England Biolabs, MA, USA) was added to samples and loaded into the wells of the agarose gel. The samples were separated by electrophoresis at a voltage of 90 V for 1 h.

3.1.7. Sanger sequencing reaction

Plasmids were isolated from transformed *E. coli* DH10α cells and sequenced to confirm that the various gene constructs had been correctly ligated into the pGEX-6P-2 plasmid. pGEX 5' and 3' primers specific to the pGEX-6P-2 plasmid were used. The following components were added to each reaction mix: 2 µL Big dye (ThermoFisher Scientific, MA, USA), 1.5 µL 5 × sequencing buffer, 1 µL primer at 3.2 µM, 1 µL plasmid at 150 ng/µL and $ddd H_2O$ to a total 10 µL. A cycle sequencing reaction was initiated as listed in Table 3.2.

Table 3.2 Sanger sequencing conditions with T100 Thermocycle (Bio-Rad, CA, USA)

Temperature (°C)	Time (s)	
96	60	
96	10	Repeat for 25 cycles
50	5	
60	240	
4	∞	

The cycle sequencing reaction was cleaned to remove excess reagents such as dNTPs. Mix A and B solutions were prepared where mix A contained 3 µL 3 M NaOAc pH 4.6, 62.5 µL 99% (v/v) EtOH and 14.5 µL $ddd H_2O$ while mix B contained 10 µL DNA sequencing reaction and an equal volume of $ddd H_2O$. The solutions were mixed in a ratio of 1:4 mix B:mix A. The mixture was centrifuged for 30 min at 10 000 x *g* and the supernatant carefully discarded. Ice-cold 70% (v/v) EtOH was added and centrifuged at 10 000 x *g* for 5 min. The supernatant was discarded and the remaining alcohol allowed to evaporate for 20 min. The samples were sent to the DNA Sanger Sequencing Facility (University of Pretoria, South Africa) for Sanger sequencing.

After the pGEX-6P-2-*lap* construct was confirmed primers were designed to integrate mutations into the gene to separate the two enzyme domains.

3.1.8. Site-directed mutagenesis

The 5' end of *lap* gene encodes the ALDH domain. To produce the ALDH domain alone the codon encoding arginine 454 was converted from CGC to TGA introducing a stop codon. Primers were designed to incorporate this stop codon into the *lap* gene. The pGEX-6P-2-*lap* plasmid was isolated and purified (section 3.1.6) for the site-directed mutagenesis. The following reagents were added to a 200 µL microtube.

Table 3.3 Reagents for site-directed mutagenesis

Reagent	Final concentration
Phusion DNA Polymerase (ThermoFisher Scientific, MA, USA)	1 unit
5 × Phusion HF buffer (ThermoFisher Scientific, MA, USA)	1 ×
DNA	<200 ng/µL
Forward Primer	0.5 µM
Reverse Primer	0.5 µM

The microtubes were placed in the T100 thermocycler under the following conditions.

Table 3.4 Site-directed mutagenesis conditions

Temperature (°C)	Time (s)	
98	30	
98	10	Repeated for 30 cycles
68	30	
72	225	
72	150	
4	∞	

After completion the samples were analysed using agarose gel electrophoresis (1% (w/v) agarose with 1 × SYBR safe). A visible band indicated successful amplification and *E. coli* DH10α cells were transformed by introducing the new plasmids. These propagated plasmids were isolated and sequenced (section 3.1.7) to confirm the mutation.

3.1.9. Cloning

To produce the ADH domain alone, the gene fragment corresponding to this domain was amplified and cloned into an empty pGEX-6P-2 plasmid. The primers were designed to specifically amplify the *adh* region of the gene. The pGEX-6P-2-l*ap* plasmid was isolated, as described (section 3.1.6) for cloning of the *adh* gene. The gene fragment was amplified by polymerase chain reaction (PCR) using the reagents in Table 3.5 and the conditions in Table 3.6

Table 3.5 Reagents for PCR

Reagent	Final Concentration
Taq DNA Polymerase (Kapa Biosystems, South Africa)	0.5 U
10 × Taq Buffer (Kapa Biosystems, South Africa)	1 ×
DNA	<200 ng/µL
Forward Primer	0.4 µM
Reverse Primer	0.4 µM
DNTP mix	1 mM
$MgCl_2$	1.5 mM

Table 3.6 Conditions with T100 thermocycle (Bio-Rad, CA, USA)

Temperature (°C)	Time (s)	
95	180	
95	30	Repeated for 35 cycles
65	30	
72	120	
72	240	
4	∞	

The amplification reaction was analysed using a 1% (w/v) agarose gel. The band of interest was isolated from the gel using the gel extraction method. The extracted and purified DNA was added to previously double digested pGEX-6P-2 plasmids and ligated (section 3.1.5). The ligated plasmids were used to transform *E. coli* DH10α cells for plasmid propagation. The propagated plasmids were sequenced to ensure the correct gene fragment had been transferred to the plasmid.

3.2. Protein-based techniques

3.2.1. Induction and expression

Pre-cultures of the transformed bacterial cells with 50 µg/mL amp were incubated o/n with shaking at 37°C. The pre-cultures were used to inoculate 1 L LB with 100 mg/mL Amp. The cultures were incubated at 37°C with shaking at 200 rpm until an OD_{600} of 0.6 to 0.8 was reached. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.1 mM to induce induce production of the target protein. The cultures were incubated o/n at 28°C with shaking at 200 rpm. The next day the cultures were centrifuged (Sorvall Lynx 6000, ThermoFisher Scientific, MA, USA) at 11 000 x *g* for 15 min. Supernatants were

discarded and the pellets resuspended in protein buffer: 100mM Tris pH 7.8, 100 mM NaCl, 1 mM DTT, 5% (v/v) glycerol. The cells were sonicated using a Qsonica Sonicator Q500 (Branson, CT, USA) at 40% for 7 min 30 s, with alternating 30 s pulses and 30 s pauses on ice. Lysed cells were centrifuged at 37 500 x *g* for 1 h. The resulting supernatants were decanted into a Falcon tube and stored at 4˚C. A representative pellet sample was solubilized using 4 M urea to check that target proteins were not produced insolubly.

3.2.2. Production optimization

The production conditions of LAP were varied to deduce the optimal temperature for induction as well as IPTG concentration. The selected temperatures were 20, 25 and 30°C. The IPTG concentrations selected were 0.1, 0.5 and 1 mM. Pre-cultures of the transformed bacterial cells were grown o/n at 37°C shaking at 200 rpm. The pre-cultures were added in a 1:10 ratio into new LB and grown until an OD_{600} of 0.4 to 0.6. The cultures were then induced with IPTG at the varying concentrations. Each IPTG concentration was grown at 20, 25 and 30°C with o/n shaking at 200 rpm. The cells were pelleted by centrifugation at 11 000 x *g* for 15 min and the supernatant discarded. The cell pellets were resuspended in protein buffer (100 mM Tris pH 7.8, 100 mM NaCl, 1 mM DTT, 5% (v/v) glycerol) for lysis. The cells were disrupted by sonication and centrifuged at 37 000 x *g* for 1 h to separate soluble and insoluble fractions. The latter were analysed by polyacrylamide agarose gel electrophoresis (PAGE) on a 10% sodium dodecyl sulphate-(SDS) gel separated by a current of 40 mA applied for 45 min.

3.2.3. SDS-PAGE analysis

Throughout the production and purification processes 10 µL samples were collected for later analysis by SDS-PAGE (Mini-PROTEAN System, Bio-Rad, CA, USA). Polyacrylamide gels with 10 and 12% (w/v) acrylamide were prepared for LAP and for ALDH and ADH, respectively. The higher concentration of acrylamide in gels for ALDH and ADH was chosen to accommodate their smaller size as higher acrylamide concentrations will retard proteins in moving through the gel. The gels have two sections, a stacking and a separating region. The stacking region has a lower acrylamide concentration (~3.7%) compared to the separating gel, resulting in larger pores such that proteins are concentrated ahead of the separating gel where they are separated by size.

Table 3.7 Reagents for 10% SDS-PAGE gel

Reagents	Stacking gel (mL)	Separating gel (mL)
1 M Tris pH 6.8	0.625	-
1.5 M Tris pH 8.8	-	1.50
40% (w/v) acrylamide	0.625	1.50
10% (w/v) SDS	-	0.060
TEMED	0.015	0.002
10% (w/v) APS	0.025	0.060
Final volume	5.00	6.00

The same reagents in table 3.7 were used for the 12% separating mix except that a 1.5 M Tris pH 8.8 buffer was used and 12% (w/v) SDS added. The separating solution was added to ¾ of the gel cassettes and covered with isopropanol to ensure the top of the separating solution is level. The solution was allowed to polymerize for 30 min. The isopropanol was removed from the gel cassettes and the stacking solution added. A comb was inserted and the stacking solution allowed to polymerize for 15 min. Upon removal of the comb the gels were ready for analysis. The gels were added to a gel tank containing running buffer (20 mM Tris pH 8.3, 3 mM SDS, 200 mM glycine). Protein samples for analysis were prepared by adding sample buffer (20 mM Tris pH 6.8, 2% (w/v) SDS, 20% (v/v) glycerol, 2 mM EDTA, 160 mM DTT and 0.1 mg/ml bromphenol blue dye) and heating the samples at 95°C for 3 min. The proteins were separated by SDS-PAGE at a current of 40 mA for 45 min per gel. The acrylamide gels were stained with Coomassie protein stain (2 mM Coomassie Blue, 60% (v/v) ethanol, 12% (v/v) acetic acid) for 30 min at room temperature, destained with protein gel destain (42% (v/v) ethanol and 10% (v/v) acetic acid), and analysed using a Gel Doc XR+ System and ImageLab software (Bio-Rad, Hercules, CA, USA).

3.2.4. 3C protease production

LB with 50 mg/mL amp was inoculated with *E. coli* BL21 cells carrying the plasmid pGEX-6P-2-*3c* and incubated o/n at 37˚C with shaking. 1 L LB was inoculated with the preculture and left to grow until an OD_{600} of 0.4 to 0.6. Target gene expression was initiated with 0.1 mM IPTG. The cultures were incubated o/n at 28°C. Cells were harvested at 17 000 x *g* for 15 min. The supernatant was discarded and the pellet was resuspended with lysis buffer (50 mM Tris pH 8.0, 300 mM NaCl, and 10% (v/v) glycerol). The cells were lysed by sonication at 40% for 10 min, 30 s on and 30 s off. Soluble and insoluble fractions were

separated by centrifugation at 35 000 x *g* for 1 h. The soluble fraction was added to glutathione agarose beads in a Falcon tube. The tubes were agitated on a roller mixer o/n at 4°C, whereupon their contents were decanted into drip columns. The glutathione agarose beads were washed extensively with lysis buffer. Elution buffer (50 mM Tris pH 8.0, 30 mM NaCl, 20 mM reduced glutathione, 10% (v/v) glycerol) was added and the columns were agitated on the roller mixer for 3 h at 4˚C. The elution was collected and concentration determined using ND-100 spectrophotometer at 280 nm. The elution buffer was replaced by the storage buffer (50 mM Tris pH 7.4, 150 mM NaCl, 10 mM EDTA, 20% (v/v) glycerol) by dialysis (see buffer exchange below). The 3C protease was stored at -20˚C.

3.2.5. Affinity chromatography

If the target protein is soluble, it may be purified by affinity chromatography. 2 mL glutathione agarose beads were added to the soluble fraction. The corresponding Falcon tube was agitated on a roller mixer o/n at 4°C. The contents were decanted into a drip column and the eluant collected. The glutathione agarose beads were washed extensively with protein buffer: 100 mM Tris pH 7.8, 100 mM NaCl, 1 mM DTT, 5% (v/v) glycerol. 20 µg 3C protease was added to the column and incubated on a roller mixer o/n at 4°C. The target protein was eluted with protein buffer and stored at 4°C.

3.2.6. Buffer exchange

The 3C protease elution buffer contains reduced glutathione which needs to be removed before use. Reduced glutathione interrupts the bond of the GST tag to the glutathione column. If the reduced glutathione is not removed and the 3C protease is added to the affinity column the bound fusion protein will elute from the stationery beads. Buffers were exchanged by dialysis. Two litres of the required buffer was prepared. The buffer was added to a 3 L beaker. 10 kDa molecular weight cut-off snakeskin dialysis tubing (ThermoFisher Scientific, MA, USA) was used for the dialysis. The snakeskin was secured on the one end and the complete volume of the protein was added. The open end was secured and floatation apparatus attached. The protein containing pouch was placed in the buffer and incubated o/n at 4°C stirring. Due to the small protein volume, a single round of dialysis in a large volume sufficed. The protein was removed the next day and concentrated. The contents of a buffer can affect crystallisation and therefore the buffer may need to be changed before crystallisation. The same process was performed in preparation for crystallisation where the proteins were exchanged from protein buffer (100 mM Tris pH 7.8, 100 mM NaCl, 1 mM DTT, 5% (v/v) glycerol) to dddH_2O.

3.2.7. Size-exclusion chromatography

Most proteins are highly enriched by affinity chromatography. However, to remove any contaminants, exchange the buffer or analyse the size distribution of the protein, the sample would be further purified by size-exclusion chromatography on an Äkta Explorer FPLC System (Cytiva, MA, USA). A calibrated Superdex 200 column (Cytiva, MA, USA) was equilibrated with three column volumes of protein buffer (100 mM Tris pH 7.8, 100 mM NaCl, 1 mM DTT, 5% (v/v) glycerol). Once the protein sample was loaded onto the column, 30 mL of protein buffer was run through the column and the elutions were collected in 2 mL fractions. The chromatogram was recorded and relevant peaks analysed by SDS-PAGE to assess protein purity and possible oligomerisation state.

3.2.8. Anion exchange chromatography

Anion exchange chromatography was executed with an Äkta Explorer (Cytiva, MA, USA). A HiTrap 5 mL Q HP column was used with a pressure limit of 0.3 MPa. The column was equilibrated with three column volumes of low salt buffer (20 mM Tris pH 7.8, 20 mM NaCl). The protein sample was loaded onto the column and the column washed with three column volumes (15 mL) of low salt buffer. Flow through fractions were collected to potentially recover unbound protein. The protein was eluted by linearly increasing the proportion of a high salt buffer (20 mM Tris pH 7.8, 1 M NaCl) from 0 to 100%. The eluted solutions were collected, the chromatogram recorded and the relevant peaks analysed by SDS-PAGE to identify the target protein peak and assess its purity.

3.2.9. Enzyme assays

Acetaldehyde dehydrogenase and alcohol dehydrogenase both use oxidised nicotinamide adenine dinucleotide NAD^+. The reduced NADH product has an absorbance peak at 340 nm, allowing the reactions to be monitored. The conditions to analyse each domain separately were adapted from Jagadeesan *et al* (Jagadeesan *et al.*, 2010). The concentration of the purified LAP was determined using ND-100 spectrophotometer (PEQLAB Biotechnologie, Germany). For ALDH, 200 µL of 80 mM Tris pH 8, 4 mM NAD^+, 0.3 mM dithiothreitol, 0.1 mM CoASH, 5 µM purified LAP were prepared in a 300 µL well. Acetaldehyde concentration varied from 0 to 100 mM in regular increments. For ADH, 200 µL of 80 mM Tris pH 8, 4 mM NAD^+, 5 µM purified LAP were prepared. Ethanol concentrations ranged from 0 to 100 mM. Reactions were monitored at 25˚C for 2 h by

recording the OD_{340} at 15 s intervals on a SpectraMax Paradigm Multi-Mode Microplate Reader (Molecular Devices, CA, USA).

3.3. Structural and biophysical techniques

3.3.1. Screening for crystallization conditions

Protein crystallisation is often challenging as conditions need to be found to produce single, diffraction-quality crystals. To rapidly screen a wide range of conditions, commercial crystallization screens PEGRX1 and PEGRX2 (Hampton Research, CA, USA) were used in 96 well sitting-drop, vapour diffusion plates. LAP was diluted to a concentration of 1 mg/mL using ddH_2O. From each screen condition 1 mL was added to a corresponding reservoir well and 1 µL of the reservoir solution was added to the ledge with 1 µL of LAP added and mixed. Once all 96 wells were complete the plate was sealed to ensure it is airtight. The plate was stored at 18°C and checked regularly using a CL 1500 HAL microscope (Zeiss, Germany). The above was repeated with a higher concentration of 2.5 mg/mL LAP.

3.3.2. Optimization of crystallisation

To improve the quality and size of initial microcrystals, the starting conditions identified from the crystal screen were optimized in hanging-drop, vapour-diffusion crystallisation experiments. Buffers, salt and/or precipitants were added to the reservoir well. 2 µL of the reservoir solution was added to an equal volume of the protein on a glass coverslip and mixed. The coverslip was inverted and placed on top of the well, the rim of which had previously been covered by silicone jelly. As the water concentration in the crystallization drop is higher than in the well, water will diffuse via the vapour form from the drop to the reservoir solution, supersaturating the protein and ideally initiating crystallization. Controlling the process by lowering the temperature allows for slow and ordered crystal formation. A concentration of 2.5 mg/mL LAP was used.

3.3.3. Crystal seeding

Protein crystallization can often be simplified by adding crystal nuclei to the crystallization droplet to eliminate the need for protein supersaturation and associated fast crystal growth. The best nuclei derive from previous rounds of crystallization – though other protein crystals and nucleants have been described. For LAP microcrystals from the initial lead identification and later rounds of crystallization were used. Appropriate crystallization drops were removed from the experiment and added to 35 µL of their respective reservoir solutions to produce a

stock solution. The solution was vortexed and a tenfold dilution series was prepared over seven steps. The procedure described in 3.3.2 was amended by the addition of 1 µL of crystal nuclei of varying concentrations to the crystallization drop. The plate was incubated at 18°C and periodically inspected on a CL 1500 HAL microscope (Zeiss, Germany).

3.3.4. Dynamic Light Scattering

Dynamic light scattering is a technique that allows the distribution of particle sizes in a solution to be analysed. This method is particularly relevant to proteins as proteins may associate to form dimers, trimers, defined higher oligomers or less defined aggregates. Solutions containing a single protein oligomerisation state are described as monodisperse and have the highest probability of crystallizing. Dynamic light scattering is a simple technique to investigate the dispersity of proteins or other solutions. Light scattered at 90° to the incident light is recorded and fluctuations analysed in terms of particle size.

For LAP, purified protein was diluted to 1 mg/mL and transferred to a disposable cuvette. The particle size distribution was investigated using a Zetasizer Nano ZS (Malvern, UK).

3.3.5. Negative staining electron microscopy

Negative staining electron microscopy can be used to visualise the morphology of subcellular components as well as macromolecules such as proteins. The technique uses a sample support, such as a grid, which is electron transparent. The protein of interest is adsorbed to the grid to ensure it is not lost during washing. The washing procedure removes contaminants such as salts which could interfere with the staining procedure. The heavy metal staining provides a contrast to view the structural aspects of the protein of interest using a transmission electron microscope.

Purified LAP, prepared as mentioned above, with a concentration of 0.01 mg/mL was used for the electron microscopy experiments. The prepared sample was delivered to the Electron Microscopy facility at the University of Cape Town where the following was completed by Dr Jeremy Woodward - grid preparation, imaging and processing of the images.

4. Results

4.1. Cloning of LAP

4.1.1. Restriction digest of pUC57-*lap*

The *lap* gene was synthesised by Gene Universal (Newark, DE, USA) and inserted in a pUC57 plasmid. As pUC57 is not an expression vector, the *lap* gene had to be transferred to an appropriate vector for protein production. As related AdhE structures indicate that the N-terminus is exterior to the protein, a fusion protein with an N-terminal GST-tag would be possible. Such a GST tag would allow for rapid protein purification by affinity chromatography and is known to improve protein solubility. The vector pGEX-6P-2 was chosen as it is designed to produce fusion proteins bearing an N-terminal GST-tag. The multicloning site of pGEX-6P-2 starts and ends with BamHI and NotI cleavage sites respectively, with the BamHI site immediately downstream of a C3 protease cleavage site encoding region 3' of the GST-encoding gene. This arrangement avoids introducing non-physiological amino acids between the GST-tag and LAP and hence to LAP after cleavage.

As the *lap* gene was designed to have BamHI and NotI restriction sites, the pUC57-*lap* plasmid was digested with BamHI and NotI to cut the *lap* gene from the pUC57 plasmid.

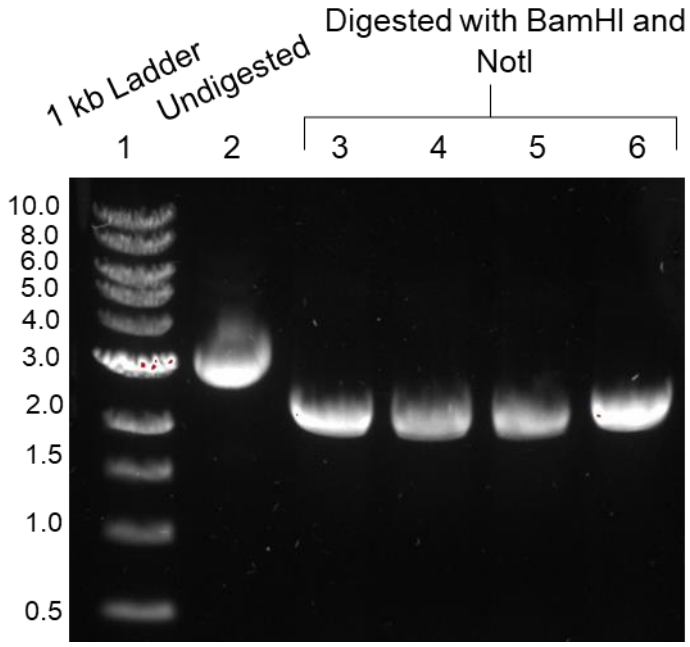

Figure 4.1 Agarose gel analysis of the double digestion of the pUC57-*lap* plasmid.
Lane 1: 1 kb DNA marker. Lane 2: Undigested pUC57-lap. Lanes 3-6: Products of the double digestion of pUC57-lap with BamHI and NotI. 1% (w/v) agarose in 1xTAE buffer with 1x SYBR safe.

The undigested and digested plasmids (Figure 4.1 lanes 2 to 6) have similar band intensities. The undigested plasmid in lane 2 is presumably exclusively supercoiled causing

it to migrate through the gel more rapidly than a more relaxed plasmid would. It thus appears to have a size of around 3 kb rather than the theoretical size of 5.3 kb. Lanes 3 to 4 (Figure 4.1) each seem to reveal single bands at around 2.5 kb. This may initially be confusing until considering that the pUC57 plasmid backbone has a size of 2.7 kb while the *lap* gene has a size of 2.6 kb. Clearly thus the restriction digest has been successful, resulting in two overlapping bands on the 1% (w/v) agarose gel.

4.1.2. Secondary restriction of pUC57

As both the plasmid and the *lap* gene fragments have identical overhanging ends, ligating them with a pGEX-6P-2 plasmid, would result in 50% incorrect plasmids. To prevent his from happening, we decided to separate the plasmid and the *lap* gene through an additional, in this case ScaI restriction site located near the middle of pUC57 plasmid. This would result in one longer *lap* gene fragment and two shorter pUC57 fragments. The DNA in, lanes 3 to 6 of Figure 4.1 was thus extracted and further cleaved with ScaI.

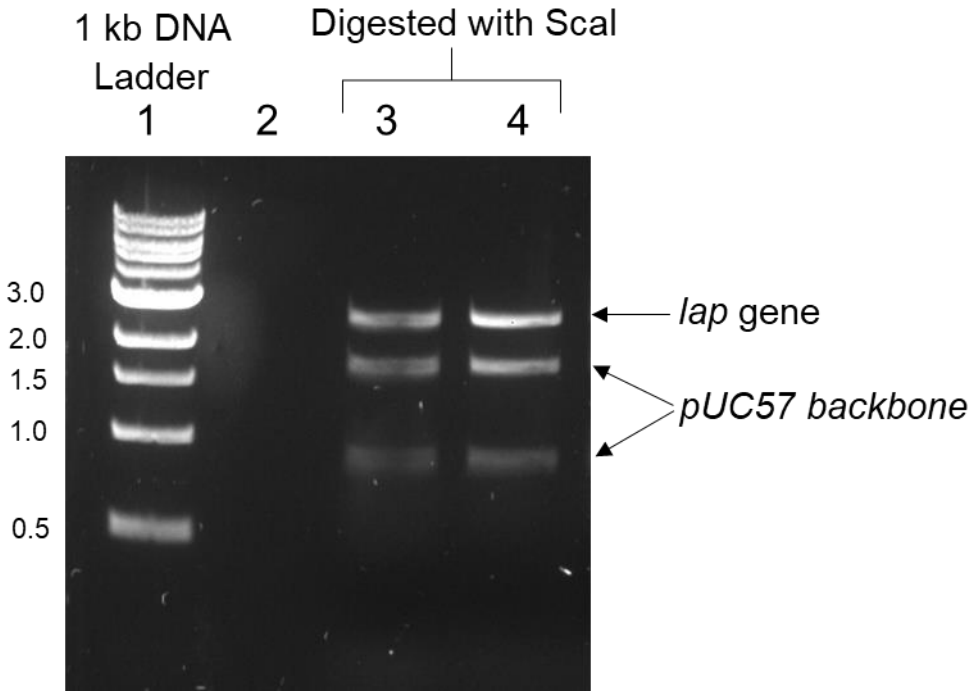

Figure 4.2 Agarose gel analysis showing the digestion of the extracted DNA.
Lane 1: 1 kb marker. Lane 2: Empty. Lanes 3-4: Products of the digestion of the previously double digested DNA with ScaI. 1% (w/v) agarose in 1xTAE buffer with 1x SYBR safe.

As expected, three bands were visible in lanes 3 and 4 in Figure 4.2 bearing the DNA fragments after BamHI, NotI and ScaI restriction digest. The largest band of ~2.5 kb matches the theoretical size of the *lap* gene of 2.6 kb. The two lower bands ~1.7 and 0.8 kb correspond to the size of the cleaved pUC57 backbone. The secondary digestion to separate the *lap* gene from the pUC57 backbone was thus successful.

4.1.3. Cloning of *lap* into pGEX-6P-2

The isolated *lap* gene was extracted from the agarose gel and purified. pGEX-6P-2 plasmids were isolated and prepared by digesting with BamHI and NotI. The plasmid backbone and insert with matching overhanging ends were mixed and ligated. The ligation product was used to transform *E. coli* DH10α cells. The transformed cells were plated and colonies were selected, propagated and plasmids isolated. The isolated plasmids were again digested with BamHI and NotI, then analysed with a 1% (w/v) agarose gel.

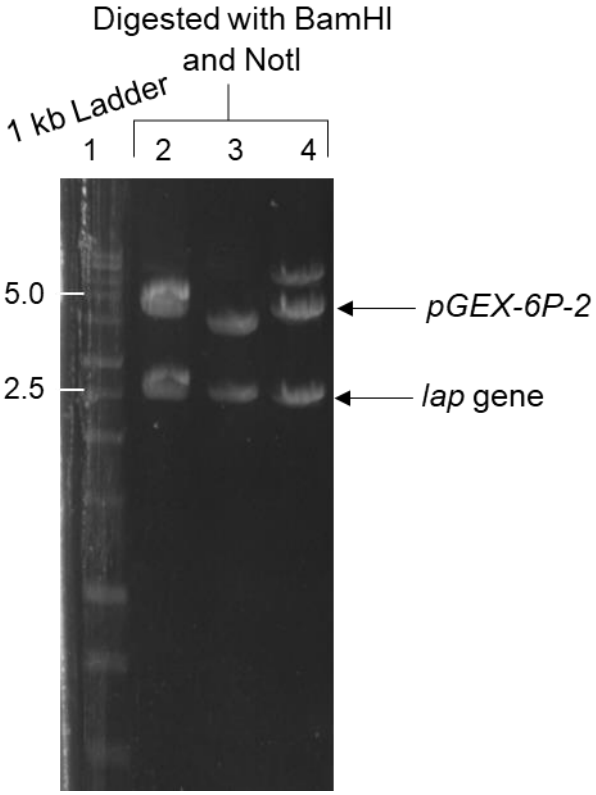

Figure 4.3 Agarose gel analysis for the digestion of the pGEX-6P-2-*lap* plasmid.
Lane 1: 1 kb marker. Lane 2-4: Products of the double digestion of pGEX-6P-2-lap with BamHI and NotI. 1% (w/v) agarose in 1xTAE buffer with 1x SYBR safe.

In Figure 4.3 two bands are present after the digest, one band at ~5 kb and one band ~2.5 kb. These sizes correspond to the size of the empty pGEX-6P-2 plasmid and the *lap* gene respectively indicating that the *lap* gene has been successfully ligated into the pGEX-6P-2 plasmid.

4.1.4. Sanger sequencing of pGEX-6P-2-*lap*

The plasmids were propagated and isolated for sequencing to further confirm the presence of the LAP gene. The forward and reverse primers, pGEX 5' and pGEX 3' respectively, were used to sequence the insert of the pGEX-6P-2-*lap* vector.

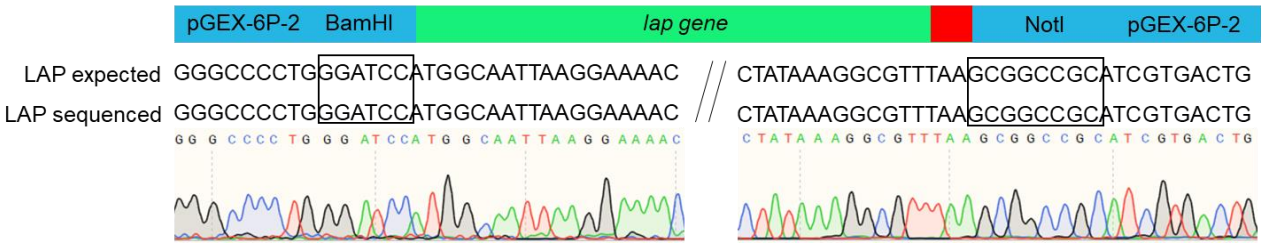

Figure 4.4 Sequence data for pGEX-6P-2-*lap*.
The gene sequence including the corresponding chromatograms. The red block indicates the stop codon.

The DNA sequence obtained from Sanger sequencing matches the expected sequence confirming that the *lap* gene was successfully ligated into pGEX-6P-2, in the correct orientation and in frame.

4.1.5. Converting pGEX-6P2-*lap* to pGEX-6P2-*aldh*

The pGEX-6P-2-*lap* construct was sequenced and was successful, it was then used to create constructs for the two functionally distinct domains. For the N-terminal ALDH domain of LAP, an arginine codon downstream of the coding sequence was converted to a stop codon eliminating the C-terminal domain. Mutagenesis and sequencing primers were designed for this purpose (supplementary data Table 8.2).

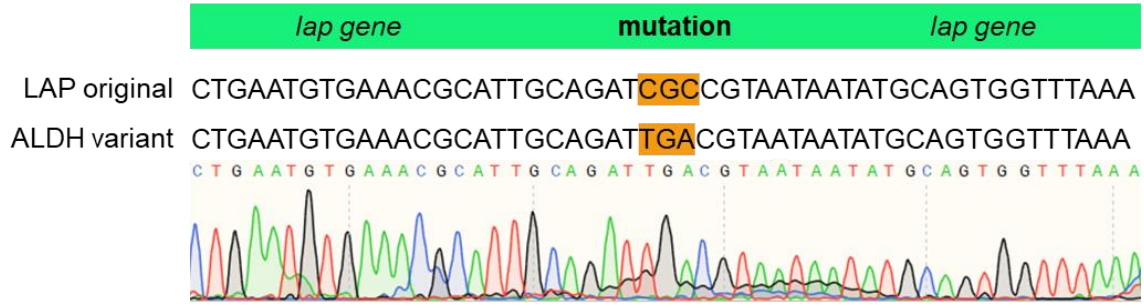

Figure 4.5 Sequence data for the pGEX-6P-2-*aldh* plasmid.
The gene sequence including the corresponding chromatograms. The orange highlight indicates the stop codon mutation. The sequence in colour is the experimentally determined sequence for the chromatogram below.

The theoretical DNA sequences before and after mutagenesis as well as the experimental sequence after mutation are aligned in figure 4.5. As the experimental sequence carries the stop codon, this confirms that the site directed mutagenesis experiment was successful. The construct was therefore used to transform *E. coli* BL21 cells.

4.1.6. Cloning of gene fragment *adh* into pGEX-6P-2

To produce a GST-ADH fusion protein of the C-terminal ADH domain of LAP, the DNA fragment encoding the ADH domain was amplified from the pGEX-6P-2-*lap* plasmid by PCR and ligated into an empty pGEX-6P-2 plasmid.

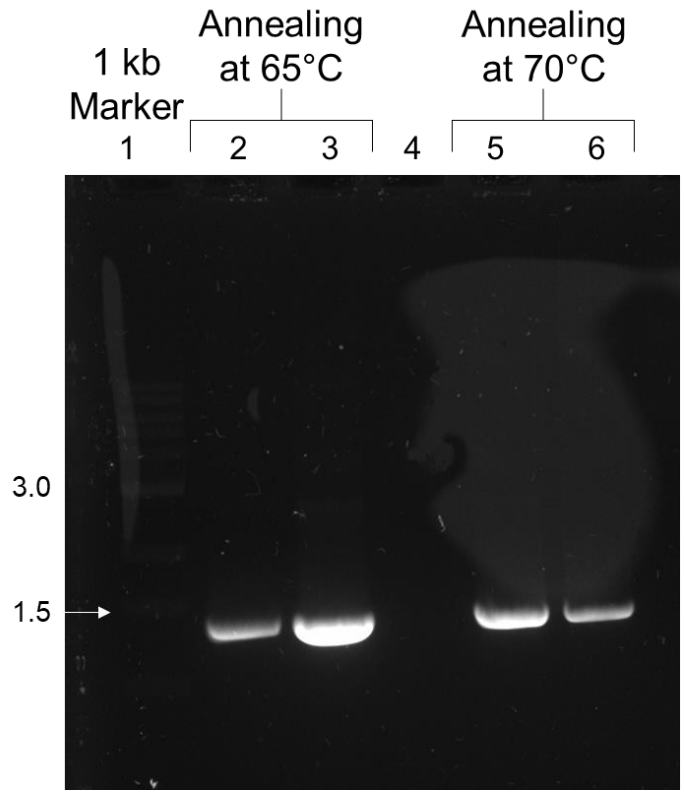

Figure 4.6 Agarose gel analysis for *adh* gene region PCR amplification.
Lane 1: 1 kb DNA marker. Lanes 2 to 3: Amplified DNA encoding the ADH domain using an annealing temperature of 65˚C. Lane 4: Empty. Lanes 5 to 6: Amplified DNA encoding the ADH domain using an annealing temperature of 70˚C. 1% (w/v) agarose in TAE buffer with SYBR safe.

The theoretical size of the *adh* gene fragment is 1.3 kb. The amplified DNA in Figure 4.6 is observed just below the 1.5 kb ladder mark indicating that a fragment of the correct size had been amplified.

The DNA was extracted from the agarose gel in Figure 4.6 and ligated into an empty, digested pGEX-6P-2 plasmid. The successfully ligated pGEX-6P-2-*adh* plasmids were used to transform *E. coli* DH10α cells for propagation. Five transformed colonies were selected and their DNA extracted. The *adh* gene fragment was PCR amplified to identify cells for which the ligation and transformation had worked correctly.

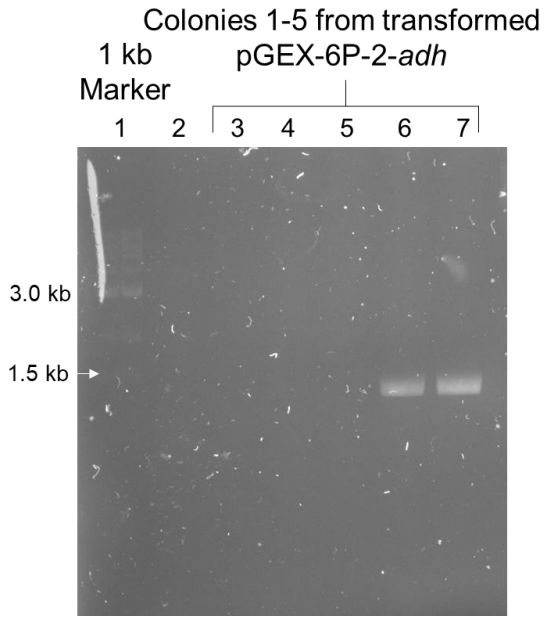

Figure 4.7 Agarose gel analysis of a PCR.
Lane 1: 1 kb DNA marker. Lane 2: Empty. Lanes 3 to 7: DNA samples after PCR amplification from the DNA extracted from the *E. coli* DH10α colonies. 1% (w/v) agarose in 1xTAE buffer with 1x SYBR safe.

Lanes 3 to 5 of Figure 4.7 display no bands after PCR amplification of the *adh* gene fragment, indicating that these colonies were transformed with non-recombinant plasmids. Lanes 6 and 7 each reveal a band at just below 1.5 kb after amplification. As the theoretical size of the *adh* gene is 1.3 kb, colonies 4 and 5 were concluded to contain the correct recombinant plasmid pGEX-6P-2-*adh*.

Plasmids from colonies 4 and 5 were sequenced by PCR and capillary electrophoresis to unequivocally confirm the presence of the *adh* gene.

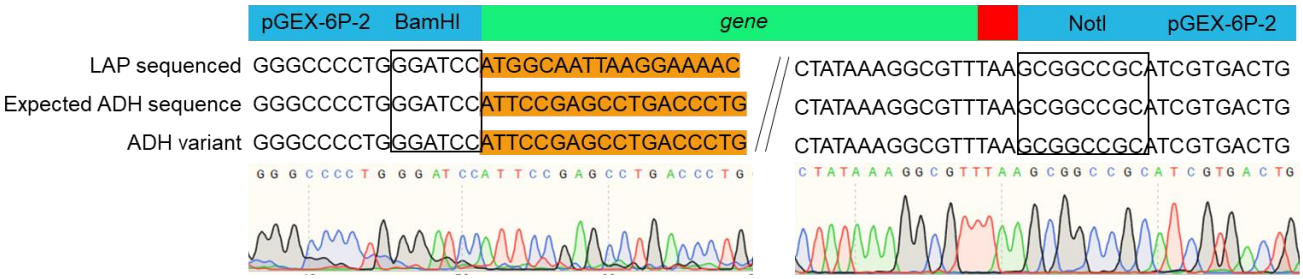

Figure 4.8 Sequence data for the pGEX-6P-2-*adh* plasmid.
The gene sequence including the corresponding chromatograms of the pGEX-6P-2-*adh* plasmid. The orange highlight indicates the start of the gene of interest.

In Figure 4.8 the DNA sequence for the ADH encoding region is compared to that of the full gene. The DNA region highlighted in orange mark the starting nucleotides of the corresponding gene inserts. Clearly the ADH encoding region differs from that of the LAP encoding construct. Instead the ADH encoding region matches the sequence for the

expected initial nucleotides for the *adh* gene. The amplification and transfer of the ADH encoding gene fragment to the pGEX-6P-2 plasmid was thus clearly successful.

4.2. Protein production and optimization of LAP and the domains

4.2.1. Test production of LAP

The pGEX-6P-2-*lap* plasmid was tested for protein production. The confirmed plasmid was used to transform *E. coli* BL21 cells. Three colonies were picked for production and grown at 25 and 37˚C respectively for 3 h.

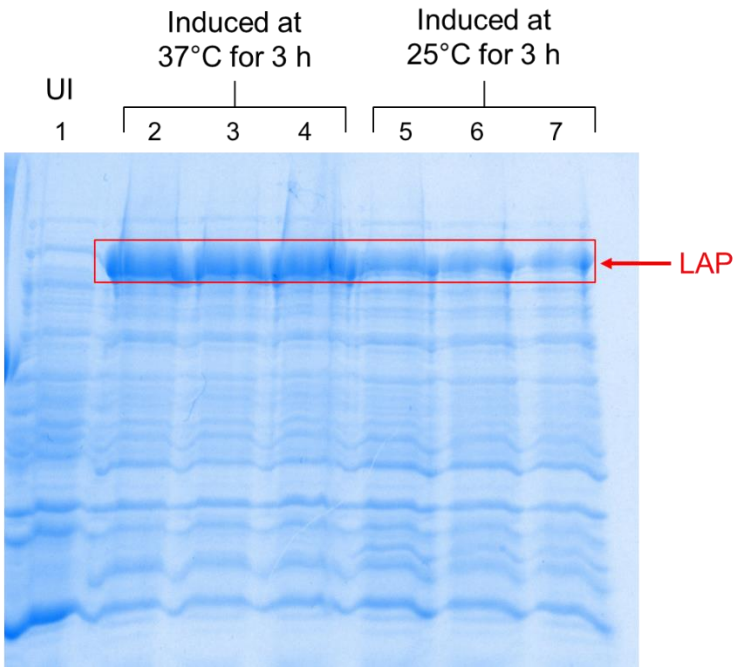

Figure 4.9 SDS-PAGE analysis of LAP production check.
Lane 1: Uninduced sample. Lanes 2-4: Induced samples at 37°C for 3 h. Lanes 5-7: Induced samples at 25˚C for 3 h. 10% (w/v) acrylamide.

The thick band visible in the induced samples in lanes 2 to 7 (Figure 4.9) is not observed in the uninduced sample in lane 1. As the band is visible near the top of the gel, and the expected size of the fusion protein is 130 kDa, we could conclude that LAP was indeed being produced – despite the absence of a corresponding marker. (Please see Figure 4.6 below for size confirmation.) The bands in lanes 2 to 4 are significantly more pronounced than those in lanes 5 to 7. As lanes 2 to 4 correspond to samples grown at 37°C rather than 25°C for lanes 5 to 7, the rate of metabolism at 37˚C is clearly higher than at the lower temperature.

4.2.2. Optimization of LAP production

To optimize protein production further – and to check for protein solubility - different temperatures and IPTG concentrations were assessed.

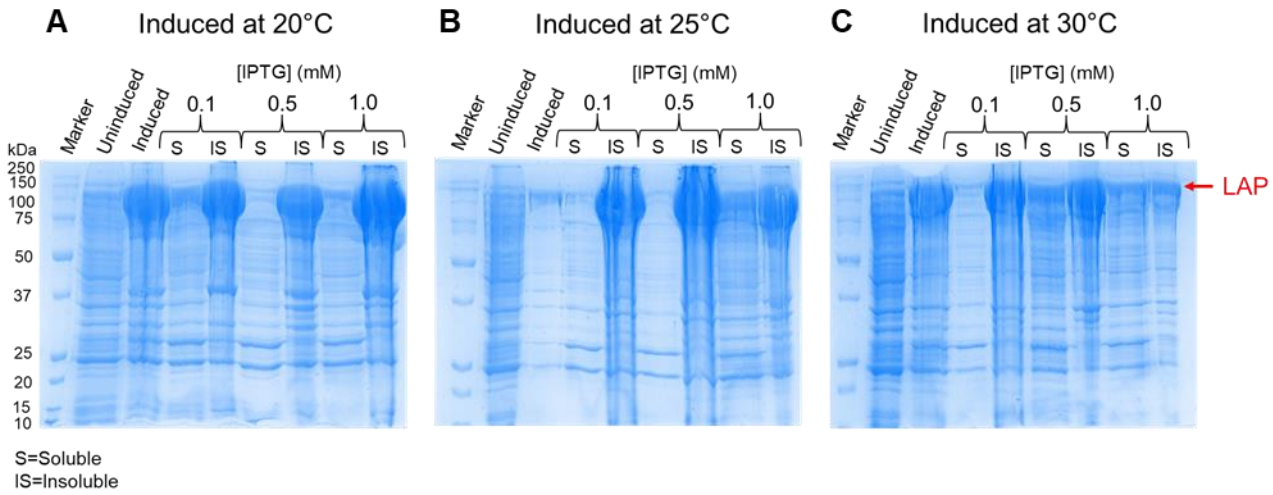

Figure 4.10 SDS-PAGE analysis of LAP production optimization.
All samples were incubated o/n at their respective temperature. 10% (w/v) acrylamide.

In figure 4.10, the size of the major band seen across all three images is between 100 kDa and 150 kDa, congruent with the theoretical size of the GST-LAP fusion protein of 130 kDa. The fusion protein is visible in both soluble and insoluble fractions, though the insoluble fractions contain more protein than soluble lanes. As soluble protein is required for downstream purification, the optimal conditions appeared to be 20°C and 0.1 mM IPTG, 25°C and 1 mM IPTG, or 30°C and 0.5 mM IPTG. Higher IPTG concentrations often impact cellular metabolism negatively producing more insoluble protein. Correspondingly, the lowest IPTG concentration with the highest soluble yield was chosen for downstream protein production i.e. 20°C and 0.1 mM IPTG.

4.2.3. Test production of ALDH domain

The pGEX-6P-2-*aldh* plasmid was tested for protein production. The confirmed plasmid was used to transform *E. coli* BL21 cells. Two colonies were picked for production and grown at 25˚C O/N.

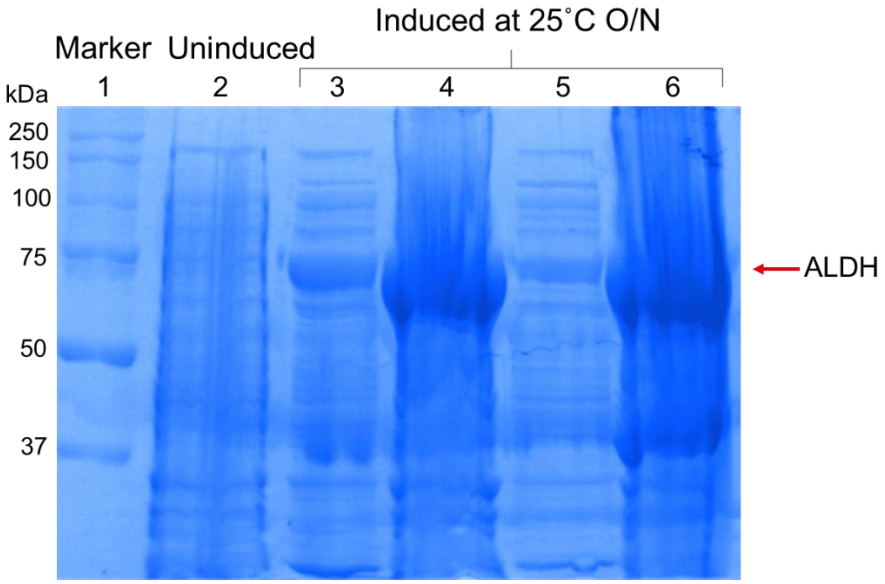

Figure 4.11 SDS-PAGE analysis of ALDH domain production check.
Lane 1: protein size marker. Lane 2: Uninduced sample. Lanes 3 to 6 induced samples. Lanes 3 and 5: induced and soluble. Lanes 4 and 6: Induced and insoluble. 10% (w/v) acrylamide.

The uninduced and induced samples of the GST-ALDH fusion protein were analysed on an SDS-PAGE gel (Figure 4.11). The induced samples in lanes 3 to 6 reveal a thick band around the size of ~75 kDa. This band is not visible in the uninduced sample. As the expected size of the GST-ALDH fusion protein is 76 kDa, the GST-ALDH protein is clearly being produced successfully.

4.2.4. Optimization of ALDH production

To further optimize protein production – and to check for protein solubility – production was assessed for various temperatures and IPTG concentrations (Figure 4.12).

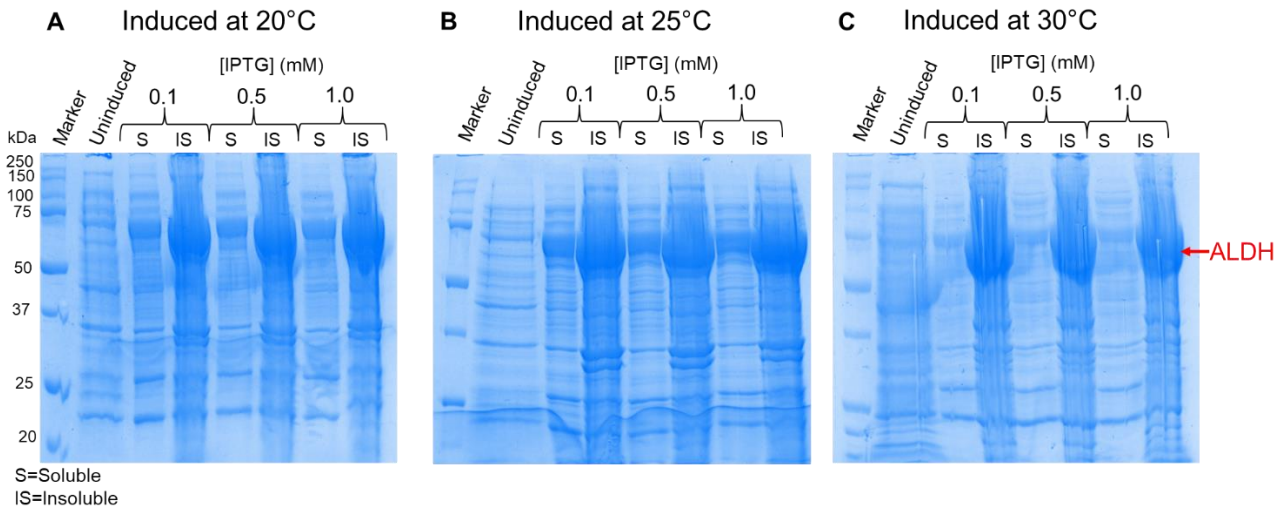

Figure 4.12 SDS-PAGE analysis of the ALDH variant production optimization.
All samples were incubated o/n at their respective temperature. 10% (w/v) acrylamide.

The GST-ALDH fusion protein has a mass of 65 kDa. The most prominent band in the induced samples in Figure 4.12 is also ~65 kDa in size, indicating that the GST-ALDH fusion protein is being produced at significant levels. The band is more prominent in the insoluble fractions but much less pronounce in the soluble fraction. The concentration of IPTG had almost no effect on the amount of soluble protein at any temperature. By contrast temperature appears critical as comparable amounts of soluble protein are produced at 20 and 25°C (with more insoluble protein at 25˚C), while the yield of soluble protein is lower at 30°C. Lanes depicting solubilised insoluble fractions have a “smeared” appearance, presumably due to lipids. High IPTG concentrations clearly impact soluble protein production. Correspondingly, the lowest IPTG concentration with the highest soluble yield was chosen for downstream production parameters: i.e. 25˚C and 0.1 mM IPTG.

4.2.5. Test production of ADH domain

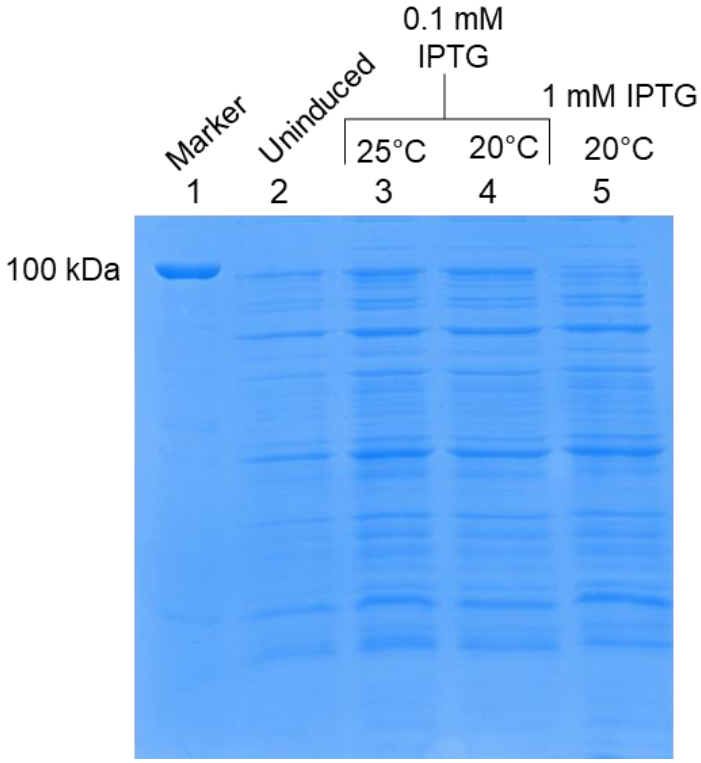

Figure 4.13 SDS-PAGE analysis of the ADH domain soluble production check.
Lane 1: Protein marker. Lane 2: Uninduced sample at 20°C. Lane 3 to 4: Soluble protein at 20 and 25°C induced with 0.1 mM IPTG. Lane 5: Soluble protein at 20°C induced with 0.1 mM IPTG.

In figure 4.13, the uninduced sample in lane 2 has less pronounced bands than the induced samples (lanes 3 to 5). The GST-ADH fusion protein has a theoretical mass of 76 kDa. The relative band sizes and thickness between the induced and uninduced samples appear to be at the same sizes for both, without a clear additional protein band being observed. A thick band at the size of ~76 kDa in the induced samples would indicate successful production of the ADH domain. Although the bands in the induced soluble fractions are significantly more intense than in the uninduced soluble fraction, no significant difference between uninduced and induced fractions is seen. This would imply that the ADH domain was not being produced. This was repeated for other colonies with different temperatures and IPTG concentration, but again no production was seen.

4.3. Purification of LAP

4.3.1. Affinity chromatography

Protein purity is often critical for various downstream experiments. The pGEX-6P-2 plasmid was used to generate GST fusion protein of full-length LAP, as well as the two individual domains to allow for rapid purification by affinity chromatography where GST functions as a purification tag.

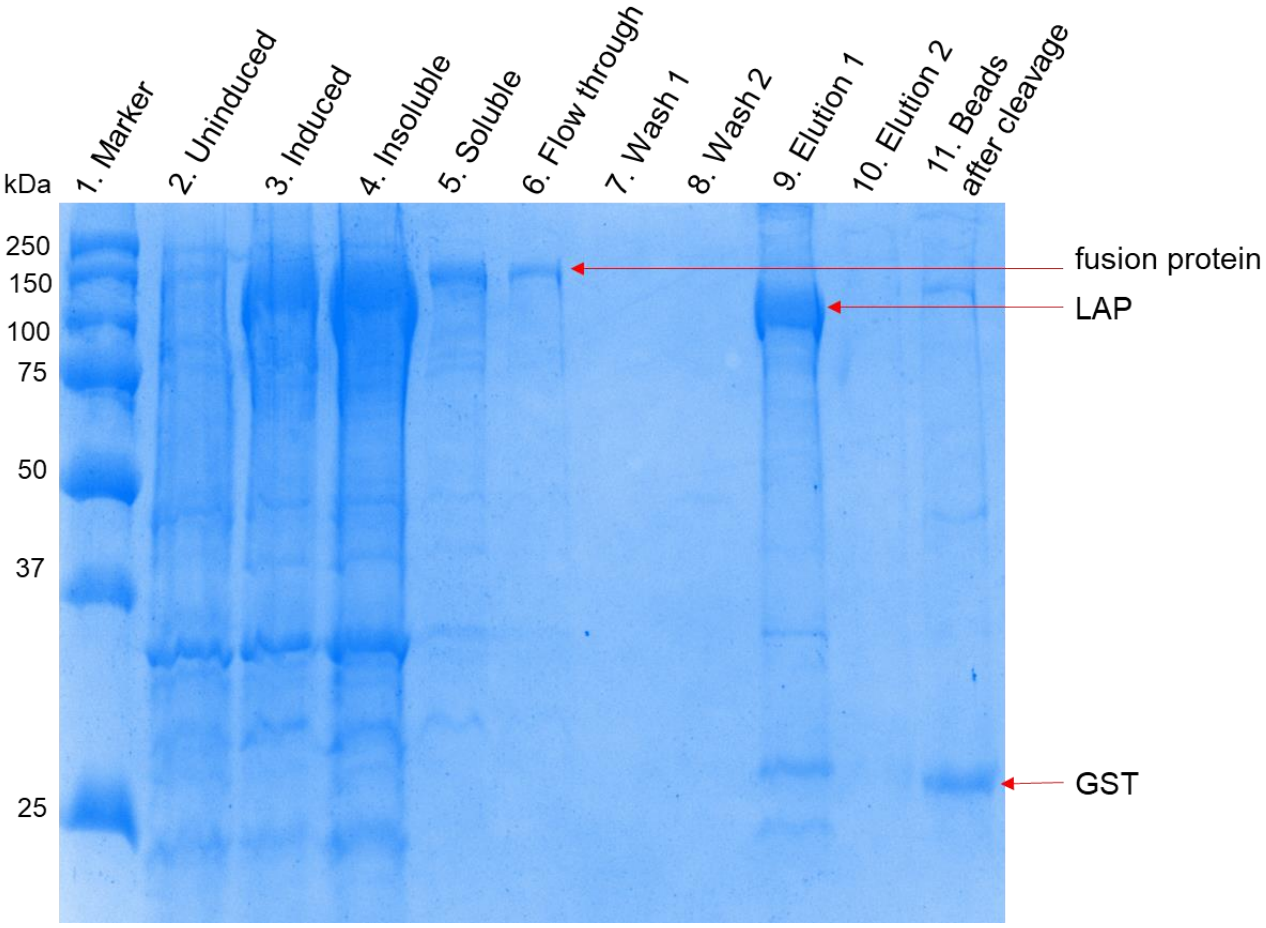

Figure 4.14 SDS-PAGE analysis of the production and purification of LAP.
Lanes 1 to 5: Production samples of LAP. Lanes 6 to 11: Affinity chromatography samples of LAP. 10% (w/v) acrylamide.

A pronounced band is visible at ~130 kDa in the induced sample (Lane 3 of Figure 4.14) compared to the uninduced sample (Lane 2) indicating successful production of the fusion protein. The soluble fraction shown in lane 5 was used for further purification by affinity chromatography. Some of the fusion protein did not bind to the column and was lost in the flow through (lane 6). Both the flow through and soluble fractions (Lanes 5 and 6) were diluted and therefore appear to have less protein than other lanes. The “wash” lanes 7 to 8 do not reveal any bands indicating that very little non-specifically bound protein could be removed at this stage. 3C protease was added to the column to cleave the GST-LAP fusion protein and releasing LAP. Correspondingly, the elution fraction (Lane 9) reveals a

prominent band at ~100 kDa, which matches the expected size of LAP at 94 kDa. Lane 9 contains another noticeable band at 25 kDa, identical in height to the GST that remains bound to the sepharose beads after elution (lane 11). This indicates that a smaller amount of GST co-eluted with the LAP protein, though a majority remains attached to the beads.

4.3.2. Size exclusion chromatography

The remaining GST is a significant contaminant of the LAP protein sample and needed to be removed to achieve a pure protein sample. In a next step size-exclusion chromatography was used to further purify LAP.

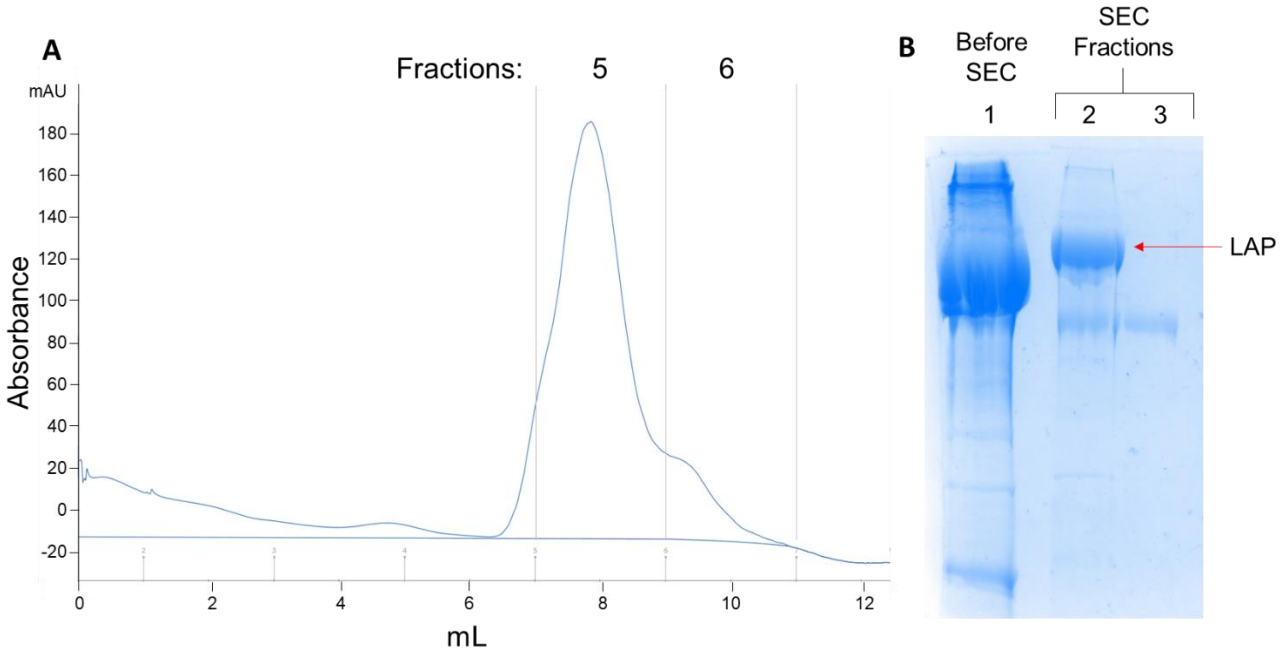

Figure 4.15 SEC chromatogram and SDS-PAGE analysis of LAP purification.
A: Chromatogram for size exclusion chromatography (SEC) of LAP. Sephadex 200 column. B: Lane 1: LAP after affinity chromatography and before SEC. Lane 2 is fraction 5 and lane 3 is fraction 6 from the chromatogram. 10% (w/v) acrylamide.

In Figure 4.15 A there is a peak between 6 and 10 mL of elution. This peak was collected and analysed using SDS-PAGE analysis. Lane 1 in Figure 4.15 B contains the concentrated LAP sample before size exclusion chromatography. Impurities can be seen above and below the major band of LAP. In lane 2 and 3 are the peak fraction samples 5 and 6 respectively. In lane 5 there is the major band corresponding to LAP but there is also a band below the LAP indicating the purification was not completely successful. Lane 6 contains only the smaller protein.

4.3.3. Anion exchange chromatography

Additional purification approach was needed to completely purify LAP. Protein is lost at each additional step of purification. The above sample was used for enzyme assays which does not require a high protein concentration nor completely pure protein.

An alternative purification method used was anion exchange chromatography. LAP is transferred into a low salt buffer with pH 8 have LAP negatively charged (pI=6.5). The sample buffer was exchanged via dialysis to low salt buffer (20 mM Tris pH 8, 20 mM NaCl) in preparation for anion exchange chromatography.

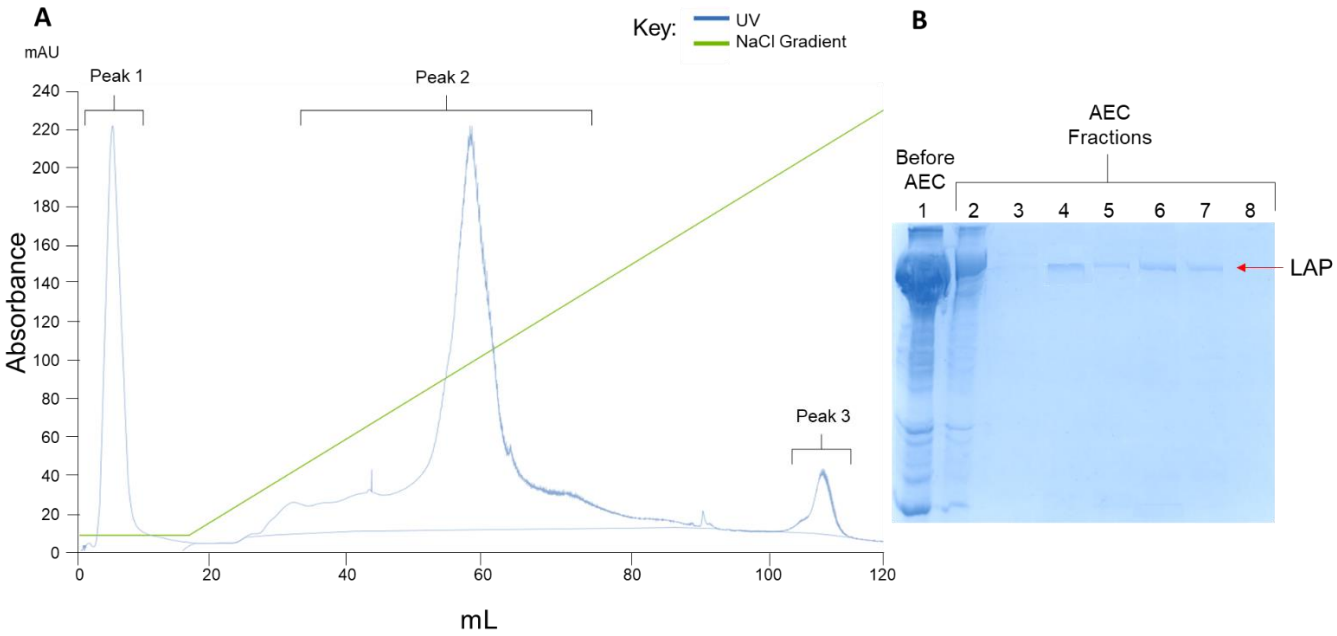

Figure 4.16 AEC chromatogram and SDS-PAGE analysis of LAP purification.
A: Chromatogram for anion exchange chromatography (AEC) of LAP. HiTrap 5 ml Q HP column. B: Lane 1: Concentrated sample after affinity chromatogram. Lane 2: peak 1. Lanes 3 to 7: peak 2. Lane 8: peak 3. 10% (w/v) acrylamide.

The samples of peaks 1 to 3 in figure 4.16 A were collected and analysed using SDS-PAGE analysis. Lane 1 is the elution after affinity chromatography. The sample was highly concentrated to be able to visualize all the impurities. The thick band at the top of the gel is the protein of interest, LAP. Lane 2 is the initial peak 1 during the wash steps. The protein visualised in lane 2 could be due to overloading the column or the incorrect charge on the LAP molecule which can both lead to the protein not binding to the column. Lanes 3 to 7 are the samples from the main peak. These lanes contain a single band at the size of LAP indicating that LAP was successfully purified for downstream experiments. Lane 8 visualizing peak 3 has no band visible, this could be due to a low concentration of the protein present.

4.4. Enzymatic characterization of LAP domains

The enzyme kinetic results in the figures below were all performed using the full-length LAP protein. The conditions to analyse each domain separately were adapted from Jagadeesan *et al* (Jagadeesan *et al.*, 2010). To determine the V_{max}, K_m and k_{cat} of the enzymes the absorbance of NADH was monitored over the reaction time. The 340 nm absorbance values were recorded every 15 s for each of the concentrations for one hour. The absorbance vs time data for each substrate concentration were analysed graphically. The initial gradient of the curve of absorbance vs time was determined and taken to represent the initial velocity at the specified concentration. Plots of initial velocities versus substrate concentrations were analysed using the Michaelis-Menton equation as outlined below to determine the V_{max}, K_m and k_{cat} of each enzyme. Each assay was repeated threefold.

4.4.1. ALDH

Table 4.1 Table displaying the initial reaction rates of ALDH

		[Acetaldehyde] (mM)										
		0,0	0,1	0,5	1,0	2,5	5,0	10,0	15,0	25,0	30,0	50,0
Initial Velocity (mM/min)	1	0,0	0,097	0,166	0,249	0,288	0,335	0,444	0,515	0,534	0,581	0,529
	2	0,0	0,124	0,098	0,278	0,302	0,406	0,466	0,458	0,514	0,541	0,587
	3	0,0	0,119	0,146	0,231	0,299	0,330	0,431	0,554	0,580	0,501	0,594
	avg	0,0	0,113	0,137	0,253	0,296	0,357	0,447	0,509	0,543	0,541	0,570

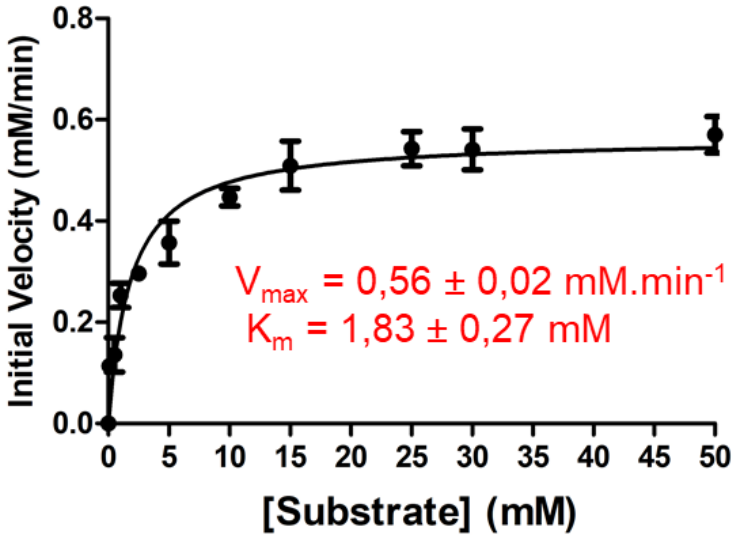

Figure 4.17 Enzyme Activity of ALDH.
The Michaelis-Menton interpretation of the enzyme activity data, n = 3. Regression analysis was used to match the data to the Michaelis-Menton equation (R^2= 0.94). GraphPad Prism Software (San Diego, CA, USA) www.graphpad.com.

Regression analysis using the Michaelis-Menton equation were used to formulate the curves with GraphPad Prism Software (San Diego, CA, USA). The K_m with acetaldehyde as the

substrate for the ALDH domain is 1.83 ± 0.27 mM using the Michaelis-Menton plot (Figure 4.17). The V_{max} value for the ALDH domain is 0.56 ± 0.02 mM.min^{-1} using the Michaelis-Menton plot (Figure 4.17). The k_{cat} for ALDH was determined to be 1.88 s^{-1}.

4.4.2. ADH

Table 4.2 Table of initial reaction rates of ADH.

		[Ethanol] (mM)												
		0,0	0,05	0,1	0,5	1,0	5,0	10,0	15,0	20,0	40,0	60,0	80,0	100,0
Initial Velocity (mM/min)	1	0,0	0,021	0,049	0,068	0,247	0,428	0,483	0,514	0,715	0,880	1,270	0,855	1,032
	2	0,0	0,048	0,057	0,099	0,199	0,538	0,491	0,689	0,668	0,765	1,066	1,216	1,199
	3	0,0	0,012	0,046	0,109	0,288	0,398	0,558	0,746	0,678	0,952	0,884	1,029	0,995
	avg	0,0	0,027	0,051	0,092	0,245	0,455	0,510	0,650	0,687	0,866	1,073	1,033	1,075

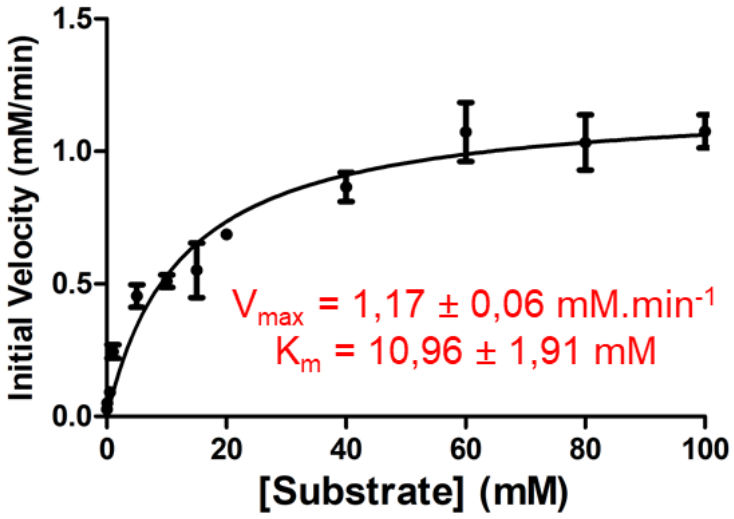

Figure 4.18 Enzyme Activity of ADH.
Panel A depicts the Michaelis-Menton plot of enzyme activity, n = 3. Regression analysis was used to match the data to the Michaelis-Menton equation (R^2= 0.93). GraphPad Prism Software (San Diego, CA, USA) www.graphpad.com.

The K_m value with ethanol as the substrate for ADH using the Michaelis-Menton plot is 10.96 ± 1.91 mM (Figure 4.18). The V_{max} value for ADH using the Michaelis-Menton plot is 1.17 ± 0.06 mM.min^{-1} (Figure 4.18). The k_{cat} for ADH was determined to be 3.91 s^{-1}.

4.5. Structural and biophysical characterization of LAP

4.5.1. Dynamic light scattering of LAP

LAP needed to be characterised to infer whether or not it is occurring as a monomer or in a higher oligomeric state. The hydrodynamic radius of LAP could be tested for using dynamic light scattering.

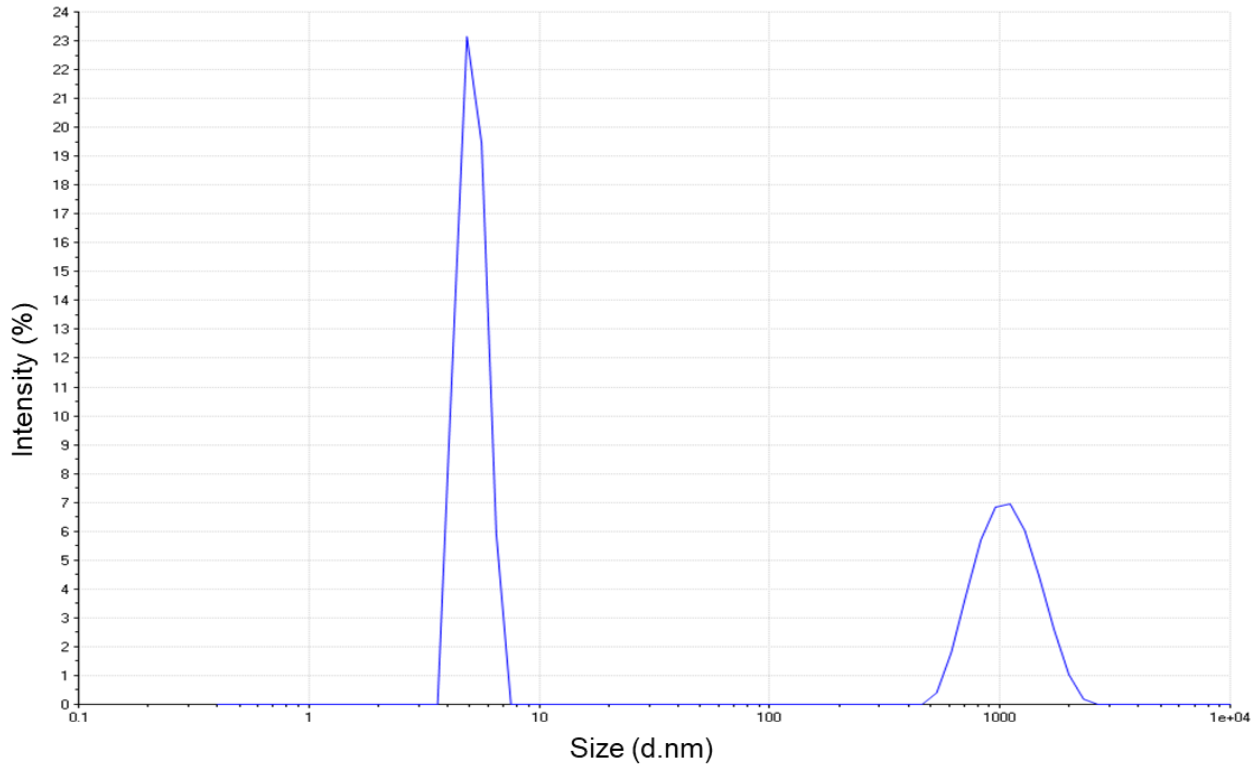

Figure 4.19 Graph of the dynamic light scattering of LAP.
Peak one is ~8 nm and peak two is ~1000 nm.

Two peaks are observed in figure 4.19. The first peak corresponds to ~60% of the molecules in solution while the second peak accounts for the remaining ~40%. The first peak has a peak maximum value of ~5.13 nm. Assuming an average protein density and standard spherical geometry this value may be converted to a molecular mass of ~200 kDa. As the known mass of monomeric LAP is 94 kDa, the first peak could be interpreted to correspond to a dimer of LAP. The second peak has a peak value of a little above 1000 nm, implying a much larger particle volume. Although LAP is known to form long helical arrangements, the indicated size may still be too large, implying that LAP had instead aggregated into large clusters. This may indicate that the sample may have aged and had possibly not been optimally prepared. Large protein aggregates could (and should) have been removed by filtration and/or centrifugation prior to DLS analysis and other experiments such as crystallisation.

4.5.2. Crystallization of LAP

As the polymerization of LAP has not been studied in much detail, the starting assumption was that LAP would potentially exist in a dynamic equilibrium of different oligomerisation states. Of these monomeric or dimeric LAP could possibly have been amenable to structure determination by X-ray crystallography. This would, however, have required that LAP be

crystallized first. One approach to protein crystallization involves vapour diffusion techniques as described (3.3.1 and 3.3.2).

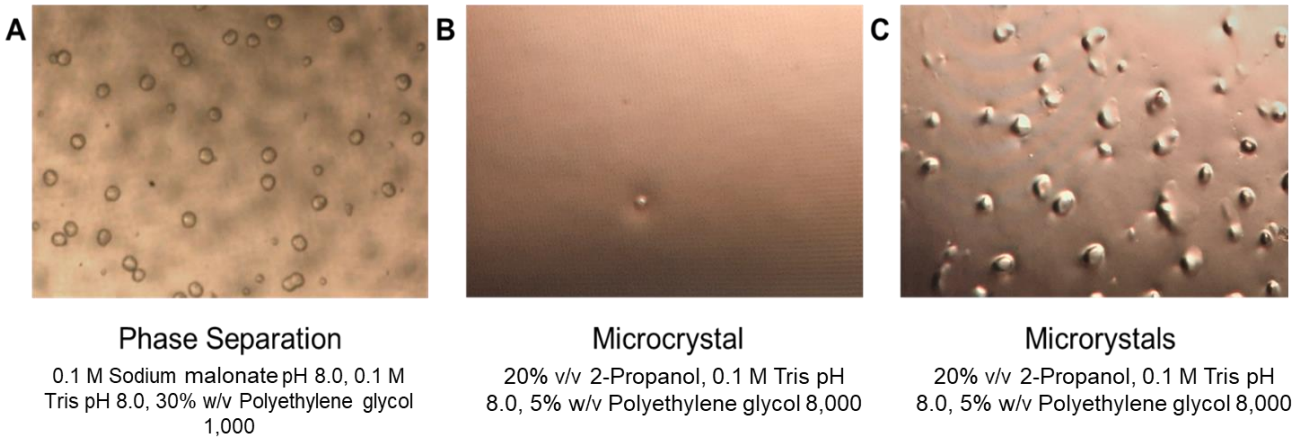

Figure 4.20 Images from crystallisation experiments.
A and B: Initial LAP screening experiment. C: LAP crystals after seeding. Imaged with digital microscope VHX-100 (Keyence, Osaka, Japan) 200x magnification.

The crystal screening experiments for LAP using the sitting drop vapour diffusion experiments covered an array of 96 conditions from the crystallization screens PEGRX1 and PEGRX2 (Hampton Research, CA, USA). After four weeks at 18°C, some crystallization experiments produced small droplet-like structures of ~9.4 µm in diameter (Figure 4.20 A). Droplet formation is a fairly frequent phenomenon in protein crystallization caused by “phase separation”, where one liquid phase (often containing the protein) separates from a second solution (possibly with a lower protein concentration). While not useful in themselves, droplets with high protein concentrations often favour protein crystallisation resulting in rosettes of many small protein crystals. While the crystals are mostly too small for any structural analysis, they may be isolated and crushed to generate crystallization nuclei for later rounds of protein crystallization. In the case of LAP, no such crystallization was observed. Another observation involved possible microcrystals of around 2.1 µm after four weeks at 18˚C (Figure 4.20 B). The microcrystals were incubated for a further four weeks but no further growth was observed.

The phase separation droplets and possible microcrystal samples were collected and used for seeding in a next round of crystal screening. The conditions of the initial screen were repeated but 1 µL of seeding solution was added to each experiment. The conditions that resulted in the presumed microcrystals in the first experiment again resulted in similar microcrystals though both the number and size of microcrystals had increased with an average longest axis of 10.4 µM. As they were not sufficiently large for X-ray diffraction

experiments on a home source, their incubation at 18˚C was extended for a further four weeks. However, no further growth was observed.

4.5.3 Homology with *E. Coli* AdhE

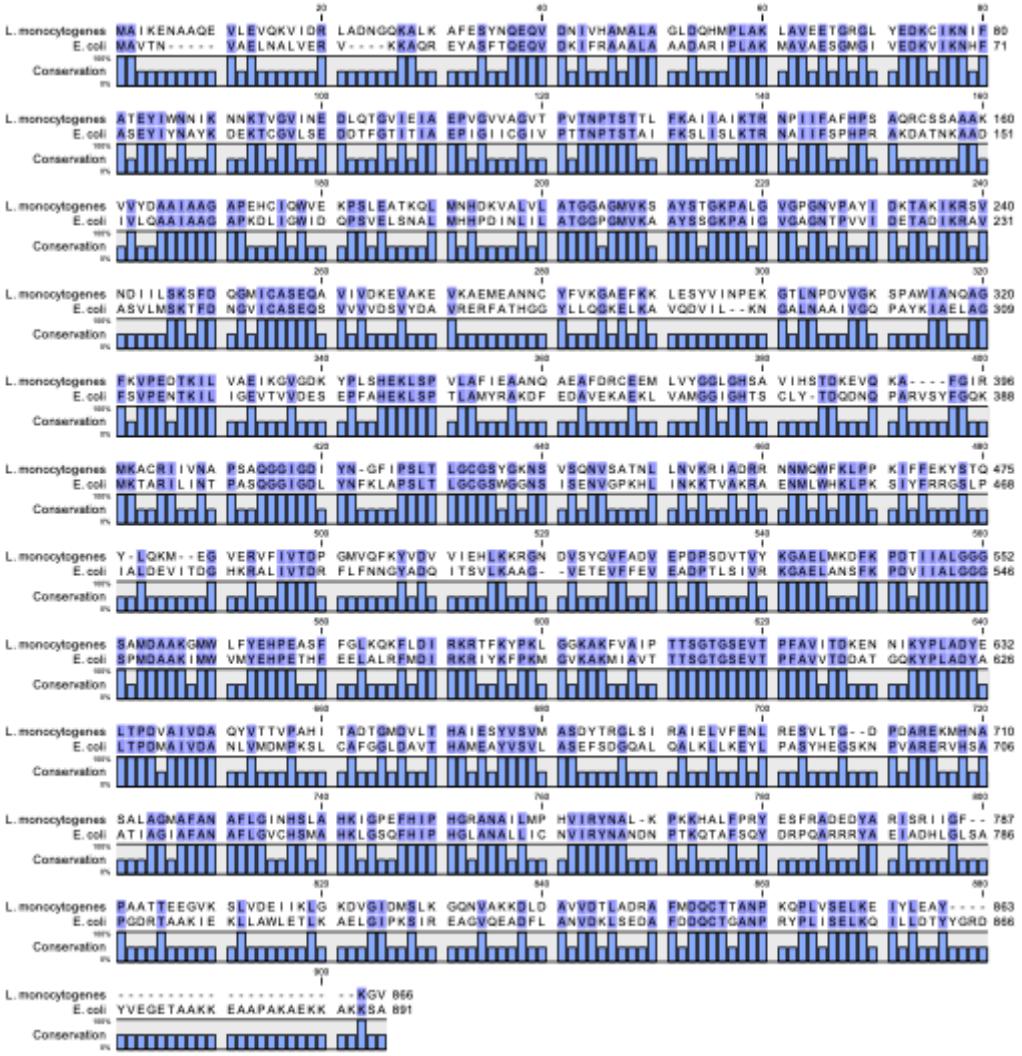

Figure 4.21 Amino acid sequence alignment of AdhE's from *L. monocytogenes* and *E. coli.* The upper sequence corresponds to AdhE from *L. monocytogenes*, the lower one to AdhE from *E. coli*. Conserved residues are marked in blue. (CLC Main Workbench, Qiagen, Germany).

AdhE proteins in different bacterial species have equivalent functions and share significant sequence identity. AdhE from *E. coli* has been studied extensively both in terms of structure and function. Its structure provides a useful template in creating a working model of AdhE from *Lm*. The amino acid sequences of AdhE from *L. monocytogenes* and *E. coli* share a sequence identity of 47.9% confirming their evolutionary relatedness. Combining the sequence alignment and the EM structure of *E. coli* AdhE (PDB: 7bvp), a homology model of LAP/AdhE was prepared using the SWISS-MODEL Server (https://swissmodel.expasy.org).

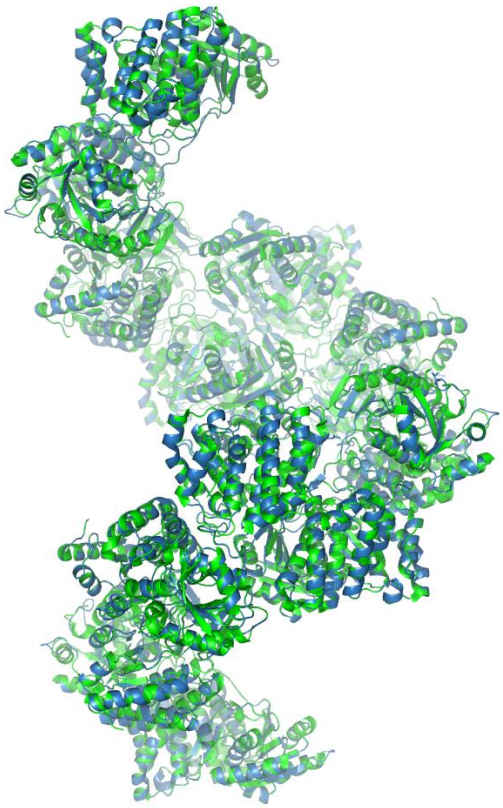

Figure 4.22 Modelled spirosome structure AdhE from *Lm*, aligned with the template AdhE model from *E. coli*.
The modelled structure of the AdhE spirosome from *Lm* (blue) is aligned with the EM structure of the E. coli AdhE (PDB: 7BVP) spirosome (green). Image generated in Pymol.

The threading model of the *L. monocytogenes* AdhE model provides a detailed starting point to further analyse this protein (figure 4.22). No major side clashes were observed in the interface region implying that *Lm* AdhE very likely forms a spirosome similar to that of *E. coli* AdhE.

4.5.3. Electron microscopy analysis of LAP

To broadly ascertain whether LAP/AdhE forms oligomeric structures similar to the *E. coli* AdhE spirosomes, an analysis by negative stain electron microscopy was attempted. Negative stain EM could provide a first low resolution visualisation, revealing whether an oligomeric state similar to that seen in figure 4.19 would be observed. For this purpose, a sample of freshly produced and purified LAP was sent to Dr Jeremy Woodward at the Electron Microscopy Facility, University of Cape Town. Dr Woodard undertook the analysis including steps such as grid preparation, imaging and processing. Images provided by Dr Woodward are shown in Fig. 4.23.

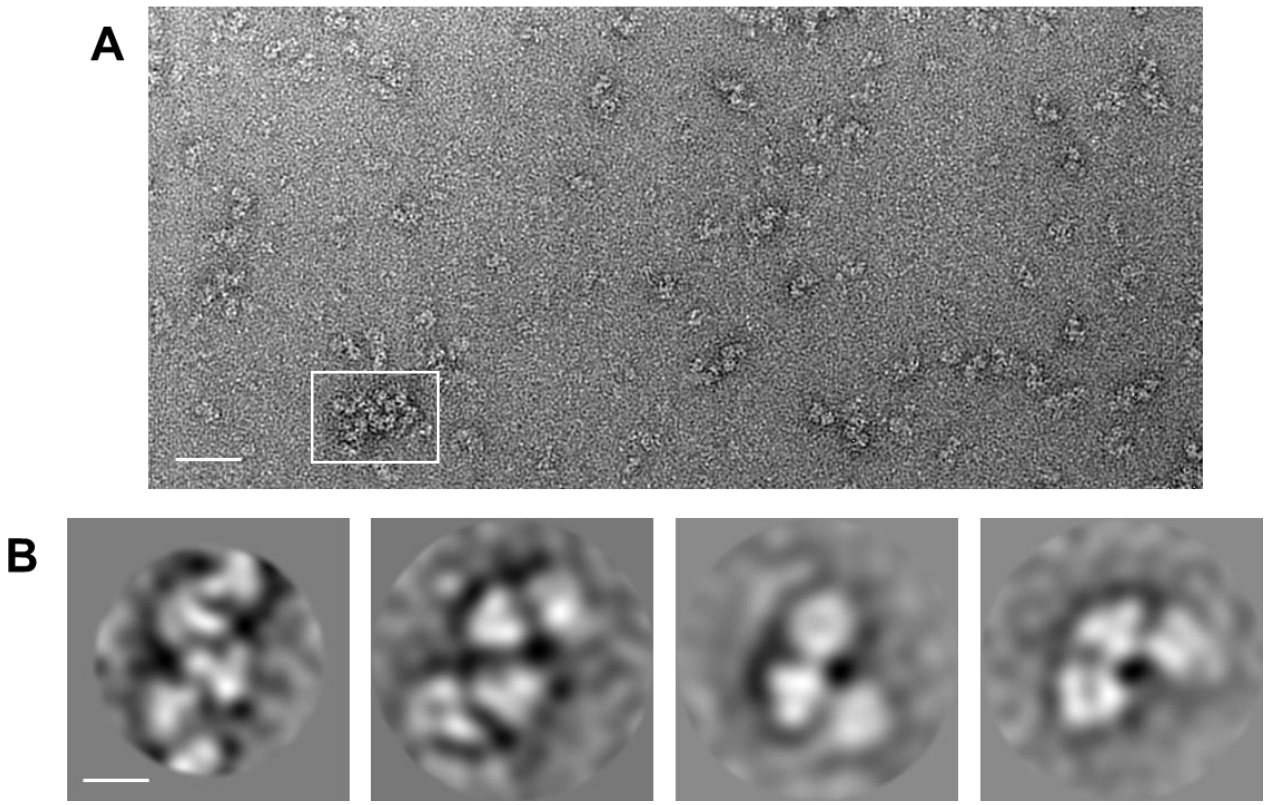

Figure 4.23 Negative stain EM images of LAP.
A. Original negative stain EM image of LAP showing filamentous or possibly helical fragments. Scale bar 120 nm. Figure B displays some 2D class averages of the fragments seen in figure A. Scale bar 9 nm.

Figure 4.23 A displays an original negative stain electron microscopy image of LAP which clearly reveal extended filaments potentially interpretable as helices of various lengths and aggregation states. While some filaments appear to consist of only one or two helical turns, others, such as emphasized by the white box near the lower, left-hand corner are much longer and quite extensive. The filaments in the white box could be beneficial when it comes to the 2D class averaging as it allows for different angles of the filaments to be facing the surface. Figure 4.23 B shows some 2D class averages of multiple aligned images of the filaments. Individual, adjacent monomers of LAP are distinctly discernible in these images.

5. Discussion

5.1. Cloning

The plasmid harboring the *adhE* gene encoding LAP as received from the commercial suppler was only ~100 bp larger than the *adhE* gene itself. A 1% (w/v) agarose gel would not be expected to separate the plasmid and gene insert – as observed in figure 4.1 where both the pUC57 backbone and lap gene overlap. A 2% (w/v) agarose gel and a reduced voltage of 50 V (Figure 8.5) similarly did not succeed in separating the plasmid backbone and gene insert.

Identifying a ScaI site restriction site unique to the pUC57 backbone located roughly opposite the multicloning site allowed the plasmid backbone to be cleaved into smaller fragments. Following BamHI, NotI and ScaI digestion, the backbone was confirmed to have been reduced to smaller fragments (Figure 4*.2*) allowing the *adhE* insert of 2.5 kb to be extracted and cloned into a pGEX-6P-2 plasmid. The resulting pGEX-6P-2-*lap* construct was confirmed by sequencing (Figure 4**.**4).

Following the transformation of *E. coli* BL21 cells, LAP protein was produced at the expected molar mass (Figure 4.9) meaning that the plasmid could be used as a starting point to genetically seperate the domains. Enzymatic activities of each domain would potentially be accessible without interference of the other. The gene fragment encoding the N-terminal ALDH domain of LAP was seperated from fragment encoding the C-terminal ADH domain by adding a stop codon creating a GST-ALDH encoding construct (Figure 4.5).

To generate a GST-ADH encoding construct, the ALDH-encoding fragment of the *lap* gene could theoretically have been removed by site-directed mutagenesis, in which primers link the region upstream of the *lap* gene with the region with the ADH encoding region while the ALDH-encoding fragment loops out as is removed from the pGEX-6P-2-lap plasmid. This approach is tricky as the primers need to anneal to and link two remote regions of the plasmid. Correspondingly, multiple attempts with different annealing temperatures were not successful. Instead a standard approach in which the ADH-encoding region was amplified by PCR proved successful (Figure 4.6). New 5’ BamHI and 3’ NotI sites added via the amplification primers ensured simple ligating into an empty pGEX-6P-2 vector following BamHI and NotI digestion. The pGEX-6P-2-*adh* plasmid was confirmed by sequencing (Figure 4.8).

5.2. Production and purification

The GST-LAP production was optimized to obtain the highest yield of soluble LAP for downstream purification and experiments (Figure 4.10). Purification of full-length GST-LAP fusion protein followed a generic protocol involving affinity purification (Figure 4.14) with glutathione sepharose followed by size-exclusion chromatography to remove impurities either larger or smaller than the LAP. LAP eluted in a broad peak around 7 mL elution volume of the Sephadex 200 column (Figure 4.15). This indicates that LAP forms large aggregates as the sooner a protein elutes from the size-exclusion column the larger it is. The smaller proteins remain in the resin of the column for longer and elute after the large proteins which move through the resin and elute near the start of the elution. The sample was further purified by anion exchange chromatography using a buffer pH of 7.5 well above the isoelectric point of LAP at 6.5, yielding essentially pure LAP (Figure 4.16). The high salt buffer after anion exchange chromatography was replaced by 100 mM Tris pH 7.8, 100 mM NaCl, 1 mM DTT, 5% (v/v) glycerol by filtration. DTT was added to maintain a reducing environment to prevent disulphide bridge formation. Glycerol was added to avoid precipitation, though full-length LAP still proved prone to precipitation at concentrations above 3 mg/mL.

Production of the two LAP domains in *E. coli* BL21 was assessed. The GST-ALDH fusion protein was produced successfully (Figure 4.11) and the optimization of the GST-ALDH production proceeded (Figure 4.12). The GST-ALDH fusion protein produced after optimization was purified using glutathione affinity chromatography. The protein, however, proved highly susceptible to degradation so that no pure, intact protein was produced. Production attempts of GST-ADH fusion protein yielded no bands of expected size in SDS PAGE analyses. Transforming competent *E. coli* BL21 cells with the pGEX-6P-2-*adh* plasmid a second time and growing transformed cells at various temperatures and using different IPTG concentrations for induction, did not provide evidence for either the fusion protein or GST alone being produced (Figure 4.13). The original plan to analyse the enzymatic properties of the separate ALDH and ADH domains was therefore aborted in favour of working with the full-length protein.

5.3. Enzyme activity

The ALDH and ADH domains of LAP were characterised enzymatically using the purified, full-length protein. The Michaelis-Menton constant (K_m) is equivalent to the dissociation constant of the enzyme and substrate. A large K_m value thus indicates weak binding and vice versa. The Michaelis-Menton plot derived K_m for the ALDH domain, with acetaldehyde

as a substrate, was 1.83 ± 0.27 mM while that for the ADH domain, with ethanol as the substrate, was 10.96 ± 1.91 mM (Figure 4.17 and Figure 4.18). V_{max} for the ALDH and ADH domains were 0.56 ± 0.02 mM.min^{-1} and 1.17 ± 0.06 mM.min^{-1} respectively (Figure 4.17 and Figure 4.18). The K_m and V_{max} for the ADH domain are almost double those for ALDH. Although the larger K_m of ADH indicates a lower substrate affinity than for ALDH, the higher V_{max} indicates a higher reaction rate. Both parameters can be combined in the k_{cat} or turnover number which describes the number of substrate molecules converted to product per unit time. The k_{cat} for ALDH and ADH were determined to be 1.88 s^{-1} and 3.91 s^{-1} respectively. ADH thus turns over substrate to product more quickly than ALDH. The k_{cat} values for AdhE of *Citrobacter* sp S77 were reported as 3.2 s^{-1} and 13 s^{-1} for the ALDH and ADH domains respectively (Tsuji *et al.*, 2016), which are somewhat higher than those of *L. monocytogenes* LAP/AdhE. Although the values are similar for ALDH k_{cat} values, the ADH domain of AdhE in *Citrobacter* sp S77 is 3-fold faster than that of LAP/AdhE. Thus indicating that LAP/AdhE has a slower alcohol production than *Citrobacter* sp S77.

5.4. Structural analysis of LAP

The three-dimensional structure of LAP/AdhE currently remains unexplored. However, closely related proteins have been analysed in some detail so that much about the structure of LAP can already be inferred. Understanding the structure of LAP especially in complex with its reported human receptor HSP60 would be highly relevant medically as this could provide a structural basis to develop strategies for the prevention of human listeriosis. This aspect, however, was not part of this dissertation. Rather, the aim was to functionally and biochemically analyse the protein LAP/AdhE.

X-ray crystallography is a well-established technique that provides high-resolution structural data about proteins. It requires that the protein in question first needs to be crystallized. Identifying crystallisation conditions is often an arduous journey simplified by using sparse-matrix crystallization screens that efficiently combine a large array of reagents. Outcomes for an initial crystallization screen for LAP yielded mainly phase separation and microcrystals (Figure 4***.20***). Although the latter were not useful in themselves for diffraction experiments, they were used to prepare nuclei for seeded crystallisation experiments. However, only similar microcrystals were observed with no improvement in size and morphology.

LAP eluted from the column at the void volume during size exclusion chromatography indicating that LAP mostly forms large oligomers (Section 5.2) that are larger than the separation capacity of the column. In principle, dynamic light scattering (DLS) can be useful in quantifying the size of the LAP aggregates. DLS analyses indicated that LAP aggregates

come in two sizes: ~60% of the protein formed small aggregates potentially consisting of LAP dimers while 40% of the protein formed large aggregates. As AdhE from other bacterial species are known to form spirosomes, it was possible that the large aggregates were not simply randomly aggregated LAP proteins - as frequently happens with aging protein preparations - but rather well-defined spirosomes of variable length. We therefore undertook to characterize LAP by electron microscopy for a first, low resolution analysis to distinguish filamentous spirosomes from random aggregates. A first round of negatively stained electron microscopy confirm that listerial LAP or AdhE also forms filaments (Figure 4.23). Similar images captured by Kim *et al* in 2019 for *E. coli* AdhE spirosomes are shown below:

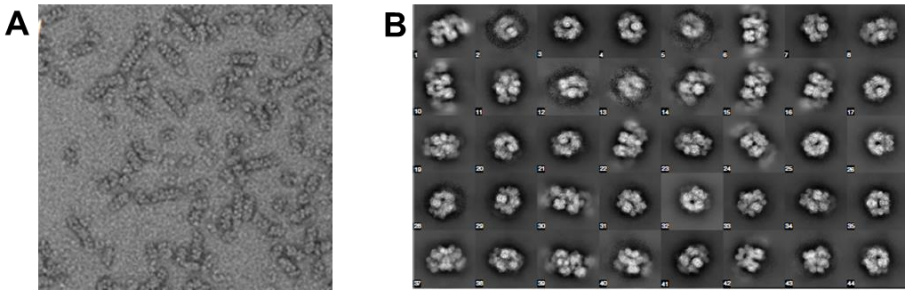

Figure 5.1 Electron microscopy analysis of *E. coli* AdhE spirosomes.
A. Negative stain electron microscopy image of the *E. coli* AdhE. B. 2D class average of the *E. coli* AdhE. Images from the 2019 publication by Kim *et al*.

Figure 5.1 A is the negative stain electron microscopy image of *E. coli* AdhE (Kim *et al.*, 2019). The image has multiple filaments that appear longer than the filaments seen in figure 4.23 which are somewhat shorter and fewer than those seen in figure 5.1 A above. This may, in part, have been due to the preparation or the transport of LAP from Pretoria to Cape Town and would need to be optimized before any additional experiments would be undertaken. The 2D class averages seen in figure 4.23 B resemble those seen in figure 5.1 B - though they appear longer and more regular in the latter. This confirms that LAP forms extended spirosomes similar to those seen in other bacterial species.

With the confirmation of spirosome formation by LAP and hence the length of individual filaments being uncontrolled, structure determination of LAP by X-ray crystallography would prove largely difficult. Smaller fragments could perhaps be generated by introducing critically positioned mutations at interaction surfaces. This could stabilize the oligomeric conformation of LAP increasing the possibility of crystal formation.

The extended filaments of AdhE from *E. coli* (Kim *et al.*, 2020) were found to help regulate enzyme activity (Pony *et al.*, 2020). Future studies of LAP could investigate whether co-factors similarly affect the filament conformation and enzyme activity.

6. Conclusion

The bifunctional acetaldehyde alcohol dehydrogenase (AdhE) or LAP of *Listeria monocytogenes* specifically forms extended filaments similar to those described for *E. coli* AdhE (Kim *et al.*, 2020). Filament conformations may potentially control enzyme activity which could be analysed in future studies. The maximal enzyme rate for the LAP ADH domain is twice that of the ALDH domain, this correlates to the high K_m of ADH and low K_m of ALDH. The fast reaction rate of ADH can be investigated further for industrial application. The LAP protein has a simple production and purification methodology which can benefit industry financially for alcohol production. The enzymes of the bifunctional protein may require their joint conformation for activity and to prevent degradation like was found by Pony *et al* (Pony *et al.*, 2020).

7. References

Brugère-Picoux, J. 2008. Ovine listeriosis. *Small Ruminant Research,* 76**,** 12-20.

Burkholder, K. M., Kim, K.-P., Mishra, K. K., Medina, S., Hahm, B.-K., Kim, H. & Bhunia, A. K. 2009. Expression of LAP, a SecA2-dependent secretory protein, is induced under anaerobic environment. *Microbes and infection,* 11**,** 859-867.

Cheng, M. Y., Hartl, F.-U. & Norwich, A. L. 1990. The mitochondrial chaperonin hsp60 is required for its own assembly. *Nature,* 348**,** 455.

Cossart, P. 2011. Illuminating the landscape of host–pathogen interactions with the bacterium Listeria monocytogenes. *Proceedings of the National Academy of Sciences,* 108**,** 19484-19491.

Cossart, P. & Helenius, A. 2014. Endocytosis of viruses and bacteria. *Cold Spring Harbor perspectives in biology,* 6**,** a016972.

De Las Heras, A., Cain, R. J., Bielecka, M. K. & Vazquez-Boland, J. A. 2011. Regulation of Listeria virulence: PrfA master and commander. *Current opinion in microbiology,* 14**,** 118-127.

De Noordhout, C. M., Devleesschauwer, B., Angulo, F. J., Verbeke, G., Haagsma, J., Kirk, M., Havelaar, A. & Speybroeck, N. 2014. The global burden of listeriosis: a systematic review and meta-analysis. *The Lancet Infectious Diseases,* 14**,** 1073-1082.

Deno, N., Peterson, H. J. & Saines, G. S. 1960. The Hydride-Transfer Reaction. *Chemical Reviews,* 60**,** 7-14.

Drolia, R., Tenguria, S., Durkes, A. C., Turner, J. R. & Bhunia, A. K. 2018. Listeria adhesion protein induces intestinal epithelial barrier dysfunction for bacterial translocation. *Cell host & microbe,* 23**,** 470-484. e7.

Echave, P., Tamarit, J., Cabiscol, E. & Ros, J. 2003. Novel antioxidant role of alcohol dehydrogenase E from Escherichia coli. *Journal of Biological Chemistry,* 278**,** 30193-30198.

Gandhi, M. & Chikindas, M. L. 2007. Listeria: a foodborne pathogen that knows how to survive. *International journal of food microbiology,* 113**,** 1-15.

Gouin, E., Welch, M. D. & Cossart, P. 2005. Actin-based motility of intracellular pathogens. *Current opinion in microbiology,* 8**,** 35-45.

Havelaar, A. H., Kirk, M. D., Torgerson, P. R., Gibb, H. J., Hald, T., Lake, R. J., Praet, N., Bellinger, D. C., De Silva, N. R. & Gargouri, N. 2015. World Health Organization global estimates and regional comparisons of the burden of foodborne disease in 2010. *PLoS medicine,* 12**,** e1001923.

Horwich, A. L., Fenton, W. A., Chapman, E. & Farr, G. W. 2007. Two families of chaperonin: physiology and mechanism. *Annu. Rev. Cell Dev. Biol.,* 23**,** 115-145.

Huberts, D. H. & Van Der Klei, I. J. 2010. Moonlighting proteins: an intriguing mode of multitasking. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research,* 1803**,** 520-525.

Itoh, H., Komatsuda, A., Ohtani, H., Wakui, H., Imai, H., Sawada, K. I., Otaka, M., Ogura, M., Suzuki, A. & Hamada, F. 2002. Mammalian HSP60 is quickly sorted into the mitochondria under conditions of dehydration. *European Journal of Biochemistry,* 269**,** 5931-5938.

Jagadeesan, B., Koo, O. K., Kim, K.-P., Burkholder, K. M., Mishra, K. K., Aroonnual, A. & Bhunia, A. K. 2010. LAP, an alcohol acetaldehyde dehydrogenase enzyme in Listeria, promotes bacterial adhesion to enterocyte-like Caco-2 cells only in pathogenic species. *Microbiology,* 156**,** 2782-2795.

Jagadeesan, B., Littlejohn, A. E. F., Amalaradjou, M. a. R., Singh, A. K., Mishra, K. K., La, D., Kihara, D. & Bhunia, A. K. 2011. N-Terminal Gly224–Gly411 domain in Listeria adhesion protein interacts with host receptor Hsp60. *PloS one,* 6**,** e20694.

Jaradat, Z. W. & Bhunia, A. K. 2002. Glucose and nutrient concentrations affect the expression of a 104-kilodalton Listeria adhesion protein in Listeria monocytogenes. *Applied and environmental microbiology,* 68**,** 4876-4883.

Jaradat, Z. W., Wampler, J. L. & Bhunia, A. K. 2003. A Listeria adhesion protein-deficient Listeria monocytogenes strain shows reduced adhesion primarily to intestinal cell lines. *Medical microbiology and immunology,* 192**,** 85-91.

Johnson, R. B., Fearon, K., Mason, T. & Jindal, S. 1989. Cloning and characterization of the yeast chaperonin HSP60 gene. *Gene,* 84**,** 295-302.

Kim, G., Azmi, L., Jang, S., Jung, T., Hebert, H., Roe, A. J., Byron, O. & Song, J.-J. 2019. Aldehyde-alcohol dehydrogenase forms a high-order spirosome architecture critical for its activity. *Nature communications,* 10**,** 1-11.

Kim, G., Yang, J., Jang, J., Choi, J.-S., Roe, A. J., Byron, O., Seok, C. & Song, J.-J. 2020. Aldehyde-alcohol dehydrogenase undergoes structural transition to form extended spirosomes for substrate channeling. *Communications biology,* 3**,** 1-9.

Kirk, M. D., Pires, S. M., Black, R. E., Caipo, M., Crump, J. A., Devleesschauwer, B., Döpfer, D., Fazil, A., Fischer-Walker, C. L. & Hald, T. 2015. World Health Organization estimates of the global and regional disease burden of 22 foodborne bacterial, protozoal, and viral diseases, 2010: a data synthesis. *PLoS medicine,* 12**,** e1001921.

Lake, R. J., Cressey, P. J., Campbell, D. M. & Oakley, E. 2010. Risk ranking for foodborne microbial hazards in New Zealand: burden of disease estimates. *Risk Analysis: an International Journal,* 30**,** 743-752.

Lambrechts, A., Gevaert, K., Cossart, P., Vandekerckhove, J. & Van Troys, M. 2008. Listeria comet tails: the actin-based motility machinery at work. *Trends in cell biology,* 18**,** 220-227.

Laurenceau, R., Krasteva, P. V., Diallo, A., Ouarti, S., Duchateau, M., Malosse, C., Chamot-Rooke, J. & Fronzes, R. 2015. Conserved Streptococcus pneumoniae spirosomes suggest a single type of transformation pilus in competence. *PLoS pathogens,* 11**,** e1004835.

Lecuit, M., Ohayon, H., Braun, L., Mengaud, J. & Cossart, P. 1997. Internalin of Listeria monocytogenes with an intact leucine-rich repeat region is sufficient to promote internalization. *Infection and immunity,* 65**,** 5309-5319.

Luong, T. T., Kim, E.-H., Bak, J. P., Nguyen, C. T., Choi, S., Briles, D. E., Pyo, S. & Rhee, D.-K. 2015. Ethanol-induced alcohol dehydrogenase E (AdhE) potentiates pneumolysin in Streptococcus pneumoniae. *Infection and immunity,* 83**,** 108-119.

Murray, E. G. D., Webb, R. A. & Swann, M. B. R. 1926. A disease of rabbits characterised by a large mononuclear leucocytosis, caused by a hitherto undescribed bacillus Bacterium monocytogenes (n. sp.). *The Journal of Pathology,* 29**,** 407-439.

Pandiripally, V., Westbrook, D., Sunki, G. & Bhunia, A. 1999. Surface protein p104 is involved in adhesion of Listeria monocytogenes to human intestinal cell line, Caco-2. *Journal of medical microbiology,* 48**,** 117-124.

Peng, H., Wu, G. & Shao, W. 2008. The aldehyde/alcohol dehydrogenase (AdhE) in relation to the ethanol formation in Thermoanaerobacter ethanolicus JW200. *Anaerobe,* 14**,** 125-127.

Pizarro-Cerdá, J., Kühbacher, A. & Cossart, P. 2012. Entry of Listeria monocytogenes in mammalian epithelial cells: an updated view. *Cold Spring Harbor perspectives in medicine,* 2**,** a010009.

Pockley, A. G. & Henderson, B. 2017. Extracellular cell stress (heat shock) proteins—immune responses and disease: an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences,* 373**,** 20160522.

Pony, P., Rapisarda, C., Terradot, L., Marza, E. & Fronzes, R. 2020. Filamentation of the bacterial bi-functional alcohol/aldehyde dehydrogenase AdhE is essential for substrate channeling and enzymatic regulation. *Nature communications,* 11**,** 1-11.

Ranford, J. C., Coates, A. R. & Henderson, B. 2000. Chaperonins are cell-signalling proteins: the unfolding biology of molecular chaperones. *Expert reviews in molecular medicine,* 2**,** 1-17.

Ranson, N. A., White, H. E. & Saibil, H. R. 1998. Chaperonins. *Biochemical Journal,* 333**,** 233-242.

Rengarajan, M., Hayer, A. & Theriot, J. A. 2016. Endothelial cells use a formin-dependent phagocytosis-like process to internalize the bacterium Listeria monocytogenes. *PLoS pathogens,* 12**,** e1005603.

Rigel, N. W. & Braunstein, M. 2008. A new twist on an old pathway–accessory secretion systems. *Molecular microbiology,* 69**,** 291-302.

Ritossa, F. 1962. A new puffing pattern induced by temperature shock and DNP in Drosophila. *Experientia,* 18**,** 571-573.

Swaminathan, B. & Gerner-Smidt, P. 2007. The epidemiology of human listeriosis. *Microbes and Infection,* 9**,** 1236-1243.

Thomas, J., Govender, N., Mccarthy, K. M., Erasmus, L. K., Doyle, T. J., Allam, M., Ismail, A., Ramalwa, N., Sekwadi, P. & Ntshoe, G. 2020. Outbreak of listeriosis in South Africa associated with processed meat. *New England Journal of Medicine,* 382**,** 632-643.

Tsuji, K., Yoon, K.-S. & Ogo, S. 2016. Biochemical characterization of a bifunctional acetaldehyde-alcohol dehydrogenase purified from a facultative anaerobic bacterium Citrobacter sp. S-77. *Journal of bioscience and bioengineering,* 121**,** 253-258.

Vazquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., Gonzalez-Zorn, B., Wehland, J. & Kreft, J. 2001. Listeria pathogenesis and molecular virulence determinants. *Clin Microbiol Rev,* 14**,** 584-640.

Wampler, J. L., Kim, K.-P., Jaradat, Z. & Bhunia, A. K. 2004. Heat shock protein 60 acts as a receptor for the Listeria adhesion protein in Caco-2 cells. *Infection and immunity,* 72**,** 931-936.

White, D. G., Zhao, S., Simjee, S., Wagner, D. D. & Mcdermott, P. F. 2002. Antimicrobial resistance of foodborne pathogens. *Microbes and infection,* 4**,** 405-412.

Yang, C.-K., Ewis, H. E., Zhang, X., Lu, C.-D., Hu, H.-J., Pan, Y., Abdelal, A. T. & Tai, P. C. 2011. Nonclassical protein secretion by Bacillus subtilis in the stationary phase is not due to cell lysis. *Journal of bacteriology,* 193**,** 5607-5615.

Zetterström, C. E., Uusitalo, P., Qian, W., Hinch, S., Caraballo, R., Grundström, C. & Elofsson, M. 2018. Screening for inhibitors of acetaldehyde dehydrogenase (AdhE)

from enterohemorrhagic Escherichia coli (EHEC). *SLAS Discovery: Advancing Life Sciences R&D,* 23**,** 815-822.

8. Appendix

Table 8.1 Bacterial strains and plasmids used in this study

	Description	Source
Bacterial Strains		
E. coli BL21	Prokaryotic expression host	SigmaAldrich*
E. coli DH10α	Plasmid propagation	SigmaAldrich*
Plasmids		
pGEX-6P-2-*3c*	Contains *3c protease* gene	SigmaAldrich*
pUC57-*lap*	Synthesized *lap* gene in pUC57	Gene Universal†
pGEX-6P-2	Expression vector with GST-tag	SigmaAldrich*
pGEX-6P-2-*lap*	Contains entire *lap* gene	This study
pGEX-6P-2-*aldh*	Contains *aldh* variant of *lap* gene	This study
pGEX-6P-2-*adh*	Contains *adh* region of *lap* gene	This study

* St. Louis, Mi, USA
† Newark DE, USA

Table 8.2 Primers used in this study

Description	Sequence	Source
Sequencing Primers		Inqaba Biotech**
pGEX 5’	GGGCTGGCAAGCCACGTTTGGTG	
pGEX 3’	CCGGGAGCTGCATGTGTCAGAGG	
SEQ LAP FP	GCAGAAAGCATTTGGCATTCGTATGAAAGCC	
SEQ LAP RP	CGGTCACATCACTAGGGTCCGGTTC	
Cloning Primers		
SDM ALDH FP	GCTGAATGTGAAACGCATTGCAGATTGACGTAATAATATGCAGTGG	
SDM ALDH RP	CCACTGCATATTATTACGTCAATCTGCAATGCGTTTCACATTCAGC	
ADH FP	ATATTGGATCC†ATTCCGAGCCTGACC	
ADH RP	CACGATGCGGCCGC‡TTAAACG	

*Underlined bases indicate a restriction enzyme site

** Pretoria, South Africa

†BamHI site

‡NotI site

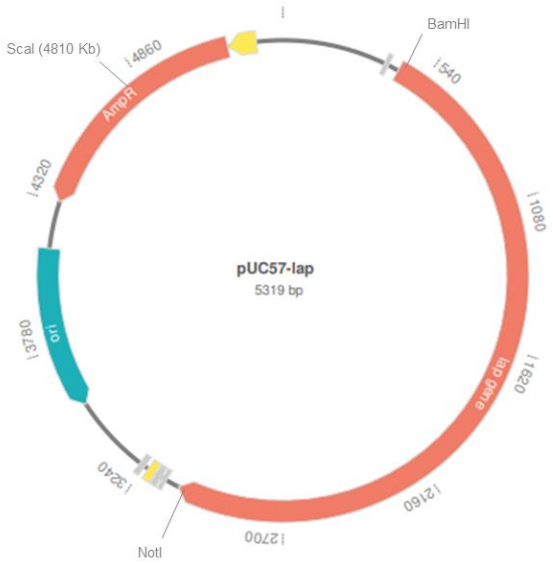

Figure 8.1 Plasmid map of pUC57-*lap.*

Figure 8.2 Plasmid map of pGEX-6P-2-*lap.*

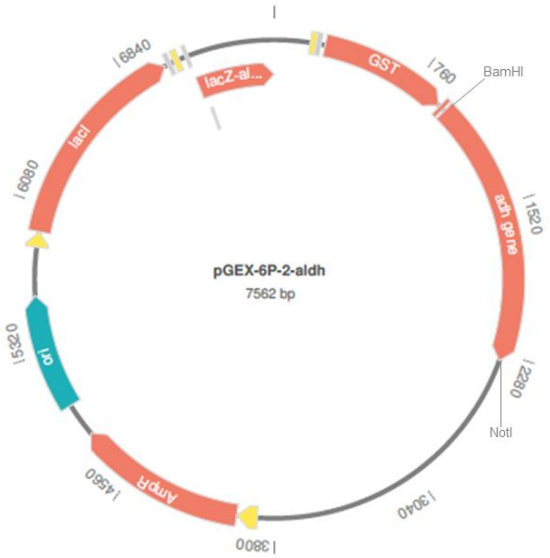

Figure 8.3 Plasmid map of pGEX-6P-2-*aldh.*

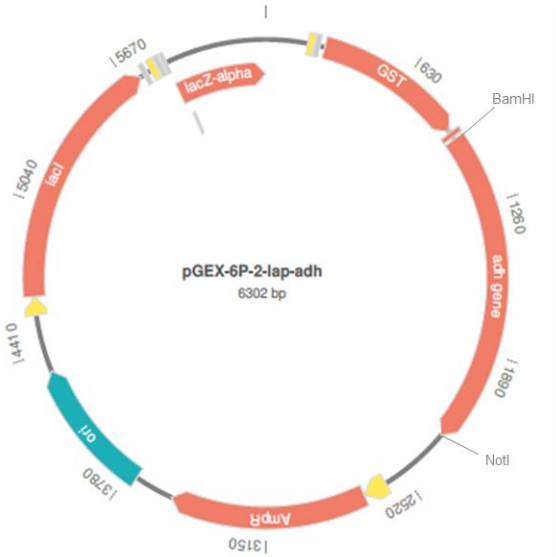

Figure 8.4 Plasmid map of pGEX-6P-2-*adh*.

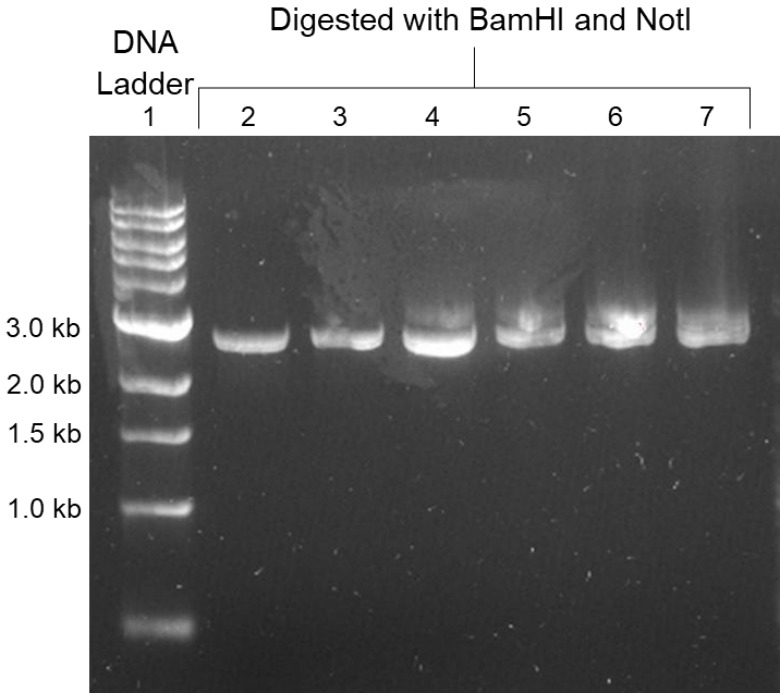

Figure 8.5 Agarose gel analysis of the separation of pUC57 plasmid and *lap*.
Lane 1: 1 kb marker. Lane 2 to 7: Products of the digestion of the double digested pUC57-*lap* with BamHI and NotI. Analysed at a voltage of 50 V for 2.5 h. 2% (w/v) agarose in 1xTAE buffer with 1x SYBR safe.

.