cloning,expression, purification and characterization of a sulfate transporter from legionella pneumophila

jayaram lamsal

Norwegian University of Life Sciences
Department of chemistry, biotechnology and food science, IKBM
Master Thesis 60 credits 2013

ACKNOWLEDGMENTS

The work described in this thesis was performed at the Centre for Molecular Medicine Norway (NCMM), University of Oslo with Professor J. Preben Morth as co-supervisors.

I wish to thank all the people who helped me in my work. Without their support, I would have not been able to bring my thesis to successful completion.

First of all, I would like to thank Harmonie Perdreau for all help, good advice and enthusiastic guidance. Her great efforts and ideas have been of great value for me, especially during the practical laboratory work.

Special thanks to J. Preben Morth, for giving me a great opportunity to worked at NCMM. His great scientific insights and ideas are very stimulating for my scientific writing.

I would like to express my sincere gratitude to my supervisor Vincent Eijsink for his positive attitude and all the support on the way.

Finally, I am deeply thankful to all the people that have helped me for providing a great working environment during my laboratory work.

Oslo, May 2013

JAYARAM LAMSAL

SUMMARY

Legionella pneumophila is a gram-negative, non-encapsulated, rod-shaped facultative intracellular pathogen with a single, polar flagellum. The SO_4^{2-} transporter protein (LPL0734) is a membrane protein of*Legionella pneumophila.*The LPL0734 protein has 12 trans-membrane helices and consists of 768 residues with the expected molecular weight of 84 kDa. LPL0734 belong to a sulfate transporter family. In this thesis, we attempt to study the characteristics of LPL0734; a suspected sulfate transporter in *Legionella pneumophilla*by cloning, expressing, purifying and crystallizing this protein.Stability testing was also conducted.

In the first part of this research, we focused on the expression of a GFP-tagged SO_4^{2-} transporter in*Escherichiacoli*strains (Rosetta-II and C43 cells) and then purification of proteins (LPL0734) by metal-affinity chromatography on Ni-column followed by size-exclusion chromatography. Protein lipidation and detergent tests were also performed in order to check the stability of the proteins. In the second part of the research, we focused on cloning of the DNA encoding the SO_4^{2-}transporter proteins of *Legionella pneumophilia* into pETM11 expression vector.

The expressed protein showed aggregation during the size-exclusion chromatography step of purification.Therefore, crystallization was not possible. Result from stability test showed the LPL0734 protein was not stable at room temperature even after one day of incubation. This protein was more stable at 4^0C independent of the detergent conditions. Increased aggregation and degradationof the protein were observed when the lipid content was high. The result suggested that there is a need to optimize the temperature and buffer-detergent composition to improve the protein stability. It is worth noting that biochemical studies of LPL0734 have not come out with any clear results.

ABBREVIATIONS

AMP	Ampicillin
β-DDM	β-n-dodecyl -D-maltoside
$C_{12}E_8$	octaethylene glycol monododecyl ether
DM	Decyl Maltoside
CAM	Chloramphenicol
CV	Column Volume
Da	Dalton
DNA	Deoxyribonucleic acid
DOPC	1, 2-dioleoyl-sn-glycero-3-phosphocholin
GF	Gel Filtration
GFP	Green Fluorescent Protein
HCL	Hydrochloric acid
HCO_3^-	Bicarbonate ion
HPH	High Pressure Homogenizer
IPTG	Isopropyl B-D-1thiogalactopyranoside
Kb	kilobase
kDa	Kilo Dalton
KNA	Kanamycin
LB	Luria-Bertani
MCS	Multiple cloning sites

MgCl2	Magnesium chloride
MW	Molecular weight
MWCO	molecular weight cutoff
NaCl	Sodium chloride
NEB	New England Biolab
ng	Nanogram
Ni	Nickel
OD	Optimal density
O/N	Overnight
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PES	polyethersulfone
PMSF	Phenylmethanesulphonylfluoride
rpm	Revolutions per minute
RT	Room temperature
SDS-PAGE	Sodium dodecyl sulphate polyacryl amide gel electrophoresis
SEC	size-exclusion chromatography
SOC	Super optimal broth with catabolite repressor
SO_4^{2-}	Sulfate ion
W/V	weight/volume

CONTENTS

1. INTRODUCTION

The cell is the basic structure and fundamental building block of all living organisms. Cells are surrounded by an impermeable lipidic membrane consists of a phospholipid bilayer (Figure 1.1). The membrane prevents leakage of internal components, and it also function as a physical barrier that protects the cell from foreign molecules(Sherwood 2012). In order to maintain life, a cell must be able to communicate with the surroundings environment and control the uptake of nutrients, waste disposal and ion conductance(Goodsell 2009).The key players in these communication processes are membrane proteins that are responsible for regulating the permeability of the membranes.

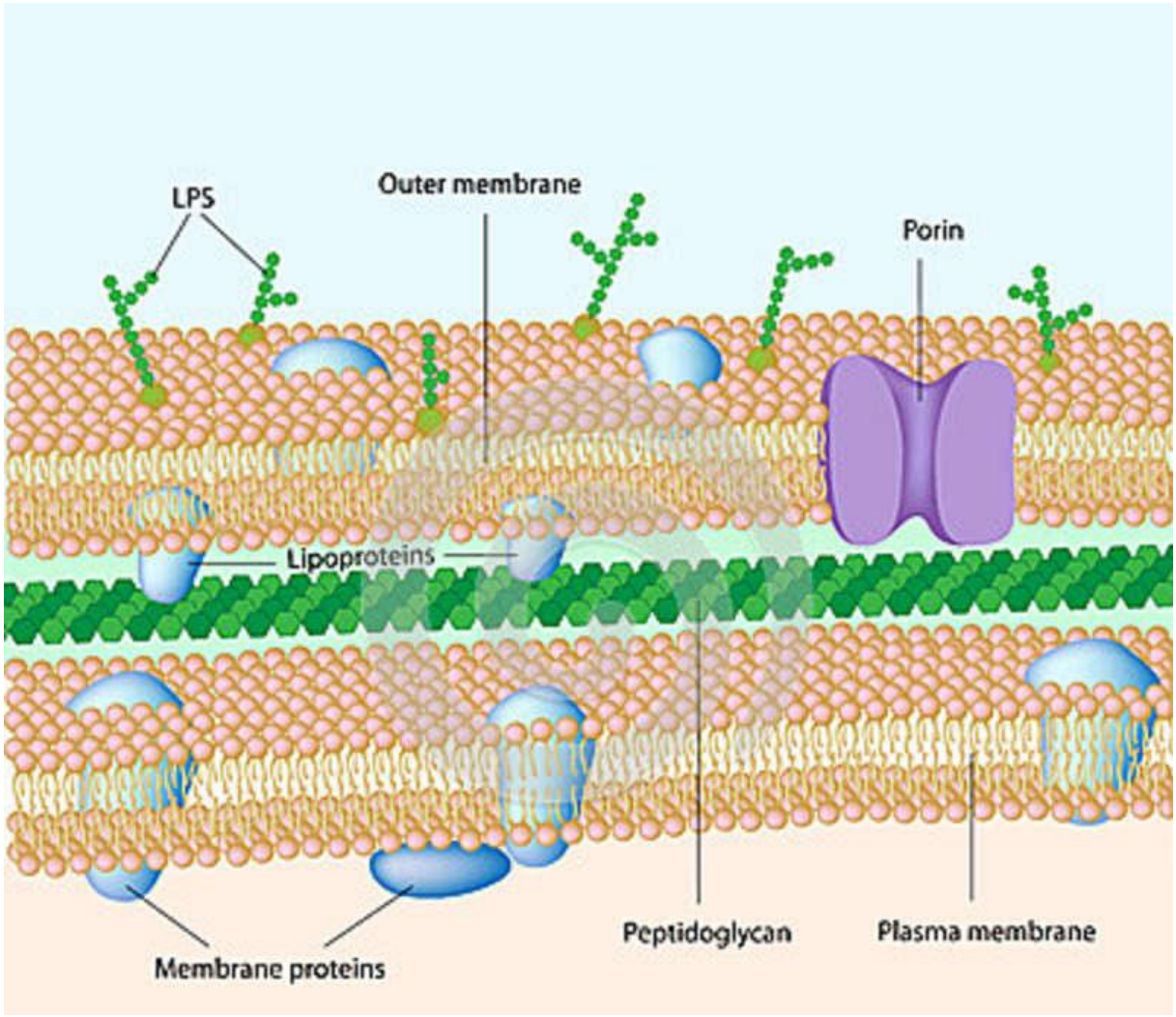

Figure 1.1: Schematic diagram of a lipid bilayer with different kinds of lipids and associated membrane proteins of gram negative bacteria. (Beveridge 1999)

1.1 Membrane protein

Membrane protein play important roles in a wide range of functions in cells including transport of chemical species into and out of cells (such as Na+- K+ channels), helps to communicate cells with their environments. Membrane proteins also helps to know immunerecognized cells, whether it is foreign or not, for hormones or other signaling molecules (Hanke and Schulue 1993). From a pharmaceutical point of view, half of all drugs target are membrane proteins. Therefore, the studies of membrane proteins are of great importance to the chemists (Psakis, Nitschkowski et al. 2007).

The diversity of function of membrane proteins was mirrored in the great variability in the three-dimensional structures. Determination of the three-dimensional structures would facilitate the assignment of the functions of the protein(Jackson and Mantsch 1995). For any protein, this is a demanding procedure but has turned out to be considerably more difficult for membrane proteins than for globular proteins. The fact is that only just above 100 structures of membrane proteins have been solved (Surade, Klein et al. 2006), which is in contrast 40000 of solved structures for globular proteins deposited in the Protein Data Bank (PDB) (Berman, Westbrook et al. 2000). Why are membrane proteins so challenging? The major reason behind this is the membrane proteins interaction withthe membrane lipids that are necessary for correct folding. Without the amphipathic lipid molecules, a membrane protein does not fold into its native structure. There are many other factors that obstruct the production and analysis of membrane protein such ashydrophobicity of membrane proteins, general instability under diverse conditions of buffer-detergent composition, pH and temperatures(Alexandrov, Mileni et al. 2008). The aim of the present study was to achieve cloning, expression, detergent screening, purification and characterization of membrane protein (LPL0734) from *Legionella pneumophila. Legionellapneumophila* is a causative agent of legionellosis; a fatal, acute respiratory infections. Gaining insight into the structure-function relationship of such target proteins can be useful tools for future drug design.

1.2 *Legionella pneumophila*

Legionella pneumophila, the causative agent of Legionnaires diseases, is a gram-negative, non-encapsulated, strictly aerobic bacillus with a single, polar flagellum (Figure 1.2). This disease was first described in July 1976, when an outbreak of severe pneumonia occurred among people attending at an American Legion convention in Philadelphia(Diederen 2008). *Legionella pneumophila* is mainly a water-bornedisease that exists in natural fresh water. These bacteria found mostly as a parasite of protozoa, its natural hosts, but it can also be associated with biofilms. Protozoa provide both nutrients and protection for *Legionella pneumophila* that can survive in harsh environment conditions and also make the bacteria more virulent.The bacterium enters the human body by aerosol droplets that exist and multiplies within human macrophages and the alveolar epithelium. These facultative intracellular parasites elude macrophages defense by inhibiting the oxygen-dependent sequelae of phagocytosis and blocking phagosome-lysosome fusion (High, Torosian et al. 1993). Legionnaires' disease cause severe pneumonia with symptoms like fever, cough, chest pain, breathlessness and diarrhea.

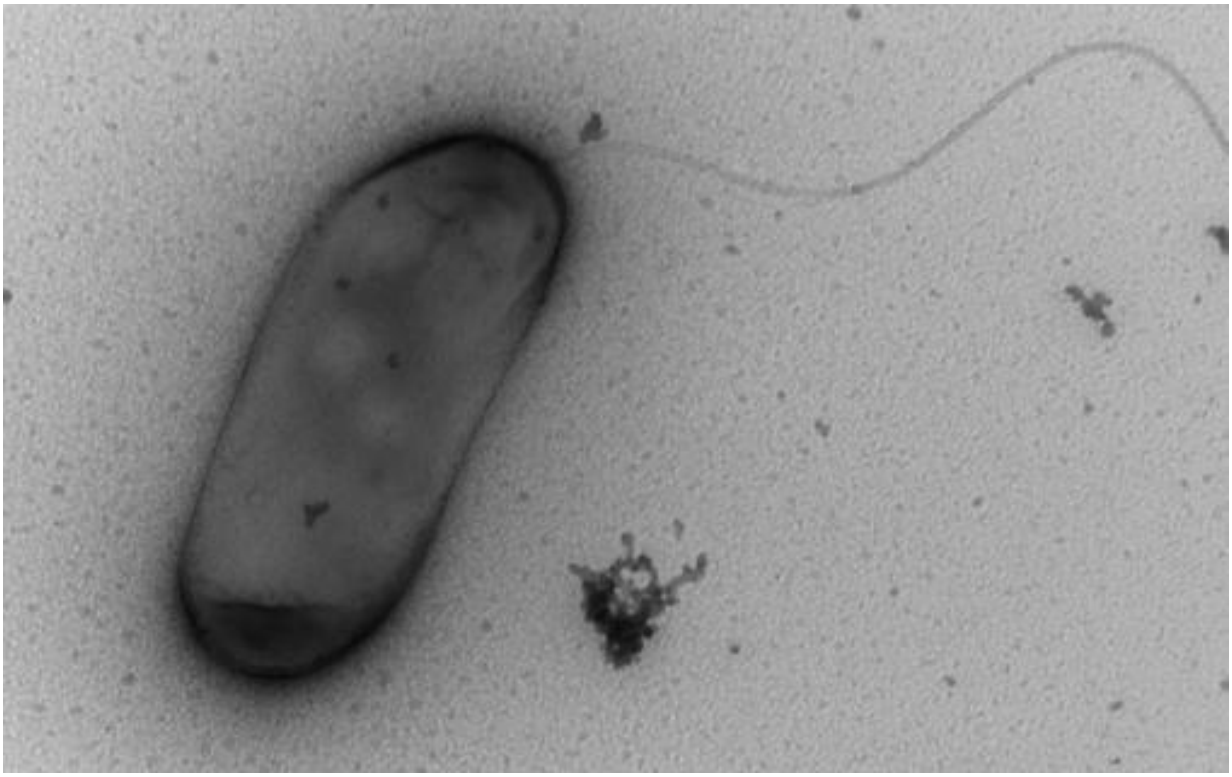

Figure 1.2: Electron microscopy showing the structure of *Legionella pneumophila* (Cazalet, Gomez-Valero et al. 2010).

The pathogenesis of the bacteria depends upon the number of infecting bacteria, the route of entry, the effects of host defense mechanisms and virulence factors. Bacterial cell surface proteins have also been associated with the mechanisms of pathogenicity of gram-negative bacteria. These observations suggest that preexisting surface proteins of the *Legionella pneumophila*may participate in the pathogenesis. However, the mechanisms governing the infectious processes of *Legionella pneumophila* are ill defined. The presence of different structure and membrane components of the bacteria may play important roles in the process of adherence and uptake(Rodgers and Davey 1982).

The genome of the *Legionella pneumophila* contains a gene encoding putative bi-functional transporter (LPL0734) (Figure 1.3).This is a membrane protein with 12 trans-membrane helices. The protein(LPL0734) consists of 768 residues with the expected molecular weight of 84,120 Da.Functionally the putative protein (LPL0734)could similar to inorganic transporter and to carbonic anhydrase (bi-functional).The LPL0734 protein belongs to a Sulfate transporter family. Sulphur is essential for some of the most vital biological activities such as genes involved in sulphur metabolism;translation initiation and redox maintenance have been implicated in virulence. Lesions in sulphur metabolism have been shown in *Brucella melitensis*(Lestrate, Delrue et al. 2000).Sulfur being essential for the biosynthesis of cysteine and methionine, and also involved in protein structure through disulphide bonds. Given the central role of sulphur metabolism, we decided to investigate this area in the *Legionella pneumophil*. However, the mechanism behind this characteristic will not be revealed until thecrystal structure of LPL0734 is solved.In order to provide insights into its biochemical function, we tried to overexpress, purify, characterize and crystallizeLPL0734. However, the biochemical properties of this protein have not been reported.

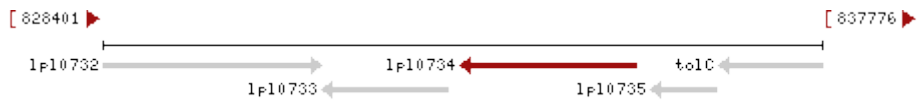

Figure 1.3: Genomic region around gene locus LPL0734

To study the potential role of the LPL0734in the physiologyof *L.pneumophila,*we required to purifyingthis protein. Here, the target genesare covalently linked to green fluorescent protein (GFP: ~27 kDa) (Kawate and Gouaux 2006). The terminuses of GFP are fused with a polyhistidine tag for affinity purification and thrombin site for proteolytic cleavage of the target protein from GFP (Figure 1.4). The covalently fused GFP constructs allow one to determine protein expression in bacterial cells by fluorescence.

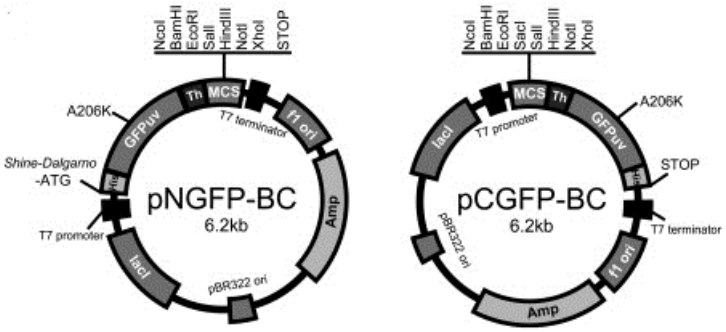

Figure 1.4: Bacterial expression vectors (pNGFP-BC and pCGFP-BC).

1.3 Green Fluorescent protein

The green fluorescent protein first isolated from jellyfish Aequorea victoria becomes one of the most widely studied and exploited proteins in biochemistry. The GFP protein composed of 238 amino acid residues with the length of 26.9 kDa. This protein gives bright green fluorescence when exposed to ultraviolet light(Prendergast and Mann 1978). In molecular biology, the GFP gene is a widely used reporter of expression. This is done by introducing the GFP gene into organisms and maintained their genome. To date, the GFP gene have been introduced and expressed in many bacteria, plant and also in mammalian cells including human. GFP has become a well-established marker for visualizing a variety of intracellular proteins, and also used for monitoring gene expression as well as for various physiological processes.The proteins can be fused to either the N-terminus of GFP protein or C-terminus of GFP.In this present study, the target genes are covalently linked to GFP protein and studied its properties.

1.4 Physiological roles of sulfate transporters

Every organismrequired essential chemicals and elements for their normal growth and functions. The different nutritional elements that are required for bacterial cells consist of Carbon, Hydrogen, Oxygen, Nitrogen, Sulphur, Phosphorus, Potassium and Calcium. These molecules are in the form of inorganic ions that serve either a structural or functional roles in the cells. Among them, sulfate is one of the most important elements in cells. Sulfate is one of the major sulphur sources in many organisms. Cells required sulfate transporters protein that is involved in moving sulfate into or out of it. This is the fact that sulphur being a hydrophilic anion that could not passively cross the lipid bilayer of cell membranes(Markovich 2011).Sulfate is required for the maintenance of cell membranes. Recently, molecules have been identified that facilitate cellular sulfate transporter to/from the extracellular environment (Markovich 2001).This type of trans-membrane movement regulates the content of cellular sulfate, required for biological processes, and they would also able to control sulfate homeostasis in the cells. With the use of molecular and cell biological techniques, many families of the sulfate transporters of bacterium have been cloned in the past few years.An energy-dependent process carries out the first step in the sulfate uptake into the cell by specific protein known as sulfate permease subsequently activated by ATP forming adenosine-5-phosphosulfate (APS), which is further phosphorylated into 3'-phosphoadenosine-5-phosphosulfate (PAPS). In bacteria PAPS subsequently reduced to sulfide ion, which is then assimilation to form cysteine and methionine(Figure 1.5)(Piłsyk and Paszewski 2009)

Figure 1.5: General view of the sulfate assimilation pathway.

Why sulphur is important? Sulphur isthe essential element that all organisms require for growth. In bacteria, sulphur accounts to 0.5-1% of the cell dry weight and is needed primarily as a component of the aminoacid(Kertesz 2000) such as methionine and cysteine. Sulphuris also involved in protein structure through disulphide bonds(Tina, Bhadra et al. 2007). In

addition to this, sulphur also involved in a wide range of enzyme cofactors, including biotin, coenzyme M, coenzyme A, lipoic acid and thiamine. Sulphur is also common in the environment that makes up 0.1% of the earth crust, but this material is much moreinaccessible to living organisms(Kertesz 2000).

1.5 Carbonic anhydrases

Among the naturally occurring chemical compounds carbon dioxide(CO_2) is a key metabolite in all prokaryotes and eukaryotes. Carbon dioxide exists in equilibrium with bicarbonate. Bicarbonate is the most physiological important and is negatively charged and is highly soluble in aqueous solution but poorly soluble in lipids. While carbon dioxide is highly soluble in both lipids and aqueous that can freely diffuse in and out of the cells but the HCO_3^- must be transported across the cell membrane. At physiological pH the inter-conversion of carbon dioxide and bicarbonate proceeds slowly, so organisms produce enzymes that facilitate the reactions. Carbonic anhydrase are zinc-containing enzymes that catalyze the reversible reaction between carbon dioxide and bicarbonate (Figure 1.6). Carbonic anhydrase plays essential roles in facilitating the transport of CO_2 and protons in the intracellular space, across biological membranes.The active site of most carbonic anhydrases contains a zinc ion that can exist in two forms; at high pH that is active in the hydration of carbon dioxide, where as at low pH that is active in the dehydration of bicarbonate(Chegwidden and Carter 2000).

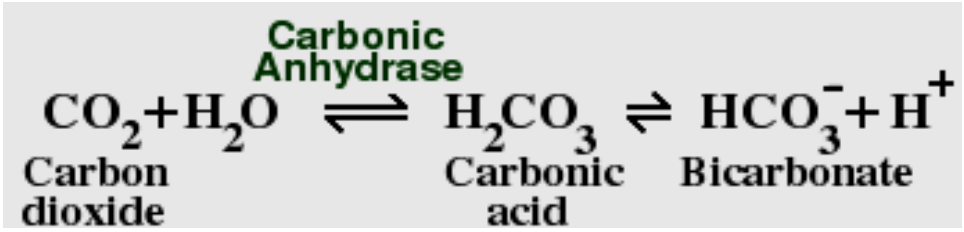

Figure 1.6: The reaction catalyzed by carbonic anhydrase.

In animals carbonic anhydrase takes part in a truly remarkable range of the physiological process including acid-base balance, respiration, biosynthetic pathway and variety of processes involving ion, gas and fluid transfer(Krishnamurthy, Kaufman et al. 2008).

The Carbonic anhydrase is a ubiquitous enzyme that nature has evolved the catalytic capacity to hydrate carbon dioxide and dehydrate bicarbonate several times.This enzyme has been found in all living organisms. Three classes of carbonic anhydrase enzymes;alpha, beta and gamma have been recognized. They have no significant sequence identity and have distinct folds in structures(Pocker 2000).Structural differences have been found in all three enzyme but the active sites of all three classes function with a single zinc atom.The different classes of carbonic anhydrase have differing in distributions. In mammals, the isoenzyme belongs to the alpha class; in plants, mainly the beta-class; where as prokaryotes encode all three classes of enzymes. In prokaryotes beta and gamma classes are predominant(Smith, Jakubzick et al. 1999).

Most of the researches have focused on plant or mammalian carbonic anhydrases with little to prokaryotic enzymes. The gamma class carbonic anhydrase has been isolated and characterized fromarchaea called*Methanosarcina thermophile* (Alber and Ferry 1994).Most of the prokaryotes have two or even all three classes of carbonic anhydrase. The presence of multiple carbonic anhydrase suggests that this enzymehas a major role in the prokaryotes(Smith and Ferry 2000). However the role of carbonic anhydrase isstill largelyunknown.

In prokaryotes, very few carbonic anhydrase has been purified from the small number of species of bacteria domain since it was first identified in *Neisseria sicca* in 1963(Gill, Fedorka-Cray et al. 1984; Shekh, Krishnamurthi et al. 2012). The first carbonic anhydrase activities were detected in strains of *Neisseria* and *Streptococcus salivarius*. Thee carbonic anhydrase activity from the *Neisseria sicca* strain was purified and found similar properties to that of human carbonic anhydrase(Shekh, Krishnamurthi et al. 2012).

1.6 Roles of Carbonic Anhydrase

There are two general roles have been suggested for Carbonic anhydrase. They are (i) transport of carbon dioxide or bicarbonate and (ii) to provide carbon dioxide or bicarbonate for enzymatic reaction. The carbonic anhydrase activity was also detected in some prokaryotes that produce acetate as an end product of fermentation(Braus-Stromeyer, Schnappauf et al. 1997). The carbonic anhydrase was also detected in cyanobacteria that play

important roles in co_2 fixation(Fukuzawa, Suzuki et al. 1992). The carbonic anhydrase isozymes played several important physiological and physio-pathological functions in different organisms.In animals; carbonic anhydrase plays several roles such as; transport of CO_2/HCO_3^-between tissues and the lungs, pH and CO_2homeostasis, biosynthetic reactions, such as the gluconeogenesis, in plant and algae; CO_2 fixation (Supuran 2003). Clinically, the presence of carbonic anhydrase enzymes in so different isoforms in different organismsthat helps to design of drugs that inhibittheir function. In the recent medicine carbonic anhydrase inhibitors are used as antitumor agents/diagnostic tools for tumors, antiglaucoma drugs, anticonvulsants and antifungal/antimicrobials. Carbonic anhydrase inhibitors drugs target to the pathogenic organisms such as *Mycobacterium tuberculosis, candida albicans* and *Helicobacter pylori*(Supuran 2007). The primary carbonic anhydrase inhibitors for bacteria are sulfonamides drugs.Many bacteria contain putative carbonic anhydrase genes from more than one class.The differences of carbonic anhydrase's activity in many bacteria underscore their importance in bacterial physiology.

1.7 Cloning

In addition, the aim of this study was also to characterize SO_4^{2-} transporter gene from *Legionella pneumophila*and cloned into expression vector. Investigations of the membrane protein of *Legionella pneumophila* are crucial to determining pathogenicity at the cellular level but have been hampered by difficulties in separating cell surface components. *C*omplete PCR-derived DNA fragments containing the structural genes for $\mathbf{SO_4^{2-}/HCO_3^-}$ transporter (LPL0734) of the *Legionella pneumophila* were cloned into an expression vector (pETM11)(Dümmler, Lawrence et al. 2005). The length of DNA fragments is 2307 bp, which encodes 84.12 kDaproteins.The clones expressing vectorincorporates His_6-tag and a TEV protease cleavage site N-terminal to the target construct.The pETM11 vectors are derived from pET (Novagene) backbone.These vectors have important features, which consists of a 6xHis-tag, a protease recognition site and the functional multiple cloning sites (MCS) starting with Ncol recognition site (Figure 1.5). The Ncol sequence has ATG codon (start codon) that can be used for the functional expression of the target protein. The functional multiple cloning sites assure that the same couple of restriction sites inserted in the PCR product, which can be used for direct sub-cloning in other vectors. The 6xHis tag is suitable for metal-affinity purification. During affinity purification process the His tagged protein binds very tightly to

nickel ions so; His tagged proteins are purified on a column.Using imidazole gradient then elutes the purified protein. Imidazole has a higher affinity for nickel than that of histamine.

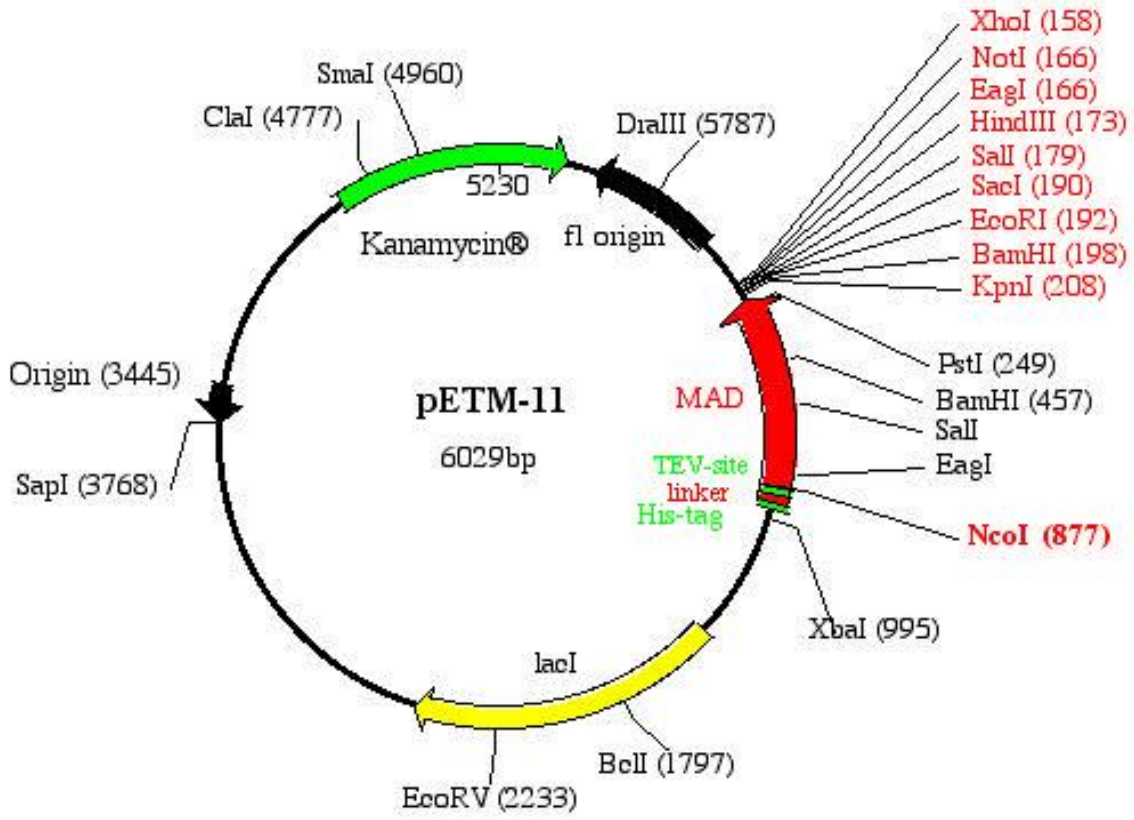

Figure 1.7: The systemic diagram of pETM11 vector:The multiple cloning sites composed of a number of restriction enzyme sites was preserved for insertion of target genes of protein aimed to be studied.

The sulphur assimilation pathway plays a role in the pathogenesis and survival of some bacteria (Hatzios and Bertozzi 2011). Therefore sulfate transporters are essential for the survival and possibly in the pathogenesis of *Legionella pneumophilla*. Given the importance of sulfate transporters in *Legionella pneumophilla*, detail knowledge of these proteins could result in the development of new drugs to treat legionellosis caused by *Legionella pneumophilla*. Hence the main objective of this thesis was to investigate the characteristic properties of LPL0734; a suspected sulfate transporter in *Legionella pneumophila.*

2. MATERIALS AND METHODS

Laboratory equipment

Automatic pipettes (thermo labsystems)

Disposable cuvettes, 1.5 m (Brand)

Ultracentrifuge tubes (Beckman specifications)

Various glass equipment's

Vortex-machine (IKA)

Water baths (julabo)

Instruments

AKTA prime system (AKTAprime^{TM}PLUS)

Centrifuge JA 25.50 (Beckman-NCMM)

Electrophoresis equipment (Bio-rad)

Agilent 1260-machine (column Tosho TSK)

Gel electrophoresis (Bio-rad)

High Pressure Homogenizer (Emulsiflex C3)

PCR-machine (Techne, TC-5000)

Table centrifuge (Biofuge)

Ti45 rotor (Beckman-Biotek)

Chemidoc XRS+ (Bio-rad)

Nanodrop (Thermo)

Chemicals

Agar (Cambrex)

Ampicillin (Sigma)

Bromophenol blue (Sigma)

$C_{12}E_8$ (Affymetrix)

DDM (Affymetrix)

Ethanol (Arcus)

HCL (Mecrk)

NaCl (Sigma)

Glycerol (Prolabs)

Kanamycin (Sigma)

$MgCl_2$(Merck)

Primers:

Primers designed for the mutagenesis and the subcloning of *Legionella* Sulfate/Bicarbonate transporter: (Harmonie)

Mut-SO-Legio-S: 5'- cagccgt actccacggt tttttcatct tg -3'

Mut-SO-Legio-AS: 5'-ca agatgaaaaa accgtggagt acggctg -3'

(Mismatch $Tm_{14/29\ (-1)}$= 59°C or Tm_{IDT}= 62°C)

Forward SO-Legio-FP: 5'- gcaccatggct attgata gaaatatagt taaccagc -3'

($Tm1_{7/25}$= 51°C; $Tm2_{14/36}$= 62°C)

Reverse: 5'- cttttag aagccaaagc aagt taactcgagacc -3'

Complement SO-Legio-RP: 5'- ggtctcgagtta actt gctttggctt ctaaaag -3'

($Tm1_{8/24}$= 52.5°C; $Tm2_{14/33}$= 62°C)

Bacterial strains

DH10B and XLl-blue chemical competent cells

Rosetta-II and C43 (DE3)competent cells.

Plasmids

PCR-Blunt-TOPO (Vectors for cloning of PCR fragments; Kanamycin resistance)

Expresssion vectors; pNGFP-BC, pCGFP-BC, and pETM11

2.1 Expression of a GFP-tagged *Legionella* $SO4^{2-}$ transporter in bacteria

The DNA encoding the*Legionella pneumophila* SO_4^{2-} transporter (LPL0734)was cloned into pNGFP-BC vector with a N-terminal fusion and into pCGFP-BC vector with a C-terminal fusion (Kawate and Gouaux 2006) (Figure 2.1) consisting ofagreen fluorescent protein (GFP), a 8 *x* His tag and a thrombin protease cleavage site (cloning done by Harmonie). The resulting constructSO_4 transporter/pNGFP-BC (30 ng) or SO_4 transporter/pCGFP-BC (30 ng) was transformed into chemically competent*Escherichia coli* Rosetta II (DE3) cells. Transformation was done using the heat-shock method with a water-bath at 42^{o}C for 45 seconds. After heat shock, 250µl of room temperature(RT)Super Optimal broth with catabolite repression (SOC) medium was added to the transformed cells andthe cellswere shaken horizontally at 200 rpm for 1 hour at 37^{o}C. 150µl of transformed cells were spread on a pre-warmed ampicillin (AMP, vector's resistance) + chloramphenicol (CAM, Rosetta II cells resistance) plate and incubated overnight at 37^{o}C.

5-10 colonies from the transformed plate were taken and grown in Luria-Bertani (LB)medium containing 100-µg ml^{-1}AMP and 100-µg ml^{-1}CAM at 37^{o}C, 200 rpm, as an overnight culture. On the next morning, LB medium with 100 µg ml^{-1}AMP and 100 µg ml^{-1} CAM was inoculated with 1:100 of the overnight culture. Inoculate was incubated at 37 ^{o}C, 180 rpm, until the OD_{600} reached 0.6. When theOD_{600}reached0.6 (t_0), the culture was cooled down on ice for 1 hour and isopropyl B-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM to induce expression. The culture was grown overnight at 18^{o}C, 180

rpm. The induced culture(t_{IN}) was pelleted down at 4000 rpm for 15 minutes and the pellet was stored at -20^{o}C.

2.2 Small-scale analysis (Lysis, electrophoresis and GFP-picture)

Samples of 1 ml were taken during the procedure for both non-induced (t_0) and induced (t_{IN}) times. The samples were centrifuged for 5 min at 13,000 rpm and the small pellets were studied to check the expression of the SO_4 transporter. The pelletswere washed by re-suspending them with 200 µl ice-cold washing buffer and spun for 5 minutes at 13,000 rpm. Then, the supernatants wereremoved; the pellets were re-suspended in (50 µl x OD_{600}) ice-cold lysis buffer and lysed by sonication for 30 min in a cold water-bath.4×Laemmle loading buffer was added to a final concentration of 1x.The samples werewell resuspended and incubated for 3 minutes at RT.The tubes were centrifuged for 15 minutes at 13,000 rpm in order to precipitate most of the DNA. 10 µl of t_0 and t_{IN} samples with 5µl of standard pre-stained protein ladder (PageRuler™ Prestained Protein Ladder,Thermo Scientific) (Figure 2.1) were loaded on a 12% SDS-PAGE gel and run at 90 V for 15 minutes until the samples run as a single line and then 1 hour 30 minutes at 120 V (we used a low-voltage to avoid the thermal denaturation of the GFP fluorescence).

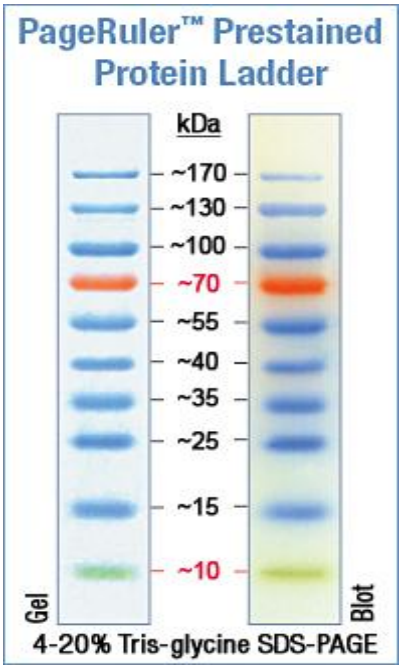

Figure 2.1: Page Ruler™ Pre-stained Protein Ladder for monitoring protein separation during SDS-polyacrylamide gel electrophoresis.

2.3 Study of thegel with a GFP-filter

The Green FluorescentProtein (GFP) is a 27 kDa protein and the SO_4^{2-} transporterfrom*Legionella* is a84 kDa membrane protein. So, the target protein band was almost equal to 110 kDa.The fluorescent bandin the gel was viewed on Chemidoc XRS+ with a GFP filter (program-GFP applic. high sens.), where the size of the protein was estimated by the comparison of its migration distance with the known molecular weight marker.

2.4 Coomassie blue staining

After the GFP detection, the SDS-gel was stained with Coomassie blue staining solution for 15-60 minutes and then destained overnight with 10%acetic acid and we took the picture of the gel. Then the size of the protein was estimated on Chemidoc XRS+.After a band of 110 kDa wasdetected in the gel, we continued for a large-scale analysis as follow:

2.5 Membrane protein isolation

Following harvesting, the cell pellet was washed in 50ml/liter (of culture) of ice-cold washing buffer(20 mM Tris-HCl pH 8, 100mM NaCl, 5% glycerol and 1mM phenylmethanesulphonylfluoride (PMSF)) by resuspension and spinning at 4000 rpm for 15 minutes (JS 4.0 rotor, Beckman-NCMM) and the supernatant was removed. The cells were further resuspended in 20ml/liter of ice-cold lysis buffer (50 mM Tris-HCl pH 8, 100mM NaCl, 5% glycerol, 5 µg/ml DNAse I and 1mM PMSF). The cells were disrupted using a HighPressure Homogenizer (HPH,Emulsiflex C3) at 15,000 psiand the crude lysate cells were centrifuged for 20 min at 18,000 rpm, 4 ^{o}C(JA 25.50, Beckman-NCMM), to remove cellular debris, inclusion bodies and DNA.The clarified supernatant was transferred to 70 ml ultracentrifuge tubes (Beckman specifications) and the membranes were isolated by 1 hour 15 minutes of ultracentrifugation at 45,000rpm,4^{o}C (Ti45 rotor, Beckman-Biotek). The ‘soft’ membrane pellet was resuspended in 70 ml ice-cold washing buffer, using a Dounce homogenizer.The membranes were ultra-centrifuged again for 1hr 15 min at 45,000 rpm, 4^{o}C, the supernatant was removed and the membrane pellet was resuspended in resuspensionbuffer

(20 mM Tris-HCl pH 8, 100mM NaCl, 10% glycerol, 5mM β-mercaptoethanol, 1 mM $MgCl_2$ and 1mM PMSF)to a total protein concentration of~ 10 mg/ml.

2.6Membrane protein solubilization

Membrane proteins were solubilized by adding 0.5 % w/v β-n-dodecyl -D-maltoside (β-DDM) detergent and incubated with mild agitation using magnetic bead and stirrer overnight at 4^oC in the cold-room. On the next morning, the solubilized membranes were poured in a 70 ml ultracentrifuge tube (Beckman specifications) and centrifuged by 1 hour 15 minutes of ultracentrifugation at 45,000 rpm, 4^oC (Ti45 rotor, Beckman - Biotek). Unsolubilized membranes were pelleted down and the supernatant (solubilized membranes) were taken for further purification.

2.7 Membrane protein purification

The histidine-tagged membrane proteins were purified using the AKTA prime system and a Ni^{2+}-column (His-Trap FF crude, 5 ml column, GE healthcare). The AKTA primesystem was washed with filtered water. Ethanol from the column was washed out8 Column Volumes (CV) of filtered water (40 ml) at 2.5 ml/min. The column was equilibrated with 8 CV of filtered buffer A(20 mM Tris-HCl pH 8, 100mM NaCl, 10% glycerol, 5mM β-mercaptoethanol, 1 mM $MgCl_2$, 1mM β-DDM, 20 mM imidazole) at 2.5 ml/min. 20 mM imidazole was added to the sample to remove unspecific binding of proteins. The sample was loaded on the Ni-column at 2.5 ml/min until the sample was almost totally finished, and the Flow-Through was collected. The column was then washedwith 8 CV filtered buffer A at 2.5 ml/min;the wash was collected in a new bottle. The His-tagged proteins were eluted with a gradient from 0 to 100 % filtered buffer B (same composition as buffer A but with 500 mM imidazole) for 8 CV at 2.5 ml/min andelutes were collected in fractions of 1.5 ml each, in 25 glass tubes.

The protein concentrations from the different elute tubes were measured on the Nanodrop (MW: 110 kDa; Ext. coefficient: 49280 $Mol^{-1}cm^{-1}$). 5 µl of each elute was run on a 12% SDS-PAGE gel. The elute tubes having a GFP-tagged protein of 110 kDa were taken and were pooled in a 15ml tube and the concentration of the protein was measured. The pool of our protein was concentrated by using a 50,000 Damolecular weight cutoff (MWCO)

polyethersulfone (PES) ultrafiltration membrane by centrifugation (4000 rpm, xtimes 30 min, 4°C) until the volume reached ~300 µl. The concentration of the concentrated purified protein was measured;the protein was flash-frozen in liquid Nitrogen and conserved at -20^{o}C.

2.8 Fluorescence Size-Exclusion Chromatography

The LPL0734 proteins were separated using molecular size through a size-exclusion chromatography (SEC) or Gel Filtration (GF). The Agilent system 1560 was washed with filtered water.An TSK column (29 ml, TSK G3000SW, Tosoh) was washed with filtered water and equilibrated with freshly prepared filtered gel-filtration buffer (20 mM Tris-HCl pH 8, 100 mM NaCl, 10% glycerol, 1 mM $MgCl_2$, 2.5 mM DTT, 5x CMC β-DDM) at a flow rate of 0.5ml/min for 58 min (2 CV) and with a maximal pressure of 48 bars. The Ni-purified protein sample was thawed on ice and ultra-centrifuged for 20 minutes at 70,000 rpm, 4°C.150 µl of clarified supernatant was transferred to an Agilent tube and we injected three times 50 µl. Each injection was run with 1 CV of buffer at 0.5 ml/min. The elute fractionsof 0.5 ml were collected between 20 and 35 min. The protein concentrations of the different peaks weremeasured on the NanoDrop (MW: 110 kDa; Ext. coeff: 49280 Mol^{-1} cm^{-1}). Running the different elutes on a 12% SDS-PAGE gel assessed the protein homogeneity and purity.

2.9 Protein Lipidation and Detergent screening (Stability test)

To lipidate our protein we used the lipid 1, 2-dioleoyl-sn-glycero-3-phosphocholin (DOPC). The lipid to protein ratio must be determined empirically for each individual membrane protein and each membrane protein batch (Gourdon, Andersen et al. 2011); the ratio ranges tested were 1:0, 1:0.2, 1:0.5, 1:0.8 (µg protein:µg lipid). DOPC is conserved in a chloroform solution and we needed to evaporate chloroform before adding the lipid to our protein, to conserve its properties.

DOPC solution was first added to a glass test tube. Blowing argon gas into the tube at a low flow rate helps to evaporated chloroform, thus preventing oxidation. Our purified protein wasconcentrated to ~ 5 mg $ml^{-1}$and then 20 µlof the proteins were added to each glass tubes and re-lipidated overnight with a small magnetic spin bar on a magnetic stirrer at 4^{o}C.

Insoluble DOPC and aggregated proteins were removed by ultra-centrifugation at 70,000 rpmfor 15 minutes, 4°C (TLA 120.1 rotor, Beckman-NCMM). Additionally, two different detergentsoctaethylene glycol monododecyl ether ($C_{12}E_8$) andDecyl Maltoside (DM), were added to the lipidated protein at a final concentration of 0.5% for a detergent screening.

2.10 Cloning of *Legionella* SO_4^{2-} transporter into pETM11 vector

Firstly, the bacterial expression vector pETM11(Dümmler, Lawrence et al. 2005) *(*Figure 1.5) was studied: Multiple Cloning Site (MCS), restriction enzyme sites available, open reading frame and other sequence sites (N-terminal tag composed of 6-His/TEV cleavage site). Secondly, we verified which restriction enzymes were cutting in the MCS of the expression vector ant not in the DNA sequence of our protein and we also determined a couple of restriction enzymes that we could use for the cloning (*NcoI* and *XhoI* were chosen, after silent mutation of the internal *NcoI* restriction site in the sequence of the SO_4^{2-} transporter). Thirdly, the primers were designed for the mutagenesis and the sub-cloning of the*Legionella* SO_4 transporter.

2.10.1 Mutagenesis of the internal *NcoI* restriction site

The DNA encoding the SO_4^{2-} transporter protein was cloned into a TOPO vector (Harmonie). The TOPO construct was amplified by apolymerase chain reaction (PCR) with a Pfu Ultra High-Fidelity DNA polymerase(NEB) and primers designedto mutate the internal *NcoI* restriction site of the SO_4 transporter (PCR realized on TC5000, MIDSCI Company).The primers were 5'- cagccgt actccacggt tttttcatct tg -3' for the sense primer and 5'- ca agatgaaaaa accgtggagt acggctg -3'for the antisense primer. The amplification product was analyzed by electrophoresis with a 1 kb DNA ladder, on a 0.8% agarose gel stained with 3x Gel Red (Biotium Inc).The PCR product wasthen treated with 2 µl *DpnI* restriction enzyme and incubated for 1 hour at 37°C to digest the parental unmuted DNA. Afterwards, 2 µl of digested PCR productwas transformed into chemically competent DH10Band XL1-blue cells, mixed gently and incubated on ice for 30 min. After the incubation, the cells weretransformed by heat shock in a water bath at 42 °C for 45 seconds. This tube was transferred back on ice

for 2 minutes to reduce damages to the cells. 250 µl of room temperature SOC medium was added to each tube and the cells were shakenhorizontally at 200 rpm for 1 hr at 37 ^{o}C. 250 µl of each transformation was spread on pre-warmed kanamycin plates (KAN = vector's resistance).

Four colonies were screened by PCR with a GoTaq polymerase (references, Promega). The primers were 5'- gca cc atg gct attgata gaaatatagt taaccagc -3' for the forward primer and 5'- ggt ctcgag tta actt gctttggctt ctaaaag -3' for the reverse primer and contained NcoI and XhoI sites respectively. In parallel of the PCR,overnight miniculturesofthe positive colonies were amplified in 5 ml LB medium +100 µg ml^{-1} kanamycin at 37°C, 200 rpm.
Plasmid DNA was purified from the overnight cultures bya MiniPrep (Miniprep kit, Qiagen) and the DNA were eluted in50 µl of sterile water and the concentration was measured on Nanodrop.
Plasmids were checked by restriction analysis with *EcoRI*and were analyzed by electrophoresis on a 0.8% agarose gel stained with 3x Gel Red.Finally,the positive plasmids weresequencedfor their *NcoI* mutated site with M13-FP and M13-RP primers (GATC company: 20 µl of DNA at 50 ng/µl).

2.10.2 PCR and addition of the *XhoI*/*NcoI* restriction sites

Addition of the *XhoI*/*NcoI* restriction sites on a positive clone of Mut-Solfate-Legionella/TOPO by PCR, using the Pfu Ultra Fusion Polymerase. The primers were 5'- gca cc atg gct attcgata gaaatatagt taaccagc -3' for the forward primer and 5'- ggt ctcgag tta actt gctttggctt ctaaaag -3' for the reverse primer and contained NcoI and XhoI sites respectively. PCR products were then purified using a gel purification kit and the DNA were eluted in 30 µl of sterile water and the concentration was measured on Nanodrop.

2.10.3 Integration of SO_4 transporter in the pETM11 vector

The TOPO-cloned insertsand pETM11 vectorwere digested for 1 hour and 30 minutes at 37^{o}C with 10x Reaction Enzyme buffer4, 10x BSA, Restriction Enzyme 1 (*NcoI*), Restriction Enzyme 2 (*XhoI*)(NEB). The restriction enzymes were then heat-denatured for 20 min at 65^{o}C. To avoid its self-ligation, the digested pETM11 vector wasdephosphorylated using 10x

Antarctic Phosphatase Reaction buffer and Antarctic Phosphatase (NEB) and incubated for 1 hour at 37^{o}C. The Antarctic Phosphatase was then heat-denaturedfor 5 min at 65^{o}C. Digested TOPO-cloned inserts and dephosphorylated vector were purified by gel purification (gel purification kit, Qiagen) and were eluted in 30µlsterile water. For the ligation, digested inserts were mixed with the digested vectors with 10x T4 DNA Ligase bufferand T4 DNA Ligase (NEB) in a 1:3 molar ratio (vector: insert ratio) and incubated overnight at 16^{o}C.The ligated products were transformed into chemically competent DH10B and XL1-blue cellsusing the same protocol as described earlier.

3.RESULTS

3.1 Small scale screening of sulfate-transporter

The main goal of doing small-scale expression and purification was to check the highest expression level of LPL0734 in different strains of *E. coli*with two different constructs. The LPL0734 C-terminally tagged construct (pCGFP-BC) (50 ng/µl) and the LPL0734 N-terminaly tagged construct (pNGFP-BC) (38 ng/µl) were transformed into *E. coli* (Rosetta II and C43) at 37 ^{0}C, numbers of colonies were obtained. It was found that C43 have very few colonies (5-7) as compared to the Rosetta cells (30-40). In order to check the expression of the LPL0734, 1 ml samples of both proteins before IPTG induction (t_o) and after IPTG inductions (t_{IN}) for all the constructs have been taken and run on a 12% SDS-PAGE gel and run at 90 V for 15 minutes until the samples run as a single line and then 1 hour 30 minutes at 120 V (we used a low-voltage to avoid the thermal denaturation of the GFP fluorescence)

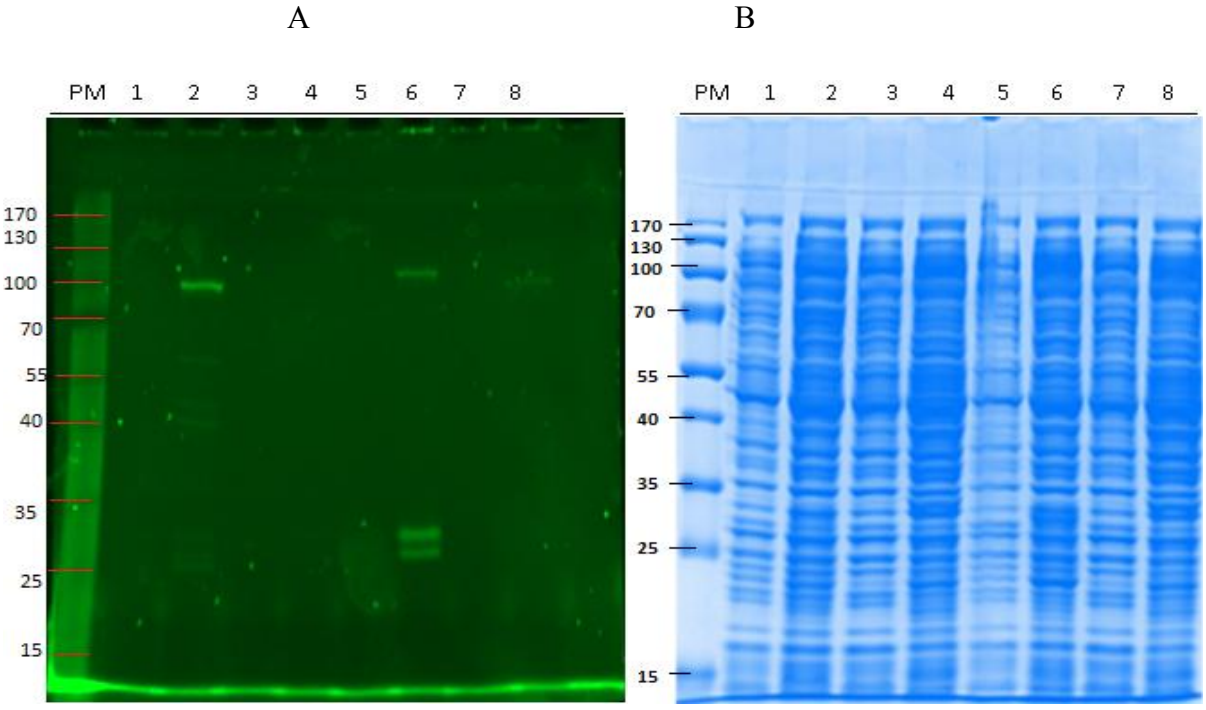

Figure 3.1: The pictures shows SDS-PAGE of LPL0734 recombinant protein before and after IPTG induction. (A) GFP picture: (B) Coomassie blue staining of the same SDS gel. PM, molecular weight markers (labeled in kDa); lane 1 and lane 2, non-induced and induced pCGFP-BC vector into Rosetta II respectively; similarly, lane 3 and lane 4, non-induced and induced LPL0734/pNGFP-BC vector into C43 cells; lane 5 and lane 6, non-induced and induced LPL0734/pNGFP-BC vector into Rosetta II cells; lane 7 and lane 8, non-induced and induced LPL0734/pCGFP-BC vector into C43.

Expression experiments showed good expression using Rosetta II cells for both constructs (in pCGFP-BC and pNGFP-BC) but there was no protein expression using C43 cells. Figure 3.1 shows successful expression of LPL0734 using Rosetta II cells. In this figure we could see the band of size about 27 kDa in both lanes 2 and 6, and this is approximately the same size of the GFP as shown in the figure. We could see the clear band of size about 110 kDa in both lanes 2 and 6, and this is approximately the same size of the LPL0734. Additional bands were also noted in lane 3.

3.2 Membrane protein isolation and solubilization

The LPL0734 protein was expressed in Rosetta II cells. Approximately 50 ml of cell pellets from both constructs (in pCGFP-BC and pNGFP-BC into Rosetta II) were taken for membrane isolation. The cells were opened with a High Pressure Homogenizer at 15,000 psi and cellular debris, inclusion bodies and DNA were removed by centrifugation. The membranes were isolated by ultracentrifugation. The isolated membrane proteins were solubilized O/N by adding 0.5 % w/v β-DDM detergent. On the next morning the solubilized membranes (supernatant) were taken by ultracentrifugation where as the unsolubilized membranes were settles down in the tube-forming pellet.

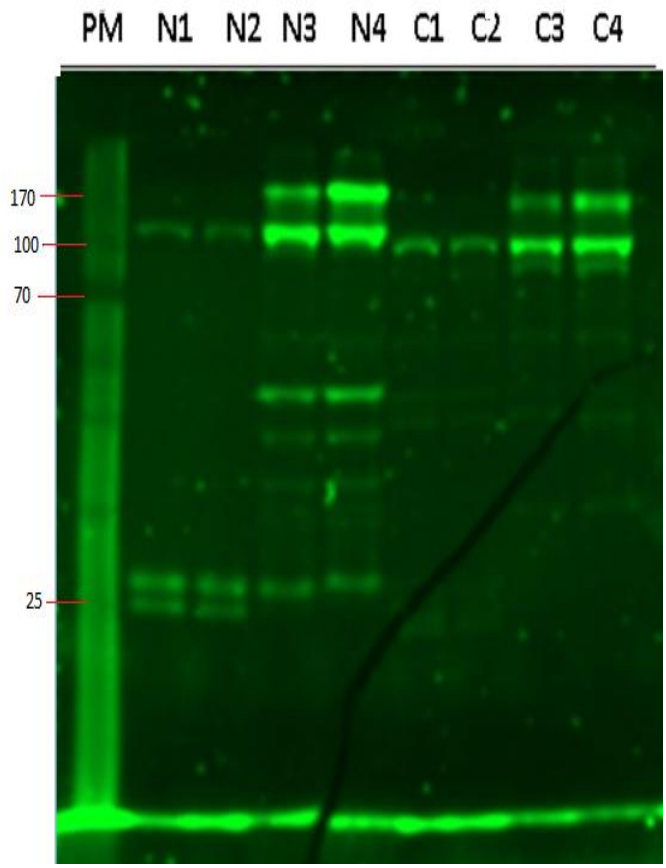

Figure 3.2: GFP picture of isolated membrane and solubilized membrane proteins. PM, molecular-weight markers (labeled in kDa); lane N1 and C1, cells before opening with HPH; lane N2 and C2, lysate cells after HPH; lane N3 and C3, isolated membrane; N4 and C4, solubilized membranes. (Note: N = LPL0735 with N-terminal tag; C = LPL0735 with C-terminal tag)

In this figure we could see the band of size about 110 kDa in all lanes,and this is approximately the same size of the LPL0734.In the lanes of N-terminal tag protein, we could see the clear band of size about 27 kDa and this is approximately the size of GPF tag. Additional bands were also noted in lane 3. From the figure we could also see that the construct LPL0734/pNGFP-BC shows higher fluorescence as compared to pCGFP-BC.For further characterization of LPL0734 protein, we decided to continue with a large-scale analysis of the construct LPL0734/pNGFP-BC as follows.

3.3 Determining protein molecular weight through SDS-PAGE

When protein mobility during SDS-PAGE was graphed against the log of the protein molecular weight, a linear relationship was observed (Figure 3.4). Thus, the rate of protein migration during SDS-PAGE was proportional to the Log of the protein molecular weight. The molecular weight of LPL0734 was determined by comparing its mobility during SDS-PAGE electrophoresis with protein markers of known molecular weights. The protein markers were used to create a standard curve.

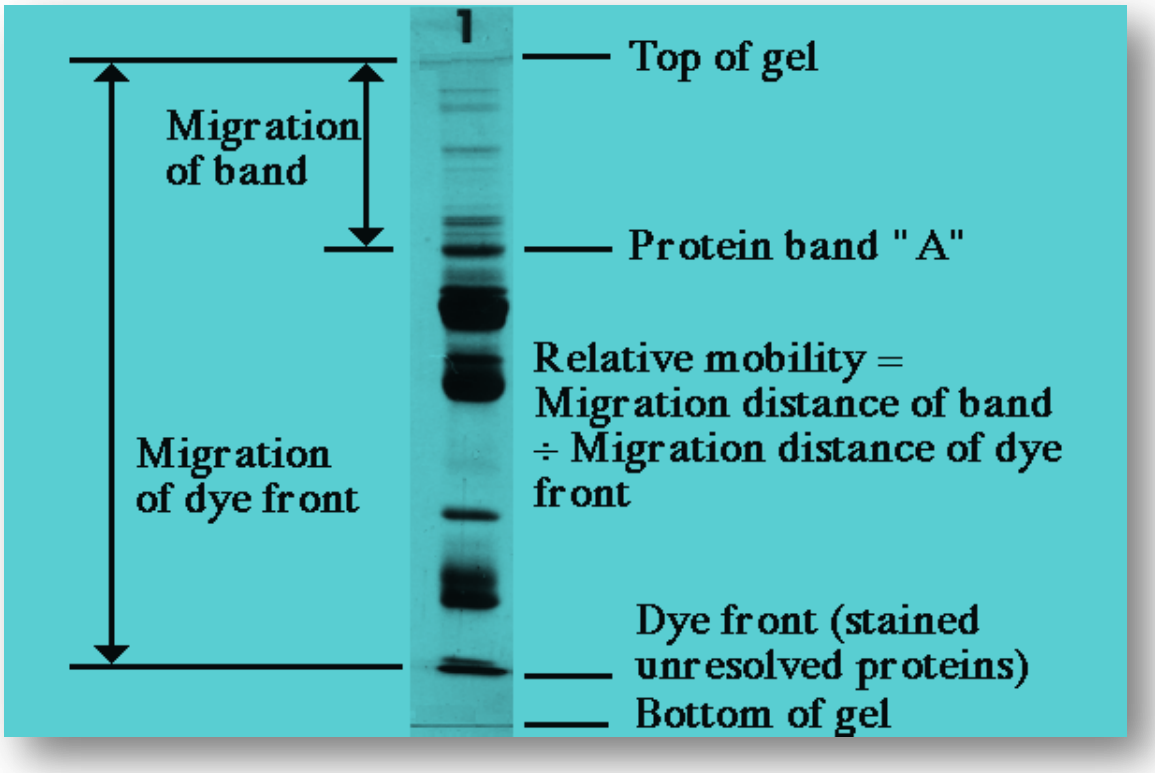

Figure 3.3: The picture shows the separation of the polypeptide chains of different proteins in gels. Relative mobility of the proteins was calculated by measuring the distance moved by the dye front and the bands. Again the electrophoretic mobility for marker polypeptide chains is plotted against the log of their molecular weights (Figure 3.4). From this mobility, a molecular weight of the individual protein can be predicted (Hames 1998).

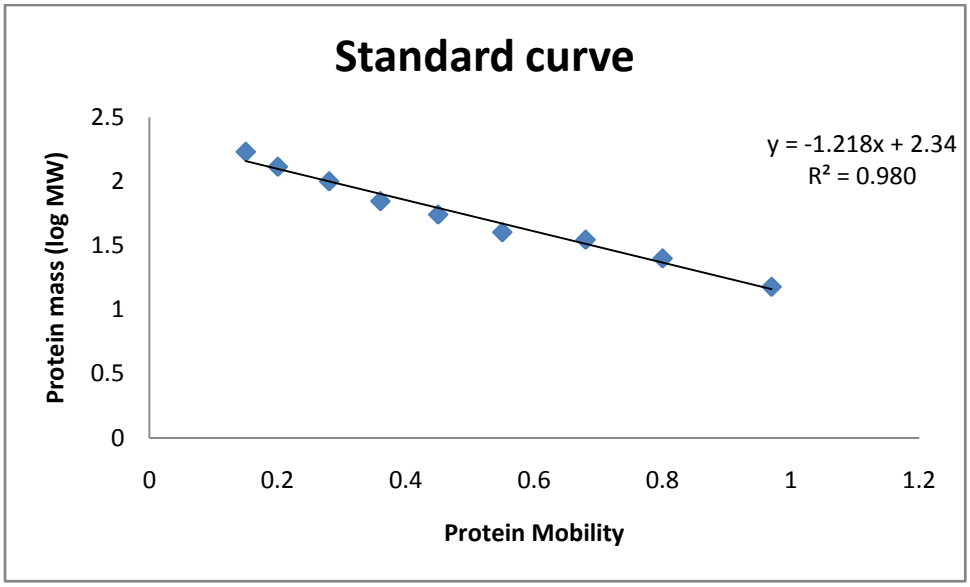

Figure 3.4: Relationship between protein molecular weight and mobility in SDS-PAGE.

From the graph we can calculate the mass of our protein. The protein mobility of the LPL0734 was found to be 0.27. So, the molecular weight of the protein LPL0734 was found to be 102 kDa. However, the expected size of our protein was 110 kDa but from the graph we got only 102.

3.4 Large scale expression screening (12 liters) of bacterial transporters

The DNA encoding *L.pneumophilia* gene lpl0734 cloned into a pNGFP-BC was taken for a large-scale analysis. The protein was expressed in *Escherichia coli* (Rosetta II cells) in 12 liters LB media with two antibiotics (Ampicillin + Chloramphenicol). Following harvesting, the cells were lysed using a high-pressure homogenizer at 15,000 psi. Membrane preparation and purification was performed as described in materials and methods.

3.5 Membrane protein purification

The solubilized membrane protein was captured on a His-Trap FF crude 5 ml column using an affinity purification protocol. The crude protein was passed through the Nickel beads. His-tagged protein bound to the beads, whereas other interfering proteins did not bind to the beads and were removed. We also added 20 mM imidazole to the samples before loading onto the column, which helped to remove unspecific binding of proteins. The Nickel has higher binding affinity to imidazole than Histidine. So, the proteins were then eluted using an imidazole gradient with an increasing concentration from 30 mM to 500 mM. Below, we can see chromatogram

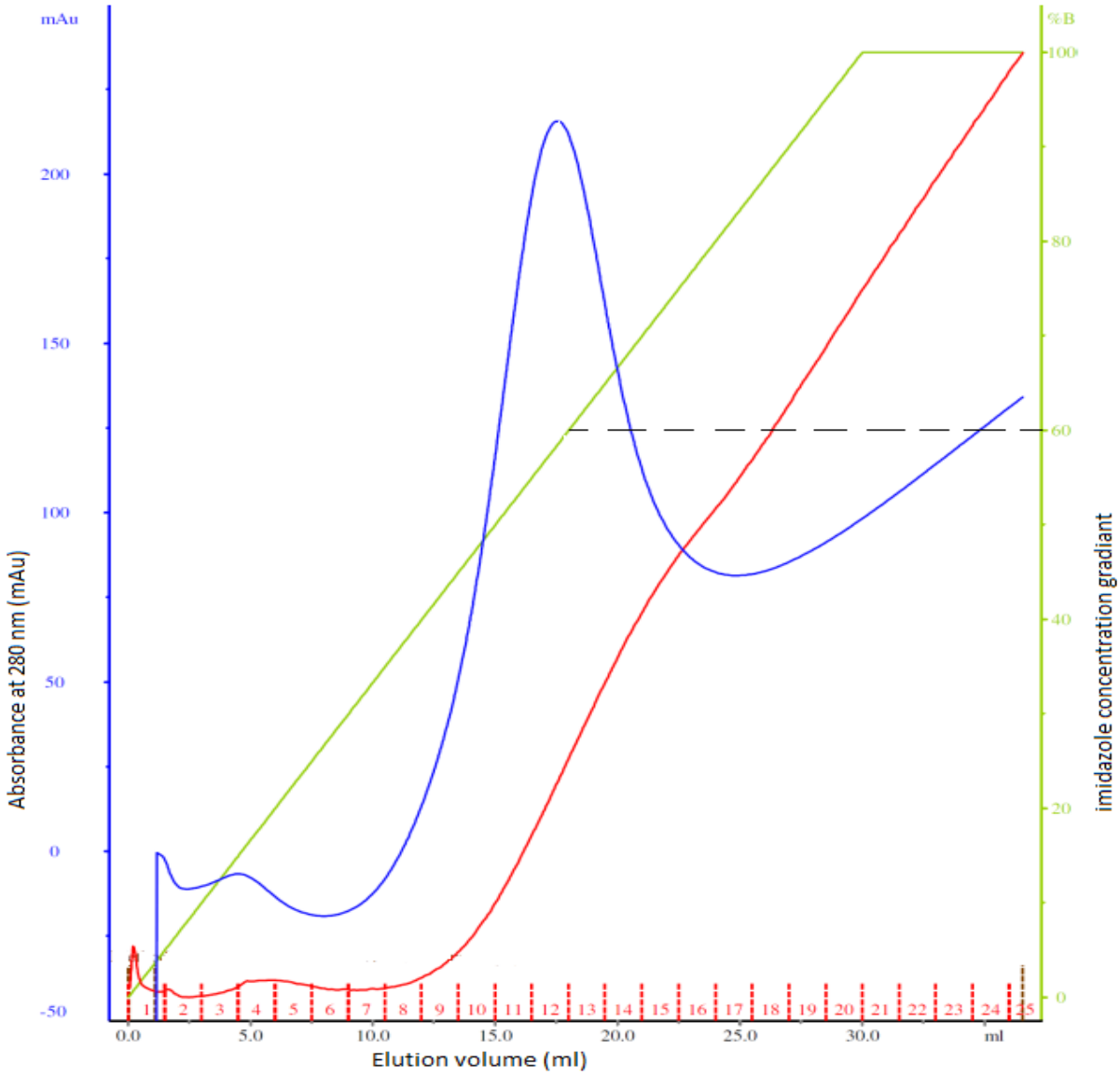

Figure 3.5: Chromatogram of the nickel purification. Protein elution was monitored by absorbance at 280 nm (blue, left y-axis in milli-absorbance units [mAu]) with the imidazole concentration gradient (green, right y-axis) over elution volume (ml, x-axis). The black dashed line shows the elution of protein at a concentration of approximately 60% buffer B (300 mM imidazole).

From the chromatogram, the LPL0734 protein was eluted at a concentration of approx. 60% buffer B (300mM of imidazole). The peak concentration of the LPL0734 protein was found in elutes tube E_{12}. After this, we measured protein concentration on Nanodrop (MW: 110kDa; Ext. coeff. 49.28L mol^{-1} cm^{-1}) in the different elute tubes (mg/ml) as shown in the table 3.1. Furthermore, the sizes of the proteins were verified using SDS-PAGE gel.

Tube selected:	3	6	8	10	12	14	16	18	22
Conc. On Nanodrop (mg/ml)	0.058	0.03	0.065	0.241	0.53	0.39	0.25	0.241	0.292

Table 1: Elutes of the His-tagged protein were collected in fractions of 1.5 ml each, in 25 glass tubes and their concentrations were measured using Nanodrop. We selected several elutes fraction to be run on a SDS-PAGE gel.

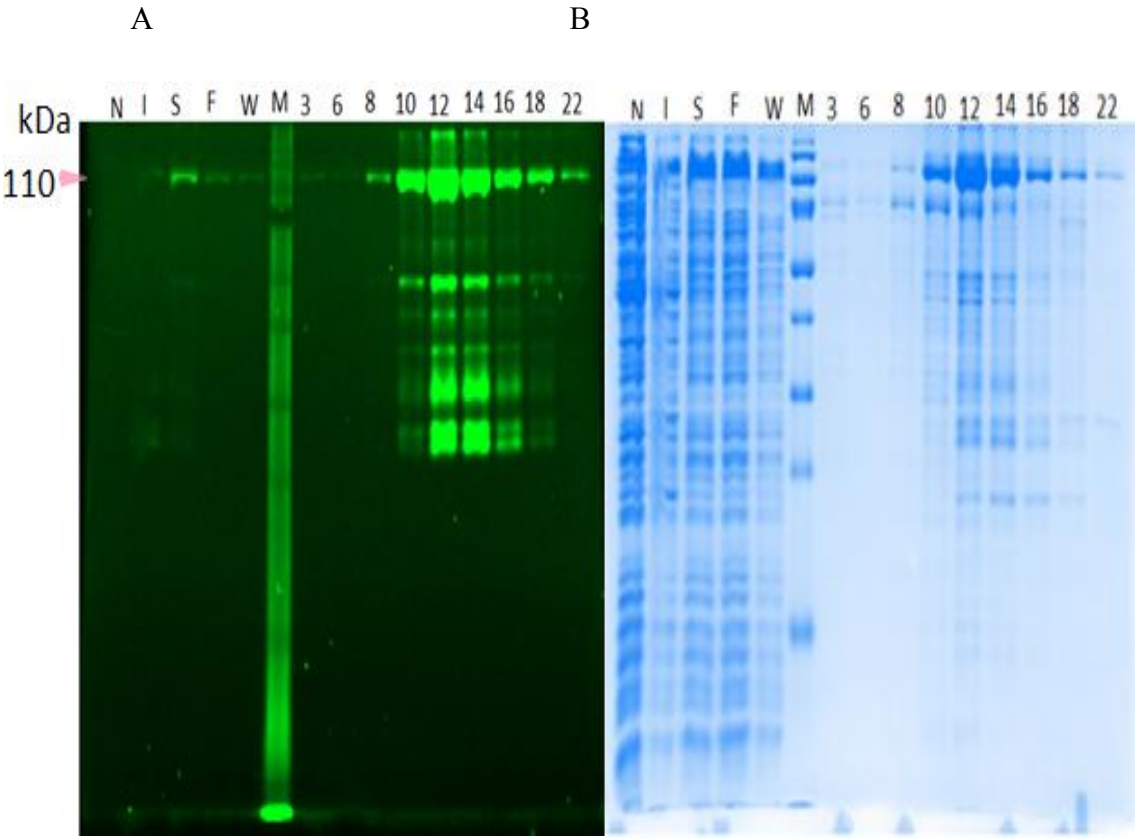

Figure 3.6: The picture shows LPL0734 protein samples analyzed by SDS-PAGE during affinity chromatography using Ni^{2+}-column. Picture (A) showing GFP picture and (B) Coomassie blue staining of the same SDS gel. LPL0734 correspond to bands around 110 kDa N; non-induced. I; induced. S; membrane solubilization.F; flow-through.W; wash.M; Molecular marker.3 – 22; elution fraction from imidazole gradient.

The gel 3.6 shows fractions from the elution peak. In the figure we could see the predominant band around 110 kDa, which is the size of LPL0734. The highest concentration of the protein band could see in lane 12. After this, seven elute tubes E_{10} to E_{16} were taken and were pooled in a 15 ml tube and the concentration of the protein was measured on Nanodrop. The pool of our protein was then concentrated by using a 50,000 Da MWCOpolyethersulfone ultrafiltrationmembrane by centrifugation until the volume reached 250 µl. Concentration of the concentrated purified protein (in **250 µl**) = 45 **mg/ml**wasrunon SDS-PAGE gel (Figure 3.7)

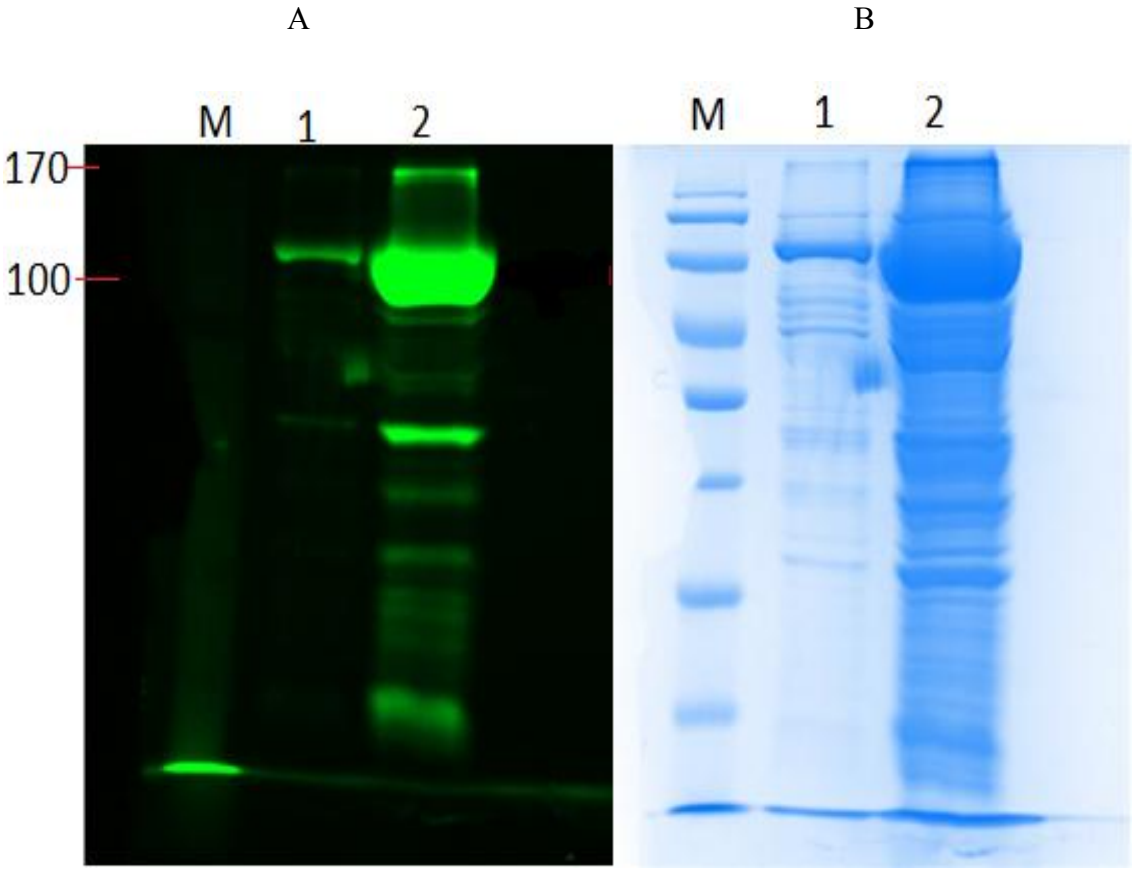

Figure 3.7: The picture shows Ni purified LPL0734 GFP fusion band (A) and (B) Coomassie blue staining of the same SDS gel,before and after concentration by using 50,000 MWCO PES.

At this step, we could see the thick concentrated band in lane 2 as compared to the lane 1 with the size of 110 kDa, However, some prominent contamination, especially from lower molecular weight proteins could also noted in the lane 2. An attempt to remove the lower molecular weight protein was continued with an FSEC purification step.

3.6 Fluorescence Size-Exclusion Chromatography

In the Size-exclusion chromatographic method, the molecules in solution are separated by their molecular size, based on the flow of the sample through porous packing. During the day of experiment, the Ni-purified LPL0734 membrane protein having a concentration of 45 mg/ml was taken and ultracentrifuge to remove precipitated and aggregated proteins. After ultracentrifugation the supernatant was taken and the protein concentration was found to be 43mg/ml on Nanodrop. This shows that we didn't lose a lot of proteins; we don't have an aggregation problem. 150 µl of the clarified protein was transferred to an Agilent tube and we injected 50 µl for three times. Each injection was run with 1 CV of buffer at 0.5 ml/min. The elute fractions of 0.5 ml were collected between 20 and 35 min and we got the following results. Running the different elutes on a 12% SDS-PAGE (Figure 3.) assessed the protein homogeneity and purity.

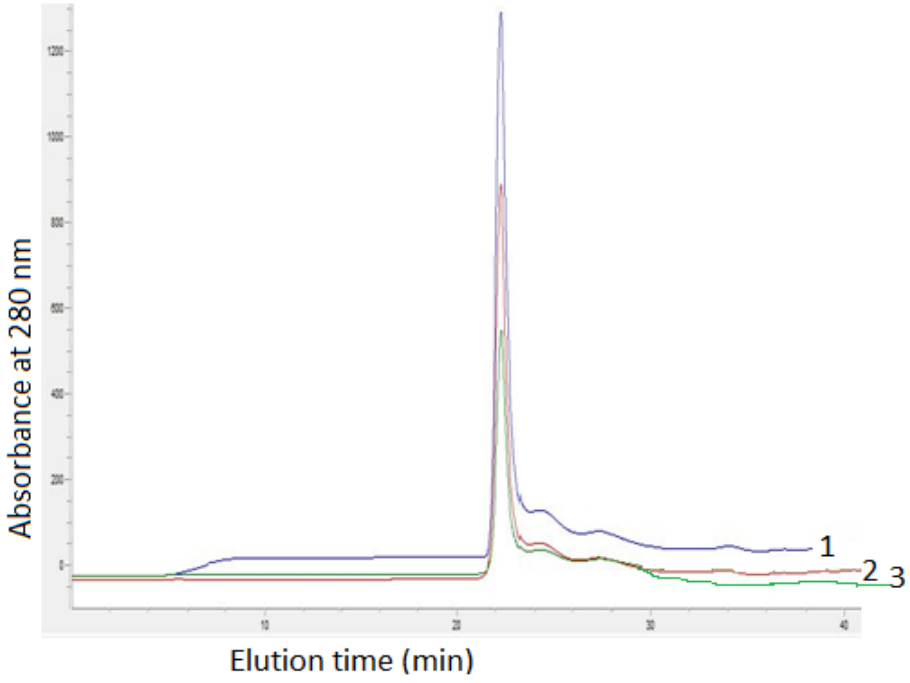

Figure 3.8: Chromatogram on Tosoh TSK showing the overlay of 3 different Runs, absorbance of protein at 280 nm. It can be seen from the figure that the absorbance pattern of the three different runs of the same amount of protein show the same profile

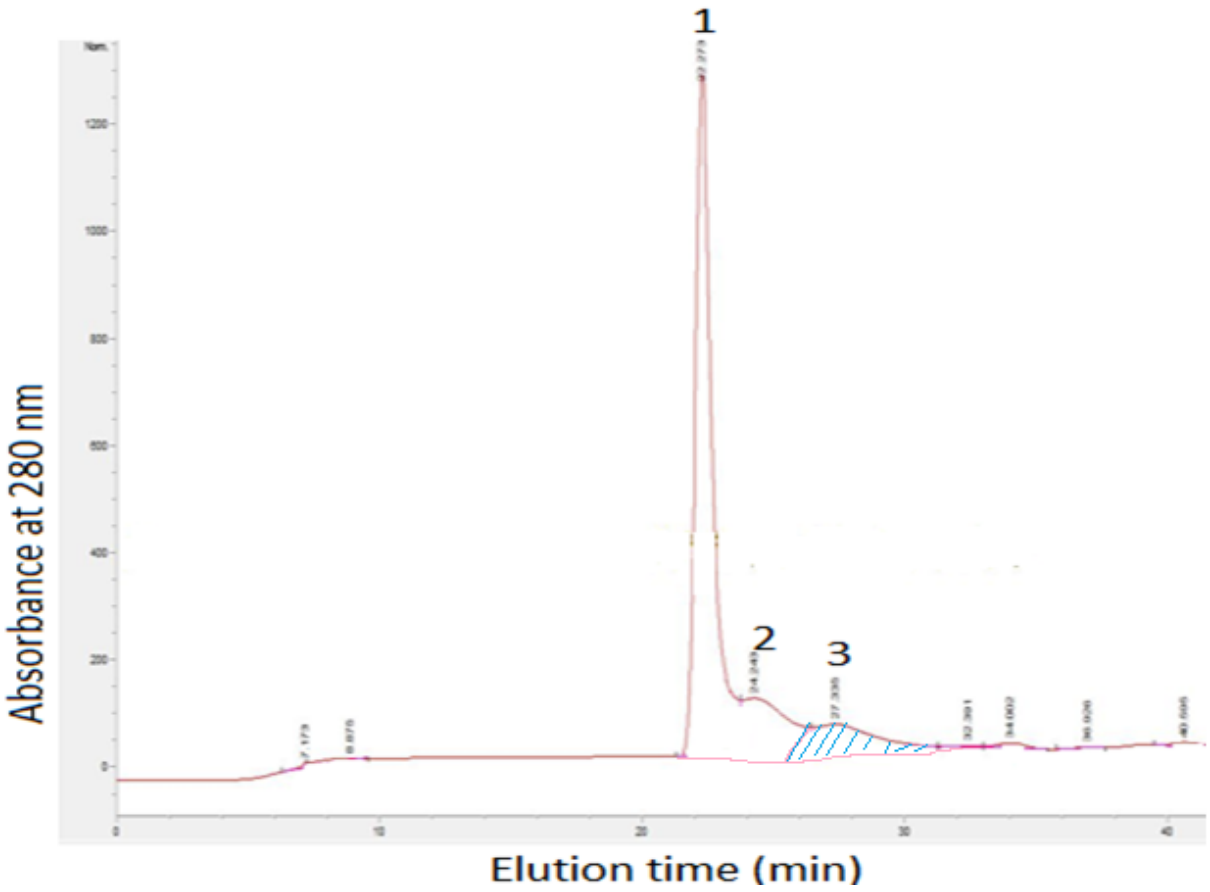

Figure 3.9: Analysis of sample purity by size-exclusion chromatography onTosoh TSK showing the absorbance of protein at 280 nm.50-µl of the samples was loaded onto column at a flow rate of 0.5 ml/min.The size of the proteins was estimated by comparing with the known molecular weight sample given by the Tosoh TSK.

The elution profile shows three peaks at 280 nm; the first two peaks (peak time: 22.278 min and 24.243 min) correspond to higher oligomeric status of LPL0734 protein. The third peak with blue dashed line (peak time: 27.335 min) might be monomeric LPL0734. Furthermore, we collected the fraction of three different peaks, and then we mixed the pool volume of peak 1 of three different run. Similarly, we did for peak 2 and peak 3. The concentrations of total volume pool before and after concentration were measured and we got the results as shown in the table 2.

Peak Time (min)	Run1	Run2	Run3	Total volume pool R1+R2+R3 (ml)	Pool concentration (mg/ml)	Pool conc. (mg/ml) after concentration
≈ 22.278	A4 (Peak 1)	A4 (Peak 1)	A4 (Peak 1)	1.5ml (Collectionof peak 1)	1.4	8.0 mg/ml in 200 µl
≈ 24.243	A6+A7 (Peak2)	A6+A7 (Peak2)	A6+A7 (Peak2)	3 ml (Collection of peak 2)	0.26	6.0 mg/ml in 100 µl
≈ 27.335	A8+A9 +B1 (Peak3)	A8+A9+B 1 (Peak3)	A8+A9+B 1 (Peak3)	4.5 ml (Peak 3)	0.16	5.4 mg/ml in 100 µl

Table 2: The table shows the collected fractions of three different peaks from three different runs, based on times. During this time we pooled all the fractions from peak 1 from three different runs, similarly for peak 2 and peak 3 and then we measured the pool concentration before and after concentrated. This profiles shows that the majority of the protein is aggregated.

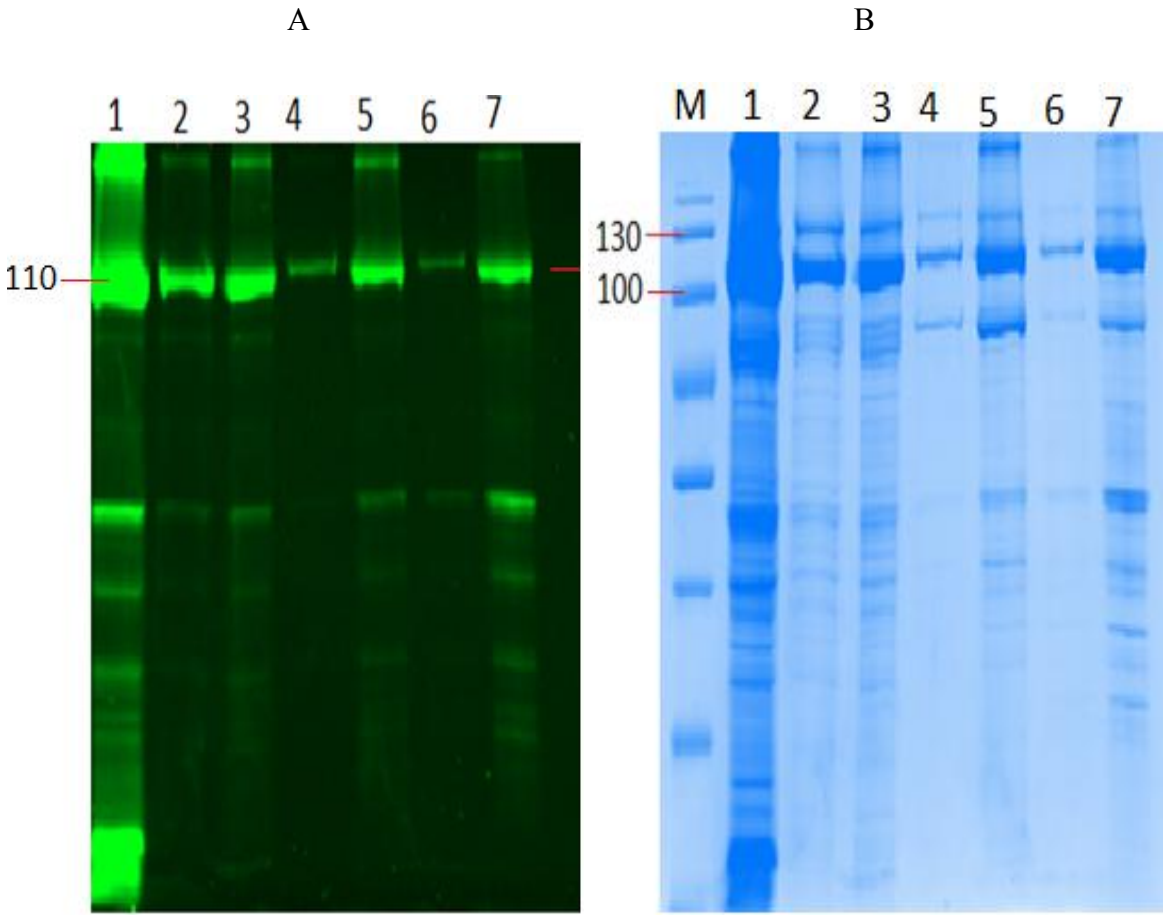

Figure 3.10:LPL0734 purification by size-exclusion chromatography on Tosoh TSK. (A) SDS-PAGE gel shows peak fractions of three different peaks with predominant bands around 110 kDa; (B) Coomassie blue staining of the same SDS gel; M; molecular marker. Lane 1; protein incubated O/N at RT, Lane 2 and Lane 3 correspond to peak 1 before and after concentration respectively. Similarly Lane 3 and Lane 4correspond to peak 2, Lane 5 and Lane 6 correspond to peak 3.

The SDS-PAGE gel shows the fractions from the elution peak, the band around 110 kDa isthe predominant band of LPL0734. At this stage we could see the bands of lower molecular size.In addition the protein was also found in the aggregated form.Protein lipidation and detergent screening methods checked further protein stability.

3.7 Protein Lipidation and Detergent screening (Stability test)

Several initial observations indicated that LPL0734 might form dimer or higher oligomers, including the elution profile during gel filtration chromatography. We therefore decided to check the stability and detergent screening test of LPL0734. For protein lipidation, we used

DOPC, whichwas synthetic phospholipid. DOPC facilitates the association of proteins with membranes and mediates protein trafficking. Recently DOPC has been appreciated as a regulator of membrane protein stability. The purified protein (45mg/ml; After Ni-purification) was diluted to a final concentration of 5 mg/ml. We used a protein to lipid ratio in the interval 1:0 to 1:0.8 (µg protein: µg lipid) as described in the table 2.3.

Protein/lipid ratio (µg:µg)	1:0	1:0.2	1:0.5	1:0.8
In µg	400:0	400:80	400:200	400:320
In µl	80 µl: 0	80:3.2	80:8	80:12.8

The following table shows the eye observation of O/N incubation of protein/lipid on a magnetic stirrer, it was found that higher the lipid composition was, higher the protein was precipitated.

Protein/lipid ratio (µg:µg)	Result (protein stability)
1:0	Ok
1:0.2	+
1:0.5	++ Precipitation
1:0.8	++++ Precipitation

In addition, the protein was ultra-centrifuged for 15 min at 70,000 rpm and then non–solubilized lipid; precipitated and aggregated proteins were removed. After ultracentrifugation the concentration of the protein was measured and we got the results as shown in the following table. From ultracentrifugation, it was found bigger pellets in the tubes containing higher lipid composition.

Protein/lipid ratio (µg: µg)	Conc. (mg/ml) after ultracentrifugation
1:0	3.250
1:0.2	2.20
1:0.5	1.699
1:0.8	0.747

3.8 Detergent screening:

Two different detergents (det.) $C_{12}E_8$ and DM were used in the detergent screening at a final concentration of 0.5%. Due to the amount of protein available, we used 7 µg of lipidated protein in each tube for the stability assay in different conditions as given in the table below. Moreover, the characteristic properties of the protein were studied on 12 % SDS-PAGE gels.

7 µg of lipidated protein at different protein:lipid ratio (µg:µg)	**T_0** (11.30 am) Wed. 23/05/12	**Day 1(D1)** Thursday morning 22 hours				**Day 2(D2)** Friday evening 52 hours				**Day 6(D6)** Tuesday 29/05/2012 146 hours			
		RT	4°C	4°C + det. (µl)		RT	4°C	4°C + det. (µl)			4°C	4°C + det. (µl)	
				$C_{12}E_8$	DM			$C_{12}E_8$	DM			$C_{12}E_8$	DM
1:0 (2.1 µl in each tube)				+0.1	+0.1			+0.1	+0.1			+0.1	+0.1
1:0.2 (3.2 µl in each tbe)				+0.16	+0.16			+0.16	+0.16			+0.16	+0.16
1:0.5 (4.1 µl in each tube)				+0.2	+0.2			+0.2	+0.2			+0.2	+0.2

Table 3: This table shows the set of lipidated protein used for the stability assay in different conditions.Two different detergents were used at a final concentration of 0.5%. We checked the stability of protein at two different temperatures. The two different conditions are: Room Temperatureand 4^{O}Cwith and without detergents. The characteristic properties of the protein were studied by running this samples on 12 % SDS-PAGE gels.

A B

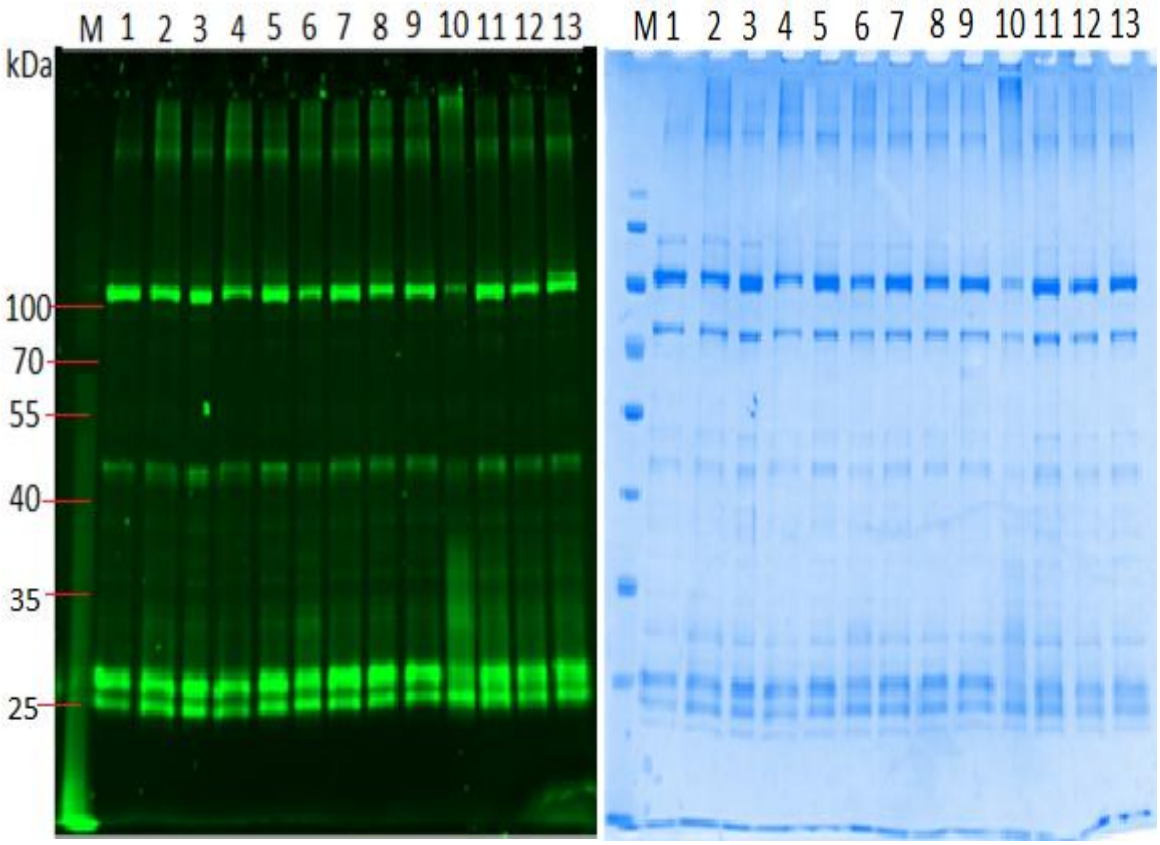

Figure 3.11: Effects of the addition of two different detergents (DM and $C_{12}E_8$) on the stability of purified LPL0734 as assayed bySDS-PAGE. GFP fusion protein bands (A) and its coomassie blue staining (B) of **7** µg of lipidated protein at **1:0**proteins: lipid ratio (µg: µg) at different conditions. M; Molecular marker.Control samples without test detergent were run in Lane 1(negative control). Lane 2; lipidated protein incubated for 1 day at RT without test detergent. Lane 3; lipidated protein incubated for 1 day at 4°C without test detergent. Lane 4; lipidated protein incubated for 1 day at 4°C with $C_{12}E_8$ test detergent. Lane 5; lipidated protein incubated for 1 day at 4°C with DM test detergent. Lane 6 (protein sample after 2 days at room temperature without detergent). Lane 7 (protein sample after 2 days at 4°C without detergent).Lane 8 (protein sample after 2 days at 4°C with $C_{12}E_8$ detergent). Lane 9 (Protein sample after 2 days at 4°C with DM detergent). Lane 10 (protein sample after 6 days at room temperature without detergent). Lane 11 (protein sample after 6 days at 4°C without detergent). Lane 12 (protein sample after 6 days at 4°C with $C_{12}E_8$ detergent). Lane 13 (protein sample after 6 days at 4°C with DMdetergent).

The SDS-PAGE gel (Figure 3.11) shows the effect of addition of two different detergents (DM and $C_{12}E_8$) in**7** µg of lipidated protein at **1:0** proteins: lipid ratio (µg: µg) at different environment conditions.In comparison with the negativecontrol, after one day we could see the aggregation of the protein at RT than in the other conditions. After 2 days, it is difficult to see a difference between the 4 conditions. In Day 6, we could see the concentrated bandat the top of the gel, incubated at RT.

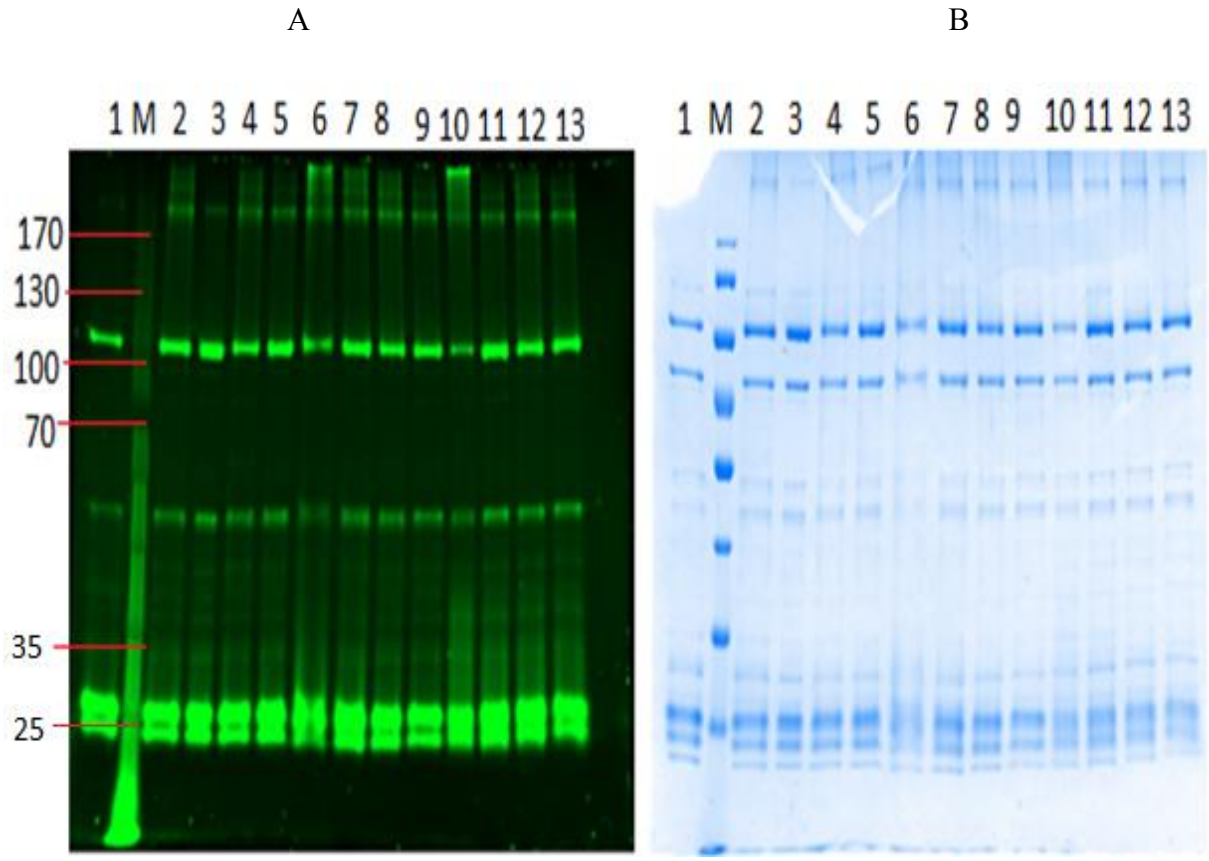

Figure 3.12: Effects of the addition of two different detergents (DM and $C_{12}E_8$) on the stability of purified LPL0734 as assayed by SDS-PAGE. GFP fusion protein bands (A) and its coomassie blue staining (B) of **7** µg of lipidated protein at **1:0.2** proteins: lipid ratio (µg: µg) at different conditions. M; Molecular marker.Control samples without test detergent were run in Lane 1(negative control). Lane 2; lipidated protein incubated for 1 day at RT without test detergent. Lane 3; lipidated protein incubated for 1 day at 4°C without test detergent. Lane 4; lipidated protein incubated for 1 day at 4°C with $C_{12}E_8$ test detergent. Lane 5; lipidated protein incubated for 1 day at 4°C with DM test detergent. Lane 6 (protein sample after 2 days at room temperature without detergent). Lane 7 (protein sample after 2 days at 4°C without detergent). Lane 8 (protein sample after 2 days at 4°C with $C_{12}E_8$ detergent). Lane 9 (Protein sample after 2 days at 4°C with DM detergent). Lane 10 (protein sample after 6 days at room temperature without detergent). Lane 11 (protein sample after 6 days at 4°C without detergent). Lane 12 (protein sample after 6 days at 4°C with $C_{12}E_8$ detergent). Lane 13 (protein sample after 6 days at 4°C with DMdetergent).

The figure 3.12 shows SDS-PAGE analysis of 7 µg of lipidated protein at **1:0.2** proteins: lipid ratio (µg: µg) at different conditions.In comparison with the negative control, we could see that there was no any difference after one day at 4°C without detergent, whereas the protein showed more aggregation in the other conditions. After day 2 and day 6, we observed an aggregation in all conditions but more at RT.

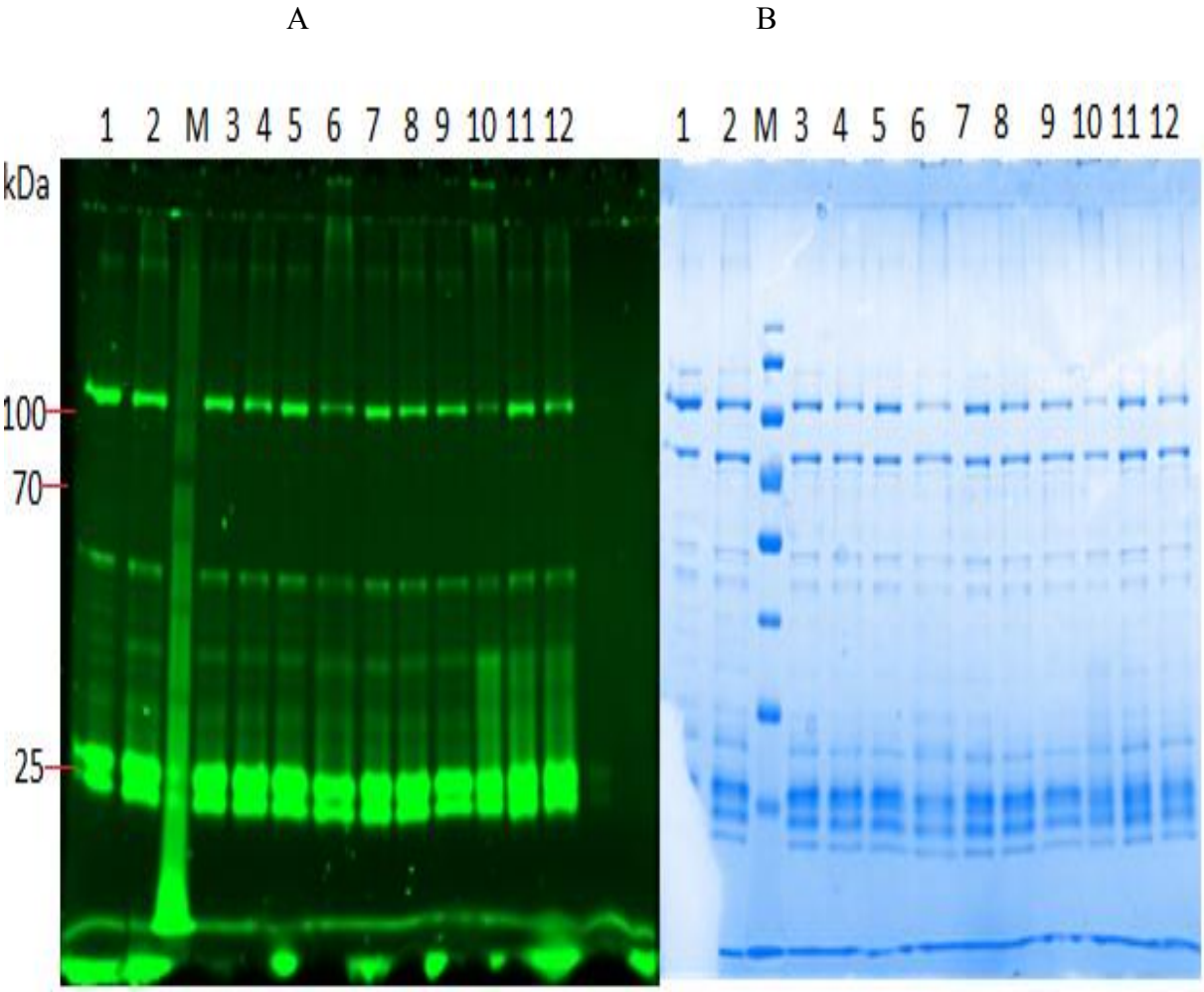

Figure 3.13: Effects of the addition of two different detergents (DM and $C_{12}E_8$) on the stability of purified LPL0734 as assayed by SDS-PAGE. GFP fusion protein bands (A) and its coomassie blue staining (B) of **7** µg of lipidated protein at **1:0.5** proteins: lipid ratio (µg: µg) at different conditions. M; Molecular marker.Control samples without test detergent were run in Lane 1(negative control). Lane 2; lipidated protein incubated for 1 day at RT without test detergent. Lane 3; lipidated protein incubated for 1 day at 4^{o}C without test detergent. Lane 4; lipidated protein incubated for 1 day at 4^{o}C with $C_{12}E_8$ test detergent. Lane 5; lipidated protein incubated for 1 day at 4^{o}C with DM test detergent. Lane 6 (protein sample after 2 days at room temperature without detergent). Lane 7 (protein sample after 2 days at 4^{o}C without detergent). Lane 8 (protein sample after 2 days at 4^{o}C with $C_{12}E_8$ detergent). Lane 9 (Protein sample after 2 days at 4^{o}C with DM detergent). Lane 10 (protein sample after 6 days at room temperature without detergent). Lane 11 (protein sample after 6 days at 4^{o}C without detergent). Lane 12 (protein sample after 6 days at 4^{o}C with $C_{12}E_8$ detergent). Lane 13 (protein sample after 6 days at 4^{o}C with DMdetergent).

The above figure shows SDS-PAGE analysis of 7 µg of lipidated protein at **1:0.5** proteins: lipid ratio (µg: µg) at different conditions. In comparison with the negative control Lane 1, after one day we observed more aggregation of the protein at RT than in the other conditions. After day 2 and day 6, we observed an aggregation in all conditions but more at RT. We also observed an increased degradation of the protein after day 6.

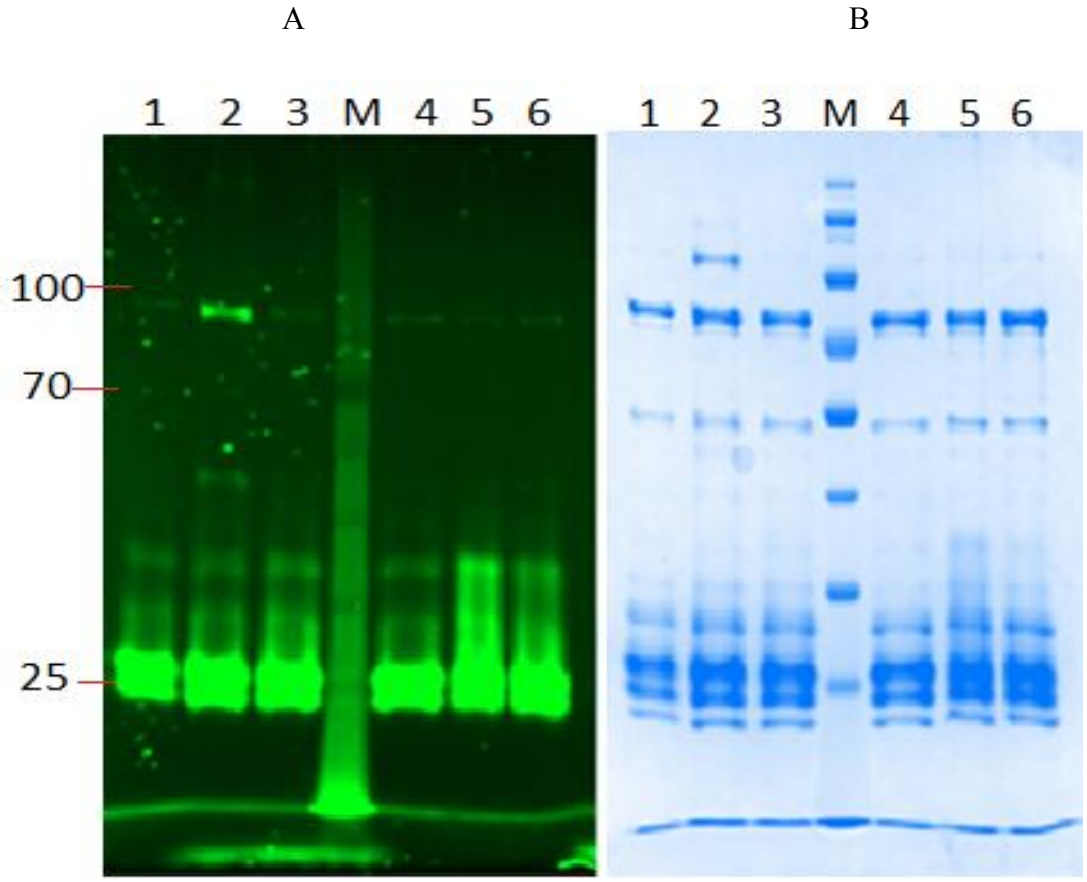

Figure 3.14: SDS-PAGE analysis of **7** µg of lipidated protein at 1:0.8 proteins: lipid ratio (µg: µg) at different conditions. M: Molecular marker. Lane 1: lipidated protein incubated for 1 day at RT without test detergent. Lane 2; lipidated protein incubated for 1 day at 4^{o}C without detergent. Lane 3: lipidated protein incubated for 2 days at RT without detergent. Lane 4; lipidated protein incubated for 2 days at 4^{o}C without detergent. Lane 5: lipidated protein incubated for 6 days at RT without detergent. Lane 6; lipidated protein incubated for 6 daysat 4^{o}C without detergent.

The above figure shows SDS-PAGE analysis of 7 µg of lipidated protein at **1:0.8** proteins: lipid ratio (µg: µg) at different conditions. In comparison with the negative control, after one day we observed less aggregation protein at 4^{o}C and higher at RT. After day 2 and day 6, we observed more aggregation and increased degradation of the protein.

3.9 Cloning of *Legionella* SO_4^{2-} transporter into pETM11 vector:

Complete PCR-derived DNA fragments containing structural genes for **SO_4^{2-}/HCO_3^-** transporter (LPL0734) of the *Legionella pneumophila* were cloned into an expression vector pETM11 (EMBL collection) using NcoI and XhoI restriction sites. This vector incorporates a His_6-tag and a TEV protease cleavage site N-terminal to the target construct. The DNA sequence of Sulfate-Legionella contains an internal recognition sequence for NcoI. So, the site-directed point mutagenesis was donein order to remove the internal NcoI restriction site with designed primers by PCR using the kit from Agilent at three different annealing temperatures as described in materials and methods. The amplification product was analyzed by electrophoresis with a 1 kb DNA ladder, on a 0.8% agarose gel stained with 3x Gel Red (Biotium Inc). From the figure we can see some band but not distinct one.

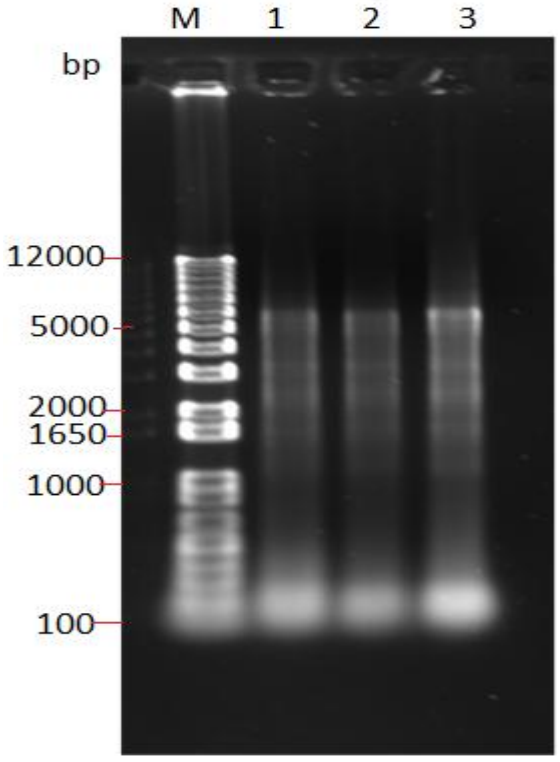

Figure 3.15: Agarose gel electrophoresis of PCR products of so-legio in TOPO vector (5.8 kb) with three different annealing temperature.Lane M; 1 kb Plus DNA Ladder. Lane 1; annealing temperature 56^{o}C. Lane 2; annealing temperature 58^{o}C. Lane 3; annealing temperature 60^{o}C

The PCR product was then treated with 2 µl *DpnI* restriction enzyme to digest the parental unmuted DNA. Afterwards, 2 µl of digested PCR product was transformed into chemically competent DH10Band XL1-blue cells. Number of colonies per plate was count from the

overnight incubation. It was found that the growth of cells depend upon the annealing temperature and the type of competent cells DH1-B and XL1-blue cells as shown in the table4:

PCR product (annealing temperature) ^{o} C	Competent cells	Number of colonies
56	DH10B	0
	XL1-blue	400
58	DH10B	8
	XL1-blue	400
60	DH10B	30
	XL1-blue	40

Table 4: the table shows the number of clonies of two different competent cells at different annealing temperature.

Colony screening: The constructs of the colonies were screened by PCR with GoTaq polymerase (Promega). Here we checked 5 colonies from each plates and named the tube as shown in table below:

PCR (annealing temp.)	Competent cells	Name of clone
56^{o}C	DH10B	----------
	XL1-blue	X11, X12, X13, X14, X 15
58^{o}C	DH10B	D21, D22, D23, D24, D25
	XL1-blue	X21, X22, X23, X24, X 25
60^{o}C	DH10B	D31, D32, D33, D34, D35
	XL1-blue	X31, X32, X33, X34, X 35

Table 5: The table shows the name of clone of two different competent cells.

Further we started overnight minicultures of all the clones in 5 ml LB medium + kanamycin at 37^{O}C.It was found that all the colonies showed positive results and then, we randomly select 4 tubes of overnight minicultures of 5 ml in LB medium + kanamycin at 37°C.(X15, X 25, D24, D 33)

3.9.1 Plasmid DNA purification and checking

Plasmid DNA was purified from the overnight cultures by a Mini Prep (Miniprep kit, Qiagen) and the DNA were eluted in50 µl of sterile water and their concentration was measured on Nanodrop and we got the result as shown in the table 6.

Tubes	Concentration (ng/µl)
X15	328.1
X 25	226.1
D24	203
D 33	187.4

Table 6: This table shows the concentration of four different cultures after purified DNA using a Miniprep kit.

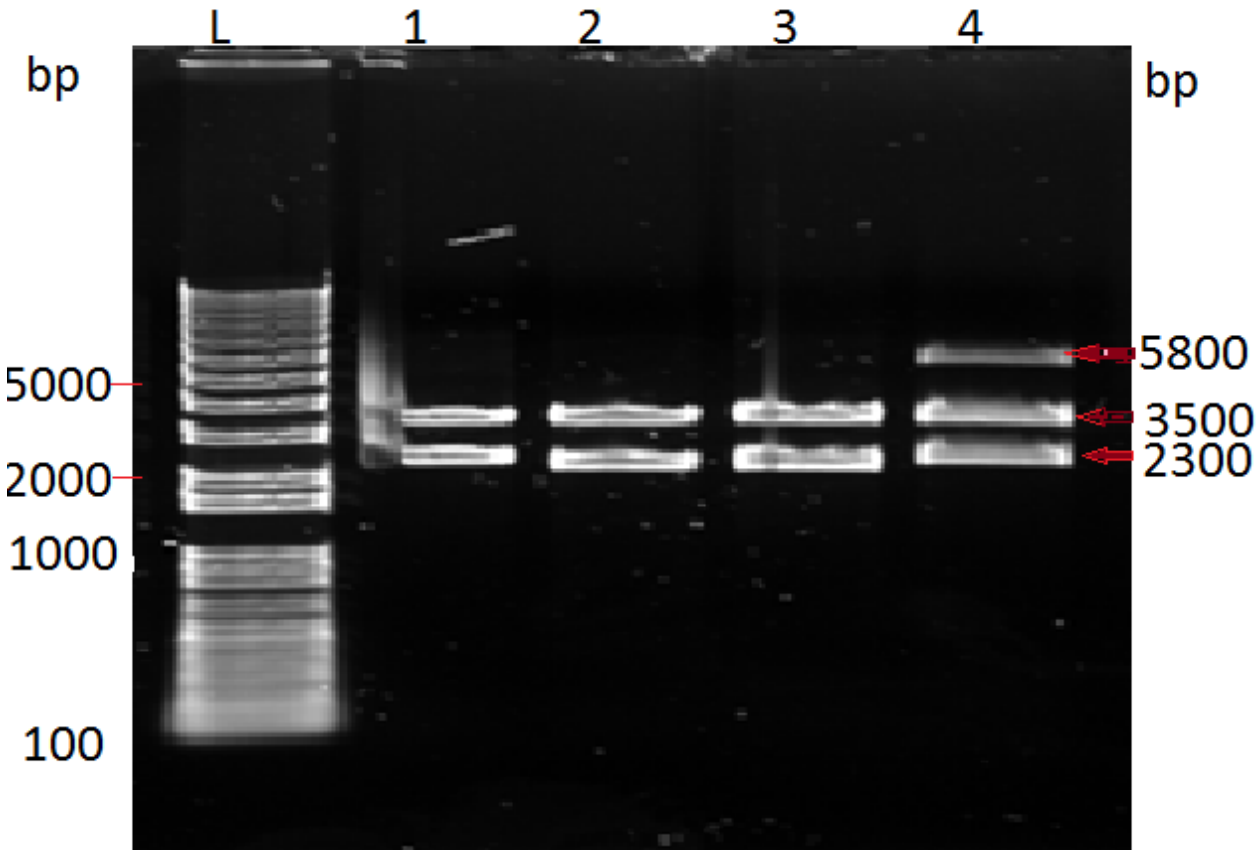

Figure 3.16: Agarose gel electrophoresis for checking of the plasmids by restriction analysis with EcoRI (restriction sites flanking the PCR product in the TOPO vector). From left: L; 1 kb Plus DNA Ladder. Lane 1: Clone X25 (annealing temperature was 58^{o}C and the competent cells was XL1-blue). Lane 2: Clone D24 (annealing temperature was 58^{o}C and the competent cells was DH10B). Lane 3: Clone X15 (annealing temperature was 56^{o}C and the competent cells XL1-blue). Lane 4: Clone D33 (annealing temperature is 60^{o}C and the competent cells DH10B

The above figure 3.16 shows the agarose gel electrophoresis for checking of the plasmids by restriction analysis with EcoRI. From the figure 3.16, we could see the bands around 2300 bp which were most likely the size of DNA and the bands around 3500 bp could be TOPO vecors. In Lane 4 we could see the addition band around 5800 bp. We could also see that Lane 1 and Lane 4 have slightly higher bands as compared to other two clones.

Finally, the positive plasmids were sequenced for their *NcoI* mutated site with M13-FP and M13-RP primers (GATC company: 20 µl of DNA at 50 ng/µl). We compared the mutated sequence of each construct with the original sequence of Solfate-*Legionella*and we obtained the following results;

- Beginnings of the sequences of all four clones were ok.
- Ending of the sequences of all four clones were also ok.
- But there was problem with clone X25 and D33 in the middle of the sequences.
- We checked the mutation and was found good from clones D24 and X15 (Figure 3.16) and then we continued with X15 for further purification

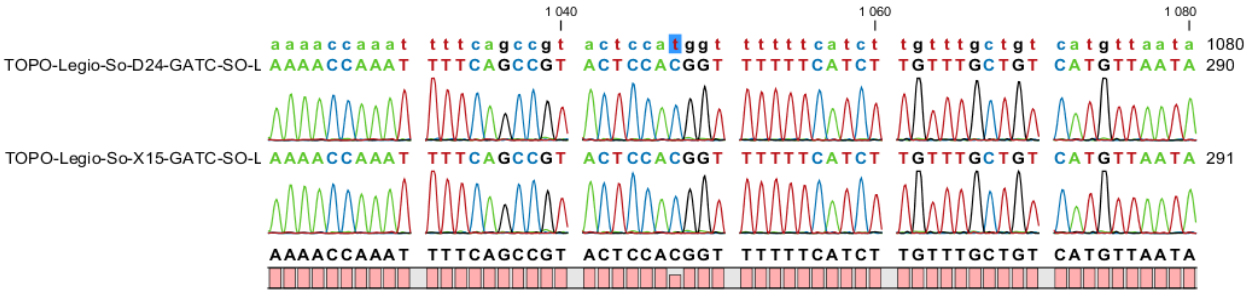

Figure 3.17: The picture shows the final sequencing of the positive plasmids for their *NcoI* mutated site with M13-FP and M13-RP primers (GATC company: 20 µl of DNA at 50 ng/µl) from clones D24 and X15.

3.9.2 Integration of SO_4 transporter in the pETM11 vector

The TOPO-cloned inserts and pETM11 vector were first digestedwith restriction enzymes. The restriction enzymes were then heat-denatured. To avoid its self-ligation, the digested pETM11 vector was dephosphorylated. Digested TOPO-cloned inserts and dephosphorylated vector were purified by gel purification. Ligation of the digested inserts with the digested vectors was done and then products were transformed into chemically competent DH10B and

Top10 cells (made by Hanne-NCMM). Ligation of the digested inserts with the digested vectors was incubated overnight at 16^{O}C.The ligated products were transformed into chemically competent DH10B and XL1-blue cells using the same protocol as described earlier. There was no any growth of bacteria after overnight incubation at 37^{O}C. We have tried these steps for many times but were unsuccessful.

4. DISCUSSION

Sulfur is one of the most important macronutrients in cells and may be vital in the pathogenesis and survival of some bacteria. LPL0734 is a sulfate transporter in the membrane of *Legionella pneumophila.* Membrane proteins play very important roles in all cells and are key drug targets.

4.1 Analysis of small-scale expression of LPL0734

In this thesis, we attempted to express and purify the LPL0734 protein from *Legionella pneumophila*. Therecombinant construct pCGFP-BC contained a polyhistidine tag at C-terminus and pNGFP-BC contained a polyhistidine tag at N-terminus that allowed the single-step isolation by Ni-affinity chromatography. Nickel has high binding affinity to histidine. The higher affinity of imidazole to nickel relative to histidine was exploited to elute the bound histidine tagged protein. GFPuv was chosenfor bacterial expression which helped to maximize the stability of the chromophore in each expression system and exploit the strong fluorescence signals(Crameri, Whitehorn et al. 1996). The level of expression was examined in two different competent *E.coli* strains (Rosetta II and C43) because not all competent strains express membrane proteins. When pCGFP-BC or pNGFP-BC was transformed into *E.coli*strains (Rosetta II and C43), very few colonies of C43 were obtained as compared to the Rosetta cells.This could mean that over expression of this protein might have been more toxic in C43 cells than Rosetta II cells. In figure 3.1,we could see successful expression of LPL0734 using Rosetta II cells, but there was no protein expression using C43 cells. The band around 110 kDa in both of the lanes 2 and 6 were most likely of LPL0734 protein because they have similar size. While the bands around 27 kDa in lane 2 and 6 were most likely GFP tag, cleaved from the proteins.Additional bands were noted in lane 3 that could be the degradation product of LPL0734.Examination of the solubilized isolated membranes using β-DDM detergent showed higher expression with the construct LPL0734/ pNGFP-BC than in LPL0734/ pCGFP-BC(figure 3.2). The choice of the N- and C-terminal domains represents an important consideration because even small differences can dramatically influence both solubility and expression(Gräslund, Nordlund et al. 2008). Additional bands were also noted in lane 3 that could be the degraded proteins of LPL0734.

4.2Protein molecular weight through SDS-PAGE

The Coomassie blue staining gel map was considered (Figure 3.1B) in order to find out the molecular weight of LPL0734 protein. The electrophoretic mobility for marker polypeptide chains were measured using the formula given in figure 3.3and plotted against the log of their molecular weights(figure 3.4).By using the relative mobility formula (Figure 3.3), the protein mobility of LPL0734 was found to be 0.27 andcalculated molecular size was found to be 102 kDa. However, expected molecular size of the GFP tagged LPL0734 is 110kDa (molecular size of LPL0734 is 84 kDa).This difference might be due to partially folded protein as a result of us not heating sample to unfold the protein. Heating was compromised due to the fact that LPL0734 is heat sensitive. Heating protein samples in the presence of SDS leads protein long linear molecules coated with SDS that might give uniformly negative ionic charge(Garcia, Magalhães et al. 2006).

4.3 Large scale expression screening (12 liters) of bacterial transporters

Functional and structural studies of recombinant proteins usually required to purification on a relatively large scale and under native-conditions. Affinity tags are widely used to facilitate, detection and purification of the recombinant proteins in the biotechnology industry(Sun, Tropea et al. 2011). Figure 3.5 shows the elution of protein at a concentration of approximately 60% buffer B (300 mM imidazole). First the concentration of different elute during affinity purification was measured on Nanodrop, and then the size of the protein was check by running the protein samples onSDS-PAGE gelbecause high concentration of elutes on Nanodrop doesnot mean that this is a good amount of protein, which might be due to imidazole contents. The pool of LPL0734 protein was then concentrated by using a 50,000 Da MWCO membranebecause the upper limit of molecular weight which membrane was permeable is 50,000 Da. So, the LPL0734 protein, which was 110 kDa,remains in the membrane, which is not permeable. Though membrane protein are difficult to purify (Lin and Guidotti 2009),we managed to purify solubilized LPL0734 protein in milligram quantities using affinity purification. The Nanodropconcentration 45mg/ml (with the ext. coefficient 49.28 L mol^{-1} cm^{-1}) was sufficient for further characterizations of LPL0734. Although we

could see the concentrated protein band around 110 KDa (figure 3.7); which is a size of LPL0734, however,there was some prominent contamination (small bands), which suggested degradation of LPL0734.

In order to separate the targeted molecular size of the proteins, the Ni-purified LPL0734 protein was taken and analyzed its purity by size-exclusion chromatography on Tosho TSK. This technique helps us to separate the protein by their size. When the samples were delivered to a column that is packed with porous particles. The larger molecules quickly pass through the column because the large molecules cannot enter the pores and elute first, whereas the smaller size of polymers enter the pores and take longer time to exit.However aggregation was an unresolved challenged as could be seen by the large void peaks observed during size-exclusion chromatography (figure 3.9). Many other studies have identified aggregation as a major challenge in the studies of membrane proteins(Seddon, Curnow et al. 2004; Borch and Hamann 2009). One should be able to avoid aggregation of the protein in order to get insights on the structures of the proteinbecause the protein should be folded in native form that determines the biological function. Aggregation also prevented the removal of the affinity tag, further complicating crystallization. According to Cromwell et al, (Cromwell, Hilario et al. 2006), the degree of aggregation of protein observed at the time of purification may be because the protein experiences a wide range of pH, temperature conditions, ionic strength and protein concentrations during the process.

4.4 Protein Lipidation and Detergent screening (Stability test)

Furthermore protein lipidation and detergent screening test was done in order to check protein homogeneity and protein stability.Several initial observationsindicate that LPL0734 might form higher oligomers. First the LPL0734 protein was taken for lipidation. For protein lipidation, we used DOPC, which was synthetic phospholipid. This might facilitates the association of protein with membrane proteintrafficking. After overnight incubation of protein with lipid on a magnetic stirrer, it was found that higher precipitation in a tube containing higher amount of lipid.Two different detergents $C_{12}E_8$ and DM wereused in the detergent screening, which are often used to check the homogeneity of membrane protein. Analysis of the effects of the addition of two different detergents was studied under different conditions.

In figure 3.11, the protein showed more aggregation at RT than at $4^{O}C$ even after one day of incubation. It could be hypothesized that temperature greater than $4^{O}C$ makes the protein unstable.After day 2, we observed an increased aggregation in all the conditions. In Day 6, we observed a high aggregation of the protein at RT as compared to other conditions.Besides this, we observed lower bands in all conditions that could be the degradation bands of LPL0734 and GFp tag, cleaved from the protein. Similar patterns of results were observed from figure 3.12 and 3.13.Furthermore the aggregation of the protein started from day 1 and was relatively stable and increased degradation of the proteinswas observed at higher lipid contents (Figure 3.14).This suggested that the DOPC does not facilitate the association of proteins, mediates protein trafficking.

For successful purification of membrane proteins the choice of detergent is one of a key factor(Hunte, von Jagow et al. 2003).The central region could be also related for the thermo-stability gain that is the temperature required for protein to stay in their native form (Vogt, Woell et al. 1997). Choosing different conditions that gives pure and stable membrane protein is a challenging task.Protein engineering could lead to increased stability (Matouschek, Kellis et al. 1989).

As mentioned above, we could not able to report crystal structure of LPL0734. In some extents our attempt to purify LPL0734 was successful. However some challenges remains to be resolved. One of the main reasons why this approach has failed must be because of protein instability and aggregation. No doubt, several factors are involved in the expression and purification of membrane proteins. The result further indicates that we need to do more research in different conditions such as temperature, buffer-detergent composition and pH that gives pure and stable membrane protein.

4.5Construction of the recombinant pETM11 vector producing *Legionella* SO_4^{2-} transporter (LPL0734)

The cloning process was challenging, but using the cloning strategy(Crabtree ; Sambrook, Fritsch et al. 1989), the pETM11 was successfully constructed. The constructed expression vectors have an antibiotic resistance gene against kanamycin toselect the transformed cells on

agar plate containing kanamycin. Cloning of Sulfate-*Legionella*was carried out for studying protein structure and function of ionic-transporter systems.

The bacterial expression vector pETM11 was first studied and then verified the restriction enzymes that were cutting in the MCS of the expression vector. In this study we used Ncol and Xhol restriction enzyme.The DNA sequence of *Legionella pneumophila*encodingLPL0734 protein have internal Ncol restriction site. So, we removed the internal restriction site by mutation usingPCR with the kit from Agilent with three different annealing temperatures and the designed primers for the mutation. Figure 3.15 shows agarose gel electrophoresis of PCR products of sulfate-*legionella* in TOPO. We used three different temperatures because of the hairpin risk at 55^{O}C.The result from figure 3.16 shows that the bands around 2300 bp were more likely he DNA of solfate-*Legionella* with restriction sites in it and the bands around 3500 bp were more likely the TOPO vectors.In Lane 4, we can see the band around 5800 bp that might be the DNA plus TOPO vectors. We also observed that lane 1 and lane 4 have slightly higher band as compared to lane 2 and lane 3 (around 2300 bp in the figure 3.16), this might be due to additional sequence base. This indicates that we managed to mutate the internal Ncol restriction site in the sequence of sulfate-*legionella.*

By sequencing the original sequence of the gene and the mutated sequences of the clones D24 and X15, we observed that by mutationthe internal Ncol site was successfully removedfrom the sequence of *Legionella pneumophila*(Figure 3.17).

Unfortunately, the integration of Sulfate transporter gene of *Legionella pneumophilia*in pETM11 vector was unsuccessful. Problem was seenafter transformation of the ligation mixinto chemically competent DH10B and XBl-blue cells. There was no any bacterial growth after overnight incubation at 37^{O}C. Although, we have tried many times, unfortunately, we could not proceed further.Problem could be either with transformation procedure or with competent cells, or with ligation reaction, such as poor ligase activity.

5. CONCLUSION

Determination of the structures of membrane proteins in bacteria is a challenging task that is essential to understand biological function at the molecular level. In order to provide insights into its biochemical properties, we tried to overexpress, purify, characterize and crystallizeLPL0734 of *Legionella pneumophila.*At present, we have not been able to produce crystal of LPL0734. It may be consider that the problem arises from protein instability and aggregation.The result from table 2 performed using size exclusion chromatography showed that we could purify the protein in milligram quantities. However, some challenges remain to be resolved. It is evident from figure 3.9 that the large void peaks suggest that the majority of the protein is in the form of aggregation. Furthermore, the SDS-PAGE gel (Figure 3.10)of three different peaks also shows degradation product of LPL0734. At this stage, we did not manage to remove the degradation product in addition toaggregation. The result from the stability test showed that the LPL0734 protein was not stable even at room temperature after one day of incubation, but the protein was more stable at 4^0C independently of the detergent content. Increased aggregations and degradations of the protein are observed at more lipid contents (Figure 3.11 to figure 3.14). The result further indicates that there is a need to optimize the temperature and buffer-detergent composition to improve the protein stability. The structure of the LPL0734 is still elusive. In order to know the exact residues that contribute the most, more extended pedigrees are needed. In addition, this thesis describes cloning of Solfate-*Legionella*(LPL0734) in pETM11 bacterial expression vector.

6. FUTURE PERSPECTS

In the present study, we describe the cloning, expression, purification, and characterization of LPL0734 from *Legionella pneumophila*. Despite expressing and purifying of LPL0734,we are unable to get the crystal structure of it. The problem arises from protein instability, aggregation and thermo-stability.Different parameters should be tested separately and together to mimic conditions in the cell membrane in order to attempt to achieve the stability of this protein.The result described here and consideration of the characteristics of LPL0734 during the time of expression and purification might be the basis for many new studies in the coming future.

7. REFRENCES

Alber, B. E. and J. G. Ferry (1994). "A carbonic anhydrase from the archaeon Methanosarcina thermophila." Proceedings of the National Academy of Sciences**91**(15): 6909-6913.

Alexandrov, A. I., M. Mileni, et al. (2008). "Microscale fluorescent thermal stability assay for membrane proteins." Structure**16**(3): 351-359.

Berman, H. M., J. Westbrook, et al. (2000). "The protein data bank." Nucleic acids research**28**(1): 235-242.

Beveridge, T. J. (1999). "Structures of gram-negative cell walls and their derived membrane vesicles." Journal of bacteriology**181**(16): 4725-4733.

Borch, J. and T. Hamann (2009). "The nanodisc: a novel tool for membrane protein studies." Biological chemistry**390**(8): 805-814.

Braus-Stromeyer, S. A., G. Schnappauf, et al. (1997). "Carbonic anhydrase in Acetobacterium woodii and other acetogenic bacteria." Journal of bacteriology**179**(22): 7197-7200.

Cazalet, C., L. Gomez-Valero, et al. (2010). "Analysis of the Legionella longbeachae genome and transcriptome uncovers unique strategies to cause Legionnaires' disease." PLoS genetics**6**(2): e1000851.

Chegwidden, W. R. and N. D. Carter (2000). "Introduction to the carbonic anhydrases." Experientia-Supplements Only**90**: 13-28.

Crabtree, J. S. "PROTOCOLS FOR RECOMBINANT DNA ISOLATION, CLONING & SEQUENCING…."

Crameri, A., E. Whitehorn, et al. (1996). "Improved green fluorescent protein by molecular evolution using." Nat. Biotechnol**14**(3): 315-319.

Cromwell, M. E., E. Hilario, et al. (2006). "Protein aggregation and bioprocessing." The AAPS journal**8**(3): E572-E579.

Diederen, B. (2008). "< i> Legionella</i> spp. and Legionnaires' disease." Journal of Infection**56**(1): 1-12.

Dümmler, A., A.-M. Lawrence, et al. (2005). "Simplified screening for the detection of soluble fusion constructs expressed in E. coli using a modular set of vectors." Microbial Cell Factories**4**(1): 34.

Fukuzawa, H., E. Suzuki, et al. (1992). "A gene homologous to chloroplast carbonic anhydrase (icfA) is essential to photosynthetic carbon dioxide fixation by Synechococcus PCC7942." Proceedings of the National Academy of Sciences**89**(10): 4437-4441.

Garcia, J. S., C. S. d. Magalhães, et al. (2006). "Trends in metal-binding and metalloprotein analysis." Talanta**69**(1): 1-15.

Gill, S. R., P. J. Fedorka-Cray, et al. (1984). "Purification and properties of the carbonic anhydrase of Rhodospirillum rubrum." Archives of microbiology**138**(2): 113-118.

Goodsell, D. S. (2009). The machinery of life, Springer.

Gourdon, P., J. L. Andersen, et al. (2011). "HiLiDe瑿Systematic Approach to Membrane Protein Crystallization in Lipid and Detergent." Crystal Growth & Design**11**(6): 2098-2106.

Gräslund, S., P. Nordlund, et al. (2008). "Protein production and purification." Nature methods**5**(2): 135-146.

Hames, B. D. (1998). Gel Electrophoresis of Proteins: A Practical Approach: A Practical Approach, OUP Oxford.

Hatzios, S. K. and C. R. Bertozzi (2011). "The regulation of sulfur metabolism in Mycobacterium tuberculosis." PLoS pathogens**7**(7): e1002036.

High, A. S., S. D. Torosian, et al. (1993). "Cloning, nucleotide sequence and expression in Escherichia coli of a gene (ompM) encoding a 25 kDa major outer-membrane protein (MOMP) of Legionella pneumophila." Journal of General Microbiology**139**(8): 1715-1721.

Hunte, C., G. von Jagow, et al. (2003). Membrane protein purification and crystallization: a practical guide, Academic press.

Jackson, M. and H. H. Mantsch (1995). "The use and misuse of FTIR spectroscopy in the determination of protein structure." Critical Reviews in Biochemistry and Molecular Biology**30**(2): 95-120.

Kawate, T. and E. Gouaux (2006). "Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins." Structure**14**(4): 673-681.

Kertesz, M. A. (2000). "Riding the sulfur cycle–metabolism of sulfonates and sulfate esters in Gram- negative bacteria." FEMS microbiology reviews**24**(2): 135-175.

Krishnamurthy, V. M., G. K. Kaufman, et al. (2008). "Carbonic anhydrase as a model for biophysical and physical-organic studies of proteins and protein–ligand binding." Chemical reviews**108**(3): 946.

Lestrate, P., R. M. Delrue, et al. (2000). "Identification and characterization of in vivo attenuated mutants of Brucella melitensis." Molecular microbiology**38**(3): 543-551.

Lin, S.-H. and G. Guidotti (2009). "Purification of membrane proteins." Methods in enzymology**463**: 619-629.

Markovich, D. (2001). "Physiological roles and regulation of mammalian sulfate transporters." Physiological reviews**81**(4): 1499-1533.

Markovich, D. (2011). "Physiological roles of renal anion transporters NaS1 and Sat1." American Journal of Physiology-Renal Physiology**300**(6): F1267-F1270.

Matouschek, A., J. T. Kellis, et al. (1989). "Mapping the transition state and pathway of protein folding by protein engineering." Nature**340**(6229): 122-126.

Piłsyk, S. and A. Paszewski (2009). "Sulfate permeases–phylogenetic diversity of sulfate transport." Acta Biocimica Polonica**56**(3): 375-384.

Pocker, Y. (2000). "Water in enzyme reactions: biophysical aspects of hydration-dehydration processes." Cellular and Molecular Life Sciences CMLS**57**(7): 1008-1017.

Prendergast, F. G. and K. G. Mann (1978). "Chemical and physical properties of aequorin and the green fluorescent protein isolated from Aequorea forskalea." Biochemistry**17**(17): 3448-3453.

Rodgers, F. G. and M. R. Davey (1982). "Ultrastructure of the cell envelope layers and surface details of Legionella pneumophila." Journal of general microbiology**128**(7): 1547-1557.

Sambrook, J., E. F. Fritsch, et al. (1989). Molecular cloning, Cold spring harbor laboratory press New York.

Seddon, A. M., P. Curnow, et al. (2004). "Membrane proteins, lipids and detergents: not just a soap opera." Biochimica et Biophysica Acta (BBA)-Biomembranes**1666**(1): 105-117.

Shekh, A. Y., K. Krishnamurthi, et al. (2012). "Recent Advancements in Carbonic Anhydrase–Driven Processes for CO2 Sequestration: Minireview." Critical Reviews in Environmental Science and Technology**42**(14): 1419-1440.

Sherwood, L. (2012). Human physiology: from cells to systems, Thomson Brooks/Cole.

Smith, K. S. and J. G. Ferry (2000). "Prokaryotic carbonic anhydrases." FEMS microbiology reviews**24**(4): 335-366.

Smith, K. S., C. Jakubzick, et al. (1999). "Carbonic anhydrase is an ancient enzyme widespread in prokaryotes." Proceedings of the National Academy of Sciences**96**(26): 15184-15189.

Sun, P., J. E. Tropea, et al. (2011). Enhancing the solubility of recombinant proteins in Escherichia coli by using hexahistidine-tagged maltose-binding protein as a fusion partner. Heterologous Gene Expression in E. coli, Springer**:** 259-274.

Supuran, C. T. (2003). "Carbonic anhydrase inhibitors in the treatment and prophylaxis of obesity." Expert Opinion on Therapeutic Patents**13**(10): 1545-1550.

Supuran, C. T. (2007). "Carbonic anhydrases as drug targets-an overview." Current topics in medicinal chemistry**7**(9): 825-833.

Surade, S., M. Klein, et al. (2006). "Comparative analysis and “expression space” coverage of the production of prokaryotic membrane proteins for structural genomics." Protein science**15**(9): 2178-2189.

Tina, K., R. Bhadra, et al. (2007). "PIC: protein interactions calculator." Nucleic acids research**35**(suppl 2): W473-W476.

Vogt, G., S. Woell, et al. (1997). "Protein thermal stability, hydrogen bonds, and ion pairs." Journal of Molecular Biology**269**(4): 631-643.

8. APPENDIX

SO_Legionella_original 2307 bp linear DNA

```
      1 atgattgata gaaatatagt taaccagcgt ctatttcgta tctactctaa acgttattta
61 aaatttgatt ttgttgctgc aattgtcgtt ttcctcgtgg cgattcctct atgcttgggt
121 attgccctgg cttctggcgc acctcttttt tctggtatct taagtggaat cattggcggt
181 attattgtgg gtatatttag cggctcacaa gttagtgtca gcgggcctgc tgcgggtatg
241 gcagccgttg tattggcggc aatctctcaa cttggtgatt ttaatacctt cttattggct
301 cttacgattg caggcctttt acaaatgata taggtgcat taagggcagg atttgttgca
361 gattatgttc cttcgaacgt agttcaaggc ttactgtgtt caatcgggat cttgctaatc
421 attaagcaat taccacttgc attcactctc tcatcagatt ttgatgagct taaaacacat
481 ttattggaaa caacagaagg gtttacagta agccctttgt tagctttgtc gcagcacatt
541 aatgaaggcg cactcatcat tactactctt tcattagcta ttttaatcta ttttgatata
601 accaaaaata aaattctgaa ggaaatccca gcccctatcc ttgtggttct ggcaggaata
661 ttgttaaatg agcttttcat ctggacagac tccagcctgg cacaaaactc gcctcaactg
721 gttaatatcc cagataccaa tggttttttc caatttttca gccatcttga ataccctgac
781 tggtctgcct ggactaatcc taaagtctat ctctatgcgc tagtcatatg tattgttgct
841 tctttggaaa cgttgcttaa tctcaaagcg tcagaacgat tggataaaaa aagaaggcac
901 agcccaacta accaggaatt agttgctcaa ggcttgggta atataacatc cggtttggtt
961 gggggtattc ctgttacatc agtcattgtc agatcatcca ttaatatcca tgctggttca
     1021 aaaaccaaat tttcagccgt actccatggt tttttcatct tgtttgctgt catgttaata
     1081 cctggtgcat tgaataaaat tcctttgtct tctctcgctg caatttaat ttacactggc
     1141 tataaactaa ataaaccagc tatctatatt aatatatttt cacaaggcag cgatcgattc
     1201 atccctttta tagtgaccgt gatcagtatt attgccttca atcttcttgc aggtatatta
     1261 ataggtctgg ctatcagtct cttttatata ttaaagtcca atagccaagc tcgtattgat
     1321 atcattaagg aaatttatcc aaatggttca acataccgtt tgatgcttcc ccaacaaatg
     1381 acttttctta acaaagctgc gttggtagct gaattggata cattgcccag acgttctcaa
     1441 ctgattatcg atgcacgtca ttctcaatac attgataaag aaattctgga cttgtttaat
     1501 gaattgaaag aagagttggc aggcagtaaa caaatttctg taaatttcac cggatttcag
     1561 gaacactata aaattcataa ctacatcgac ttcattacgg ttacaaccta tgatgttcaa
     1621 actaacctga cgcccgcaaa agtattaaac atactgcttg agggaaatca gcgcttttta
     1681 agcgataacc ggattcatcg atcaaaccaa atcgatataa aatacactgc aaaaacacaa
     1741 catccaattg ctgtggtttt agcctgtatc gattcaagag ttcctgtaga aaccatcttt
     1801 gatatgagtt tcggtgatct tttttgtgtc cgcatagcag gcaatgtgat taatgacgat
     1861 attttggcca gcatagaata cgcctgtaac gtagttggtg caaagcttat catggttctg
     1921 ggacataccc gatgcggggc cattcaatct gcctgtgatg gtattgaaaa aggtcatatc
     1981 acccaactgc tctctaaaat taagcctgca gttaacgcag aaaaagaaac tacaacagaa
     2041 cgtaacggaa aaaatcaaac ttttgtaaat catgtcactg aattaaatgt tgcgaatacc
     2101 ttacagaaca tttataaaaa aagtgatatt ttacggacta tgatagacag taatgagata
     2161 gggatggttg gtgctattta tgatgtaagc agcggaaaag tgagctgtaa aacatatacc
     2221 gaagagttat ccactttgga tggggacgaa aataagctcc tggctcaaaa atttgaaacc
     2281 attcttttag aagccaaagc aagttaa
```

Figure A.1; The original DNA sequence of *Legionella pneumophilia*that encodes sulfate-transporter protein (LPL0734). The middle of the sequence contains NcoI restriction site. The aim of this cloning is also to check the silent mutation of the internal NcoI restriction site in the sequence of gene encoding sulfate transporter in *Legionella pneumophilia*.

Cloning sequence SO_Legionella 2323 bp

```
gcaccatggct attgata gaaatatagt taaccagcgt ctatttcgta tctactctaa acgttattta
61 aaatttgatt ttgttgctgc aattgtcgtt ttcctcgtgg cgattcctct atgcttgggt
121 attgccctgg cttctggcgc acctcttttt tctggtatct taagtggaat cattggcggt
```

```
181 attattgtgg gtatatttag cggctcacaa gttagtgtca gcgggcctgc tgcgggtatg
241 gcagccgttg tattggcggc aatctctcaa cttggtgatt ttaatacctt cttattggct
301 cttacgattg caggcctttt acaaatgata ataggtgcat taagggcagg atttgttgca
361 gattatgttc cttcgaacgt agttcaaggc ttactgtgtt caatcgggat cttgctaatc
421 attaagcaat taccacttgc attcactctc tcatcagatt ttgatgagct taaaacacat
481 ttattggaaa caacagaagg gtttacagta agccctttgt tagctttgtc gcagcacatt
541 aatgaaggcg cactcatcat tactactctt tcattagcta ttttaatcta ttttgatata
601 accaaaaata aaattctgaa ggaaatccca gcccctatcc ttgtggttct ggcaggaata
661 ttgttaaatg agcttttcat ctggacagac tccagcctgg cacaaaactc gcctcaactg
721 gttaatatcc cagataccaa tggttttttc caatttttca gccatcttga ataccctgac
781 tggtctgcct ggactaatcc taaagtctat ctctatgcgc tagtcatatg tattgttgct
841 tctttggaaa cgttgcttaa tctcaaagcg tcagaacgat tggataaaaa aagaaggcac
901 agcccaacta accaggaatt agttgctcaa ggcttgggta atataacatc cggtttggtt
961 gggggtattc ctgttacatc agtcattgtc agatcatcca ttaatatcca tgctggttca
     1021 aaaaccaaat tttcagccgt actccacggt tttttcatct tgtttgctgt catgttaata
     1081 cctggtgcat tgaataaaat tcctttgtct tctctcgctg caattttaat ttacactggc
     1141 tataaactaa ataaaccagc tatctatatt aatatatttt cacaaggcag cgatcgattc
     1201 atccctttta tagtgaccgt gatcagtatt attgccttca atcttcttgc aggtatatta
     1261 ataggtctgg ctatcagtct cttttatata ttaaagtcca atagccaagc tcgtattgat
     1321 atcattaagg aaatttatcc aaatggttca acataccgtt tgatgcttcc ccaacaaatg
     1381 acttttctta acaaagctgc gttggtagct gaattggata cattgcccag acgttctcaa
     1441 ctgattatcg atgcacgtca ttctcaatac attgataaag aaattctgga cttgtttaat
     1501 gaattgaaag aagagttggc aggcagtaaa caaatttctg taaatttcac cggatttcag
     1561 gaacactata aaattcataa ctacatcgac ttcattacgg ttacaaccta tgatgttcaa
     1621 actaacctga cgcccgcaaa agtattaaac atactgcttg agggaaatca gcgcttttta
     1681 agcgataacc ggattcatcg atcaaaccaa atcgatataa aatacactgc aaaaacacaa
     1741 catccaattg ctgtggtttt agcctgtatc gattcaagag ttcctgtaga aaccatcttt
     1801 gatatgagtt tcggtgatct tttttgtgtc cgcatagcag gcaatgtgat taatgacgat
     1861 attttggcca gcatagaata cgcctgtaac gtagttggtg caaagcttat catggttctg
     1921 ggacataccc gatgcggggc cattcaatct gcctgtgatg gtattgaaaa aggtcatatc
     1981 acccaactgc tctctaaaat taagcctgca gttaacgcag aaaaagaaac tacaacagaa
     2041 cgtaacggaa aaaatcaaac ttttgtaaat catgtcactg aattaaatgt tgcgaatacc
     2101 ttacagaaca tttataaaaa aagtgatatt ttacggacta tgatagacag taatgagata
     2161 gggatggttg gtgctattta tgatgtaagc agcggaaaag tgagctgtaa aacatatacc
     2221 gaagagttat ccactttgga tggggacgaa aataagctcc tggctcaaaa atttgaaacc
     2281 attcttttag aagccaaagc aagttaactcgagacc
```

Figure A.2: Cloning sequence of *Legionella pneumophilia*after removal of internal NcoI restriction sites present in the DNA sequence. Finally, addition of with restrictions sites at the both end of the sequence.The sequenceis 2323 bp long.