Functional Analysis of
Vitelline Membrane Outer Layer 1 (VMO1)

A thesis
submitted in partial fulfilment
of the requirements for the degree

of

Master of Science in Biological Sciences

at

The University of Waikato

by

PRANJALI R. SAGVEKAR

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2018

Abstract

Hearing loss is an auditory impairment affecting millions of people world-wide. It has considerable impact psychologically, socially, and physically in adults; with the affected numbers increasing with age, and those with learning disabilities are at a particular risk. With the current advancements in audiological technologies and research related to regeneration of inner ear cells, it is important to understand hearing loss at a molecular level and the proteins involved in the mechanism.

Because of the many complications involved in isolating genes from the actual human ear, the use of animal/bacterial models and commercial mammalian cell lines have become important and informative. In addition, bioinformatics analysis makes it possible to converge the similarities between the human genome and genes from other species in real time. This aids in screening the genetic causes which underlie hearing loss and improve our understanding to unravel diagnostic strategies.

The Vitelline Membrane Outer Layer 1 homolog (*Vmo1*) gene is of particular interest as it was uniquely and highly expressed in the inner ear of mouse. The function of this gene is currently unknown. The aim of this research is to induce recombinant human VMO1 protein in a prokaryotic system (*E. coli*) and ascertain its expression using a range of molecular and cellular assays. To do so, this research focuses on four objectives.

Our first objective was to use bioinformatics to examine for any predicted or known function. Studying the gene structure and amino acid sequence gave us better understanding of how the gene is conserved across many organisms. For example, *VMO1* gene in humans have structural similarities with 176 other organisms and highly conserved through evolution. A plasmid containing *VMO1* under the control of T7 promoter was extracted using a Zyppy™ Miniprep method and later sequenced. With the help of Geneious® software (version 7.1.8 and 10.2.3) we assembled the sequenced reads and generated a 423 nucleotides consensus sequence with a predicted molecular weight of 15 kDa. This sequence was hydrophobic in nature with conserved cysteine residues. A basic local

alignment tool matched this sequence with Human *VMO1* mRNA Isoform 3 with 100% identity at the nucleotide and amino acid level. In addition, the plasmid stability test indicated that the plasmid was indeed stable.

The second objective was to optimise growth parameters to induce the recombinant human VMO1 protein in *E. coli*. The protein induction was carried out using two methods: a traditional method and an ethanol-based method. The protein expression was assessed by resolving the cell lysates on 12% SDS-PAGE gels, followed by Coomassie Blue staining and antibody detection by His-tag antibody and VMO1 antibody. After optimising the parameters, the ethanol method was found to be the most suitable to implement since it enhanced the expression of recombinant protein with only 1% ethanol at 37°C with continuous shaking. It was easy to perform and was less time consuming. This method gave best results with terrific broth. An insoluble protein of 15 kDa was thus produced along with other multiple bands seen at 30 and 40 kDa. All of which were identified using both His-tag or VMO1 antibodies. These additional bands are hypothesized to be present because of the hydrophobic and sticky nature of the VMO1 protein or presence of start codons in their ORFs.

The third objective was to purify the protein using a trial scale purification affinity column. The protein lysate was surrendered to pre-purification steps that included treatment with DNase I, Protease inhibitor and Binding buffer. While the actual purification was performed with buffer systems without Imidazole and with Imidazole. Denaturing conditions of Imidazole and Urea allowed for proper refolding of the proteins and its subsequent identification using western blot. We were able to purify the VMO1 recombinant protein, but as multiple bands (15, 30, 40 kDa). The presence of these bands was consistent from protein expression to protein purification.

Lastly, VMO1 antibody was validated for its specificity and examined for tissue specific protein expression using immunohistochemistry on the mouse ear (P240), mouse cochlea (P5), and rat lung (L5); where Vmo1 signals were detected. Our data suggest that Vmo1 is a secreted protein rather than being localised to the Reissner's membrane in the inner ear. Thus, the IHC data is in agreement with bioinformatics analysis that show presence of *VMO1* in the lungs, kidney, brain,

etc. However, inclusion of additional control slides (i.e., no primary antibody) would further strengthen this data.

Thus, in the future it is recommended, to excise the three observed bands and send it for Mass spectrometric analysis and use molecular weight cut off (MWCO) that would retain proteins (e.g., 20 kDa and above) on a membrane. This will further purify the 15 kDa protein. In addition, a large-scale purification is to be undertaken after optimising the wash buffer and resin to protein ratio in order to achieve a single purified band (15 kDa). To further understand the specificity of VMO1 antibody, proteins extracted from mouse ear, rat lung, human tear, chicken vitelline membrane, etc., can be subjected to western blot detection. To understand the antimicrobial role of VMO1 protein, the protein lysate could be tested on medically important bacterial isolates like *S. aureus*, *Salmonella spp* etc., and look for its inhibition. Finally, protein-protein interaction between VMO1 and Lysozyme, transferase activity can be studied using Pull-down assays, Yeast two hybrid screens, etc.

Acknowledgements

My sincere thanks to the University of Waikato for granting me an offer of place to study MSc in New Zealand and for my International scholarships. I take this opportunity to express my gratitude towards my supervisor Dr. Linda Peters who with her endless patience, altruistic help, affectionate attitude, views, and fore thoughts has guided me throughout the research. Her understanding, encouraging and personal guidance have provided a good basis for my present thesis.

I am highly obliged to C.2.03, for the cordial atmosphere present throughout the period of my work, and for their constant support. Thank you Dr. Sari Karppinen, Olivia Patty, Dr. Judith Burrows, and Dr Barry O'Brien for all the technical advice and training. I also thank Dr. Greg Jacobson, Dr. Steve Bird, Dr. Ray Cursons, and Mr Tony Cecire for always addressing my doubts and concerns. My warm regards to the administration of the University for their Co-operation, advice, and suggestions throughout my academic stay. Also, sincere thanks to Cheryl Ward for all the help and support with using EndNote.

It would have been an uphill task for me to accomplish what I have, in little time if I did not have the blessings as well as mental/financial support, unbound love, and care from my family. For which, I express my gratitude to my parents (Manisha Sagvekar and Ravindra Sagvekar), brother (Varad Sagvekar), and my partner (Nikhil Chaudhari).

My heartfelt thanks to my dearest group of classmates/lab mates for being patient and ever so caring. Thank you Bhavna, Emily, Zehan, Mannar, and Andrew. Finally, I thank God for his blessings and helping me overcome my difficulties.

Table of Contents

List of Figures

List of Tables

Abbreviations

3'	Three prime DNA end
5'	Five prime DNA end
α	Alpha
β	Beta
°C	Degrees Celsius
µg	Microgram
µL	Microlitre(s)
µM	Micromolar
aa	Amino acid(s)
AHL	Acquired hearing loss
APS	Ammonium persulfate
BIS N,N'	Methylene-bis-acrylamide
BLAST	Basic local alignment search tool
bp	Base pair(s)
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
Conc	Concentration
Cl^-	Chloride
Da	Dalton(s)
dB	Decibel(s)
DDH2O	Double distilled water
DEPC	Diethyl dicarbonate
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
Et	Ethanol
EtBr	Ethidium Bromide
et al	And others
e.g.	Example
FASTA	File format for nucleotide and protein sequences
g	centrifugal force (centrifugation)
HCl	Hydrochloric acid
HL	Hearing loss

Hr(s)	Hour(s)
Hz	Hertz
IHC	Immunohistochemistry
IHCs	Inner hair cells
IPTG/IP	Isopropyl β-D-1-thiogalactopyranoside
K	Potassium
kb	Kilobase
kDa	Kilodalton(s)
L	Litre(s)
LB	Luria base/broth
M	Molar
Ma	Milliampere(s)
mg	Milligram(s)
MΩ	Mega Ohms
MIDD	Maternally inherited diabetes and deafness
Min(s)	Minute(s)
mL	Millilitre(s)
mM	Millimolar
MPSS	Massively parallel signature sequencing
Milli. Q H2O	Millipore Q ultra-purified water
mRNA	Messenger RNA
MTT	Mitochondrial mutation deafness loci
mV	Millivolt(s)
N terminal	free amine group (-NH2) at the end of a polypeptide
Na	Sodium
NC	Negative control
NaOH	Sodium hydroxide
NCBI	National centre for biotechnology information
NEB	New England biolabs
ng	Nanogram(s)
nt	Nucleotides
OHC	Outer hair cell
OOC	Organ of Corti
O/N	Overnight
ORF	Open reading frame

P	Pellet
P5	Postnatal day 5
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PC	Positive control
PCR	Polymerase chain reaction
pI	Isoelectronic point (pH at which the aa does not migrate in an electric field)
PVDF	Polyvinylidene fluoride
RefSeq	Reference sequence
RM	Reissner's membrane
mRNA	messenger Ribonucleic acid
rpm	Revolutions per minute
RT	Room temperature
S	Supernatant
SDS	Sodium dodecyl sulfate
Sec(s)	Second(s)
SNP	Single nucleotide polymorphism
TBS	Tris buffered saline
TBS-T	Tris buffered saline with Tween 20
TEMED	Tetra-methyl ethylene di-amine
U	Enzyme units
UTR	Untranslated region
UV	Ultra violet
V	Volt(s)
VM	Vitelline membrane
VMO1	Human and chicken vitelline membrane outer layer 1 gene
Vmo1	Mouse vitelline membrane outer layer 1 gene
Vmo1	Mouse vitelline membrane outer layer 1 protein
VMO1	Human, mouse, and chicken vitelline membrane outer layer 1 protein and antibody
WC	whole cells
~	Approximately
%	Percentage

CHAPTER ONE

LITERATURE REVIEW AND INTRODUCTION

1.1 Hearing loss

The ability to detect vibrations and perceive sound through both ears is termed as “Hearing”. Any individual unable to hear within the normal hearing threshold (≥ 25 decibel) is said to have the condition of hearing loss (World Health Organisation, 2017). Hearing loss can be further classified as slight, mild, moderate, severe, or profound (Table 1.1).

1.1.1 Classification of the severity of hearing impairment

Audiologists have often used various adjective descriptors to assign the degree of hearing disabilities. These descriptors are usually based on a pure tone average (PTA) of frequencies that is variable from low frequency of 500 Hertz (Hz), to middle frequencies of 1000 and 2000 Hz. These thresholds are measured on decibel (dB) hearing loss (HL) scale called the audiometric "zero" (average hearing level at each frequency in healthy young adults). However, the PTA has undergone several changes amongst investigators based on acoustic signals (Wedenberg's system 1954), audiological tests (American Speech-Language-Hearing Association, 1964), audiogram configurations, frequency, and intensities (Risberg and Martony, 1972). Table 1.1 provides a modified scale (Goodman, 1965) for hearing impairment in a better standardised way (NIH, 2011; Clark, 1981).

Table 1.1 Classification of hearing loss (Clark, 1981).

Classification	PTA range in dBHL
Normal hearing	-10 to 15
Slight hearing loss	16 to 25
Mild hearing loss	26 to 45
Moderate hearing loss	46 to 55
Moderately severe hearing loss	56 to 70
Severe hearing loss	71 to 90
Profound hearing loss	91 and more

HL can be classified as conductive, sensorineural, and mixed loss (The American Speech-Language-Hearing Association (ASHA). Conductive HL is caused by lesions in the external auditory canal or middle ear and prevents the reception of sound to the inner ear (Figure 1.1). Lesions caused by sensorineural HL are present either in the inner ear or the auditory nerve. Mixed HL is caused by chronic infection, extreme head injury, genetic disorders, or during a state of transient HL accompanied with sensorineural HL (The National foundation for the Deaf, 2017).

According to WHO, HL is the most frequent sensory impairment affecting 1 in 500 new-borns and 1 in 300 children by the age of 4. Genetically, 70 percent of it is non-syndromic (i.e., HL is the only phenotype), while 30 percent is syndromic (i.e., additional clinical findings). In addition, it can be age-related. Table 1.2 illustrates the scenario of HL in New Zealand, where 330,269 people aged ≥14 years have been affected and this number is postulated to be high by 2061. These Figures are based on the fact that older population will increase, predominantly more in the rural areas (Northland, West Coast, Tasman) than urban areas (Auckland, Wellington, Christchurch and Otago), where hearing loss will be prominent.

Table 1.2 The prevalence of hearing loss among the New Zealand population aged ≥65 years, by region.

Adapted from (Exeter. *et al.* 2015)

Region	2011	2012	2013	Change (%) 2011–2013	Rural areas 2013 (%)
Northland	5,084	7,185	9,422	85.33	50.19
Auckland	30,839	45,326	65,348	111.90	3.98
Waikato	11,130	15,527	20,435	83.60	22.88
Bay of Plenty	8,794	12,033	15,841	80.13	18.33
Gisborne	1,158	1,610	2,140	84.80	25.19
Hawke's Bay	4,652	6,340	8,088	73.86	12.69
Taranaki	3,455	4,554	5,732	65.90	22.98
Manawatu-Wanganui	7,067	9,364	11,817	67.21	19.29
Wellington	11,974	16,371	21,475	79.35	3.84
Tasman	1,512	2,297	3,082	103.84	41.21
Nelson	1,413	1,983	2,650	87.54	1.90
Marlborough	1,688	2,414	3,082	82.58	23.07
West Coast	1,021	1,453	1,884	84.52	43.42
Canterbury	16,627	23,360	31,231	87.83	16.61
Otago	5,987	8,068	10,325	72.46	20.82
Southland	2,827	3,749	4,711	66.64	30.57

In New Zealand, HL is relatively common as it affects one in six New Zealanders. In 2016, the frequency was estimated to be 880,350 people, which represents ~18.9% of the general population. In addition, its prevalence among males was greater (472,961 people) than females (407,388) (The National foundation for the Deaf, 2017).

1.2 Anatomy of the human ear

The ear functions as a sensory system to detect sound. It is unguarded to thermal injury not only due to its protruding morphology and site but also of the subcutaneous tissue layers and thin skin (Bos *et al.*, 2016). The human ear is anatomically composed of three parts: the outer, middle, and the inner ear (Figure 1.1). The outer and the middle ear transmit sound waves while the inner ear transduces sound stimuli to nerve impulses via the auditory nerve (Alters, 2000).

1.2.1 The outer, middle, and inner ear

The outer ear represents the external portion of the ear. The elastic cartilage called the Pinna extends into the auditory canal to form the outer ear. The four-centimetre-long ear canal ends with the outer lining of the eardrum, also called as the tympanic membrane. Thus, the outer ear first transmits sound to the eardrums. In comparison, the middle ear is made up of three ossicles (malleus (hammer), incus (anvil), and the stapes (stirrup) filled with air spaces and conduct sound from the eardrum to the inner ear. The eardrum forms the outer wall of the middle ear, while the cochlea forms the inner wall. The inner ear consists of a bony, spiral shaped (two and half turn) cochlea that convert vibrations into nerve impulses and a vestibular system, which is the organ for balance (Alberti, 2001).

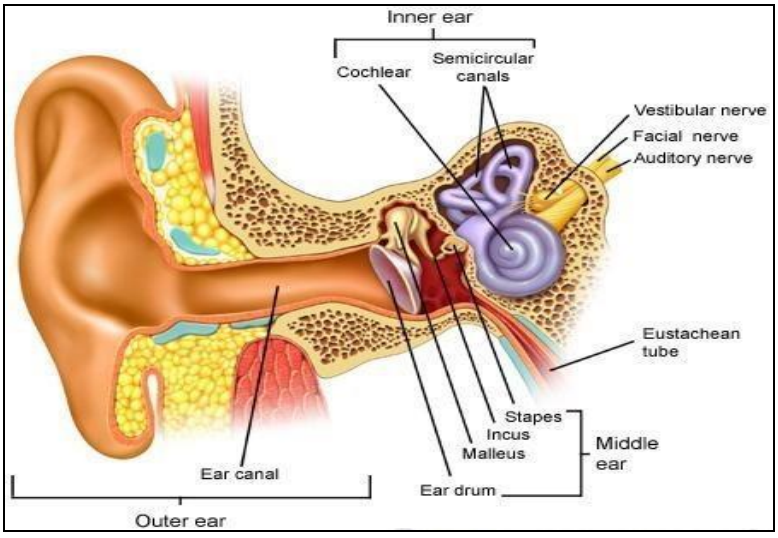


Figure 1.1 Anatomy of the human ear.

The inner ear represents the sensory organ required for hearing (cochlea) and balance (semi-circular canals). [Adapted from www.keywordsuggest.org].

In the snail shaped mammalian cochlea, the Scala media (SM) is separated from the Scala vestibuli (SV) by the Reissner's membrane (RM) (Figure 1.2). The Scala tympani (ST) is separated by the epithelial cells on the basilar membrane (also contains the Organ of Corti) and the epithelial cells of the inner and outer sulcus (Eckhard *et al.*, 2014). It is also filled with two fluids; 'Endolymph' contained in the SM, and 'Perilymph' contained in the SV, ST. The sensory receptor, known as the Organ of Corti (OOC) is embedded inside the cochlea and contains sensory

hair cells. The RM was first documented by Reissner's in 1851 and is composed of monolayer of epithelial cells. Although the function of the RM is unknown, it is speculated that it allows fluid transportation owing to its permeable nature. The pinocytotic vacuoles, microvilli and the cells present in the endolymphatic region may accentuate this function by providing high levels of adenosine triphosphatase (ATPase) activity from the basolateral and apical membranes. Cyclic Adenosine mono phosphate (AMP) may also act as a secondary regulator of ion transport across RM (Harada, 2012; Alberti, 2001; Dallos *et al.*, 1996).

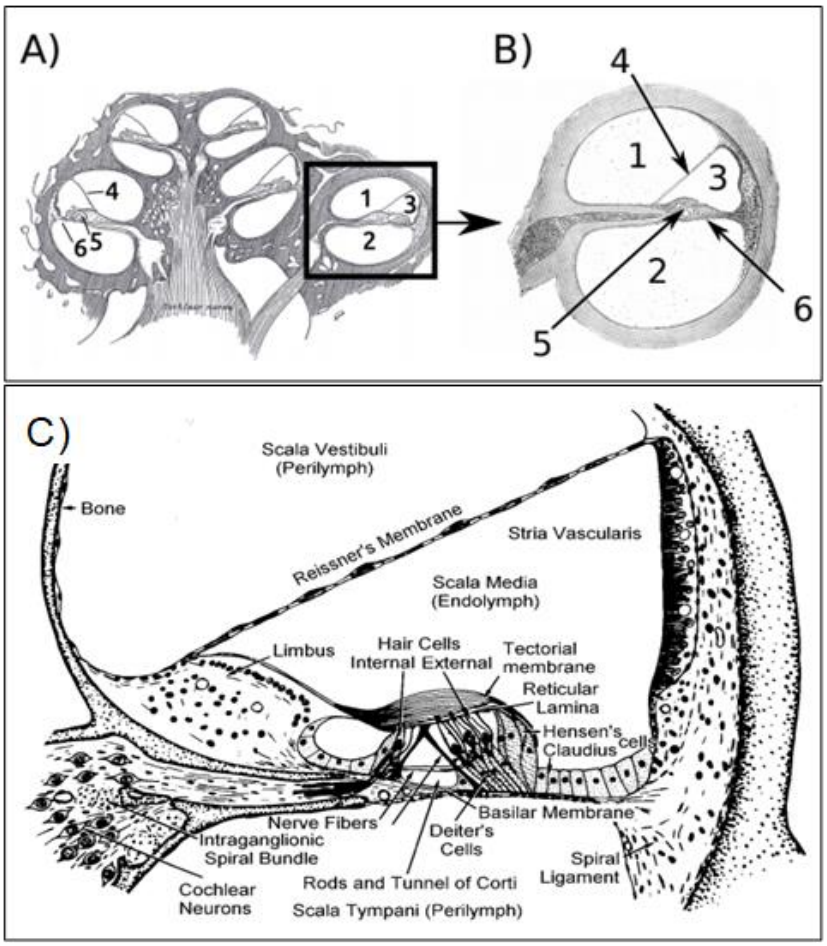


Figure 1.2 Anatomy of Organ of Corti.

A) Section of the inner ear. B) Enlargement of one cochlear duct [Adapted from (Meyer et al., 2017)]. C) One turn of the cochlea showing details of the membranous labyrinth [Adapted from (Alberti, 2001)].

There are several intercellular junctions that tightly seals the cochlear duct epithelium (CDE) and form a cochlear perilymph–endolymph barrier (PEB). Thus, allowing many transmembrane proteins to facilitate the exchange of water

(aquaporins: AQP2, AQP4 and AQP6) and ionic concentrations of Potassium (K^+), Sodium (Na^+) and Chloride (Cl^-) (Eckhard *et al.*, 2014). K^+ channels KCNQ1 and KCNJ10 have been observed in the middle ear epithelial cells (cytoplasm) of BALB/c mice (Morris *et al.*, 2012).

A remarkable difference yet lies between the perilymph and the endolymph. Since the former is thought of containing cerebrospinal fluid or blood filtrates, it establishes a high Na^+ concentration (~145 mM), a low K^+ (~5 mM), and Calcium (Ca^{2+}, 1– 2 mM). Uniquely, the latter contains only 1 mM Na^+, K^+ (155 mM) and Ca^{2+} in nanomoles. Endolymph dysregulation has also been documented in the past through Meniere's disease (Beitz *et al.*, 2003). The cochlear K^+ recycling is believed to establish endo-cochlear potential and is supported majorly by Stria Vascularis (Figure 1.3) (Hibino *et al.*, 2010).

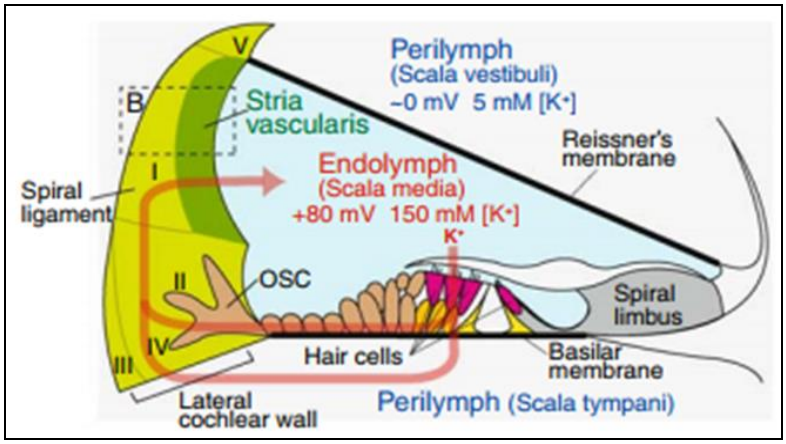


Figure 1.3 The ionic composition and potential of the endolymph.
The red arrow shows two possible routes of K+ transport from the endolymph, one through the gap-junctions of epithelial cells resting on the basilar membrane and the other through the perilymph of ST. [Adapted from (Hibino et al., 2010)].

The OOC rests itself on the basilar membrane, in the SM of the cochlea. Here, it provides a base for mechano-sensory hair cells, and is composed of three rows of outer hair cells (OHCs) and one row of inner hair cells (IHCs) (Figure 1.4). These hair cells are made up of stereocilia at the apical end that protrude outwards and attach the basilar membrane from the bottom, while the stereocilia are in contact with an overhead tectorial membrane (Figure 1.2 C). As a physiological function, the basilar membrane vibrates upon reception of sound waves and

creates shearing forces between itself and the tectorial membrane. This leads to internal changes within the stereocilia (driven by Myosin and Ca^{2+}) to create electrical signals (Watson *et al.*, 2012; Alberti, 2001; Marcus, 2011). The IHCs have a sensory function due to formation of synaptic connections with the auditory nerve and the OHCs are believed to be involved with an active motile role that possess sensitivity, frequency selection and amplification of vibrations in the OOC (Russell, 1987; Etournay *et al.*, 2010).

Other noticeable features that differentiate IHCs from OHCs are shape of the cells (wine-bottle shaped vs. cylindrical), extra cell wall membrane (poorly developed or absent vs. well developed), endoplasmic reticulum organization (poorly developed vs. well developed), organization of mitochondria (dispersed vs. closely associated with ER), etc. (Lim, 1986).

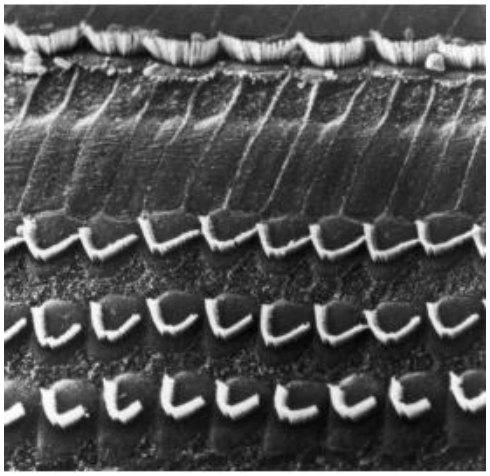


Figure 1.4 A surface view of the top of the hair cells.
The inner hair cells (top row) and outer hair cells (3 bottom rows) are seen in the scanning electron microscopic image. [Adapted from (Alberti, 2001)].

1.3 Underlying causes of HL

The etiology and causes of HL can be classified as genetic or environmental; congenital or acquired; sudden or progressive; and temporary or permanent (WHO, 2015). Congenital causes can be due to hereditary or environmental factors such as prenatal exposure *in utero* to maternal disease or inappropriate drug use such as aminoglycosides, cytotoxic drugs, antimalarial drugs, and diuretics (Campo *et al.*, 2013; Ding *et al.*, 2016). Acquired causes are due to noise

exposure, ageing (presbycusis), diseases/disorders (measles, otitis media, cancer) (Cohen *et al.*, 2014; Dhooge, 2003; Oliveira *et al.*, 2016). Other cause includes physical trauma (Patel & Groppo, 2010), or build-up of cerumen (earwax) in the ear canal (Subha & Raman, 2006). Mechanical, biological, and molecular stresses (inflammation, oxidative stress, energy exhaustion and excitotoxicity) can further aggravate cell death by apoptosis and necrosis (The National foundation for the Deaf, 2017; WHO, 2017).

1.3.1 HL at the genetic level

Non-syndromic HL follows a Mendelian inheritance pattern for a single gene mutation and can be further categorized depending on the inheritance as autosomal dominant (DFNA, a delayed onset), autosomal recessive (DFNB, usually congenital), mitochondrial, or X-linked (DFNX). Currently, 125 deafness loci are documented that includes 58 DFNA loci, 63 DFNB loci, and 4 X-linked loci (Camp & Smith, 2018; Egilmez & Kalcioglu, 2016).

Fifty percent of hearing impairment arising due to disruption of gap junctions are correlated to a gene called Connexin 26 (*GJB2*) (Shalit & Avraham., 2008). Autosomal recessive inheritance accounts for eighty percent of non-syndromic genetic HL, while autosomal dominant inheritance for the remaining twenty percent and typically leads to progressive sensorineural hearing loss (SNHL) with unreliable severity that usually begins around 10 to 40 years (Peters *et al.*, 2002). Mitochondrial inheritance patients tend to incur progressive SNHL due to rearrangements caused by point mutations and begins at 5-50 years, with variable degree of HL (Camp & Smith, 2018; Chang, 2015). Example, gene encoding components of the hair cells and the nerves (PMCA2 and otoferlin), cytoskeleton proteins (myosin VI, myosin VIIA, and myosin XVA), and proteins significant for potassium recycling in the OOC (connexin, KCNQ4, Pendrin, and Claudin 14) (Shalit & Avraham., 2008). Table 1.3 presents additional examples of HL types.

Table 1.3 Causes of hereditary hearing loss and deafness.

[Adapted from (Shearer AE et al., 1999; Camp & Smith, 2018; Song et al., 2012)].

Type of hearing impairment	Locus	Gene	Disease/Syndrome
Mitochondrial syndromic	-	*MTTL1*	Mitochondrial encephalopathy, lactic acidosis, and stroke-like episodes (MELAS) and maternally inherited diabetes and deafness (MIDD)
Mitochondrial nonsyndromic	-	*MTRNR1*	Aminoglycoside induced/ worsened, Palmoplantar keratoderma
Autosomal dominant syndromic	-	*PAX3* *MITF* *EDNR,* *EDN3* *SOX10*	Waardenburg syndrome
Autosomal recessive syndromic	DFNB18, DFNB 23, DFNB 48	*USH1C* *PCDH15* *CIB2*	Usher syndrome type I
Autosomal dominant nonsyndromic	DFNA3	*GJB6*	Progressive sensorineural hearing impairment
Autosomal recessive nonsyndromic	DFNB9	*OTOF*	Prelingual non-syndromic HL, temperature-sensitive nonsyndromic auditory neuropathy
X-linked syndromic	DFN1	*TIMM8A*	Deafness-dystonia-optic neuronopathy syndrome (Mohr-Tranebjaerg syndrome)
X-linked non-syndromic	DFNX5	*AIFM1*	Auditory neuropathy spectrum disorder

So far, many of the deafness genes are protein-coding genes. Moreover, this list includes expression of transcription factors, gap junctions, ion channels, membrane transporters, adhesion proteins, myosins, extracellular matrix proteins and cytoskeletal proteins. Apart from these, mutations in the transfer RNA (tRNA) and Ribosomal RNA (rRNA) coding genes and intron region of chromosomes have also found to cause HL (Jing *et al.*, 2012; Xing *et al.*, 2007; Park *et al.*, 2010).

Recruiting large families and undertaking a linkage and positional cloning has identified genes involved in HL. With the improvement in sequencing technology, genes expressed in discrete cell types can be identified to investigate the mechanism of inner ear physiology and development.

1.4 Discovery of gene specific transcripts in the mammalian inner ear

Massively parallel signature sequencing (MPSS) libraries from micro-dissections of 3 inner ear tissues were prepared and compared to 87 other tissues online, in the Mouse Reference Transcriptome (MRT) database (Peters *et al.*, 2007). Interestingly, the second-most abundant signature was for Vitelline membrane outer layer 1 homolog (*Vmo1)* named after chicken VMO1.

Vmo1 was shown to be expressed specifically in the mouse inner ear by reverse transcription polymerase chain reaction (RT-PCR) of cDNA (liver, kidney, pancreas, retina, brain, testes, and inner ear) and genomic DNA from mouse. The evidence was further advocated by performing an *in situ* hybridisation to detect *Vmo1* mRNA expressed exclusively in Reissner's membrane (Figure 1.5 A). The function of this gene in the inner ear is currently unknown. Out of the top ten most abundant MPSS signatures unique to one or more of the mouse inner ear libraries, seven have been shown to be associated with HL by evidence in humans or in a mouse model (Peters *et al.*, 2007). Thus, the role of *Vmo1* in the inner ear is of special interest.

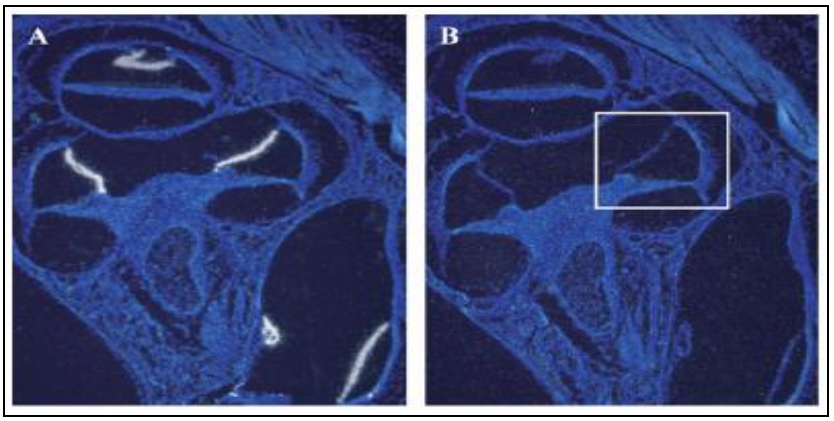


Figure 1.5 In situ hybridization in cross sections of the mouse inner ear.
A) Antisense probe for Vmo1 showing signal localization in the Reissner's membrane. B) Control (sense) probe for Vmo1. [Adapted from (Peters et al., 2007)].

1.5 Biology of the vitelline membrane

The vitelline membrane (VM) was first documented in the hen's (*Gallus gallus)* egg by X ray crystallography (Shimizu *et al.*, 1994). It is composed of two distinct layers; outer layer (OVM-*lamina extravitellina*) and the inner layer (IVM-*lamina perivitellina*) (Figure 1.6 B). The VM separates the yolk from the egg white (Figure 1.6 A) (McNally, 1943; Kido *et al.*, 1992; Shimizu *et al.*, 1994). The IVM is formed in the ovary before ovulation, whereas the OVM is formed in the upper oviduct after ovulation. The IVM has a dense fibrous network whose major glycoproteins are GP-I to GP-IV. While, the OVM is in close contact with the albumin and has a lattice fibril network that comprises of ovomucin, lysozyme and VMO1 (Kido *et al.*, 1992).

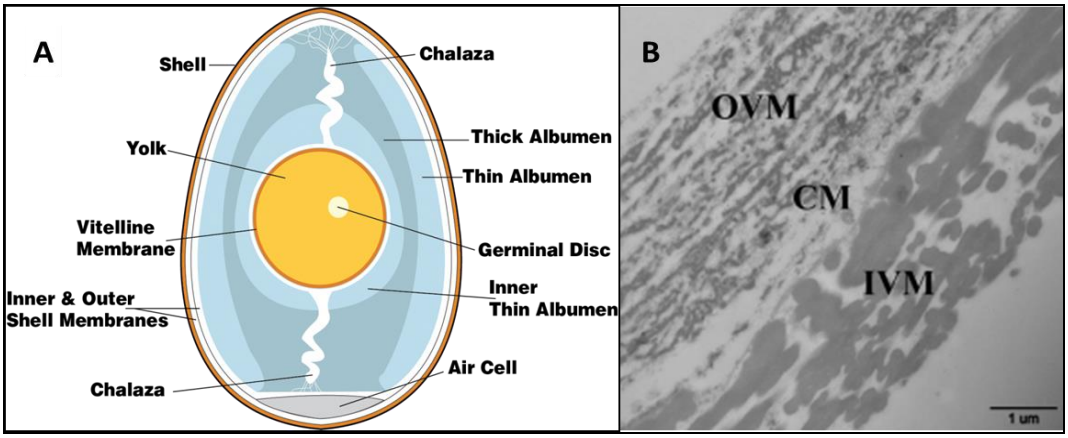


Figure 1.6 Chicken egg and VM.

A) A chicken's egg. [Adapted from (ThermoBlog, 2018)]. B) Transmission electron microscopy of vitelline membrane. OVM: outer VM; CM: continuous membrane; IVM: inner VM. [Adapted from (Li et al., 2017)].

The VM is also found in the nematode (*Caenorhabditis elegans*) (Stein & Golden, 2015) and fruit fly (*Drosophila melanogaster*) egg (Pascucci *et al.*, 1996) and possess a similar function and three distinctive layers depicted in the chicken egg. The embryo of *C. elegans* is surrounded by a concealed inner VM and a distinguished outer chitinous eggshell. When the VM was gently beamed with laser, the eggshell had the property of resealing over time. However, this membrane also lost the resealing agility when bombarded with larger holes. This scenario led to an impaired gastrulation that rendered the guts precursor cells

inefficient to migrate towards an embryonic halt. This emphasizes another critical role of the VM for pattern formation of its microenvironment to safeguard the embryo (Schierenberg & Junkersdorf, 1992). Thus, showing the importance of VM in the embryonic development.

In another study, the VM domain was characterized by the presence of three promptly spaced cysteine residues (CX7CX8C). The VM proteins along with VM domains are integrated into large disulphide linkages during late oogenesis, which are frequently utilised in extracellular matrices to stabilize other non-covalent interactions. The regulation of disulphide bond formation for balancing early elasticity and late stability in the extracellular region may be beneficial for morphogenesis of proper vitelline membrane (Wu *et al.*, 2010).

In *D. melanogaster*, the eggshell is composed of a VM that undergoes cross-linking during the oogenesis and disulphide linkages bind the VM proteins. It was also observed that when the VM was treated with reducing agents it leads to membrane solubilization. Besides the VMs structural role, it may constitute a repository for products of follicle cells involved in embryonic patterning. Vitelline membrane defects have also been detected in some mutants of the germ-line dependent genes; *fs (1) polehole [fs (1) ph]*, and *fs (1) Nasrat [fs (1) Nas*. Mutations in these genes fall into various phenotypic classes: one having either a fragile eggshell, early developmental arrest of the embryo, or defects only at the termini of the embryo (genes involved in the Tor signalling pathway) (Cernilogar *et al.*, 2001).

Alföldi *et al.* (2011) analysed whole genome of *Anolis. carolinensis* (arboreal lizard) with 93 other species of anoles to establish phylogenetic relationship to gain insights of its evolution and migration. Interestingly, they also studied egg protein genes and found 276 proteins in immature *A. carolinensis* egg. Upon mass spectrometry, 50 of those genes were found to be present in the chicken's egg. These included vitellogenins, protease inhibitors, lysozyme, VMO1 paralogues, nothepsin and natterin. With pairwise alignment of the egg protein genes between *A. carolinensis* and chicken, a rapid evolution in the egg proteins was observed compared to the non-egg proteins. Apart from that, three VMO1 paralogues (α, β, and γ) were also identified and can be inferred of being present in the last

common ancestor of all reptiles and mammals. *A. carolinensis* genome contains all three paralogues and at least one of each has been hypothesized of being lost in all other amniote genomes. Moreover, *A. carolinensis* VMO1-α family has experienced positive selection of amino acid substitutions within a negatively charged, substrate-binding cavity. These changes presumably modify its lysozyme-like transferase activity. Here, the Anole serves as a good non-avian example where the evolutionary conserved function of VMO1 could be seen.

1.6 History and biology of Vitelline Membrane Outer Protein 1

Comparative analysis of chicken VMO1 protein-coding sequence (163 amino acids, having a molecular mass of 17.979 kDa) against the mouse, human, rat, and bovine using multiple sequence alignment tool revealed high degree of homology of 53%, 55%, 48%, and 54%, respectively (Lee *et al.*, 2015). Back et al. first identified VMO1 as one of the proteins in the outer layer of egg VM in 1982. This protein was also accompanied with other proteins such as lysozyme, ovomucin, and a second vitelline membrane outer protein (VMO2) (Kido *et al.*, 1992). This was later found to be 100% identical to mature β-defensin-11. All the VM proteins participate in maintaining its structural requirements (Guérin-Dubiard & Nau, 2007; Mann, 2008).

Lysozyme (N-acetyl-muramic-hydrolase) is a bacteriostatic and a bacteriolytic enzyme and accounts for 3.5% of the total egg white proteins (Cegielska-Radziejewska *et al.*, 2008). Ovomucin accounts to 1.5-3.5% of the total egg white solids and is attributed to have jellying property of the egg white. It has also shown to inhibit hemagglutination by viruses (hen Newcastle disease virus) (Abdou *et al.*, 2013). Avian β-defensin (AvBDs) expression has been documented in many tissues, but in the genital tract, it is speculated to have an important antimicrobial role during embryonic development. It also ensures that the hen's egg is protected from pathogens. AvBDs 1-12 mRNAs showed moderate expressions in different segments of the oviduct. For example, magnum, infundibulum, isthmus, uterus, and the vagina of the oviduct (laying-hen). Out of the 12 AvBDs, only three have been identified by proteomic approaches. AvBD11 was expressed in the egg white, the eggshell, and the VM; whilst AvBD10 was

expressed in the eggshell, gallin, and in the egg white (Hervé-Grépinet *et al.*, 2010).

Schäfer et al. (1998) indicated that eggs stored under non-refrigerated conditions disintegrate VMO1 and VMO2 proteins, leading to VM distortion. This scientific finding can now be an appropriate assumption as to why Guerin-Dubiard et al. (2006) were fortunate to detect VMO1 for the first time. Along with his colleagues, they determined the molecular weight of VMO1 as 17.6 KDa, and was consistent with Schäfer et als' (1998) SDS-PAGE analysis. Also, VMO1 protein was suggested to have an isoelectronic point (pI) near 10 since it was found to be in the alkaline area of 2-D gel analysis (Guérin-Dubiard & Nau, 2007). Chromatographic analysis showed VMO1 as a spot in conjunction with lysozyme. However, there is no experimental evidence for interactions between these two proteins in the egg white (Guérin-Dubiard *et al.*, 2006).

The crystal structure of chicken VMO1 was determined by the multiple isomorphous replacement anomalous scattering (MIRAS) method at 3-Å resolution. The parent chain of this protein consists of a peculiar three β-sheets folds that forms Greek key motifs (Figure 1.7 B) (Shimizu *et al.*, 1994). The Greek key motif, was first defined by Richardson (1977) and are topological, structural signatures that contain several beta-barrels in the protein structure. They are formed by three consecutive antiparallel beta-strands connected by hairpin loop (Zhang & Kim, 2000; Hutchinson & Thornton, 1993). Sequence analyses revealed a 53 residue, a 3-protein component of Greek key motifs. Also, there are three disulphide bonds (Cys26-Cys57, Cys79-Cys110 and Cysl33-Cysl61) that connects the three β sheets together in the VMO1 protein (Figure 1.7 B and 1.7 C).

VMO1 could also synthesize N-acetyl chito-oligosaccharides (n = 14 or 15) from hexasaccharides of N-acetylglucosamine, an activity found to be similar in comparison to the transferase activity of lysozyme without the hydrolysis activity. The true physiological function remains to be poorly understood (Shimizu *et al.*, 1994). Kido *et al.* (1995), observed that VMO1 inhibited hemagglutination of wheat germ agglutinins and had a glycan synthetic activity like the transferase activity of lysozyme. He also hypothesized in an unpublished data that *N*-

acetylchitooligosaccharides (n=14-15) were synthesized from hexasaccharides of *N*-acetylglucosamine. The transferase activity was observed from (Raftery & Rand-Meir, 1968) research. Here the group studied lysozyme mediated cleavage of chitobiose in presence of methanol trans-glycosylation. For example, *p*-nitrophenyl β-D-glucoside and *p*-nitrophenyl *N*-acetyl-β-D-glucosaminide were used as trans-glycosylation acceptors that yielded β-1-4-glycosidic bonds because of the enzymatic transfer reaction.

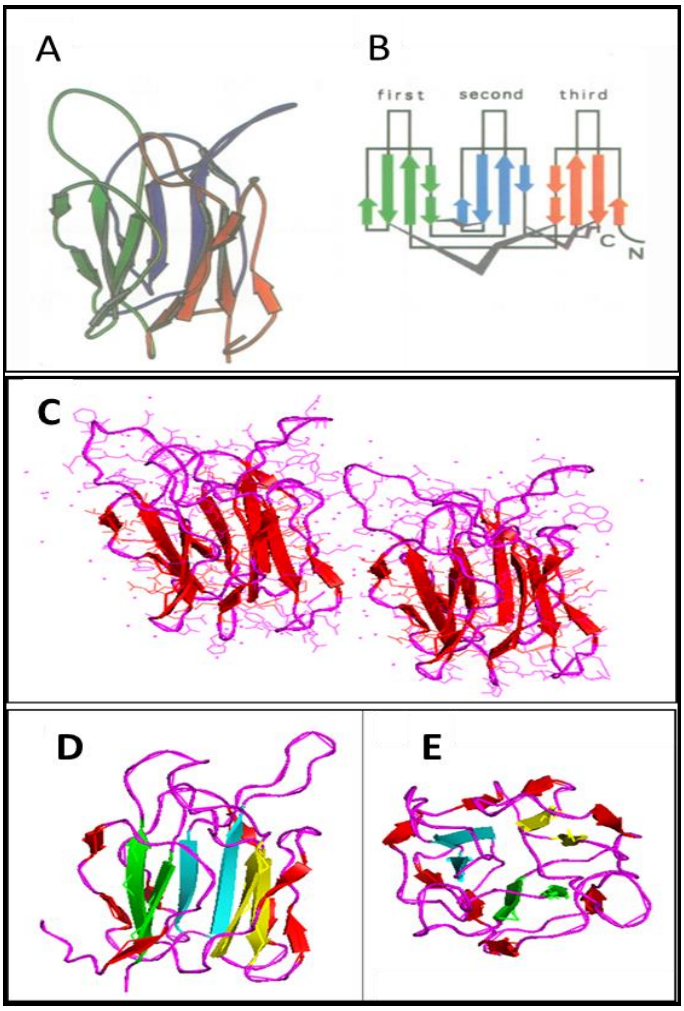


Figure 1.7 Molecular structure of chicken VMO1.

A) Ribbon representation of VMO1 in MOLSCRIPT. B) Topology of the three internal repeats of the Greek key motifs. [Adapted from (Shimizu et al., 1994)]. C) Cartoon structure designed in (PyMOL^{TM}). D) Monomer having Greek key motifs coloured in green, cyan, and yellow. E) Top view of the VMO1 protein assembly.

The reason VM and its proteins are yet to be fully understood is their poor solubility in aqueous solutions. Bioactive properties of the VM were elucidated recently by enzymatic membrane hydrolysis (i.e., Flavourzyme and trypsin treatment). This method produced peptides with anti-inflammatory, free radical scavenging (antioxidant), and metal chelation properties (Lee *et al.*, 2017).

1.7 Avian Reproductive System

Morphological development of the chicken is a gradual process lasting about 25 hours in the oviduct. Food eaten by the hen is assimilated and forms the part of an egg.

1.7.1 The formation of egg

The ovary unlike many animals in the hen is one functional entity. The female chick can have about 4000 ova, at the time of hatching, where some become full sized yolks and may develop into mature hens. The yolk (ovum) is held in a thin-walled sac, also called the follicle, which is attached to the ovary that hold the lipids and proteins in an aqueous phase.

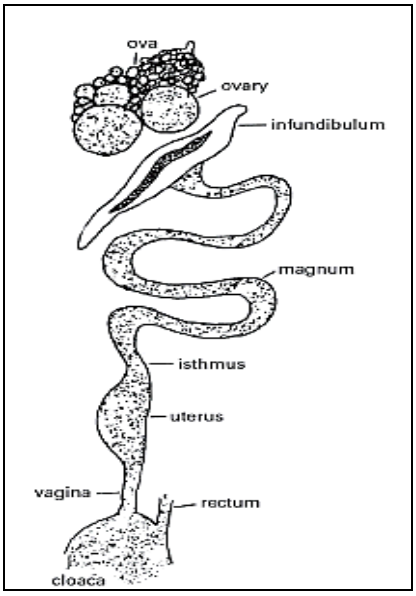


Figure 1.8 Reproductive organs of the hen showing parts of the oviduct.
[Adapted from (www.thepoultrysite.com)].

The oviduct is a muscular, coiled tubular structure where the ovum travels through and reaches the uterus. There are two oviducts, a functional left oviduct

where the mature yolk is released and held, and a right oviduct that is non-functional. The left oviduct is further divided into five distinct Sections (Figure 1.8). The Infundibulum is the place that receives the yolk after maturation and fertilization. In the Magnum, secretion of albumin (egg white) takes place by sub-epithelial tubual glands. After 3 hours inside it, the egg mass enters the Isthmus where the shell membrane is formed within 2-3 hours. In the shell gland or the uterus, the egg stays for the greatest time (19-20 hours). Water, shell materials, minerals and pigments are added here to form eggshells. Lastly the Vagina, where the egg pass through before its being laid (Ottinger & Bakst, 1995).

Morphologically, development of the chick embryo comprises of three stages. Firstly, days 0 to 7 denotes the inception of germ layers; the ectoderm, mesoderm, and the endoderm are responsible for tissue and organ differentiation during gastrulation (Lemaire & Kessel, 1997; Sheng, 2014). During days 3 to 7, functional organs and extra-embryonic membranes are developed (Mellor & Diesch, 2007). The second phase, (days 8 to 18) marks the development of chorioallantoic membrane (CAM) and embryo completion. At day 10, the chick embryo is fully formed (Moran, E.T., 2007) and the final phase (days 19–21) represents emergence. Nutrients are recovered from albumin and yolk are important for body assembly until complete embryo is formed by 14 days. Glucose becomes vital for fatty acid oxidation. Majority of the remaining albumin mixes with amniotic fluid while a part of it is absorbed for glycogen storage. Apart from these, any unabsorbed enzymatic residual remains enter the yolk sac. Calcium translocation from the shell supply minerals to the skeletal system and later by calcified granules (Tong *et al.*, 2013; Moran, E.T., 2007).

1.8 Gender specific expression of VMO1 protein

Cordeiro & Hincke. (2015) analysed a total of 228 eggshell membranes (ESMs) of fertilized eggs during chicken embryo development and identified high expression of VMO1 at days 11 and 15. In addition, a heat map analysis showed an increase in expression of VMO1 from day 3 to 19, with day 4 being the most abundant (Figure 1.9). Thus, emphasizing its role in extracellular structure organisation during developmental stages 2 and 3. VMO1 was also absent in the

blood and CAM while only being present in the ESM from fertilized eggs on day 19. Also, Actin was found in all the above-mentioned samples while lysozyme was present in CAM and ESM. These findings further stress the importance of the VMO1 protein in membrane stabilization with other proteins.

Lee *et al.* (2015) analysed the mRNA expression pattern and functional activity of *VMO1* in the laying hen oviduct using three molecular techniques; RT-PCR, quantitative RT-PCR (qRT-PCR), and RNA interference (RNAi). It was evident that *VMO1* was highly expressed in females' oviduct, more specifically only in the magnum. Their data suggests that three microRNAs (1651-3p, 1552-3p, and gga-miR-1623) had an influential role on the expression of *VMO1* through its 3′-UTR. When *VMO1* gene knockdown experiments were conducted, it revealed repression of ovomucin due to *VMO1* silencing. Apart from that, oestrogen was found to induce *VMO1* mRNA expression *in vitro*. It must not be forgotten, that ovomucin binds to VMO1 to form the foundation of vitelline membrane. Thus, it is interesting that VMO1 is expressed in oviduct and not in any other organs (Figure 1.10).

Lim & Song. (2015) proposed that *VMO1* holds a crucial contribution in the morphogenesis of the oviduct in the presence of oestrogen and moulting. They also affirmed that the onset of *VMO1* expression is associated with carcinogenesis of laying hens. A CA-125 biomarker for diagnosing early stage of ovarian cancer in women was cross-reactive with biomarkers for ovarian cancer in laying hens.

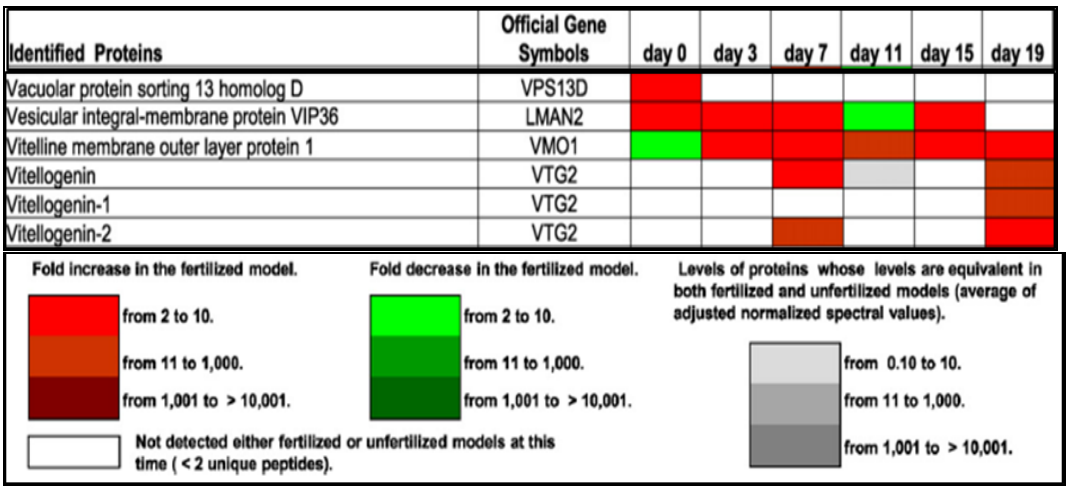


Figure 1.9 Heat-map illustrating the comparative analysis of ESM proteins.

Figure on the top show colour intensities with ESM levels in the fertilized versus unfertilized eggs at various time points of incubation. Figure at the bottom provides a key about colour intensities corresponding to a particular ESM fold increase indicated in red or fold decrease indicated in green in ESMs. [Adapted from (Cordeiro & Hincke, 2015)].

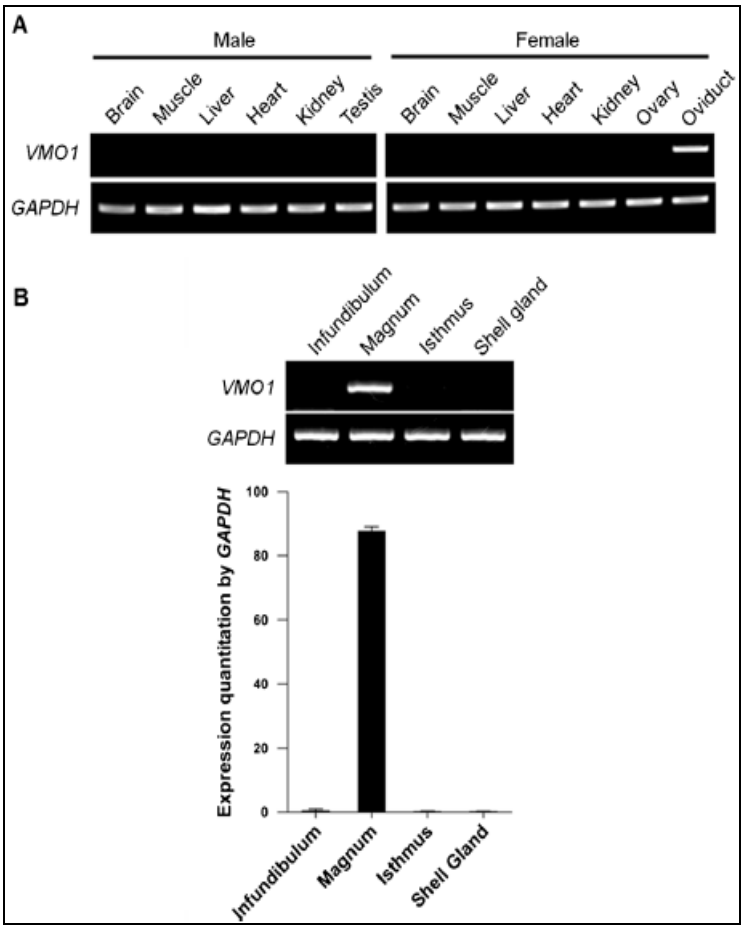


Figure 1.10 mRNA expression of *VMO1* in the chicken oviduct and magnum.

[Adapted from Lee.*et.al* (2015)].

Li *et al.* (2016) used high throughput shotgun proteomics and quantitative multiple reaction monitoring (MRM) approaches to study and compare individual plasma proteins between two genders of zebra fish (*Danio rerio*). The team identified 666 proteins in female plasma and 624 in male plasma, of which 289 were present in both the genders. Here, vmo1a was found to be a female biased protein with 6.38-fold change. Vitellogenins (1-7) were amongst the top of this expression system. Apart from these two, the protein profile in both the genders remained more or less the same. This included Apolipoproteins A, hemoglobin subunit α/β, fibrinogen, etc. Thus, suggesting the unique expression of VMO1.

1.9 Gene expression and functional analysis of VMO1 in mammalian tissues

Presence of *VMO1* have been reported in various exocrine glands and/or secretions. For example, pancreas, urine, breast, cerebrospinal fluid, respiratory secretions, and in minuscule quantities in human tears while comparatively high in camel tears (Wang *et al.*, 2014).

In a study (Alves *et al.*, 2013), proteins form urine samples were screened in patients suffering from Renal Cell carcinoma (RCC). This included 61 urine samples from Clear Cell RCC and Papillary RCC from untreated patients. They were further compared with 29-control urine sample on a CM10 chip by SELDI-TOF (Surface Enhanced Laser Desorption Ionization - Time of Flight) technique. Proteomic analysis of controls included several proteins such as albumin, immunoglobulin light chains, mannan-binding lectin-associated serine protease-2 (MASP-2) secreted and transmembrane 1 precursor (protein K12), and included VMO1. The urine samples were mixed with RP resin and purified using elution buffer containing 10-100% acetonitrile. A fraction was further resolved on 12% SDS-PAGE gel and stained with Blue safe stain. They isolated the band of interest from the fraction eluted with 60% acetonitrile that produced 2 bands; a 9,770 Da and 23,600 Da. Their VMO1 protein had 53 amino acids with an average mass of 22,034 Da.

In another study, 1543 unique tear proteins from healthy humans were identified with VMO1 found in trace amounts (Zhou *et al.*, 2012). Proteins present in the tear film often play an important defensive role to remove pathogens, while maintaining the integrity of the tear film, and modulating wound healing (Ohashi *et al.*, 2006; Flanagan & Willcox, 2009). The primary function of the tear film is to protect the cornea and conjunctiva, which forms the ocular surface and is directly in exposed to the outer environment. In humans, secretory IgA, lysozyme, lipocalin and lactoferrin are the top proteins in the tear; where lactoferrin and lysozyme possess antimicrobial activity (Zhou *et al.*, 2004; Masinick *et al.*, 1997).

Camels (*Camelus dromedarius*) often habitat under harsh environmental conditions and have been demonstrated for the presence of large amount of VMO1 homolog in their tears (Chen *et al.*, 2011). Tears were collected from 50 healthy camels in summer and winter from Saudi Arabia. These samples were resolved on 13% SDS-PAGE gel that separated into 13 bands for each season (Figure 1.11 A).

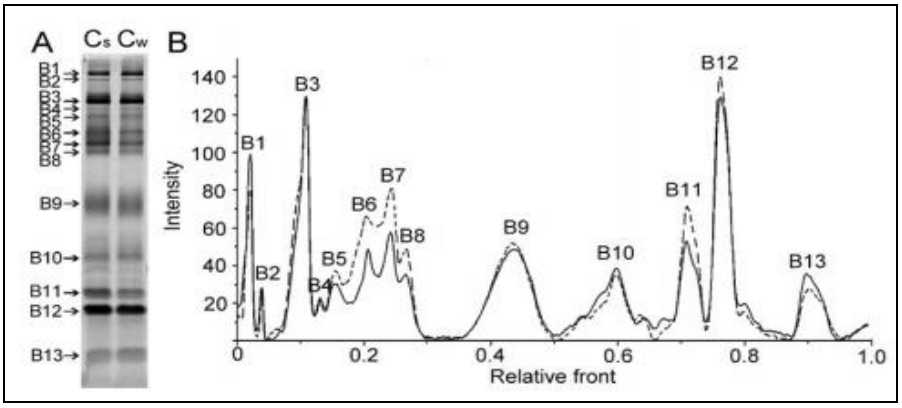


Figure 1.11 Comparison of SDS–PAGE gel patterns of protein in camel tear fluids between summer and winter.

A) PAGE gel analysis of tear proteins. B) Show graphics of lane comparison of camel tear proteins between the summer (dotted line) and the winter (solid line). [Adapted from (Chen et al., 2011)].

2-DE protein spots (seven spots on the gel for each season) revealed differential expression of proteins between summer (78 kDa) and winter (79 kDa). Also, one of the spots (spot 7) showed higher intensity in summer group than in winter. The spot was later characterized as VMO1 homolog using western blot, as a 21 kDa protein present during both the seasons. Thus, it is apparent that seasonal variation

led to the reduced expression of lactoferrin in the summer group proteins while the expression of VMO1 homolog increased compared to the winter group.

Similar studies (Shamsi *et al.*, 2011) were conducted that studied proteomic profiles of human, sheep, cow, and camel tears; and the presence of VMO1 was confirmed by western blot (Figure 1.12). Here, species-specific variation in VMO1 could be the result of structural and physiological differentiation of the ocular surface, arising due to the evolutionary changes. These changes might have benefitted camels to keep the ocular surface healthy in desert conditions. The group also affirmed that even though VMO1 was absent in cow tears during western blot analysis, the sequence variability of VMO1 homolog in cow would have prevented anti-human VMO1 antibody binding to it.

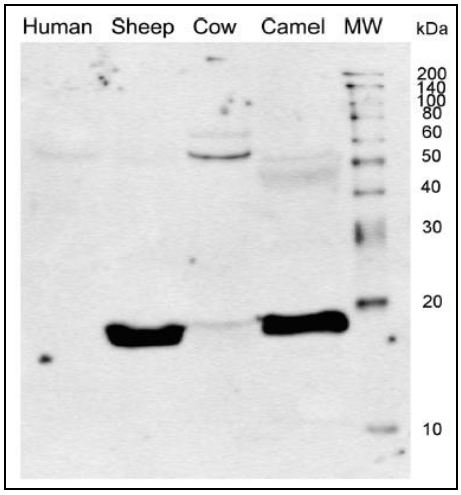


Figure 1.12 Western blot immunoassay of human, sheep, cow, and camel tears using anti-human VMO1 antibody.

A 20 kDa protein band being detected in sheep and camel tear [Adapted from (Shamsi et al., 2011)].

Apart from proteins, 18 elements were also found in camel tears using inductively coupled plasma mass spectrometry (ICP-MS). High concentrations of leucine, valine, and norvaline amino acids were detected that would have a stimulant effect on the tear film and prevent the muscle breakdown (Lu *et al.*, 2008). While sugars like D-glucohexodialdulose, allose, glucopyranose present in the eyes provide energy to the tear film muscles to maintain equilibrium (Ahamad *et al.*, 2017).

Wang *et al.* (2014) performed a dot-blot assay (incubating His-VMO1 with air dried contents of lysozyme on a nitrocellulose membrane and detection with His-tag antibody) to assess the interaction of human recombinant VMO1 with Lysozyme C (LYSC), lipocalin 1 and lactoferrin. Here, His-VMO1 was only found in the proximity of LYSC (Figure 1.13 A). Computer modelling revealed significant amino acids interactions between VMO1 [Glutamine (Q) 153, Glutamic acid (E) 110, and Q 70] and lysozyme [Arginine (R) 59, or Serine (S) 100, and Aspartic acid (D) 85] through hydrogen bonding (Figure 1.13 B).

To further advocate these interactions, His pull-down assay (eluting His-tagged VMO1 and LYSC with Dynabeads at pH 8.0, using imidazole) and immunoprecipitation assays (precipitating His-VMO1 and lysozyme together and detection by western blot using vmo1 or lysozyme antibody) were performed. It was observed that the concentration of LYSC decreased with the decrease in VMO1 and protein mixture containing both the components were subsequently detected by anti LYSC antibody. This indicates specificity between the two proteins. His-VMO1 did not show any antimicrobial activity against *Staphylococcus aureus* and *Escherichia coli*. However, it led to long period of tear film integrity when added to ocular surfaces of mice.

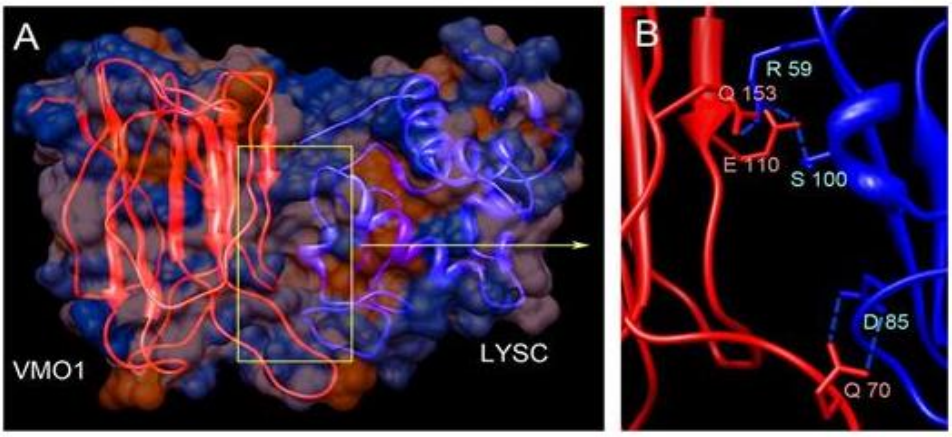


Figure 1.13 Protein interaction of VMO1 and Lysozyme.

A) Interaction model between VMO1 (red) and lysozyme (blue). These interactions take place from the non-active sites of LYSC and Greek key motif of VMO1. B) Enlarged view of interaction between important amino acid residues [Adapted from (Wang et al., 2014)].

In a canine study, RNA profiling analyses of dog olfactory epithelium showed VMO1 homolog as the most transcribed gene in two samples from dog: Bichon fries and golden retriever at 3548 and 2723 FPKM (fragment per kilo base of exon per million of fragments mapped), respectively (Galibert *et al.*, 2016). In human olfactory epithelium, *VMO1* represented one of the overexpressed genes amongst 40 others (Olender *et al.*, 2016).

1.10 Aim and objectives

The aim of this study was to assess the functionalities of VMO1 through molecular techniques. This including an in-depth analysis of the *VMO1* (human/chicken) and *Vmo1* (mouse) gene. Next, we aimed to induce and express recombinant VMO1 protein in *E. coli*. After production of the recombinant protein, we purified it using His affinity column and confirmed it by Western blot detection. Lastly, we determine the suitability of VMO1 antibody for studying protein localisation. The said aim was achieved using the following four objectives.

To determine the possible function of VMO1, a bioinformatics study was undertaken (Chapter 3) and will be discussed in depth in Chapter 3. In brief, bioinformatics was carried out to:

a) Confirming the presence of *VMO1* gene in pET28b (+) using DNA sequencing
b) Analyse if the Open reading frame (ORF) was in frame with the His-tag.
c) Study structural and chemical composition of nucleotide (*VMO1*) and protein (VMO1) sequences in chicken, human, mouse
d) To predict molecular structure of VMO1

To induce and optimize recombinant VMO1 protein, two different methods were utilised (Section 2.4 and 2.7):

a) Traditional method using IPTG (0.4 mM and 1mM)
b) Ethanol-based using IPTG (1mM) and Ethanol (1%, 2%, 3%)

A range of microbial culturing and processing was performed along with molecular biology techniques.

Purification of induced protein was achieved using (Section 2.10):

a) Purification buffer without Imidazole

b) Purification buffer with Imidazole

To determine VMO1 protein localization in mammalian tissues, molecular methods were used to (Section 2.11):

a) Validate specificity of VMO1 antibody

b) Analyze VMO1 antibody binding using fluorescent microscopy.

d) Test and validate VMO1 antibody on frozen Sections of the mouse inner ear, mouse cochlea and rat lung using IHC

CHAPTER TWO

MATERIALS AND METHODS

2.1 General materials

All laboratory practices were carried out within the C.2.03 Molecular Genetics Lab (unless otherwise stated) at the University of Waikato, Hamilton, New Zealand. All experiments were carried out aseptically on bench tops swabbed with 70% ethanol. All glassware was washed in the dishwasher and all solutions were prepared using autoclaved 15-18 mega ohm-cm double distilled deionised water (DDH2O) (Barnstead double distilled/deionization system) or Milli. Q H2O by autoclaving at 121 psi for 20 minutes. Milli. Q H2O was obtained when the purification system reached resistivity of 18.2 MΩ.cm at 25°C. All solution and buffer recipes can be found in Appendix One. All chemicals and solvents (salts, buffers, and organic solvents) were obtained from Sigma-Aldrich® Co and all molecular biology reagents (enzymes and buffers) were obtained from Invitrogen unless otherwise stated. In addition, the antibody solutions were obtained from Abcam/ GeneTex (Table 2.9).

Regarding the plastic ware, sterile aerosol tips were Sorenson™ low binding MultiGuard aerosol barrier tips supplied by Sigma-Aldrich® Co, or alternatively, MULTIMAX Low Retention Pipet Tips from BioExpress with Eppendorf auto pipettes. All 1.5mL and 2 mL micro centrifuge tubes was supplied by Raylab New Zealand ltd, whilst all 0.6 mL tubes were supplied by Axygen®. The falcon tubes were Cellstar® branded supplied by Greiner Bio-one.

All centrifugations were carried out using Eppendorf bench top centrifuge model 5415R or Biofuge *pico* HERAEUS with a maximum speed of 16.0xg or 13 rpm at room temperature (RT) unless otherwise specified. All thermocycler incubations were carried out in 1.5 ml Eppendorf Thermomixer comfort machine, while PCR was carried out in a BIO-RAD T100™ Thermal Cycler. All electroporation was carried out using Gene Pulser® and Pulse Controller set at 25 capacitance, 200 resistance, and 250 volts. The 100bp DNA ladder was sourced from Solis

BioDyne while PAGE-MASTER and Western Blot- MASTER Protein standards were sourced from GenScript.

PAGE gel electrophoresis was carried out using ATTD CROSSPOWER 1000. All plates were made using Luria Bertani and/or Terrific Broth agar media. Protein staining and protein transfer were carried out using eStain® 2.0 Protein staining System and eBlot™ Protein Transfer system from GenScript. Gels were viewed on Omega Lum™ G Imaging System (Aplegen). Tissue sectioning was performed on LEICA CM1850 UV, Germany (230V~ and 50 Hz). Microscopic observations were performed using a Leica Microscope DMRE fitted with an Olympus DP30 camera. The liquid blocker super PAP PEN was from Daido Sangyo Co., Ltd. Japan.

2.2 Growing transformed cultures

A -80°C glycerol stock of the *E. coli* Rosetta™ (DE3) pET28b (+) pLysS and BL21 (DE3) pET28b *VMO1* strains were thawed on ice and inoculated into LB broth containing 30 mg/kl Kanamycin and 34 mg/ml chloramphenicol for selective growth of bacteria. The cultures were grown at 37°C under shaking overnight (O/N). Hardie. (2015) determined that the ideal bacterial strain of choice was Rosetta since they are designed to enhance the expression of eukaryotic proteins that possess rare codons. To select for the pLysS plasmid, chloramphenicol was used and kanamycin selected for the pET28b (+) vector. All experiments were conducted in PC1 containment facility under the HSNO approval code: GMD101146. Because mammalian specimens were donated, Animal Ethics approval was not required.

2.2.1 Plasmid extraction

The Zyppy™ Plasmid Miniprep Kit: A Pellet-Free™ modified alkaline lysis method developed by Zymo Research (version 1.2.6) was used for the extraction of plasmid DNA extraction. Briefly, an O/N bacterial culture of 2 mL was pelleted in a 1.5 ml micro centrifuge tube by centrifugation at 13,000 rpm and resuspended in 100 µl of 7X Lysis Buffer. Next, 350 µl of cold Neutralization

Buffer was added, mixed thoroughly by inverting the tube 2-3 times to precipitate the *E. coli* chromosomal DNA. The sample was then centrifuged at 13000 x g for 4 minutes. The supernatant (~900 µl) was carefully transferred into the Zymo-Spin™ column and centrifuged for 15 seconds. The flow-through was discarded and 200 µl of Endo-Wash Buffer was then added to remove denatured proteins from the column, and centrifuged for 30 seconds. An additional, 400 µl of Zyppy™ Wash Buffer was transferred to the column to remove any contaminating salts and then centrifuged for 1 minute at 13000 x g. Finally, the column was then transferred into a clean 1.5 ml microcentrifuge tube and 30 µl of Zyppy™ Elution Buffer was added directly to the column matrix and incubated for one minute at RT. The sample was then centrifuged for 30 seconds at 13000 x g to elute the plasmid DNA. The extracted DNA was stored at -20°C.

The quality and quantity of Plasmid DNA was measured using a Thermo Scientific™ NanoDrop™ 2000 spectrophotometer. In addition, agarose gel electrophoresis was used for the identification and quantification of Plasmid DNA based on the electrophoretic separation based on nucleotide size. Therefore, this approximately allows 0.1–25 kbp (kilo base pairs) of nucleic acid to be analysed. The plasmid sample (2 µL) was mixed with 4 µL of 6X loading dye and loaded into the well of 1% Hydragene HyAgarose™ TAE agarose gel containing 2 µL of intercalating dye Ethidium Bromide (EtBr) and run at 90 volts for 30 minutes utilizing the EC250-90 power (E-C Apparatus Corporation).

The confirmed plasmid DNA was identified, and the concentration was determined by comparing the plasmid band to a 100 bp molecular ladder. The plasmid was then stored at -20°C until DNA sequencing.

2.3 Confirmation of *VMO1* in Plasmid DNA

2.3.1 DNA sequencing

Sequencing was carried out for the extracted *VMO1* plasmid DNA using four 5 µM primers using Applied Biosystems (3130xl) Genetic Analyzer available at the DNA sequencing facility at the University of Waikato.

Table 2.1 mentions the primers that were used for sequencing. pET28b (+) vectors bearing *VMO1* gene inserts, at a concentration of 80 ng/µL (determined by a NanoDrop 2000 Spectrophotometer) were sent for sequencing. The sequences obtained were then analyzed using Geneious version 7.1.8 and 7.10.2 (Biomatters, 2018; www.geneious.com). Once the sequences were obtained the reverse complement of the reverse primer was taken and the four sequences were aligned pair-wise and edited to generate a consensus identity. These were visually examined for any gaps and edited as per the software's prediction. Indistinguishable peaks in the electropherograms were trimmed, which mostly appeared at the beginning of the sequences.

Table 2.1 Primers used for *VMO1* plasmid DNA sequencing

Primer	Nucleotide sequence (5'-3')
SH3F	GGCGACTGGTTTCACATGTGC
SH3R	CCAAAGTCTCCCCAGCTCAG
T7 promoter	TAATACGACTCACTATAGGG
T7 terminator	GCTAGTTATTGCTCAGCGG

2.3.2 Restriction digest

Alternative to DNA sequencing, restriction digests were set up to determine if pET28b (+) vector contained an insert of the expected size using restriction enzymes. These enzymes are molecular snipping tools that identifies specific nucleotides on the double stranded DNA and cut the DNA at specific nucleotides at specific temperature, pH, salt composition, enzyme cofactors, and ionic

strength. The two restriction enzymes EcoR1 (Boehringer Mannheim) and BamH1 (NEB) were used for performing plasmid digestion. EcoR1 cleaves G/AATTC sequence while BamH1 cleaves G/GATCC sequence.

Table 2.2 Protocol for restriction digest set up.

A 10 µL restriction digest was set up on ice with controls. The tubes were then inserted into Thermomixer machine overnight for digestion.

Sr. No	**Contents**	**Tube 1**	**Tube 2**	**Tube 3**	**Tube 4**
1	Plasmid DNA 80 ng/ µL	5 µL	5 µL	5 µL	5 µL
2	10X Buffer 2	1 µL	1 µL	1 µL	1 µL
3	10X BSA	1 µL	1 µL	1 µL	1 µL
4	mQH_2O	3 µL	2 µL	2 µL	1 µL
5	EcoR1 10U/ µL	0 µL	1 µL	0 µL	1 µL
6	BamH1 10U/ µL	0 µL	0 µL	1 µL	1 µL

2.3.3 Polymerase chain reaction (PCR)

PCR is the most common *in vitro* technique for amplification of DNA templates or analyse short sequences (addgene, 2018b). The basic PCR steps includes: thermal denaturation, primer annealing, and extension by heat stable DNA polymerase. The aim was to use the original sequencing primers (Table 2.1) to determine if *VMO1* could be amplified from the extracted Plasmid DNA with an expected product size of ~450 bp using the protocol mention in Table 2.3. The PCR products were run on 1% agarose gel (Section 2.6.1) or stored at -20°C.

Table 2.3 Protocol for PCR set up.

Sterile PCR tubes were used for amplifying *VMO1* gene insert. The 20 µL PCR reaction tubes were placed inside the PCR machine and run at 94° C, 2 min. 30X [94°C, 20 secs; 60°C, 10 secs; 72°C, 30 secs] 72°C, 5 min.

Sr. No	Contents	Tube 1	Tube 2	Tube 3
1	Rosetta pET28b (+) *VMO1* plasmid DNA 100 ng/ µL	1 µL	0 µL	0 µL
2	BL21 pET28b (+) *VMO1* plasmid DNA 75 ng/ µL	0 µL	1 µL	0 µL
3	2X PCR Master Mix	10 µL	10 µL	10 µL
4	SH3F primer (10 µM)	1 µL	1 µL	1 µL
5	SH3R primer (10 µM)	1 µL	1 µL	1 µL
6	Milli. Q H2O	7 µL	7 µL	8 µL

2.3.4 Plasmid stability test

In order to successfully express recombinant VMO1 protein, it was essential that the plasmid harbouring the *VMO1* gene must be present in the bacterial system. The plasmid stability test, thus determines the proportion of cells maintaining the target plasmid or loss of plasmid using four different agar plates (Table 2.4). This test, quantitatively measures plasmid stability based on antibiotic resistance through positive selection.

Table 2.4 Requirements for plasmid stability assay

Sterile Agar Plate	Bacterial Cell Growth
LB	All viable cells
LB+ Kanamycin + Chloramphenicol	Cells carrying plasmid
LB+IPTG 1mM	Cell that have mutated or lost plasmid and cannot express the target gene
LB+ Kanamycin + Chloramphenicol + IPTG 1mM	Mutant cells that have retained the plasmid but lost the ability to express the target gene

A glycerol stock of Rosetta *E. coli* containing the pET28b (+) pLysS VMO*1* plasmid was thawed on ice. Loopful of the culture was spread onto the required

agar plates and incubated upside-down O/N at 37°C. The colonies were then counted and recorded the next day.

2.4 Recombinant protein expression

2.4.1 Induction of λDE3 Lysogens with IPTG to determine total whole cell protein using a small-scale time course

A single colony was selected from a LB agar plate containing 34 μg/μL chloramphenicol and 30 μg/μL kanamycin to inoculate in 5 mL of starter LB broth incorporated with chloramphenicol (34 μg/μL), kanamycin (30 μg/μL) and 20% filter sterilized glucose. Antibiotic solutions were added to all of our cultural media's by placing them in a water bath maintained at 50°C for optimal activity of the antibiotics. Glucose was used to control non-specific expression. The starter broth was grown O/N shaking at 200 rpm at 37°C in a Ratex Orbital Mixer Incubator. On the following day, three 14 mL culture tubes containing 5 mL of LB broth, the antibiotics, and glucose were taken. To this, 200 μL of the bacterial cells were added and incubated at 37°C at 200 rpm shaker incubator until the OD_{600} measured between 0.4 and 0.6.

To two of the culture tubes, 100 mM IPTG was added to the final concentrations of 0.4 mM and 1 mM. IPTG binds to the lac repressor, allowing genes under the lac operon to be transcribed. One tube was reserved as a control for un-induced cells where no IPTG was added. The tubes were placed again in a 37°C shaker incubator at 250 rpm with 500 μL aliquots taken at different time points (0 hr, 3 hr, 6 hr, 24 hr) into 1.7 mL microcentrifuge tubes for whole cell extraction. The bacterial cells were then spun for five min at 5000 x g at 4°C. The media was removed and transferred to a new tube (this served as your soluble fraction) while 50 μL of Cracking Buffer was added to the cell pellet (this served as the insoluble fraction). All the samples were stored at -20°C until required further. For PAGE gel analysis, 50 μL of 2X SDS Loading Buffer was added to 50 μL of soluble and insoluble fraction samples. Each sample was then boiled for five min at 99°C in a thermomixer. 5-10 μL samples were then loaded on to 12% PAGE gels to observe recombinant protein bands in non- induced and induced samples.

2.4.2 Determination if recombinant protein is soluble or insoluble

To determine if the expressed recombinant protein is soluble or insoluble, the above experiment was repeated but for one time point only i.e. 24 hours. One mL of whole cell sample extract (at a specific time point showing optimum growth) was centrifuge at 5000 x g for five min at 4°C. The media was removed and the cel pellets were resuspend in 100 µL of 5X TE buffer (pH 8). Samples were sonicated using a Misonix Sonicator® Ultrasonic Liquid Processor XL2020 with the microtip probe. The sonicator was set to Continuous mode and power setting adjusted between 2.0-3.0. Sonication was carried out six times for 20 second pulses with 20 seconds between each pulse on ice to prevent overheating of the protein samples.

Later, the samples were spun at 4°C for five min at 16000 x g. Both pellet (Insoluble fraction) and supernatant (soluble fraction) were reserved for further analysis. To process the soluble fraction, the supernatant was transferred to new tube (store at -20°C for long term storage). For PAGE gel, a 50 µL aliquot of the supernatant was transferred to a new 1.7 mL microcentrifuge tube, to which 50 µL of 2X SDS Loading Buffer was added. The pellet was resuspended in 50 µL each of Cracking Buffer and 2X SDS Loading Buffer. The soluble and insoluble fraction samples were then boiled for five min at 99°C prior to being run on 12% SDS-PAGE gels. They can be stored at -20°C or stained with Coomassie blue.

2.5 Bradford protein assay

The Bradford assay is based on the principle that, when protein molecules bind to Coomassie Blue dye it results in a colour change from brown to blue, under acidic conditions (Bradford, 1976; BIO-RAD, 2017). This method measures the presence of the basic amino acid residues; arginine (R), lysine (K) and histidine (H), which contributes to the formation of a protein-dye complex. Thus, this assay is a simple spectrometric assay to determine the concentration of total proteins in the given sample. Firstly, Bovine Serum Albumin (BSA) standards must be prepared as outlined in Table 2.5.

Table 2.5 Preparation of BSA standards.

Prepare the standard solutions as per the table below by mixing appropriate concentration of BSA with PBS and store at -20°C.

Concentration (mg/mL)	**0**	**0.1**	**0.25**	**0.5**	**1**	**2**	**4**	**6**	**8**	**10**
10 mg/mL BSA	0	1	2.5	5	10	20	40	60	80	100
1X PBS (pH 7.4)	100	99	97.5	95	90	80	60	40	20	0

A 96-well flat bottom plate was set up for protein estimation assay. Each well was loaded with 100 µL of diluted dye reagent (QuickStart™ Bradford Protein Assay, catalog n# 5000205) depending on the sample size (in triplicates). Each standard and test sample (1 µL) was pipetted into the wells. The sample and reagents were mixed thoroughly using the pipet tips. The plate was then incubated at RT for no more than 5 minutes to allow color reaction to be developed. Protein concentration was then determined by comparing the test wells against standards, depending upon the color change by visual observation. Absorbance measurement for whole cells (595 nm) was then done using "Thermoscientific Multiskan Go" instrument and software version 3.2.1.4 RE accompanying it.

2.6 Casting electrophoresis gels

2.6.1 Agarose gels

Protocol: HyAgarose™ LE Agarose (Hydragene, USA) powder was used to prepare 1% agarose gels. 0.4 grams of the powder was weighed and added to 40 mL of 1X TAE buffer in a conical flask. The solution was dissolved in the microwave at medium heat setting (700W). The amount of water evaporated was readjusted using double distilled water by measuring the weight of the flask before and after the boil.

After dissolution, the mixture was allowed to cool at RT (~22°C) and 2 µL of 10 mg/mL EtBr was added. This mixture was then poured into an Owl™ Gel casting

system, fitted with one or two combs (twelve wells 0.75 cm wide) and allowed to set at RT.

2.6.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

This method utilizes SDS, which is a strong protein-denaturing agent that binds to the protein backbone. This not only results in protein unfolding but also coats a negative charge proportional to its molecular mass. Due to this, the proteins migrate towards the positive charge and separates according to their molecular masses. PAGE analysis uses two types of gels: a stacking gel and a resolving gel. The stacking gel being mildly acidic and lower in acrylamide concentration gives a porous gel matrix. This allows the protein to form sharp bands. The alkaline resolving gel has a higher concentration of polyacrylamide, this makes the gel pores narrower. Proteins in the resolving gel are therefore separated based on molecular size.

Depending upon the VMO1 protein size (15 kDa), 12% gels were made. Gels were then casted by pouring them between two glass plates. The glass plates were thoroughly cleaned with 70% ethanol and allowed to air dry. The gel casting moulds were assembled using a caster. The resolving gel (pH 8.8) was then poured up to 60% height using a 3.5 ml transfer pipette and overlaid with Isopropyl alcohol to prevent contact with air (oxygen), and to avoid inhibition of polymerization. After this stage, the acrylamide was left to polymerise for 25-35 minutes until the gel was set. The overlaid alcohol was then removed.

Next, the stacking gel solution (pH 6.8) was poured into the mold and a clean comb was inserted. The acrylamide was left to polymerize again for another 25-35 minutes. After this stage, the gels were stored at 4°C. PAGE gels prepared using this method have a shelf life of ~3-4 weeks.

Commercially available 4-20% PAGE gels were used initially from GenScript, Cat number M42012 or MO1215 having Lot no C30121511 or C30351509 respectively. They were stored at 4°C.

In order to run PAGE gels, protein samples were mixed with 2X loading dye in a thermomixer at 85°C for 3 min or 99°C for 1 minute. Next, the casted gels were

placed on the PAGE electrophoresis apparatus. The comb was removed, and the electrophoresis tank was filled with 1X MOPS running buffer till the top of the gel casting plates (were the comb was positioned). Lastly, 5-20 µL of samples and PAGE Master Ladder was loaded into the wells. The gel was run at 140V for 45 minutes at RT. After the gel ran evenly, it was either stained with protein staining pads (Coomassie blue staining) to observe the protein bands or was transferred on to a PVDF membrane for western blot detection (Section 2.9).

The 1X MOPS running buffer was prepared by reconstituting the MOPS buffer powder (GenScript, Cat No. M00138) in 1000 mL of Milli.Q water.

2.7 Optimize recombinant protein expression using ethanol induction

This method was first tried as a pilot screen for studying protein expression in the presence of different ethanol concentration ranging from 0 to 3%, and using LB and terrific broths as described by Gaurav et al, 2015.

Starter cultures were grown as per Section 2.2.
The next day five culture tubes were taken and were labelled as: -

i. Control (un-induced).
ii. Induced in the absence of ethanol.
iii. Induced in the presence of 1% ethanol.
iv. Induced in the presence of 2% ethanol.
v. Induced in the presence of 3% ethanol.

These five tubes were then inoculated with (each containing 5mL LB/TB medium with appropriate antibiotics) with 200 µL of the overnight grown starter cultures. At the time of inoculation 1%, 2% and 3% ethanol was added to the LB/TB medium and grown at 37°C for a few hours (approx. 3–4h.) with vigorous shaking, until the OD_{600} reached 0.5–0.6 for LB media, and 1.0–1.5 for TB media. Protein expression was induced by adding IPTG to a final concentration of 1.0 mM. Finally, the cultures were grown overnight at 37°C or 22°C with continuous shaking. The two temperatures were selected to check what temperature enhanced bacterial growth, which in turn would enhance recombinant protein expression.

The next day, 1 mL of cells were harvested from each culture tube and then centrifuged at 12,000 rpm; discarding the supernatants. Samples were stored at -20°C for PAGE (Section 2.6.2) and WB analysis (Section 2.9).

2.8 Large scale overnight recombinant protein expression using Terrific Broth and 1 mM IPTG

TB media (500 mL) containing 500 µL antibiotic (30 mg/mL Kanamycin, 34 mg/mL Chloramphenicol) was inoculated with 2.5 mL starter culture (grown to an OD_{600} at 1.0-1.5 overnight). For good aeration, 25% of 2.8 L baffled flask was covered with 500 mL of TB and enclosed by cotton wool and an aluminum cover. The flask was incubated at 37°C with shaking at 250 rpm, until the OD_{600} reached 0.5-1.0. One mM IPTG (5 mL of 100 mM IPTG to a final concentration of 1 mM) was added to the media to induce expression of recombinant VMO1 protein. The flask was once again incubated at 37°C with shaking at 250 rpm overnight.

2.8.1 Extraction of Insoluble Inclusion Bodies

Following the step above, cells were harvested from the liquid culture (transferred in 150 ml bottles) by centrifugation at 10,000 g for 15 minutes at 4°C, using a pre-weighed centrifugation tube. The liquid was decanted, and the pellet was allowed to drain, removing as much liquid as possible. The pellet was freeze dried for approximately 1 hr or O/N at -20°C to aid lysis. Later, it was resuspended and mixed with 2.5 mL of Lysis buffer and 2.5 mL of 2X Protease Inhibitor – EDTA free making it a total volume of 5 mL. The samples were transferred to a 50 mL Falcon tube.

The next step was to sonicate the samples with power settings adjusted between 4–5, at 40–50% duty for 15–20 bursts. This step was performed by placing the falcon tubes in a glass beaker containing ice to avoid heat denaturation of proteins. The sonicated samples were later centrifuged at 16,000g for 20 minutes at 4°C. Both soluble and insoluble protein fractions were reserved for confirming presence of recombinant protein, if any; by SDS PAGE gel analysis.

For SDS PAGE gel analysis of the soluble fraction, 100 µL of supernatant was combined with 100 µL 2X Laemmeli Sample Buffer (Biorad) in a 1.5 mL Bioproof tube. It was heated for 3 min at 85°C to denature proteins and then stored at –20°C until further analysis. Samples about 5-10 µL were then loaded into wells along with PAGE master ladder (5µL) and ran at 138 V for 45 minutes. The gel was then stained using Coomassie Blue staining pads in the Blot machine to observe separation of protein bands.

After this stage, the pellet was stored at -20°C before proceeding with purification. To the final reserved pellet in 495 µL of DNAse (Deoxyribonuclease I) reaction buffer, 5 µL of DNase I (1 unit/µL) (Thermo Scientific, cat n# 89836) was added. The mixture was then incubated for 30 min at 37°C in the Eppendorf Thermomixer®. Cell lysis releases its nucleic contents into the solution causing it to be viscous, hence enzymatic digestion of the interfering DNA/RNA becomes important. Thus, it cleaves DNA and gets rid of any existing bacterial and contaminating DNA. DNase I, is a cell lysis reagent that's reduces viscosity caused by DNA lysates in the protein extracts. It therefore, not only facilitates pipetting but also improves protein extraction efficiency.

The mixture was later centrifuged at 16,000g for 20 minutes at 4°C. The supernatant was removed and stored at -20°C, and the final pellet was resuspended in 5 mL of 1X Binding buffer and left on ice for 1 hr, later vortexed and stored at 4°C. Supernatant and pellet (100 µL) were reserved for PAGE analysis to detect recombinant protein. The samples were combined with 100 µL 2X Laemmeli Sample Buffer (Biorad) in a 1.5 mL Bioproof tube. It was immediately heated for 3 min at 85°C to allow protein denaturation and then loaded onto the PAGE gel to observe presence of protein bands.

2.9 Western blot detection

A western blot involves the electrophoretic transfer of proteins from a pre-run PAGE gel onto a nitro-cellulose or polyvinylidene fluoride (PVDF) membrane using electric field that is directed perpendicular to the surface of the gel. This causes the proteins to move out of the gel and embed onto the membrane. PVDF membranes offer better mechanical support and allows the blot to be re-probed and stored. The proteins are detected by antigen-antibody reaction, using a chemiluminescent enzyme (Mahmood & Yang, 2012).

For Western blot detection, the protein samples were separated by PAGE gel electrophoresis as describe in Section 2.6.2. The PVFD membrane (20 µM) was activated by soaking it in 100% methanol (biological grade) for 30 seconds, followed by soaking in equilibrium buffer on the shaker for 5 minutes (recommended by the manufacturer). An eblot^{TM} Genescript electric machine was used to perform the protein transfer. The machine was turned on, the flap was opened, and the Anode Pad was placed on to the area marked for protein transfer. This was followed by placing the PVDF membrane. Air bubbles if any were removed using tongs. The pre-run gel was then placed onto the membrane and then placing the Cathode Pad over it. The flap was closed, and the machine was run for 7 minutes.

The transferred membrane was then removed and placed in an enclosed box. The membrane was then incubated in 10% non-fat dried milk blocking solution O/N or 1 hour at 4°C shaking condition. This step prevents nonspecific antibodies attaching to the membrane. The membrane was washed three times with 1X TBS (pH 7.6) for 3 minutes each, while rocking at RT. Primary antibody (1:1000), His-tag or VMO1 was added on to the membrane and was incubated overnight under shaking. The next day, the membrane was washed three times for 15 minutes each with TBST (pH 8). HRP Goat-anti rabbit secondary antibody 1:5000 (Ab 97051 Lot no# GR209629-4) was added on to the membrane and incubated at RT for 1 hour. The membrane was washed three times for 15 minutes each with TBST (pH 8). Lastly, the membrane was rinsed with TBS (pH 7.6) for 15 minutes, under shaking at RT. The blot is then developed using Western bright ECL spray

(Advansta) and chemiluminescence was observed using an aplagen, Omega Lum G Imaging system and the Image capture software accompanying it.

2.9.1 Ponceau staining

Ponceau Staining is a rapid, reversible staining method used to observe the presence of protein bands on western blots (Goldman *et al.*, 2016). Ponceau S solution, is pink colored sodium salt of diazo dye that sticks to the protein in the membrane. The western blot membrane was incubated in 0.1% Ponceau S solution for 5 minutes on shaker at RT or until the pink bands become visible.

2.9.2 Ponceau de-staining

The Ponceau Staining solution was discarded and incubated with de-staining solution and swirled until the pink colour faded away. The membrane was stored in TBS (pH 7.6) at 4°C until required for reuse.

2.9.3 Mild stripping of western blot membrane

One of cost-effective method of reusing the same blot membrane multiple times is to strip the membrane after each western blot analysis. This means removal of primary and secondary antibodies. The method allows stripping and re-probing a single membrane multiple times to test a range of antibodies and save on running and blotting many gels, samples, reagents, and time.

A previously developed PVFD Membrane was incubated at RT for 5-10 minutes by using adequate volume of mild stripping buffer (pH 2.2) that would cover the membrane. The buffer was then discarded. And this step was repeated twice. The membrane was then washed with PBS (pH 7.4) for 10 minutes twice, followed by a final wash step with TBST (pH 8) for 5 minutes. The membrane was developed to check for proof of complete stripping and observed for presence of any bands. After confirming absence of protein bands, the membrane was ready for blocking in 10% Blocking buffer.

2.10 Insoluble protein expression with Histidine tag purification column

The principle behind affinity column purification lies in the ability of a Histidine (His) tagged protein having 6 amino acids present at the N or C terminus, which binds to the ion, Nickel (Ni). The Ni is bound to nonporous agarose beads of 1-3 μm in size (His·Bind resin). The protein sample is added to a prepared His·Bind resin column, mixed, and allowed to absorb. High salt concentrated buffers such as urea or imidazole are used to elute the His-tagged protein, while low concentrations of phosphate and imidazole are used to wash away proteins bounded with less affinity.

2.10.1 Mini prep: trial purification of protein using urea denaturing His·bind column-without Imidazole

All steps were performed at RT. Well mixed, 100 μL of His Bind aliquot of resin was pipetted into a 1.7 mL microcentrifuge tube and spun at 5000 x g for 30 secs at 4°C. The liquid was removed, and the resin was washed with 250 μL of milli Q. H2O and spun at 5000 x g for 30 secs at 4°C. The step was repeated once again. The liquid was then removed and 200 μL of Equilibration Buffer (Di $NaPO_4$+ NaCl, p H 8) was added and mixed by inverting the tube. It was then spun at 5000 x g for 30 secs at 4°C. The supernatant was removed and discarded. The resuspended insoluble fraction or the final pellet about 100 μL was added to this solution, mixed well and centrifuged at 5000 x g for 30 secs at 4°C. The supernatant was saved from this step for PAGE gel analysis. Next, the affinity gel was washed with 500 μL of Wash Buffer (same as Equilibration buffer, pH 8) by mixing briefly then centrifuging at 5000 x g for 30 secs at 4°C. The supernatant was collected and saved for PAGE gel analysis. The protein was then eluted from the affinity gel using 50 μL of Elution Buffer at pH 6.0. This was mixed well with the gel and centrifuged at 5000 x g for 30 secs at 4°C. The supernatant was collected again, and the step was repeated to collect another supernatant sample.

Following this, the protein was eluted again for the final time using 50 μL of Elution Buffer (pH 4.5). Most recombinant proteins with a histidine tags do not

elute at pH 6.0, so it is advisable to try a lower pH such as pH 4.5. The solution was mixed well before centrifuging at 5000 x g for 30 secs at 4°C. The supernatant was collected, and the steps were repeated to collect another supernatant sample. All the supernatant samples were processed for analysis with PAGE gel and western blot analysis.

2.10.2 The affinity gel stripping method for reuse

Wash the affinity gel with 2 column volumes of deionized water, followed with another wash of 3 column volumes of 30% ethanol. Add Milli. Q H2O and store at 4°C for reuse.

2.10.3 Mini Prep: trial purification of protein using urea denaturing His·bind column- with Imidazole.

Strip the HIS-Select Nickel Affinity Gel (Section 2.10.2).
The following sequence of washes were undertaken to charge and equilibrate the column:

a) 3 volumes sterile deionized DDH2O
b) 5 volumes of 1X Charge Buffer (pH 7.9)
c) 2 volumes of Equilibration Buffer (pH 8.0)

About 200 μL of Equilibration Buffer (1X Binding buffer+6M Urea, pH 8.0) was added and mixed well. It was centrifuged for 30 seconds at 5,000 x g discarding the supernatant. Next, 100 μL of clear recombinant protein solution (final pellet) was added and gently mixed for a minute. It was then centrifuged for 30 seconds at 5,000 x g and saving the supernatant. The affinity gel was washed two times with at least 500 μL of Wash Buffer (same as Equilibration buffer but pH 6.3). The mixture was gently mixed with the affinity gel for 10 seconds, then centrifuged for 30 seconds at 5,000 x g. The wash-through was reserved. The target protein was eluted using 50 μL of Elution Buffer (1X Elute buffer+6M Urea, pH 6). The buffer was added and mixed with the affinity gel. It was later centrifuged for 30 seconds at 5,000 x g and the supernatant was reserved. Most of

the protein will elute off in the first 50 μL fraction, but some residual protein can be eluted in the second cycle. The two fractions were saved as a single pool or separate fractions. Lastly, elute the target protein for the final time with 50 μL of Elution Buffer pH 5.0 and pH 4.5; collecting their respective supernatants. Later, Bradford (Section 2.5) and PAGE gel analysis was continued (Section 2.6.2)

2.11 Immunohistological evaluation using VMO1 antibody

2.11.1 Preparation of microscope slides

To observe tissue sections for physiological and anatomical observations under the microscope, it was important that the tissue specimen was prevented from falling off the glass slide during staining and washing steps. Thus, grease and dust free slides are generally coated with adhesive compounds, such as poly L- Lysine, 3- amino propyl tri-ethoxysilane (APES), paraffin, etc (SIGMA-ALDRICH, 2018; Thomas Scientific, 2018). Apart from these compounds, Gelatine is the most common and frequently used adhesive material that forms a crosslinking sticky polymer for tissue attachment and Chromium potassium sulphate makes the slide positively charged to attract negatively charged tissue (Systems, 2018).

Gelatin coated microscope slides (Fronine, 76 mm x 26 mm x 1-1.2 mm with plain ground edges; catalog no. 7101WT) were prepared by placing the slides in the glass racks and soaking them in a container with hot tap water and detergent for 15 mins at RT. The slides were rinsed under running hot tap water for 5 mins at RT, followed by another rinsing with cold water for 5 mins at RT. The slides were air dried on the rack by placing it upright for one hour at RT or in the incubator at 40°C O/N. Once the slides were dried, 8 steps were undertaken as per Table 2.6 to continue ahead.

Table 2.6 Protocol for coating slides with Gelatin

Sr No	Particulars	Time (Min)	Temperature (°C)
1	Take 500 mL of sterile distilled water in a 1 L beaker		90
2	Add 1.0 g of Gelatin type A. Make sure to stir the solution until the gelatin is completely dissolved		50
3	Add 0.05 g Potassium Chrome III Sulfate $KCr(III)SO_4$		35
4	Transfer the cooled gelatin solution into the staining dish, and dip the side rack containing clean slides into it for 2 mins and remove the rack	1	RT
5	Allow the rack to dry on its side (with the frosted portion or where the label would be down)	5	RT
6	Repeat steps 4 and 5		
7	Dry the slides	O/N	40
8	Store the slides and label (G=Gelatin coated)		RT

2.11.2 Preparation of mammalian tissue

Paraffin embedded specimens of Mouse ear (P240), mouse cochlea (P5) and rat lung (L5) tissues were used that were kindly donated by Zehan Li and Tony Cecile using the methods published by Forrester-Gauntlett, B. K. E. (2013). In brief, 10 µM of specimens were cut into thin sections and collected on the Gelatin coated slides (~six sections per slides) inside Leica cryostat microtome. The tissues were rapidly frozen while sectioning inside the chamber (www.leicabiosystems.com). The slides were stored at -20°C.

2.11.3 Haematoxylin and Eosin Staining

Application of dyes or stains have often been applied to observe the morphology of the specimen. Haematoxylin (H) and Eosin (E) stain, a combination of acid and basic dyes which is commonly employed to stain and demonstrate the presence of nucleus (purple) and cytoplasmic inclusions (red-pink) in the tissue (Fischer *et al.*, 2008). To prepare slides for H and E stain, the sectioned slides stored at -20°C were brought to RT. Next, they were fixed with 4% Paraformaldehyde dissolved in PBS (pH 7.4) for 5 minutes. A washing step was then employed were the slides were put in a glass chamber and washed with running tap water (not cold) for a minute, to remove the rest of the formaldehyde. The slides were drained (not dried)

on a tissue before continuing. The actual staining was then carried as per Table 2.7 that outlines 7 steps to complete this procedure.

Table 2.7 Protocol for H and E stain

Sr No	Steps	Treatment
1	Hematoxylin Staining	Incubate the slides with Mayer's hematoxylin solution for 5 minutes to stain the nuclei dark
2	Washing	Use water to rinse all the excess hematoxylin
3	Bluing	Rinse in Scott's tap water substitute for 2 min
4	Eosin staining	Add Eosin solution and incubate the slides 10 minutes to stain the fibers red
5	Washing	Put the slides in a glass chamber and wash 3x for 1 minute with DDH2O to remove excess of Eosin
6	Dehydration	Drain successively in - 70% ethanol for a minute - 90% ethanol for 30 seconds - 100% ethanol for 30 seconds - Xylene for 30 seconds
7	Mounting	Mount the slides with 1-2 drops of a Fluoroshield™ and cover with 22x60mm cover slides, avoiding air bubbles

2.11.4 Immunohistochemistry (IHC) with Antigen Revival & Autofluro removal for protein localization

IHC demonstrates the distribution and localization of antigen of interest or specific cellular components in the tissue specimen using antigen-antibody interactions. This interaction can be detected by enzymes, such as Horseradish Peroxidase (HRP), Goat anti-Rabbit (abcam, 2018) that catalyses a coloured reaction.

Requirements: Antigen revival-citrate buffer pH 6.0, 4% PFA, 1X PBS (pH 7.4), 1X PBS-T (pH 7.2), 0.3M Glycine, Blocking solution for antibodies, tissue specimens, and fluorescence microscope. Antigen revival-citrate buffer pH 6.0, recommend by VMO1 antibody manufacturer. This buffer helps to break the protein cross-links, therefore unmasks the epitopes in formaldehyde-fixed embedded tissue sections. It therefore enhances staining and intensity of antibodies.

The 10 µM frozen sections collected on Gelatin-coated microscopic slides (stored at -20°C) were brought to RT. The Antigen Revival step was followed next that performed using Microwave method. Here, the slides were briefly boiled in antigen revival buffer, pH 6 for 10 minutes at 700W in the microwave with medium low setting. The slides were checked every 2 min to ensure sufficient solution was covering the sections and were also retained on the slides. Next, they were allowed to cool at RT. They were then washed in 1X PBS for 5 mins twice. About 2-3 drops of 0.5% Triton X-100 solution was then added on the sections and incubated for 30 minutes to permeabilize the cell membranes. Sections were rinsed with 1 PBS and 1X PBST for 5 mins each and drained. Once drained, each section was encircled with a PAP pen, that creates a hydrophobic barrier and prevents mixing of reagents. The sections were then treated using the 8 steps mentioned in Table 2.8.

Table 2.8 Protocol for IHC

Steps	Treatment	Reagent and Purpose	Treatment
1	Block (Block Aid) Thermofisher (Molecular Probes, catalog no B-10710) Few drops	Reduces the background signal due to non-specific interaction of primary and secondary antibodies with proteins	60 minutes or O/N
2	300 mM Glycine, few drops	Binds to the free aldehyde groups and reduce high background.	2 hr at 4ºC
3	Primary Antibody in 5% Goat Serum	Few drops	1 hr at RT or O/N at 4ºC
4	1X PBS-T		10 minutes
5	1X PBS-T shaking		10 minutes x2
6	Secondary antibody	Add 1:2,000 dilution made up in 1X PBS.	Incubate for 1 hr in dark at RT
7	1X PBS		15 minutes x3
8	Dry in the dark and mount using Fluoroshield™ with DAPI		Store in the dark at 4°C

Table 2.9 List of antibodies used.

Antibodies	**Concentrations used**	**Supplier**	**Catalogue Number**
Anti VMO1 Antibody Rabbit polyclonal to VMO1	1:200, 1:400	Abcam	ab126510
His-tag antibody	1:1000	GenTex Inc.,	GTX115045
HRP-Goat Anti Rabbit	1:2000	Abcam	Ab97051
Goat Anti-Rabbit IgG H&L (Alexa Fluor® 488, Excitation: 495nm, Emission: 519nm).	1:200, 1:400	Abcam	ab150077

CHAPTER THREE

BIOINFORMATICS

3.1 Introduction to Bioinformatics

Bioinformatics is conceptualising biology in terms of molecules and applying "informatics techniques" to understand and organize the information associated with these molecules on a large scale (Luscombe *et al.*, 2001). These techniques are derived from computer science and mathematics and includes a collection of nucleotide (DNA or mRNA) and/or amino acid sequences that can translate molecular data into patterns to predict function and three-dimensional (3D) structures of nucleic acids, proteins, and protein–nucleic acid complexes. Mining this information provides the potential to identify and discovery novel clinical applications for the prevention or treatment of disease. In comparison, the field of "comparative genomics" analyses similarities and differences in the genetic content from a variety of organisms ranging from bacteria to humans. By comparing the sequences, distinguishing factors can be revealed at molecular level to study evolutionary changes in functional and/or diseased genes (Touchman, 2010).

3.2 Objectives for studying bioinformatics

There were four main objectives for implementing bioinformatics in this study. Thus, this chapter will discuss the outcomes of our four objectives.

3.2.1 Objective one

The first objective was to confirm that the insertion of human *VMO1* gene in the provided plasmid construct [pET28b (+), Appendix Two, 1] was present in *E. coli* Rosetta (DE3) pLysS. For which, the plasmid DNA was extracted and sequenced using SH3F, SH3R primers, T7 promoter and T7 terminator (Table 2.1). The sequencing reads obtained from the coverage of their primers were assembled in Geneious 7.1.8 and 7.10.2 (Figure 3.7). This resulted in a 423 nucleotides consensus sequence and analysed thereon (Table 3.1).

3.2.2 Objective two

The second objective was to confirm that the open reading frame of the *VMO1* gene was in-frame with the His-tag (Figure 2.9) so the recombinant protein to be produced later could be easily purified and detected. We confirmed the presence of N-terminus His-tag by scanning our sequences in Geneious. Also, the recombinant protein produced later was detected by a His-tag antibody which affirmed our second objective.

3.2.3 Objective three

The third objective was to investigate the *VMO1* gene with respect to exon structure, predicted molecular weight, molar absorbance and extinction coefficient, isoelectric point, hydrophobicity index, characterisation of any motifs or binding domains, and conservation in other species by comparative genomics. Various database/software/tools both online and offline (including Geneious, Expasy, atg.me, National Centre for Biotechnology Information (NCBI), Uniprot, The European Bioinformatics Institute, Mouse Genome Informatics, Human Genome Browser, etc.) (Appendix Two, 2) helped in achieving these goals. Most of which have been around for more than 25 years and contributed to genomics in their own way. Basic local alignment tool, BlastP suit, pair-wise/multiple alignment; were used to understand how this gene was conserved across other organisms. These tools compare individual sequences present in the database (NCBI) and produces a match based on number of hits (identity with each amino acid or nucleotide) with Z scores. Objective three proved to be our foundation in understanding the *VMO1* gene.

3.2.3.1 Chicken (*Gallus gallus*) *VMO1*

Initially we understood the *VMO1* gene from chicken, where it was first discovered and the protein (420 kDa) was crystallized by (Shimizu *et al.*, 1994). *VMO1* in Chicken is the component of Chromosome 1 and had 2 exons (Figures 3.1 and 3.2). Its UniProtKB number is P41366 (VMO1_CHICK). The protein is

of 183 amino acids long with mass (Da) of 20,234. The exact function remains unknown; however, it represents a component of the outer membrane of the VM of the egg. It also synthesizes N-acetylchito-oligosaccharides (n=14-15) from hexasaccharides of N-acetylglucosamine surprisingly similar to the transferase activity of the lysozyme (Raftery & Rand-Meir, 1968). VMO1 binds to ovomucin fibrils in the egg yolk membrane or the vitelline membrane. The structure consists of three beta-sheets that forms Greek key motifs. Furthermore, the structure also has strong resemblance with delta-endotoxin, and a carbohydrate-binding site.

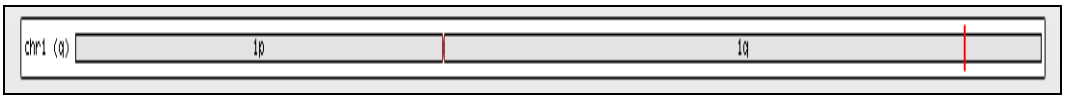


Figure 3.1 Location of chicken *VMO1* gene on Chromosome 1 indicated by the long red line.

Figure 3.2 Structure of chicken *VMO1* gene.

It is a 2626 bp linear mRNA that encodes three exons.

3.2.3.2 Human (*Homo sapiens*) *VMO1*

The *VMO1* gene in humans is mapped to chromosome 17, on the short arm p13.2 according to Entrez Gene (Figure 3.3). It covers 5.45 kb fragment, from 4694029 to 4688577 and is in the 5’ to 3’ direction (NCBI 37, August 2010) on the reverse strand. Its UniprotKb Accession number is Q7Z5L0.

VMO1 gene in humans exists in four isoforms (Figure 3.4), generated through evolutionary splicing events. RNA-sequencing data on tissue samples representing 27 different tissues, from 95 human individuals helped determine tissue-specificity of all protein-coding genes. These tissues include, adrenal, appendix, bone marrow, brain, colon, duodenum, endometrium, oesophagus, fat, gall bladder, heart, kidney, liver, lung, lymph node, ovary, pancreas, placenta, prostate, salivary gland, skin, small intestine, spleen, stomach, testis, thyroid, urinary bladder. The highest expression was found to be in spleen (5.346 RPKM),

followed by lung (3.307 RPKM), appendix (2.692 RPKM), and liver (1.855 RPKM) (NCBI Gene, 2018).

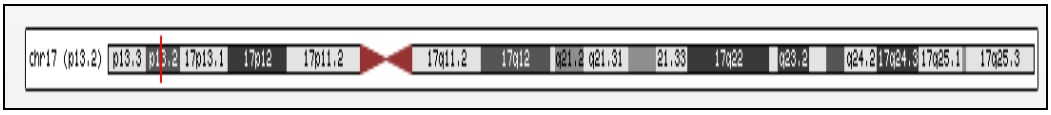


Figure 3.3 Cytogenetic location of *VMO1* gene on human chromosome 17.
The red vertical line shows *VMO1* location at Chr 17q13.2

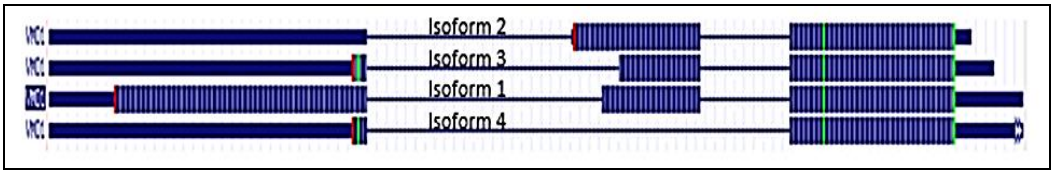


Figure 3.4 Gene structure of the four mRNA isoforms of Human *VMO1*.
The thin horizontal line indicates the intron, and the larger blue boxes are the coding exons.

Human *VMO1* gene is found to have similarities as follows; 98% with chimpanzee *VMO1*, 78.9% identity with cattle *VMO1*, 78.9 % with rat *vmo1*, 71.6 % identity with mouse *vmo1*, 59.1 % identity with chicken *VMO1*, 57% with western clawed frog *LOC101731130*, and 53% identity with *LOC793369* gene in zebrafish.

3.2.3.3 Mouse (*Mus musculus*) *Vmo1*

Mouse *Vmo1* gene has three codons (Figure 3.6) and mapped to chromosome 11, at 11 B3 (Figure 3.5) amongst 3505 other genes. Here it covers 1.11 kb, from 70328119 to 70327013 (NCBI 37a, Aug 2007), on the reverse strand. The protein has a mass of 21,957 Da. Its UniprotKb Accession Primary (citable) accession number is Q5SXG7.

This gene found to be expressed low levels, about 5.8% of the average gene. It also contains two different gt-ag introns that produces only one mRNA upon transcription (NCBI AceView, 2018b). *Vmo1* mRNA is 673 bp long, which was reconstructed from 6 different cDNA clones. The pre-messenger RNA has 3 exons that covers 1.10 kb. It has a predicted protein of 206 aa (22.6 kDa, pI 6.2).

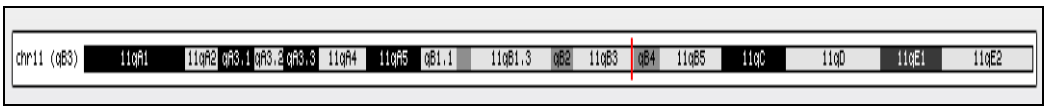


Figure 3.5 Location of mouse *Vmo1* gene on Chromosome 11 indicated by the long red line.

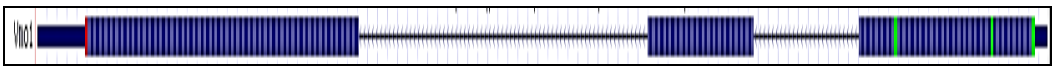


Figure 3.6 Structure of mouse *Vmo1* gene

Vmo1 gene also contains six distinct gt-ag introns (Alternative intron 1-6). On transcription, it produces seven different mRNAs, six alternatively spliced variants (VMO1 homolog CRA a- bAug10, complete CDS mRNA of 1457 bp; VMO1 homolog precursor- cAug10, complete mRNA of 773 bp; VMO1 homolog CRA b family member- dAug10, complete CDS mRNA of 813 bp; VMO1 homolog precursor family member- eAug10 of, complete CDS mRNA of 850 bp; VMO1 homolog CRA a family member- fAug10, complete mRNA of 651 bp; VMO1 homolog precursor family member- gAug10, complete mRNA of 758 bp) and one unspliced form (a Putative protein VMO1 alternative variant- aAug10, complete mRNA of 4264 bp). These variants have been reconstructed from different cDNA clones from brain, foetal eye, thymus, spleen, stomach, etc. AceView also enables to view proteins that have been identified with their respective mRNA variants.

The VMO1 proteins are expected to possess a molecular function (being a structural constituent of the VM) and to also localize in various cellular compartments (e.g. cytoplasmic region, extracellular space, extracellular region). For instance, Protein ‘d’ associated with ‘dAug10’ mRNA variant, contains two exons in the CDS region and is found in the cytoplasmic region and is of a good quality. Protein ‘g’ associated with ‘gAug10’ has very good protein quality and is either secreted or found extracellular (NCBI AceView, 2018a).

Comparative analysis of chicken VMO1 protein-coding sequence and the human, mouse, rat, and bovine VMO1 proteins using multiple sequence alignment tool revealed high degree of homology of 55%, 53%, 48%, and 54%, respectively (Lee *et al.*, 2015).

3.2.4 Objective four

Finally, our fourth objective was to predict a 3D structure for our induced, recombinant VMO1 protein. A blast search of our consensus sequence with the four human isoforms, resulted in a 100% identity with isoform 3 CDS region. The nucleotide sequence was then translated to amino acid sequence in Geneious, to get a predicted protein size, which was found to be ~15 kDa in size. The recombinant VMO1 isoform 3 protein was 140 amino acids peptide (Table 3.1, refer B and Figure 3.8). The 3D structure contained the beta folds (Greek key motifs) (Figure 3.23) and a secondary DNA fold structure (Figure 3.22).

3.3 Sequencing analysis of plasmid DNA (pET28b (+) pLysS) containing the *VMO1* insert

A consensus sequence of 423 nucleotides (Table 3.1, refer A) was generated from four overlapping contigs in Geneious following the DNA sequencing (Sections 2.3.1). All reads obtained from the three sequences of T7 promoter, SH3F, and SH3R had good peaks, except for the beginning of the sequences, while the T7 terminator failed to work.

The resulting consensus sequence (Figures 3.8 and 3.11) was blasted in NCBI. The BLAST data and Geneious software predicted four ORFs in the predicted recombinant protein sequence (Figure 3.9). Frame 2 contained the His-tag.

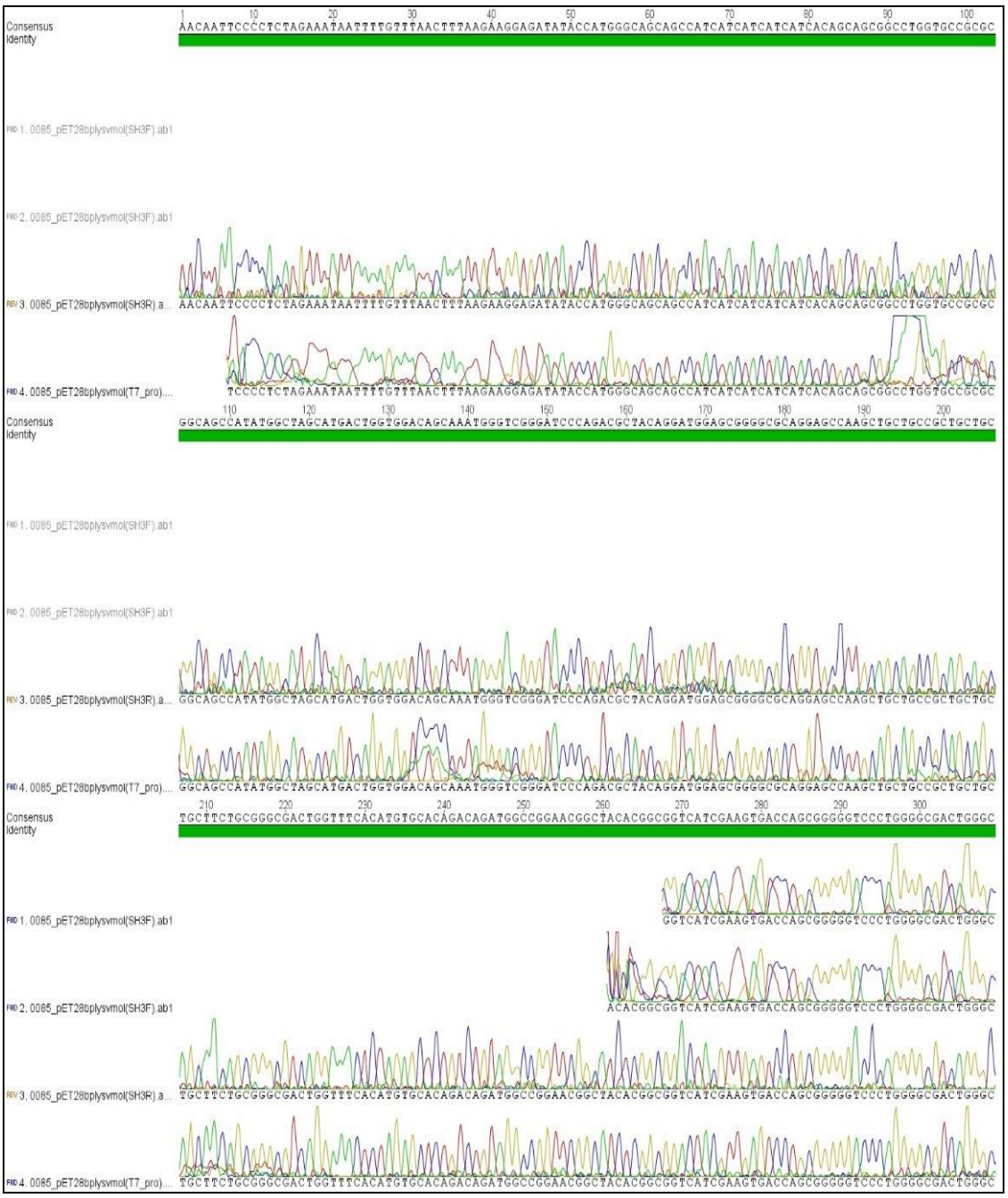


Figure 3.7 Electropherogram of the sequenced pET28B (+) pLysS - *VMO1* plasmid.

The green bar confers 100% identity.

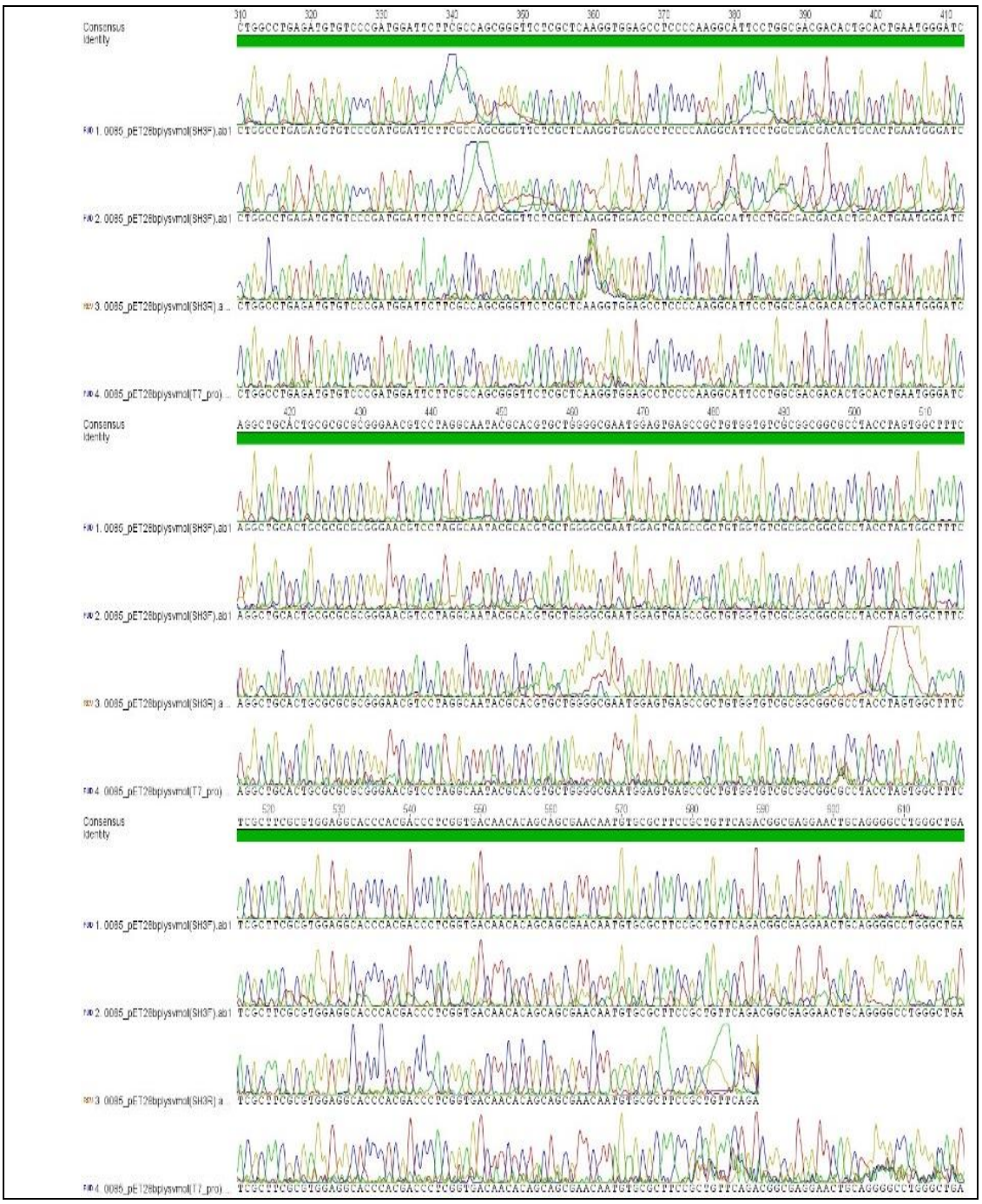


Figure 3.7: Electropherogram of the sequenced pET28B (+) pLysS - *VMO1* plasmid.

The green bar confers 100% identity (Continued).

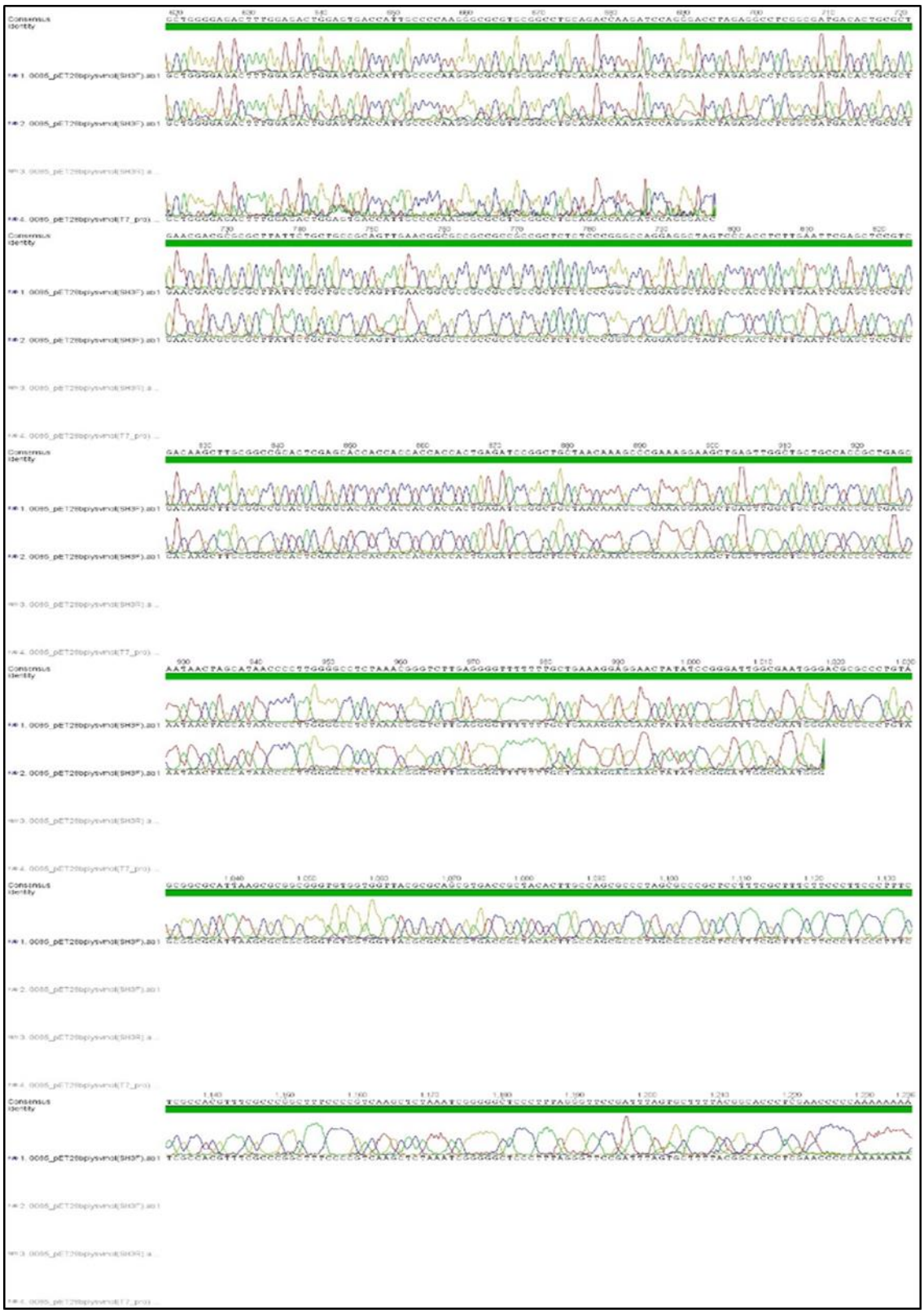


Figure 3.7: Electropherogram of the sequenced pET28B (+) pLysS - *VMO1* plasmid.

The green bar confers 100% identity (Continued).

The sequences below are obtained after ORF extraction and sequence trimming in Geneious® 7.1.8. They represent the final output after sequencing. The amino acid sequence was obtained after back translating the nucleotide sequence. The nucleotide sequence has 423 nucleotides. It shows identity from nucleotides 99-423 of the FASTA sequence with 67-391 of human *VMO1* isoform 3 mRNA (NM_001144940.1). The amino acid sequence has 140 amino acids. It showed identity from amino acids 39-140 with 1-102 amino acids of isoform 3 protein sequence (back translated from NM_001144940.1). The bold letters represent the location of human VMO1 isoform 3 in these sequences.

Table 3.1 Consensus VMO1 sequence.

A) Nucleotide sequence/ FASTA format
ATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGACTGGTGGACAG CAAATGGGTCGGGA**TCCCAGACGCTACAGGATGGAGCGGGGCGCAGGAGCCAAGCTGCTGCCGCTGCTGCTGCTTCTGC GGGCGACTGGTTTCACATGTGCACAGACAGATGGCCGGAACGGCTACACGGCGGTCATCGAAGTGACCAGCGGGGGTCC CTGGGGCGACTGGGCCTGGCCTGAGATGTGTCCCGATGGATTCTTCGCCAGCGGGTTCTCGCTCAAGGTGGAGCCTCCC CAAGGCATTCCTGGCGACGACACTGCACTGAATGGGATCAGGCTGCACTGCGCGCGCGGGAACGTCCTAGGCAATACGC ACGTGCTGGGGCGAATGGAGTGA**
B) Amino acid sequence/ Protein sequence
MGSSHHHHHHSSGLVPRGSHMASMTGGQQMGRDPRRYR**MERGAGAKLLPLLLLLRATGFTCAQTDGRNGYTAVIEVTSGGPW GDWAWPEMCPDGFFASGFSLKVEPPQGIPGDDTALNGIRLHCARGNVLGNTHVLGRME**

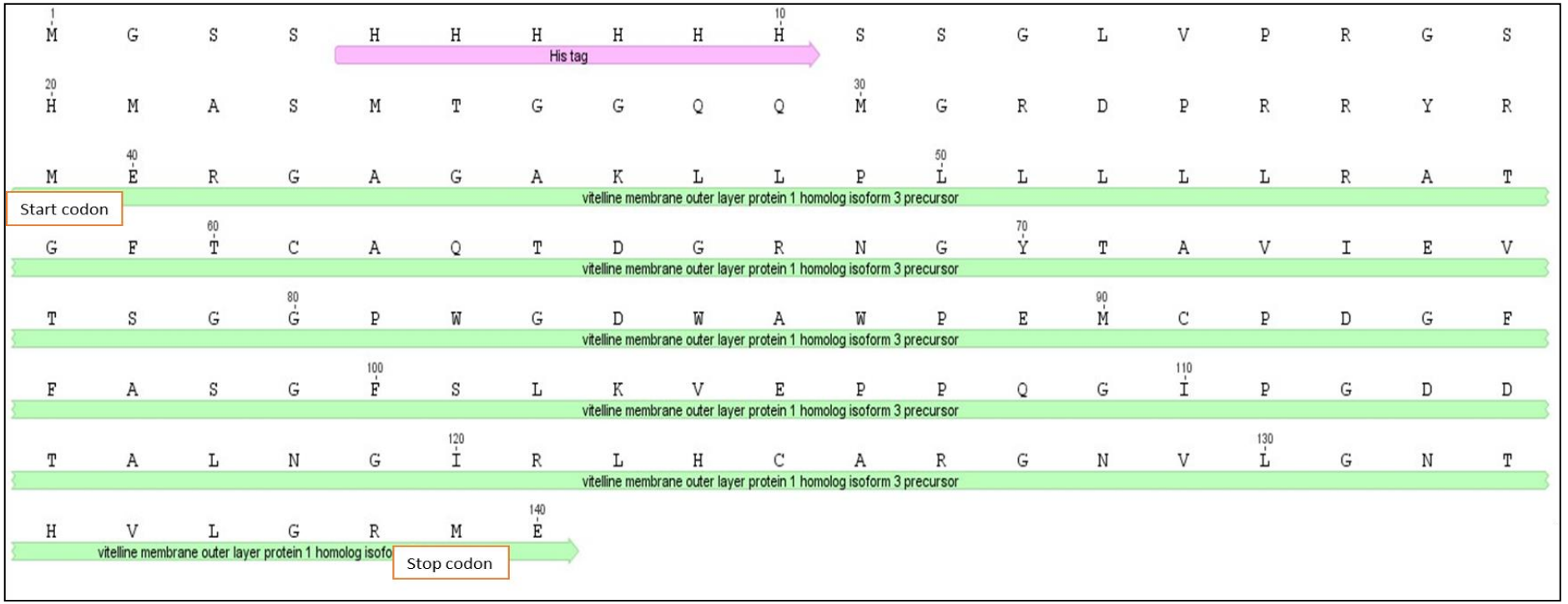


Figure 3.8 Open reading frame of recombinant protein.

The Figure represents the consensus ORF sequence that was produced from alignment of sequences generated by T7 promoter, T7 terminator, SH3F, and SH3R primers (Figure 3.7). The consensus amino acid sequence (1-140) was annotated using Geneious 7.10.2. The start codon is located at the 39th position while the stop codon is at the 139th position. The green arrow indicates the position of the recombinant human VMO1 isoform 3, which is in-frame with the upstream N terminal (position 5-10) His-tag coloured in purple.

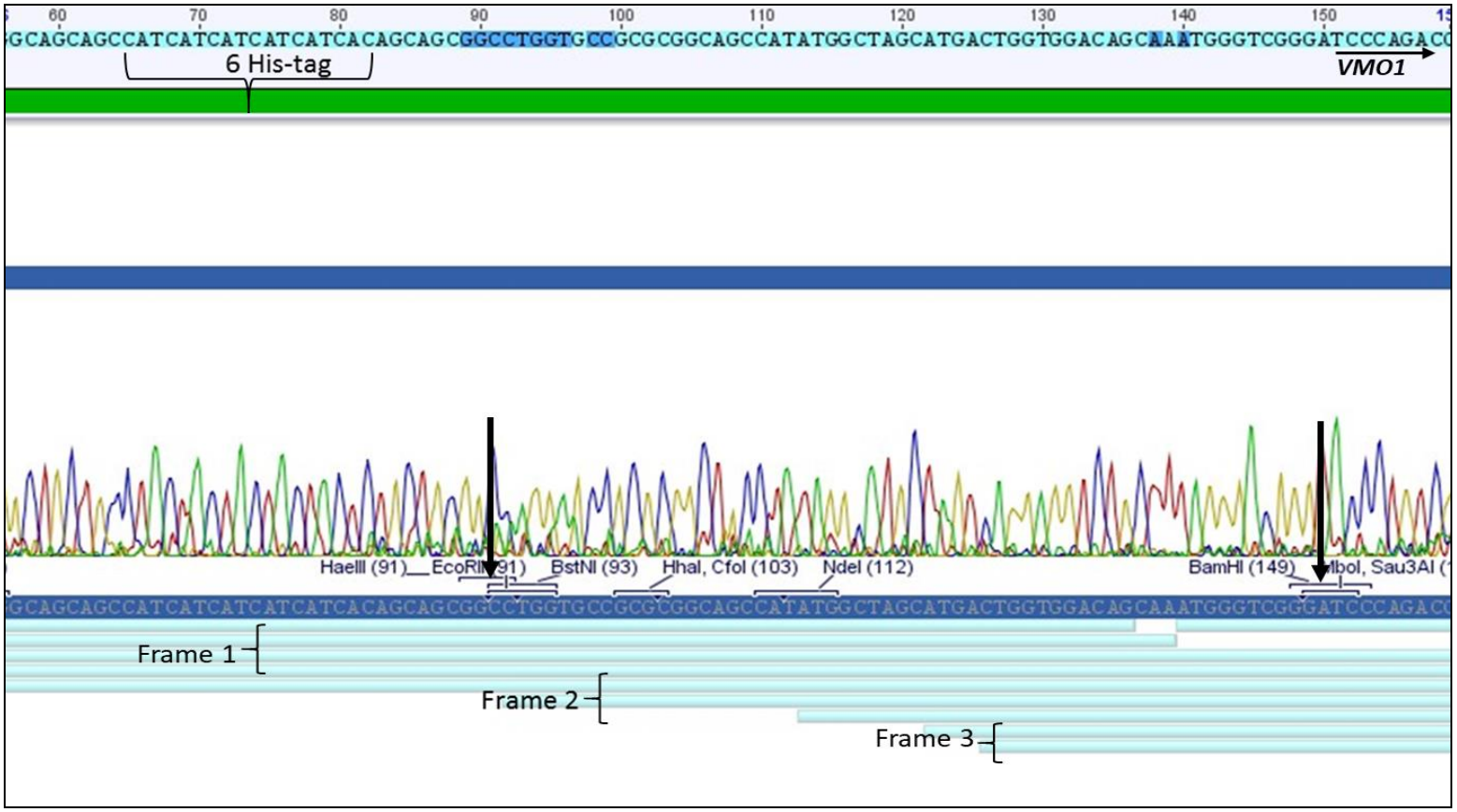


Figure 3.9 His-tag in ORF 2

The figure represents a part of an unedited assembled sequence (green bar) showing three open reading frames (light blue). The six histidine tags can be found from nucleotide position 65-82 (top left). *VMO1* gene can be found starting at nucleotide position 151 (horizontal arrow, top right) and continues thereafter in the sequence. All of which are shown to be in Frame 2 that also contains the restriction enzyme sites for EcoR1 (position 91) and BamH1 (position 149) indicated with bold, black arrows.

3.4 Confirmation of *VMO1* isoform 3 and comparative analysis.

To confirm the structural variant of human *VMO1* gene, the consensus nucleotide sequence (Table 3.3, refer A) was blast searched online in NCBI. Once the nucleic acid or FASTA sequence was uploaded as a "query" it gave a resultant output that showed higher degree of similarities. This online tool searches reference sequences previously available in NCBI database and finds matched sequences as a comparison.

Table 3.2 Top hits of BLAST search using consensus sequence.

Sr. no	***VMO1* of Organism and mRNA transcript.**	**Max Z Score**	**Query Coverage (%)**	**Identity (%)**	**Nucleotide Range**	**GenBank Accession number**
1	*Homo sapiens* variant 3	595	76	99	67-391	NM_001144940.1
2	PREDICTED: *Pan panics* variant X3	579	76	99	67-391	XM_008962565.1
3	PREDICTED: *Pan troglodyte* variant X	573	76	98	703-1027	XM_016931282.1
4	PREDICTED: *Gorilla gorilla* variant X3	562	76	98	169-493	XM_004058335.1
5	*Homo sapiens* variant 2	562	72	99	67-373	NM_001144939.1
6	*Homo sapiens* variant 1	562	72	99	67-373	NM_182566.2
7	*Nomascus leucogenys* variant X3	529	76	96	169-493	XM_003277871.2
8	*Pongo abelii* variant	507	76	95	67-391	XM_009251164.1
9	*Colobus angolensis palliates*, variant X3	462	74	93	175-490	XM_011958357.1
10	*Papio anubis*	427	71	92	172-499	XM_003912150.4
11	*Macaca fascicularus*	422	69	92	168-470	XM_00558281.2
12	*Homo sapiens* variant 4	387	50	99	67-281	NM_001144941.1
13	*Saimiri boliviensis*, variant X3	375	72	89	102-406	XM_010342015.1
14	*Felis catus*	289	74	84	1094-1404	XM_003996152.5
15	*Chinchilla lanigera* variant X1	211	69	80	1041-1324	XM_005399599.2
16	*Balaenoptera acutorostrata* variant X1	200	76	79	32-335	XM_007166495.1

3.4.1 Confirmation of *VMO1* isoform identity and Pair-wise local alignment.

3.4.1.1 *VMO1* isoform 3 identity by nucleotide alignment

To confirm that pET28b (+) had human *VMO1* isoform 3 variant in its construct, the consensus nucleotide sequence (Table 3.1, refer A) was blasted with human *VMO1* Isoform 3 mRNA (NM_001144940.1). Figure 3.10 shows 99% identity between the said sequences with zero gaps.

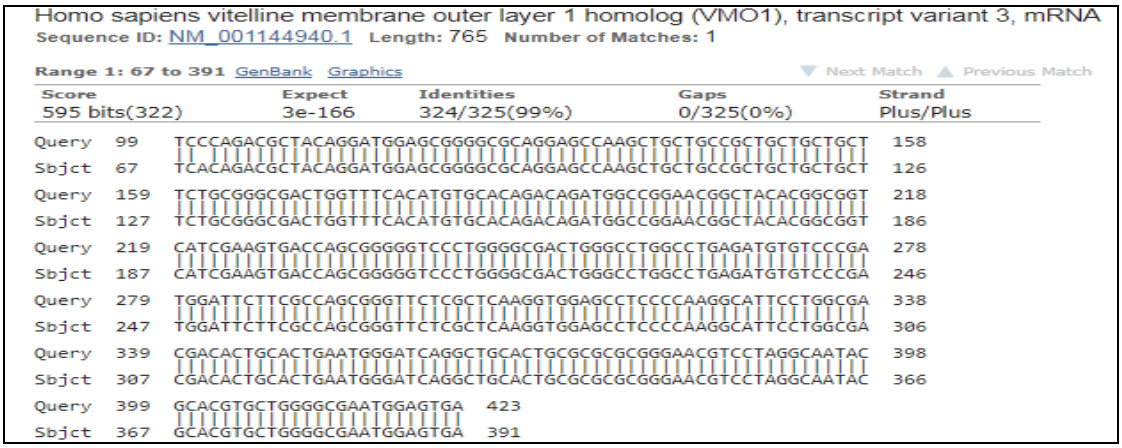
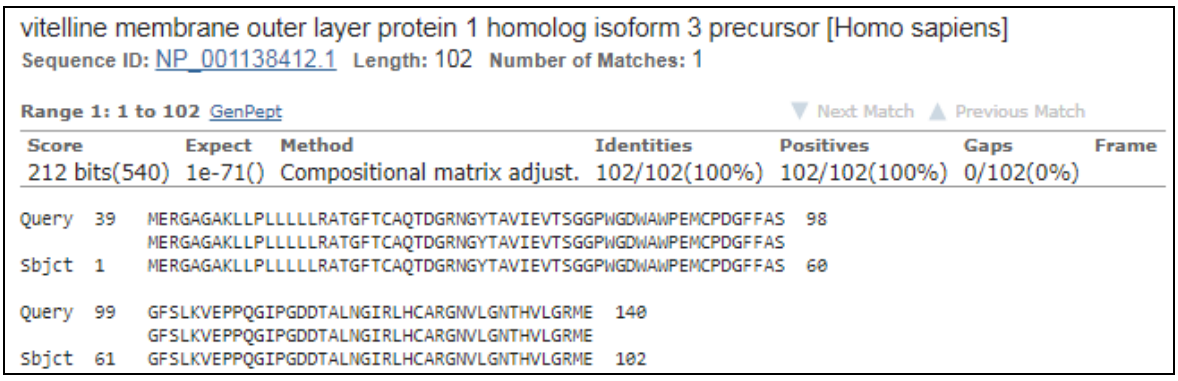


Figure 3.10 Confirmation of presence of human *VMO1* 3 isoform.

The blast search was carried using BLAST®, blastn suite.

3.4.1.2 *VMO1* isoform 3 identity by amino acid sequence alignment

The consensus protein sequence (Table 3.1, refer B) was blasted with human VMO1 Isoform 3 amino acid sequence (NP_001138412) to find regions of similarities. Figure 3.11 shows 100% identity between the said sequences with zero gaps.


vitelline membrane outer layer protein 1 homolog isoform 3 precursor [Homo sapiens]
Sequence ID: NP_001138412.1 Length: 102 Number of Matches: 1

Range 1: 1 to 102 GenPept ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps	Frame
212 bits(540)	1e-71()	Compositional matrix adjust.	102/102(100%)	102/102(100%)	0/102(0%)	

```
Query  39   MERGAGAKLLPLLLLLRATGFTCAQTDGRNGYTAVIEVTSGGPWGDWAWPEMCPDGFFAS  98
            MERGAGAKLLPLLLLLRATGFTCAQTDGRNGYTAVIEVTSGGPWGDWAWPEMCPDGFFAS
Sbjct  1    MERGAGAKLLPLLLLLRATGFTCAQTDGRNGYTAVIEVTSGGPWGDWAWPEMCPDGFFAS  60

Query  99   GFSLKVEPPQGIPGDDTALNGIRLHCARGNVLGNTHVLGRME  140
            GFSLKVEPPQGIPGDDTALNGIRLHCARGNVLGNTHVLGRME
Sbjct  61   GFSLKVEPPQGIPGDDTALNGIRLHCARGNVLGNTHVLGRME  102
```

Figure 3.11 Protein alignment between consensus plasmid sequence and human VMO1 isoform 3.

The two sequences were pair-wise alignment using the SMARTBLAST tool in NCBI.

In conclusion, sequencing step revealed that the plasmid pET28b (+) had *VMO1* isoform 3 mRNA as an insert. It can also be predicted that isoform 3 translates into 102 amino acid with a calculated molecular mass of 15 kDa in *E. coli*. This variant has three exons but contains an alternative splice site at the 3' end of the second exon, resulting in a frameshift and a smaller protein as seen in Figure 3.4.

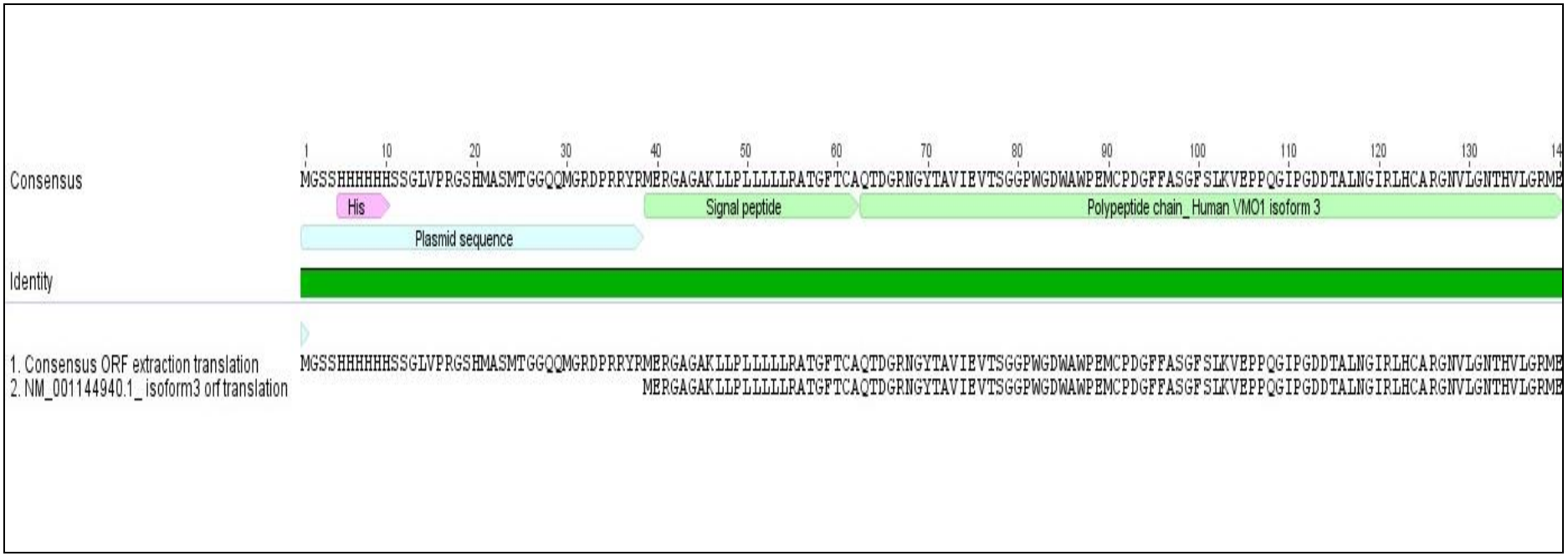


Figure 3.12 Alignment of consensus ORF extraction of plasmid pET28b (+) with isoform 3

The Consensus ORF of plasmid pET28b (+) sequence (start codon 1 to 140) shows that the predicted recombinant protein had 100% similarity (green bar) with the Human VMO1 variant 3 (39-140, GenBank Accession Q7Z5L0-3). Amino acid sequence from 1 to 38 shows plasmid sequence including the His-tag (pink bar), 39-62 represents the signal peptide (green bar), and 63-139 presents the recombinant polypeptide chain (purple) after post-translational modification.

3.5 Structural attributes and comparative analysis of VMO1 protein

The next step was to study the structural similarities and differences between the VMO1 across human, chicken, mouse. This includes pairwise local alignment, blast search, multiple alignments.

3.5.1 Chicken VMO1 protein

Chicken VMO1 has the following characteristics as described in Table 3.3. The sequence was later back translated from protein to nucleotide to perform pair-wise alignment.

Table 3.3 Canonical sequence of Chicken VMO1

Signal peptide 1-20	Denotes the presence of an N-terminal signal peptide.	MKVLTPAALILLFFFYTVDA
Chain 21 – 183	describes the extent of a polypeptide chain in the mature protein following processing	RTREYTSVITVPNGGHWGKWGIRQFCHS GYANGFALKVEPSQFGRDDTALNGIRLR CLDGSVIESLVGKWGTWTSFLVCPTGYL VSFSLRSEKSQGGGDDTAANNIQFRCSDE AVLVGDGLSWGRFGPWSKRCKICGLQT KVESPQGLRDDTALNNVRFFCCK

3.5.1.1 Back translation of amino acid sequence of chicken VMO1 to nucleotide sequence using Back translation tool

> Chicken *VMO1*_FASTA format

ATGAAGGTGCTGACCCCCGCCGCCCTGATCCTGCTGTTCTTCTTCTACACCGT
GGATGCCAGAACCAGAGAGTACACCAGCGTGATCACCGTGCCCAACGGCGGC
CACTGGGGCAAGTGGGGCATCAGACAGTTCTGCCACAGCGGCTACGCCAACG
GCTTCGCCCTGAAGGTGGAGCCCAGCCAGTTCGGCAGAGATGATACCGCCCT
GAACGGCATCAGACTGAGATGCCTGGATGGCAGCGTGATCGAGAGCCTGGTG
GGCAAGTGGGGCACCTGGACCAGCTTCCTGGTGTGCCCCACCGGCTACCTGG
TGAGCTTCAGCCTGAGAAGCGAGAAGAGCCAGGGCGGCGGCGATGATACCG
CCGCCAACAACATCCAGTTCAGATGCAGCGATGAGGCCGTGCTGGTGGGCGA
TGGCCTGAGCTGGGGCAGATTCGGCCCCTGGAGCAAGAGATGCAAGATCTGC
GGCCTGCAGACCAAGGTGGAGAGCCCCCAGGGCCTGAGAGATGATACCGCCC
TGAACAACGTGAGATTCTTCTGCTGCAAG

3.5.1.2 Pairwise sequence alignment of consensus plasmid sequence and chicken *VMO1* sequence

The consensus plasmid (top sequence) (Table 3.1 A) was aligned with back translated chicken *VMO1* (bottom sequence) in EMBL-EBI. The two sequences show identity of 263/524 (50.2%), and having Gaps 193/524 (36.8%).

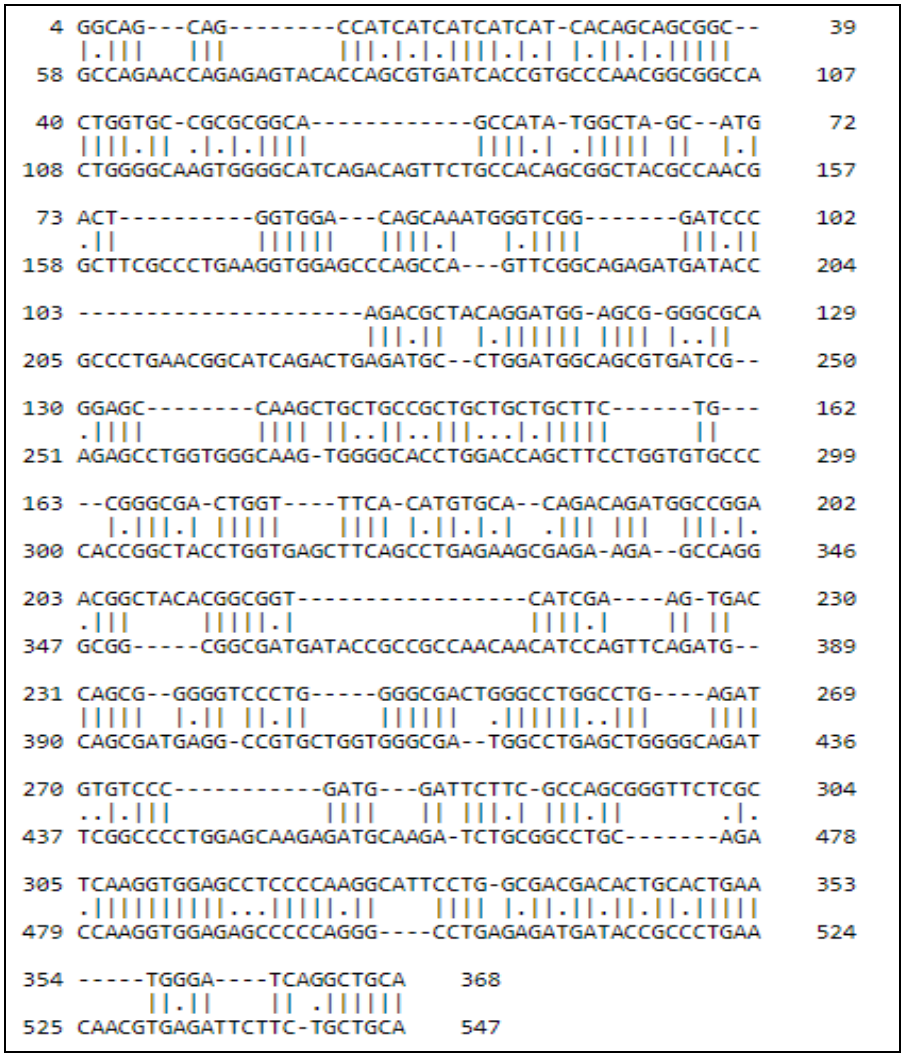


Figure 3.13 Nucleotide sequence alignment of chicken VMO1 with the plasmid sequence.

The vertical lines indicate identical nucleotides while the dots represent non-identical nucleotides. The dashes represent absence of nucleotides.

3.5.2 Human VMO1 protein

Human *VMO1* contains four isoforms (Figure 3.4) which can be translated to four amino acid sequences mentioned in Table 3.4.

Table 3.4 Structural variants of Human VMO1 isoforms amino acid sequences.

Isoform	GenBank Accession Number	Length (aa)	MW (Da)	Description	Amino Acid Sequence
1	Q7Z5L0-1	202	21,534	Referred to as the "canonical" sequence	MERGAGAKLLPLLLLLRATGFTCAQTDGRNGYTAVIEVTS GGPWGDWAWPEMCPDGFFASGFSLKVEPPQGIPGDDTALN GIRLHCARGNVLGNTHVVESQSGSWGEWSEPLWCRGGAY LVAFSLRVEAPTTLGDNTAANNVRFRCSDGEELQGPGLSW GDFGDWSDHCPKGACGLQTKIQGPRGLGDDTALNDARLFC CRS
2	Q7Z5L0-2	70	7,524	differs from the canonical sequence. 66-70: VEPPQ → LGRME 71-202: Missing.	MERGAGAKLLPLLLLLRATGFTCAQTDGRNGYTAVIEVTS GGPWGDWAWPEMCPDGFFASGFSLKLGRME
3	Q7Z5L0-3	102	10,827	differs from the canonical sequence. 98-202: VESQSGSWGE...NDARLF CCRS → LGRME	MERGAGAKLLPLLLLLRATGFTCAQTDGRNGYTAVIEVTS GGPWGDWAWPEMCPDGFFASGFSLKVEPPQGIPGDDTALN GIRLHCARGNVLGNTHVLGRME
4	Q7Z5L0-4	114	11,966	differs from the canonical sequence 104-202: SWGEWSEPLW...NDARLF CCRS → RWGAGVEDPLG	MERGAGAKLLPLLLLLRATGFTCAQTDGRNGYTAVIEVTS GGPWGDWAWPEMCPDGFFASGFSLKVEPPQGIPGDDTALN GIRLHCARGNVLGNTHVVESQSGRWGAGVEDPLG

3.5.2.1 Pairwise Sequence alignment of consensus plasmid sequence and human *VMO1* isoform 3 sequence

The human *VMO1* isoform 3 (top sequence) was pair-wise aligned with consensus plasmid (bottom sequence) (Table 3.1 A) in EMBL-EBI. The two sequences showed 43.9% identity. This was due to presence of gap after 397th nucleotide giving 448 gaps out of 818 nucleotides (54.8%).

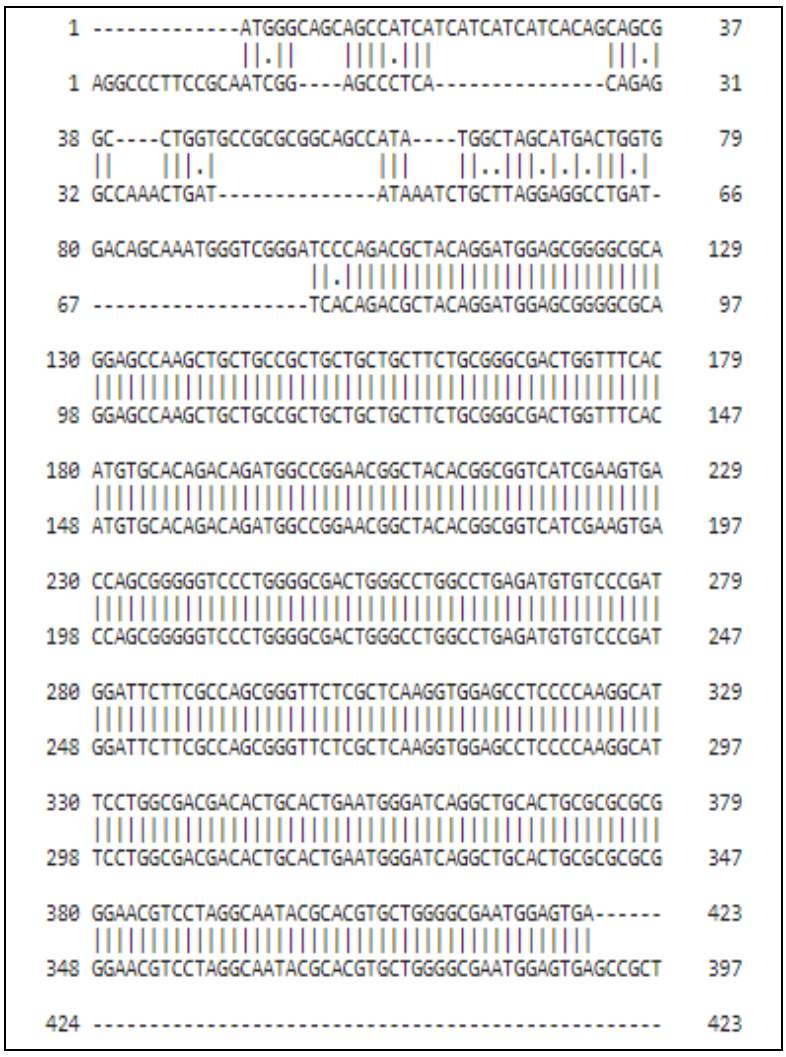


Figure 3.14 Nucleotide sequence alignment of human VMO1 isoform 3 with the plasmid sequence

The vertical lines indicate identical nucleotides while the dots represent non-identical nucleotides. The dashes represent absence of nucleotides.

3.5.3 Mouse Vmo1 protein

Mouse Vmo1 has the following characteristics as described in Table 3.5. The sequence was later back translated from protein to nucleotide to perform pair-wise alignment.

Table 3.5 Canonical sequence of Mouse Vmo1 protein

Signal peptide 1-21	'PTM / Processing' denotes the presence of an N-terminal signal peptide.	MELQAGARLLLLLGVMCYGHA
Chain 22 – 201	PTM / Processing' describes the extent of a polypeptide chain in the mature protein following processing	QIQVHVEPRYASIVDVTNGGTWGDWA WPEMCPDGYFASGFSVKVEPPQGIPGDD TALNGIRLHCTRGNSQKNTHVVESQSS WGSWSEPLWCPGTSFLVAFCLRVEPFTF PGDNTGVNNVRFCSDGVELEGPGLNWG DYGEWSNSCPKGVCGLQTKIQKPRGLR DDTALNDIRIFCCAS

3.5.3.1 Back translation of amino acid sequence of mouse Vmo1 to nucleotide sequence *Vmo1* using Back translation tool.

The back translated nucleic acid sequence from UniProt computes to 597 bp compared to 672 bp mRNA Ref sequence from GenBank (NM_001013607.1). This difference arises due to the difference in the databases, curation and annotations of sequences.

> Mouse *Vmo1*_FASTA format

ATGGAGCTGCAGGCCGGCGCCAGGCTGCTGCTGCTGCTGGGCGTGATGTGCT
ACGGCCACGCCCAGATCCAGGTGCACGTGGAGCCCAGGTACGCCAGCATCGT
GGACGTGACCAACGGCGGCACCTGGGGCGACTGGGCCTGGCCCGAGATGTGC
CCCGACGGCTACTTCGCCAGCGGCTTCAGCGTGAAGGTGGAGCCCCCCCAGG
GCATCCCCGGCGACGACACCGCCCTGAACGGCATCAGGCTGCACTGCACCAG
GGGCAACAGCCAGAAGAACACCCACGTGGTGGAGAGCCAGAGCAGCTGGGG
CAGCTGGAGCGAGCCCCTGTGGTGCCCCGGCACCAGCTTCCTGGTGGCCTTCT
GCCTGAGGGTGGAGCCCTTCACCTTCCCCGGCGACAACACCGGCGTGAACAA
CGTGAGGTTCTGCAGCGACGGCGTGGAGCTGGAGGGCCCCGGCCTGAACTGG
GGCGACTACGGCGAGTGGAGCAACAGCTGCCCCAAGGGCGTGTGCGGCCTGC
AGACCAAGATCCAGAAGCCCAGGGGCCTGAGGGACGACACCGCCCTGAACG
ACATCAGGATCTTCTGCTGCGCCAGC

3.5.3.2 Pairwise Sequence alignment of consensus plasmid sequence and mouse *Vmo1* sequence

The consensus plasmid (top sequence) (Table 3.1 A) was aligned with back translated mouse *Vmo1* (bottom sequence) in EMBL-EBI. The sequences show identity of 228/352 (64.8%), and Gaps 72/352 (20.5%).

```
115 ATGGAGCGGGGCGCAGGAGCCAAGCTGCTGCCGCTGCTG--------CTG    156
    |||||||.|...||.||.||||.|||||||||.|||||||        .||
  1 ATGGAGCTGCAGGCCGGCGCCAGGCTGCTGCTGCTGCTGGGCGTGATGTG     50

157 CTTCTGCGGGCGACTGGTTTCACATGTGCACAGA--CAGATG--------    196
    ||.|.||            |||    ||.||||  |||.||
 51 CTACGGC------------CAC----GCCCAGATCCAGGTGCACGTGGA     83

197 GCC--GGAACGGCTACACGGCGGTCATCGAAGTGACCAGCGGGGGTCCCT    244
    |||  ||.|||.|..||      ||.|.||.|||||||.|||.||..|||
 84 GCCCAGGGTACGCCAGCA------TCGTGGACGTGACCAACGGCGGCACCT    127

245 GGGGCGACTGGGCCTGGCCTGAGATGTGTCCCGATGGATTCTTCGCCAGC    294
    ||||||||||||||||||.||||||||.|||||.||.|.||||||||||
128 GGGGCGACTGGGCCTGGCCCGAGATGTGCCCCGACGGCTACTTCGCCAGC    177

295 GGGTTCTCGC-TCAAGGTGGAGCCTCCCCAAGGCATTCCTGGCGACGACA    343
    ||.||| .|| |.||||||||||.|||||.|||||.||.||||||||||
178 GGCTTC-AGCGTGAAGGTGGAGCCCCCCCCAGGGCATCCCCGGCGACGACA    226

344 CTGCACTGAATGGGATCAGGCTGCACTGCGCGCGCGGGAACGTCCTAGGC    393
    |.||.|||||.||.||||||||||||||.|..|.||.|||     ||.|
227 CCGCCCTGAACGGCATCAGGCTGCACTGCACCAGGGGCAAC-----AGCC    271

394 -----AATACGCACGT-----------------GCTGGGGCGAATGGAGT    421
        ||.||.|||||                 ||||||||...|||||.
272 AGAAGAACACCCACGTGGTGGAGAGCCAGAGCAGCTGGGGCAGCTGGAGC    321

422 GA    423
    ||
322 GA    323
```

Figure 3.15 Nucleotide sequence alignment of mouse *Vmo1* with the plasmid sequence

The vertical lines indicate identical nucleotides while the dots represent non-identical nucleotides. The dashes represent absence of nucleotides.

3.5.4 Multiple alignment of chicken, human, and mouse VMO1 protein sequences in UniProt using multiple sequence alignment tool.

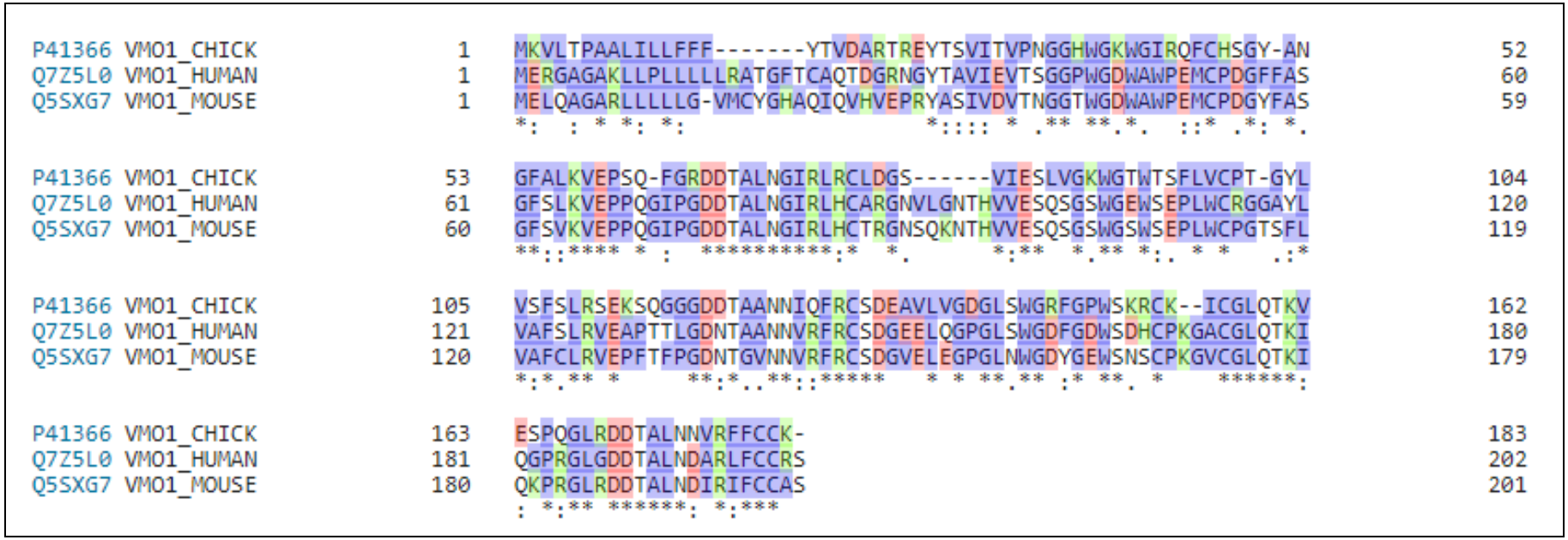


Figure 3.16 Alignment of Mouse, Human and Chicken VMO1 protein sequences.

The open reading frames of VMO1 for the mentioned organism shows more than 85% of similarity. The colour scheme represents the nature of the amino acids in the sequences as Hydrophobic (purple), Negative (Pink), Positive (Green) charges. (*) shows similar amino acids while the (:) indicates a substitute amino acid at a particular position in the sequence.

3.5.5 Multiple alignment of chicken, human, mouse, and consensus VMO1 protein sequences in UniProt using multiple sequence alignment tool.

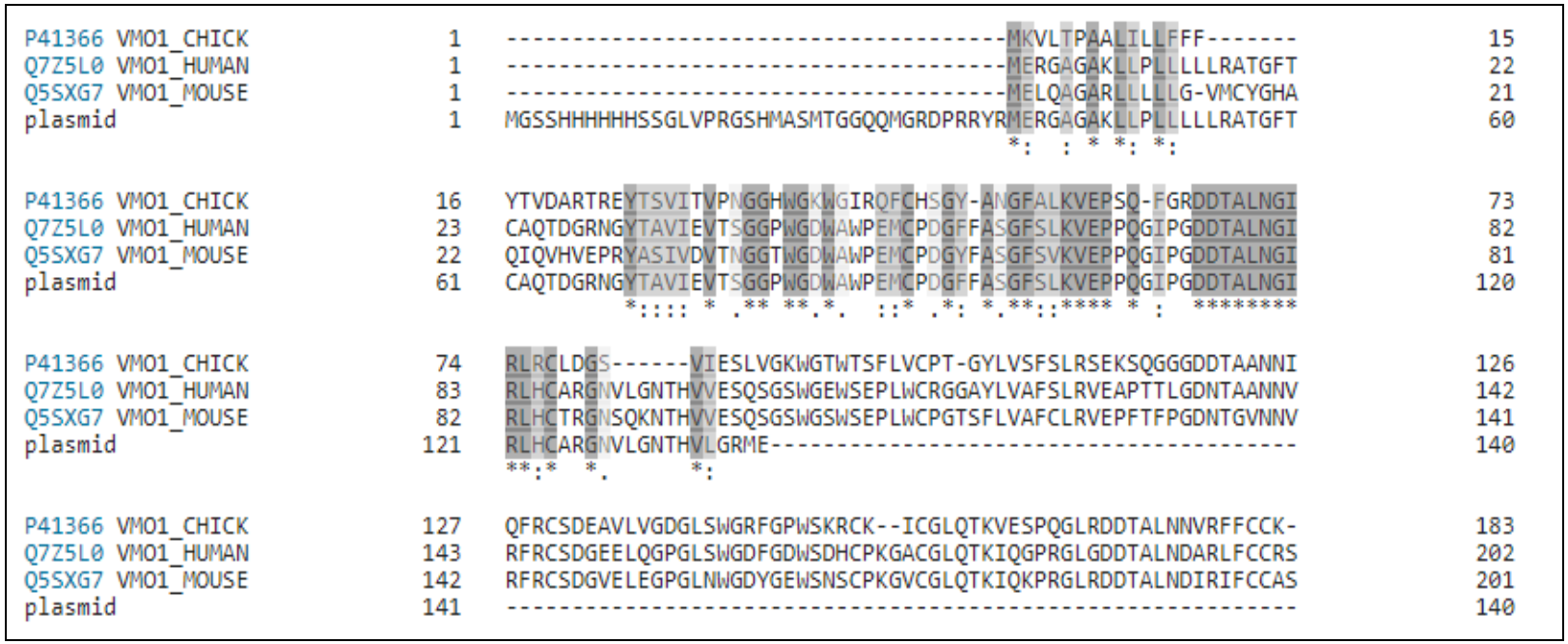


Figure 3.17 Alignment of Chicken, Human and Mouse VMO1 protein sequences with the consensus plasmid sequence.

The open reading frames of VMO1 for the mentioned organism shows and that of the plasmid sequence shows more than 75% of similarity. (*) shows similar amino acids while the (:) indicates a substitute amino acid at a particular position in the sequence.

3.6 Taxonomic tree for comparative analysis

In order to study the VMO1 gene conservation, multiple sequence alignment was carried out in UniProt that generated a taxonomic tree. Human VMO1 protein appear to be a sister clade with gorilla VMO1, as is for mouse Vmo1 and rat VMO1 with cat VMO1. Chicken VMO1 is closely related to anoles and together they form sister clade to rat, cat, gorilla and human VMO1.

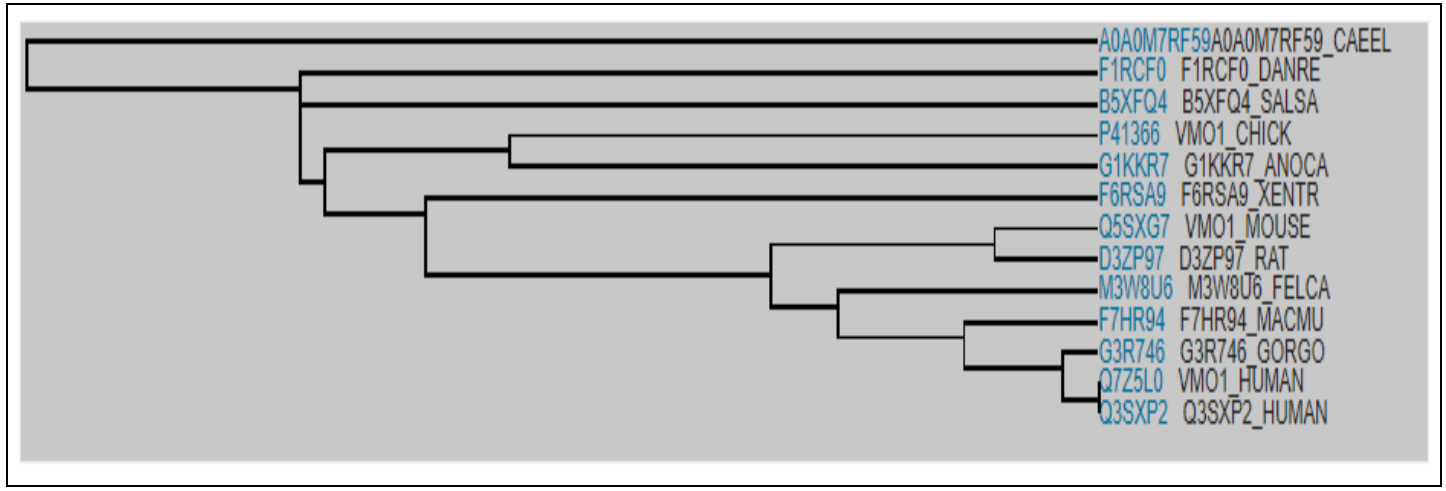


Figure 3.18 Phylogenetic tree showing VMO1 evolution.

Organisms aligned on the right are as follows CAEEL: *C. elegans* (round worm); DANRE: *Danio rerio* (Zebrafish); SALSA: *Salmo salar* (Atlantic salmon); CHICK: Chicken; ANOCO: *Anolis carolinensis* (Green anole); XENTRE: *Xenopus tropicalis* (Western clawed frog), Mouse, Rat, FELCA: *Felis catus* (Cat); MACMU: *Macaca mulatta* (Rhesus macaque); GORGO: *Gorilla* (Western lowland gorilla); Human isoform 1 and isoform 2.

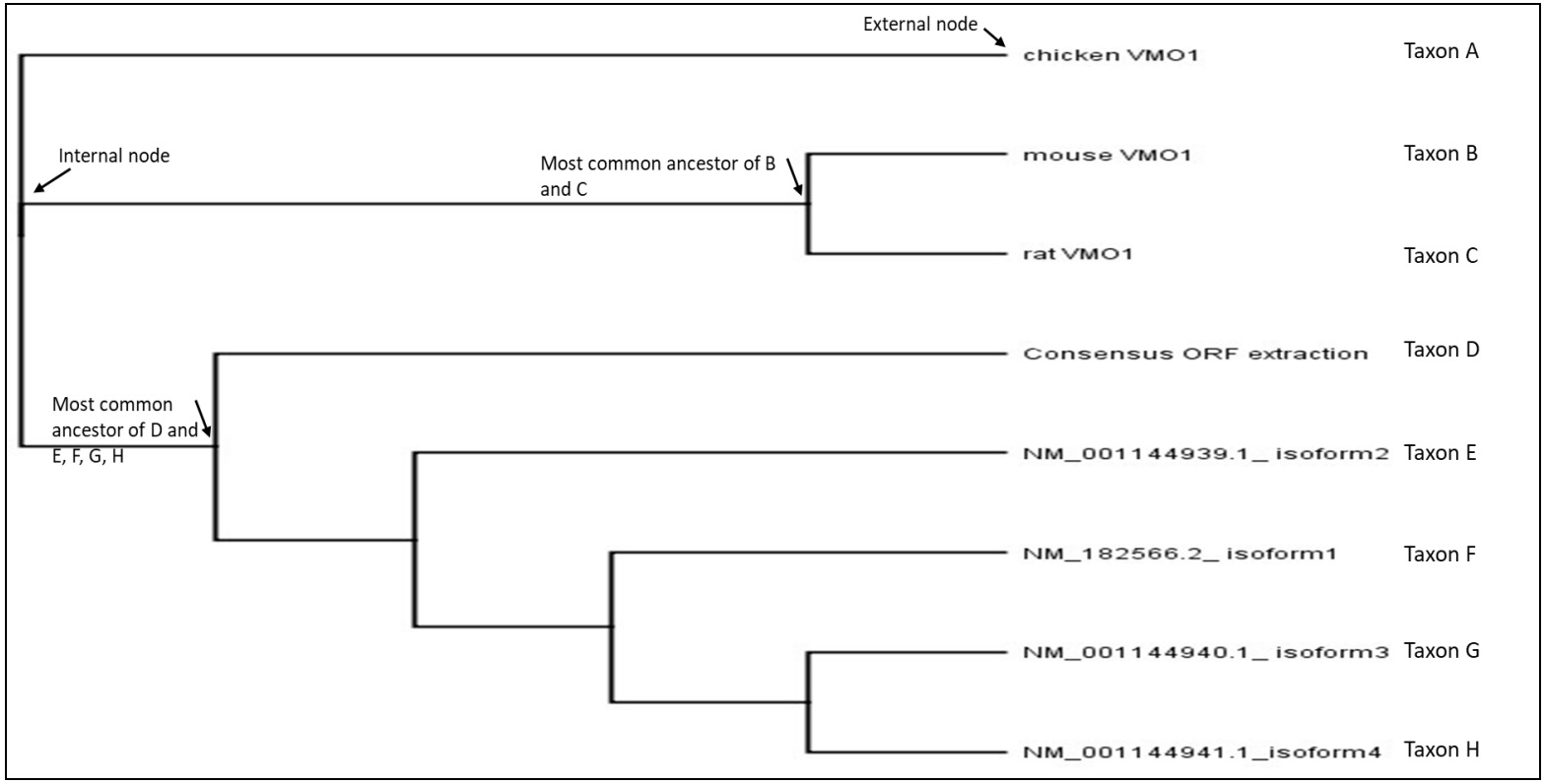


Figure 3.19 Phylogenetic tree of chicken, mouse, rat, consensus plasmid, and human *VMO1*

The polyphyletic tree was generated in Geneious Tree Builder using Global alignment with free end gaps. The Refseq was from NCBI for the organisms mentioned on the right-hand side of the figure. The Genetic Distance Model was Tamura-Nei and the tree was built using Neighbor-Joining method. This tree has 8 taxa's and 8 tips without an out group with 14 nodes. The Cost Matrix setting was set to 93% similarity.

3.7 VMO1 analysis in *E. coli*

The main aim of analysing VMO1-like proteins was to determine if similar proteins were being produced by the bacteria. Because multiple bands were evident on 12% SDS-PAGE gels when whole cell lysates were resolved and proteins were detected by Western blot, we wanted to confirm the possibility if they came from the proteins produced by *E. coli* (E.g. Figure 4.5, Figure 4.9). When the consensus amino acid sequence of the plasmid was blasted in NCBI and the search was restricted to *Escherichia coli* (taxid: 562), the following output was generated, shown in Table 3.6.

Table 3.6 List of *E. coli* strains producing VMO1-like proteins

Sr. No	Description of *E. coli* proteins	Total score Z value	Query coverage (%)	Identity (%)	Accession Number
1	GTPase Bipa/typa	48.1	22	100	4ZCI A
2	GTPase Bipa/typa Complexed with Ppgpp	48.1	22	100	4ZCM A
3	C-terminal Fragment of Bipa/typa	45.4	22	100	AZCK A
4	Ycfd	43.1	27	79	4LIT A
5	Methyltransferase Rsmh	35.0	15	76	3TKA A
6	BamA Potra 4-5	34.7	15	76	3Q6B A
7	Prib	33.5	15	71	1V1Q A

3.7.1 Multiple alignment of consensus plasmid sequence with *E. coli* VMO1 like proteins

The consensus plasmid sequence was blasted in NCBI, using the blastP suite with the amino acid sequences of all the strains mentioned in the Table 3.6, in the same order (GTPase, Bipa/ typa, GTPase Bipa/ typa Complexed with Ppgpp, C-terminal Fragment of Bipa/typa, Ycfd, Methyltransferase Rsmh, Bama Potra 4-5, Prib).

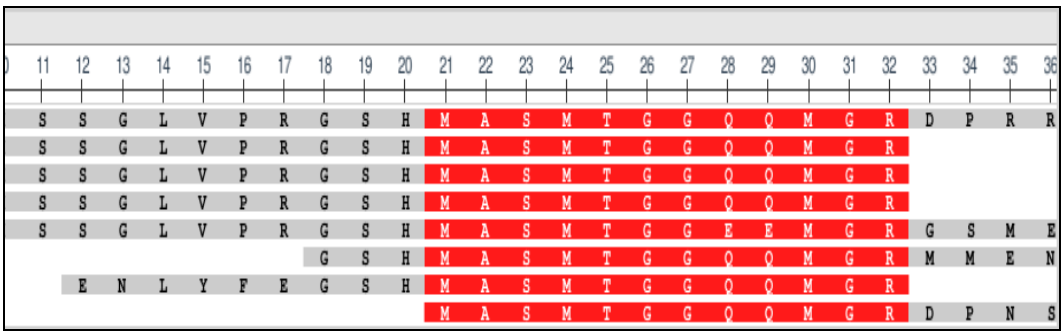


Figure 3.20 Multiple pair-wise alignment of consensus plasmid sequence and *E. coli* strains.

The top sequence represents the sequenced plasmid (predicted translated amino acid from nucleotide sequence in Geneious) followed by the same order of strains having specific genes mentioned in the table 3.6. The red colour indicates identity between the sequences, from amino acids 21 to 32.

3.7.2 Pair-wise alignment of consensus plasmid with one protein.

As an example, Chain A, Crystal Structure of *Escherichia coli* GTPase Bipa/typa (Table 3.6, Sr. No 1) was aligned with consensus plasmid.

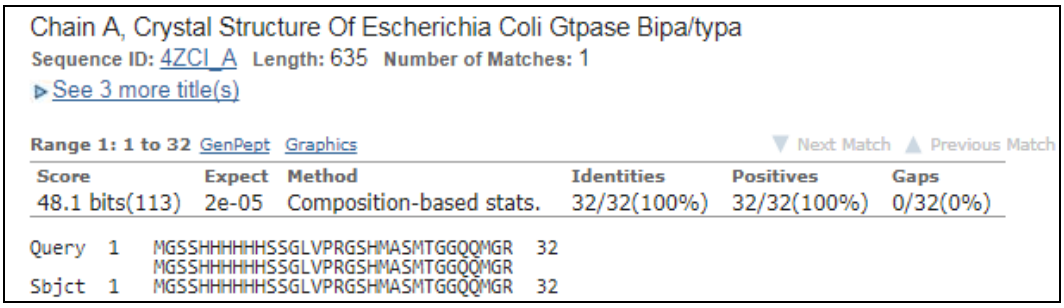


Figure 3.21 Alignment of plasmid sequence with Chain A, Crystal Structure of Escherichia coli GTPase Bipa/typa.

In conclusion, this data shows that *E. coli* does not produce a full length VMO1 protein and thus you should not expect any background signal using the VMO1 antibody (assuming it is specific to its antigen and no bacterial mutations) upon western blot analysis.

3.8 DNA fold and 3D predicted structure

In order to observe the DNA fold symmetry (Figure 3.22), the consensus sequence was selected with the following options on Geneious 7.1.8 :- Compute Options: Partition Function, Avoid isolated base pairs, Assume circular molecule; Dangling

Ends: Coaxial stacking, Energy Model: DNA (Matthews 1999), Temperature : 37°C. However, there is limited information on how this model gets generated. On the other hand, an online tool helped generate the 3D structure of the 15 kDa recombinant protein, by creating an alignment with chicken *VMO1* in Pymol. Our structure shows one and half of the characteristic Greek key motif coloured in red and cyan in Figure 3.23.

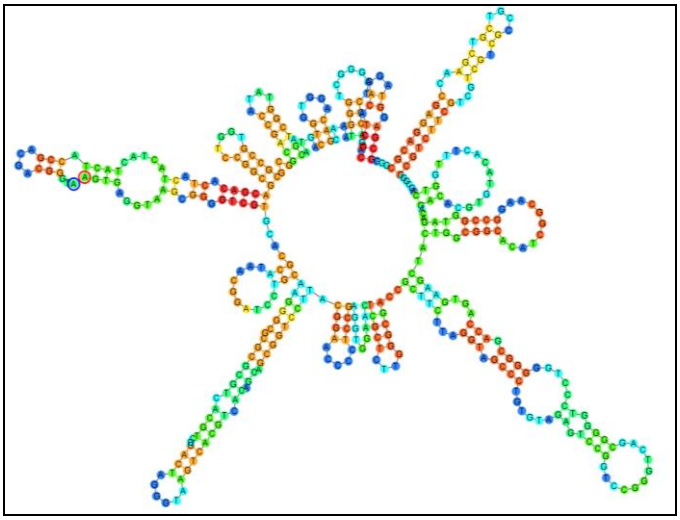


Figure 3.22 DNA fold of consensus sequence.

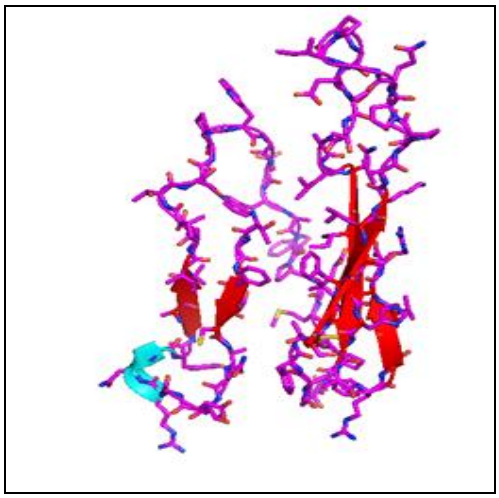


Figure 3.23 Predicted 3D structure in Pymol.

The red colour indicates the Greek key fold in a helix: loop cartoon structure developed in Pymol.

CHAPTER FOUR

RESULTS

This chapter outline the outcomes while assessing our four objectives: of understanding *VMO1* gene, inducing human recombinant VMO1 protein in bacterial system, purifying the recombinant protein, western blot detection and validating the VMO1 antibody.

4.1 Culturing transformed cells

4.1.1 Bacterial revival and growth of recombinant pET28b (+) construct

Glycerol stocks of *E. coli* Rosetta ™ (DE3) pET28b (+) pLysS and BL21 (DE3) pET28b (+) *VMO1* were streaked onto LB agar plates supplemented with antibiotics and incubated at 37°C O/N. Pinpoint, translucent colonies ranging from 1-2 mm (by scale) were seen on the agar plates (Figure 4.1). A colony was selected and used to inoculate media for extracting the plasmid DNA to confirm that the identity of the recombinant *VMO1* construct prior to further studies.

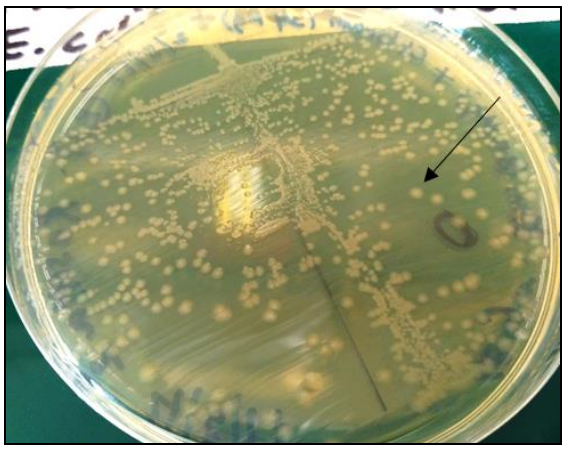


Figure 4.1 *E. coli Rosetta* ™ (DE3) pET28b (+) pLysS colonies on LB agar supplemented with Kanamycin and Chloramphenicol.

The arrow shows a representative 1-2mm sized translucent, buffy colony that was selected for bacterial growth.

4.2 Evaluation of extracted plasmid DNA constructs

The plasmid DNA was extracted through a series of lysis and purification steps using a commercial Zyppy™ kit. Briefly, the culture was initially treated with Lysis Buffer, and a blue coloration developed showing complete lysis. This was followed by addition of a yellow coloured, cold Neutralization Buffer. The solution turned yellow indicating complete neutralization, giving a yellowish precipitate of *E. coli* chromosomal DNA. The solution was then spun, and the supernatant was allowed to pass through a spin column. The sample collected on the matrix of the column was rinsed with Wash Buffer. Finally, the plasmid was eluted from the column using Elution Buffer and then subjected to a spectrophotometric analysis using Thermo Scientific™ NanoDrop 2000 spectrophotometer, PCR, Restriction digest and DNA Sequencing.

Nucleic acids and proteins have maximum absorbances at 260 and 280 nm, respectively. For DNA, 260/280 ratio of ~1.8 is considered pure. Also, absorbance at 230 nm represents contaminants; and therefore, the ratio of A260/ A230 is calculated which must be in the range of 2.0-2.2 (Matlock, 2018). In Table 4.1, the concentration of plasmid DNA was found to be higher for Rosetta than BL21. The 260/280 ratio was found to be 1.7 for both the strains, with some low impurities absorbing at 260/230 (table 4.1). These results demonstrate that we have isolated DNA and that the quality and yield is still sufficient for downstream applications.

Table 4.1 Determination of pET28b (+) nucleic acid concentration and quality

	Bacterial Strain	**Concentration (ng/µL)**	**Absorbance (nm)**			
	Plasmid DNA		260	280	260/280	260/230
1	Rosetta	113.4	2.269	1.302	1.74	1.29
2	BL21	76.8	1.537	0.901	1.71	1.20

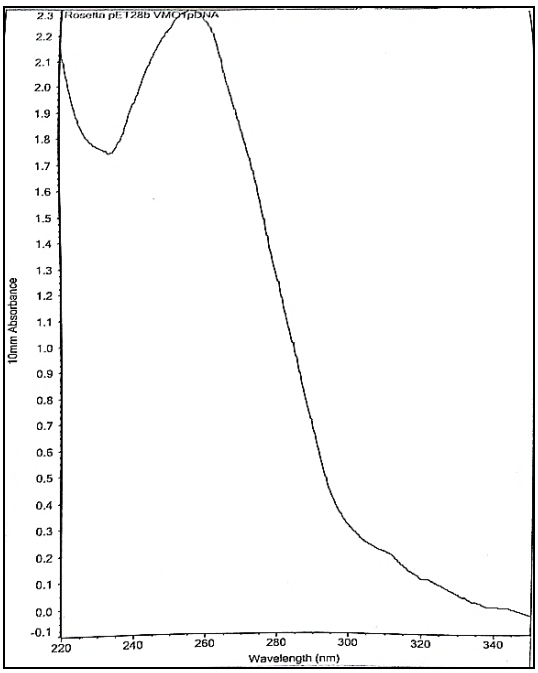


Figure 4.2 Absorbance spectra of plasmid DNA extracted from a single Rosetta colony.

The graph shows two peaks, one at 260 nm and a small inverted peak at 230 nm for the plasmid DNA extracted from Rosetta strain.

Additionally, with the absorbance readings, the software also predicted their corresponding peaks. A good peak was seen for the extracted DNA at 260 nm as expected (Figure 4.2). Since the concentration of BL21 was found to be low and the fact that Rosetta supported additional codons; all the experiments were conducted with *E. coli* Rosetta ™ (DE3) pET28b (+) pLysS.

4.3 Confirmation of *VMO1* gene insert

4.3.1 Confirmation by DNA sequencing

The plasmid was sequenced using four primers yielded four sequencing reads. These were made into one noise free, consensus sequence of 432 nucleotides in Geneious (section 2.3.1). This sequence was also confirmed of being *VMO1* isoform 3 (Section 3.4). Therefore, the next step was to initiate recombinant protein expression by culturing the bacteria.

4.3.2 Restriction digest of pET28b (+).

The Circular, pET28b (+) plasmid DNA (5 µL from total DNA of 113 ng/ml) is 5368 bp and contains a gene insert of either 765 bp (isoform 3) or 785 bp (isoform 1) (Hardie, 2015). It was extracted from Rosetta strain and digested with two restriction enzymes, EcoR1 and BamH1. The restriction digests (RD) were performed as single and double digests. Here, BamHI cuts from 198-203 bp while EcoR1 cuts from 192-197 bp (addgene, 2018a). Four samples (RD 1-4) were set for digestion (Section 2.3.2) and then resolved on 1% agarose gel.

The resulting separation on the gel (Figure 4.3). was viewed in the imager by auto exposure settings for EtBr under U.V light. The 1 Kb+ ladder separated from 12,000 bp to 500 bp, along with a 1500 bp fragment which had increased intensity that served as reference band. The samples resolved from 12,000 bp to 1500 bp Sample RD 1 (lane 2) did not have any enzyme, as a result its corresponding lane shows separation of the plasmid DNA into three visible bands that are three forms of plasmid DNA; Supercoiled, Relaxed Circle and Full-length Linear (Simple Cloning Lab, 2018). This also represents an uncut DNA and served as our negative control. When the plasmid was digested with either EcoR1 or BamH1 (RD 2-lane 3 and RD 3-lane 4, respectively), a single band was seen at ~5000 bp. This represents complete digestion of the plasmid and formation of linearized conformation. It also reveals the full size of the plasmid. RD3 (lane 4) was however seen as an incomplete digested product, evident by another band at ~10,000. RD 4 (lane 5) was double digested and as a result three bands were seen at 5000bp, 1500 bp and ~850 bp.

Based on these results we next amplified the plasmid DNA to obtain the gene of interest (isoform 3 or isoform 1) by SH3F and SH3R primers using PCR.

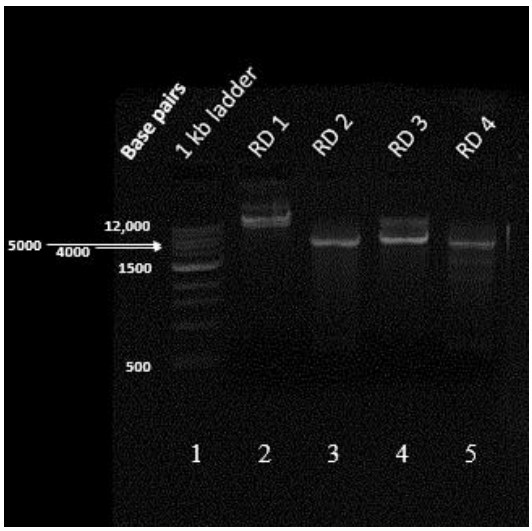


Figure 4.3 Gel Electrophoresis of Restriction digest.

Restriction digested (RD) samples (5µL) of plasmid DNA from a single Rosetta colony were loaded onto 1% agarose gel containing EtBr. Gel representation from left to right, Lane1: 1Kb plus ladder; Lane 2: RD 1 treated with no restriction enzyme; Lane 3: RD 2 treated with EcoRI only; Lane 4: RD 3 treated with BamH1 only and, Lane 5: RD 4 treated with both EcoR1 and BamH1.

4.3.3 PCR

Following restriction digestion, further confirmation on gene size was required. Thus, Plasmid DNA extracted from two bacterial strains, BL21 colonies (B1 and B2), and Rosetta colonies (R1 and2) was used. The aim was to amplify the *VMO1* gene (isoform 3 or isoform 1) (Section 2.3.3) expected to be ~438 bp in size and perform electrophoresis on 1% agarose gel. The resulting gel image (Figure 4.4) was developed in the imager with the same settings as for Restriction digest. The PCR ladder (lane 6) separated into 7 bands from 2000 bp to 100 bp. All PCR products were sized between 300 and 500 bp i.e., 400 bp, when compared to the PCR ladder. Absence of band in the negative control (lane 5) indicates that the PCR reaction set up, was indeed free from external DNA interference or contamination.

Since the resultant bands were observed around 400 bp, it was concluded that the primers successfully amplified the *VMO1* gene. The plasmid was then sent off for sequencing and bioinformatics was used to determine the presence of human *VMO1* isoform in it (Section 3.4, Section 3.4.1).

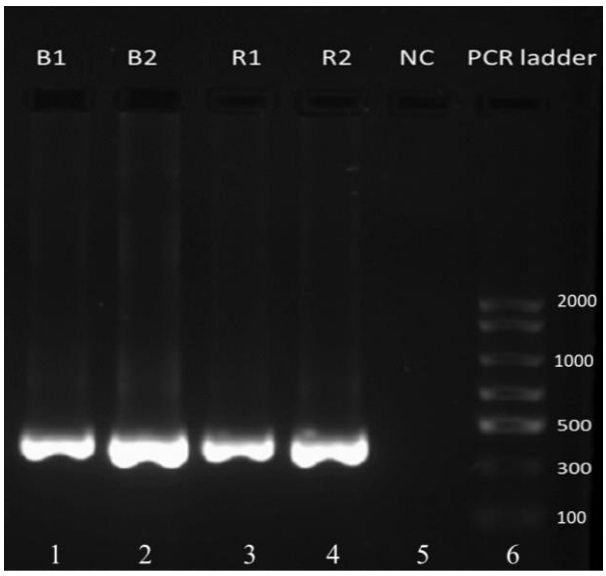


Figure 4.4 PCR analysis.

PCR samples (10 µL) were loaded onto 1% agarose gel containing EtBr. Gel representation from left to right, Lane 1: PCR product from BL21 colony 1; Lane 2: PCR product from BL21 colony 2; Lane 3: PCR product from Rosetta colony 1; Lane 4: PCR product from Rosetta colony 1; Lane 5: PCR ladder.

4.3.4 Plasmid stability test

Before inducing cells to produce any protein, it is best to test whether the bacterial cells still carry the plasmid harbouring the gene of interest. This can be done using a conventional plate method, which allows growth of the bacterial cells on a selective medium.

An O/N uninduced starter culture of *E. coli* Rosetta PLysS was serially diluted up to 10^6 and streaked onto four agar plates that differ in their antibiotics and IPTG concentration. All viable cells without the plasmid were seen growing on LB agar plates (Table 4.2, Sr. No 1), while those that retained the plasmid were seen on LB+antibiotics agar plates (Table 4.2, Sr. no 2). A small fraction of the cells also represented amongst those who lost the plasmid but were still able to express the *VMO1* gene, observed on LB+IPTG agar plate (Table 4.2, Sr. no 4). On the contrary, there was no evidence of mutant cells that retained the plasmid but lost the ability to express *VMO1* gene; as seen on LB+IPTG+antibiotics agar plates (Table 4.2, Sr. no 3).

Table 4.2 Plasmid stability test

Sr. No	Media	Dilution	Colony forming unit (cfu)	Dilution factor	Total cfu	Total average cfu
1	**LB agar without antibiotics**					
		10^4	11	0.1	$11x10^3$	$3.5x10^3$
		10^5	5	0.1	$5x10^4$	
		10^6	No growth	-	-	
2	**LB agar + Chloramphenicol + Kanamycin**					
		10^4	11	0.1	$11x10^3$	$3.5x10^3$
		10^5	5	0.1	$5x10^4$	
		10^6	No growth	-	-	
3	**LB agar + IPTG+ Chloramphenicol + Kanamycin**					
		10^4	No growth	-	-	-
		10^5	No growth	-	-	
		10^6	No growth	-	-	
4	**LB agar + IPTG**					
		10^4	2	0.1	$2x10^3$	Too less to consider
		10^5	1	0.1	$1x10^4$	
		10^6	No growth	-	-	

It can thus be inferred from our observations, that the pET28b (+) vector harbouring *VMO1* gene was stable. The next step was to induce protein expression using two methods; traditional and ethanol-based.

4.4 Recombinant protein expression using Traditional Method

4.4.1 Induction of λDE3 Lysogens with IPTG to determine total whole cell protein using a small-scale time course.

Rosetta *E. coli* pET28 (+) PLysS was cultured in LB medium having chloramphenicol and kanamycin in duplicates, one set incorporated 0.4 mM IPTG and another with 1 mM IPTG. This was done to determine what concentration of IPTG would yield best cellular expression, which in turn would lead to higher protein expression. Induced cell samples were collected at various timepoints; 0, 3 6, and 24 hours, respectively. Additionally, an uninduced or control was set up

that did not contain any IPTG. Whole cells were harvested for every sample and then the concentration was measure using the Bradford assay. Later the whole cells were either resolved on 12% SDS-PAGE gels and stained to veiw separated proteins or sonicated and then resolved on the gels.

4.4.2 Bradford assay

Protein standards were set up using BSA and Dye reagent (Section 2.5) ranging from 0 mg/ml to 10 mg/ml, into a 96 well plate. The concentrations were estimated visually (by colour change) and/or using the Multiskan™ GO Microplate Spectrophotometer.

4.4.2.1 Spectrophotometric estimation.

On visual assessment, the BSA standards showed a range of colour changes from pale orange, pale blue, moderate blue, and to a darker blue (Figure not shown, refer Table 4.3). The whole cells were then compared and matched with the colours from BSA standards. The samples appeared as moderate to dark blue in colour indicating the protein range to be between 4 to 8 mg/ml, respectively (Table 4.3 Sr No 7 to 9).

Table 4.3 Visual estimation of protein standards.

Sr No	BSA mg/ml	Coloured reaction
1	0.00 BLANK	Pale orange (+)
2	0.1	Pale orange (++)
3	0.25	Pale orange (+++)
4	0.5	Pale orange (++++)
5	1	Pale Blue (+)
6	2	Pale Blue (++)
7	4	Moderate Blue
8	6	Dark Blue (+)
9	8	Dark Blue (++)
10	10	Dark Blue (+++)

On spectrophotometric assessment, the standards showed a gradual increase in absorbance (Table 4.4), which indicates that the concentration of BSA is directly proportional to the concentration of the dye reagent. The highest concentration of samples observed with using 0.4 mM IPTG was 7.5 mg/mL after 6 hrs and with

1mM IPTG was 6 mg/mL at 0 hrs (Table 4.5, Sr no 3 and 5). It is to be noted that the Bradford assay only helps to estimate total proteins in the cells and not recombinant proteins.

Table 4.4 Absorbance for BSA standards

Sr No	**Particulars**	**Original [Abs]**	**Fitted [Abs]**
	BSA mg/ml		
1	0.00 BLANK	-	-
2	0.1	0.150	0.155
3	0.25	0.143	0.159
4	0.5	0.155	0.165
5	1	0.176	0.177
6	2	0.215	0.200
7	4	0.265	0.248
8	6	0.317	0.296
9	8	0.344	0.344
10	10	0.372	0.392

Table 4.5 Absorbance for whole cell protein lysates.

Sr No	**Particulars**	**Original [Abs]**	**Fitted [Conc]**
	Protein samples mg/ml		
1	0.4mM/0 hour	0.292	5.79
2	0.4mM/3 hour	0.287	5.76
3	**0.4mM/6 hour**	**0.333**	**7.54**
4	0.4mM/24 hour	0.297	6.02
5	**1 mM/0 hour**	**0.297**	**6.02**
6	1 mM/3 hour	0.286	5.57
7	1 mM/6 hour	0.294	5.91
8	1 mM/24 hour	0.289	5.70

Following the Bradford assy, the whole cell lysates collected at various time points were loaded onto 12% SDS-PAGE gel and stained with Coomassie blue stain (Section 2.6.2). The resulting bands were compared with the PAGE master laddder to determine the size of proteins. This ladder separated into 9 bands, while 50 kDa and 15 kDa bands were pre-stained. The recombinant VMO1 protein size was expected to be at 15 kDa from the sequencing data (Figure 3.8). So any band around this size was signifiant.

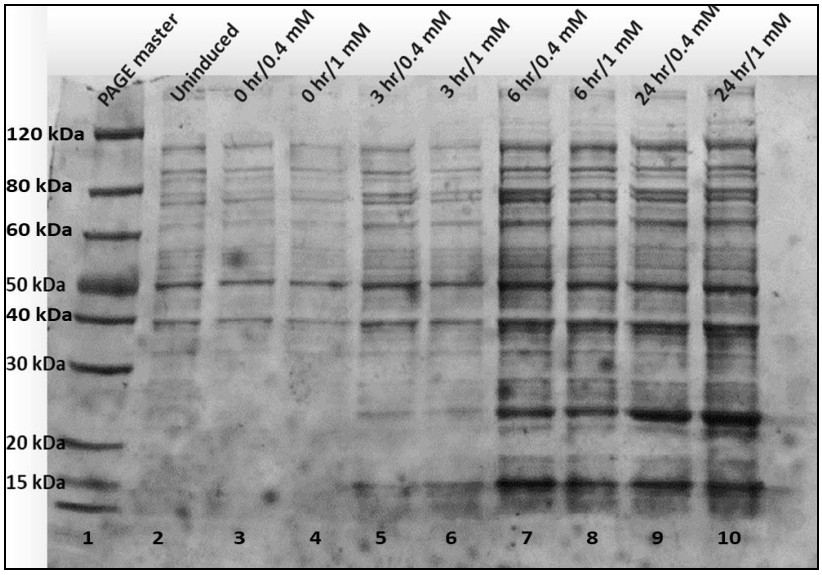


Figure 4.5 Protein induction of *E. coli* Rosetta pLysS in LB broth using Traditional method.

Samples were resolved on 12% commercial SDS-PAGE gel and stained with Coomassie blue. Two final concentrations of IPTG, 0.4 mM and 1.0 mM were used along with an uninduced control. Gel representation from left to right, Lane 1: PAGE master ladder; Lane 2: uninduced control (no IPTG); Lane 3: 0-hour/0.4 mM; Lane 4: 0-hour/1 mM; Lane 5: 3-hours/0.4 mM; Lane 6: 3-hours/1 mM; lane 7: 6-hours/0.4 mM; Lane 8: 6-hours/1 mM; Lane 9: 24-hours/0.4 mM; Lane 10: 24-hours/1 mM.

Figure 4.5 show the total separated proteins expressed in *E. coli* Rosetta cells from 120 kDa to 15 kDa. We could also see an increased protein expression with time from 0 hr to 24 hrs (e.g., lanes 3, 5, 7, and 10). The highest expression for the expected recombinant protein size was observed as a darker band at 15 kDa with 1 mM IPTG after 24 hours (lane 10) compared to 0.4 mM IPTG (lane 9).

4.4.3 Determination if recombinant protein is soluble or insoluble

To determine the fate of the recombinant protein, either being soluble (i.e., present in the supernatant) or insoluble (present in the pellet), the whole cells (24 hours/1 mM IPTG) were sonicated (Section 2.4.2). The samples were then centrifuged, denatured, and separated on 12% SDS-PAGE gel.

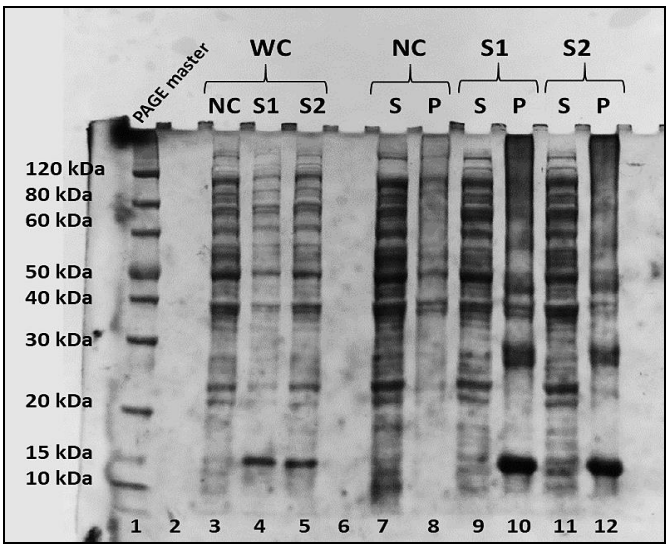


Figure 4.6 Determination of the solubility of the recombinant protein in LB broth after 24hr induction with 1 mM IPTG.

Samples were resolved on 12% commercial SDS-PAGE gel and stained with Coomassie blue. The PAGE gel represents a comparison between proteins expressed in the un-sonicated whole cells (lanes 3, 4, and 5); and sonicated whole cells (lanes 7, 8, 9, 10, 11, and 12). Gel representation from left to right, NC: negative control (no IPTG); S1 and S2: two bacterial colonies from the same agar plate that were used for protein induction experiments; S and P represents supernatant and pellet, respectively.

In Figure 4.6, lanes 4 and 5 show that prior to sonication, a 15 kDa recombinant protein is expressed in both the colony replicates of S1 and S2. This band is absent in the NC, suggesting that the recombinant protein is induced in the presence of IPTG. After sonication, a highly intense band was observed once again at 15 kDa but in the pellet (Figure 4.6, lane 10 and 12). This indicates that the recombinant protein is insoluble in nature. Interestingly, higher bands were also observed from 120 to 30 kDa.

Faint bands were also seen in the supernatant at 15 kDa. So, to confirm that the recombinant protein was indeed VMO1, which was in frame with a reporter tag, a western blot detection was conducted using a His-tag and VMO1 antibody. Protein transfer was carried out on a pre-activated PVDF membrane (Section 2.9). Ponceau staining was performed on the membrane (Section 2.9.1) to confirm protein transfer. Figure 4.7 show reddish-pink coloured bands ranging from 120 to 15 kDa in size. Thus, indicating a successful protein transfer.

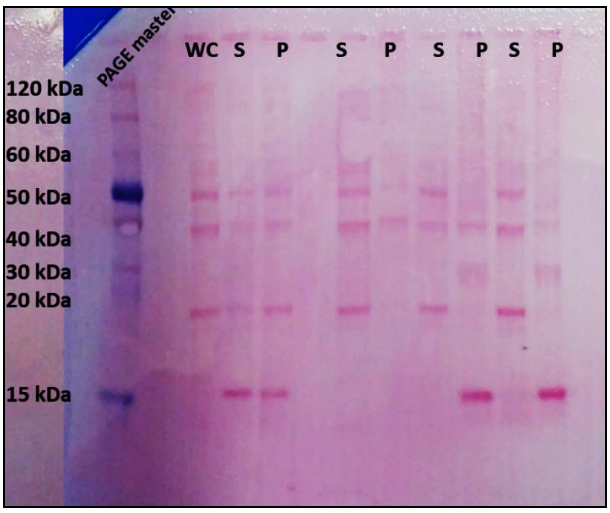


Figure 4.7 Ponceau staining of PVDF membrane.

Pink colour indicate stained protein bands.

The membrane was de-stained (Section 2.9.2), washed, blocked O/N, and incubated with His-tag antibody for western blot detection (Section 2.9). The His-tag antibody detected 15 kDa bands seen as bright bands only in the pellet (Figure 4.8, lanes 10 and 12) when compared to the supernatant (Figure 4.8, lanes 9 and 11). Other high molecular weight proteins of 60, 40, and 30 kDa were also detected. Thus, it can be inferred that the recombinant proteins are expressed as intracellular, insoluble proteins containing the Histidine tag.

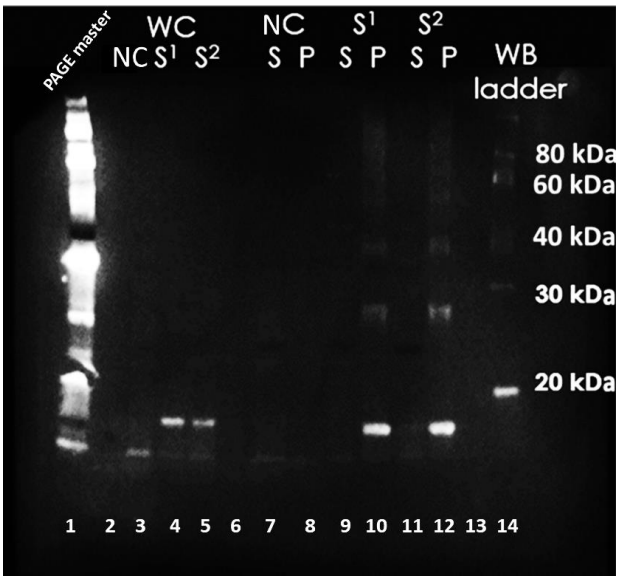


Figure 4.8 Western blot analysis of recombinant protein using His-tag antibody.

The blot represents a comparative His-tag antibody (1:1000) detection between proteins expressed in the un-sonicated whole cells (lanes 3, 4, and 5); and sonicated whole cells (lanes 7, 8, 9, 10, 11, and 12). Gel representation from left

to right, NC: negative control (no IPTG); S1 and S2: two bacterial colonies from the same agar plate that were used for protein induction experiments; S and P represents supernatant and pellet, respectively.

To confirm that the recombinant protein was expressed from *VMO1*, a commercial VMO1 primary antibody (Abcam 126510) was used for western blot detection. The membrane was stripped and reprobed (Section 2.9.3) with this antibody. The same resultant bands were also detected by the VMO1 antibody (Figure 4.9) following chemiluminescence.

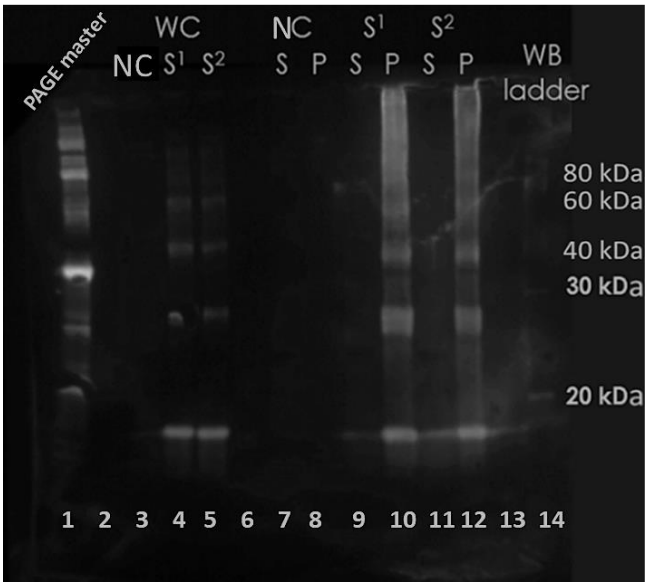


Figure 4.9 Western blot analysis of recombinant protein using VMO1 antibody.
The blot represents a comparative VMO1 (1:1000) antibody detection between whole cell proteins (WC - lanes 1, 2, 3, 4, and 5); Lanes 6 and 7: negative control (C, without IPTG); and sonicated samples (lanes 9, 10, 11, and 12). S1 and S2 represent two bacterial colonies from the same agar plate that were used for protein induction experiments. S and P represents supernatant and pellet, respectively.

Thus, it can now be inferred that the recombinant protein expressed by *E. coli* Rosetta ™ (DE3) pET28b (+) pLysS having *VMO1* gene, expressed a recombinant VMO1 protein having a His-tag when grown in presence of 1mM IPTG after 37°C.

4.5 Recombinant protein expression using a new method

To further enhance protein expression and solubility, varying concentrations of ethanol (1-3%) and induction time (5hrs and 24 hrs) were used as recommended by Gaurav et al, 2015 (Section 2.7). Here, time and temperature dependent analysis were carried out for optimization. Later, LB broth was switched to TB broth for further optimisation. All the samples were collected and processed simultaneously.

4.5.1 Time dependent comparative analysis

Overnight grown starter culture of Rosetta *E. coli* pET28 (+) pLys was grown in LB having chloramphenicol and kanamycin and 1 mM IPTG. Under the influence of IPTG, the culture was grown until the O. D_{600} reached 0.4-0.6 (LB broth). Spectrophotometric estimation was used to check the O.D at 600 nm. Next, absolute ethanol was added in the final concentrations of 1, 2, and 3%, respectively. The cells were grown at 37°C under shaker incubation and the samples were collected after 5 hours and 24 hours. The samples were reserved as whole cell or sonicated fractions for PAGE analysis or western blot analysis.

Figure 4.10 show separation of whole cell proteins after using ethanol. It can be seen that 1% ethanol (lane 13) gave the highest expression of *E. coli* Rosetta after 24 hours in LB broth compared to 5 hours. A brighter band at 15kDa can be seen. Both NC and PC worked successfully for IPTG and ethanol, for e.g., NC (lanes 9 and 11) showing no band due to absence of IPTG and ethanol respectively, while PC (lanes 10 and 12) shows a less intense band in presence of IPTG and ethanol respectively. The same range of protein separation can be observed as before (Figure 4.5) on introducing ethanol.

The next step was to sonicate the whole cell samples reserved at -20°C in order to observe protein solubility either in the supernatant (soluble fraction) or the pellet (insoluble fraction). This was later found to be in the pellet. and observed for all concentrations of ethanol. However, a high intense band was only seen for 1% ethanol after 24 hours at 15 kDa (Figure 4.11, lane 11). Apart from that, very low

intensity bands were also observed in the supernatant (Figure 4.11, lanes 5,8, and 12).

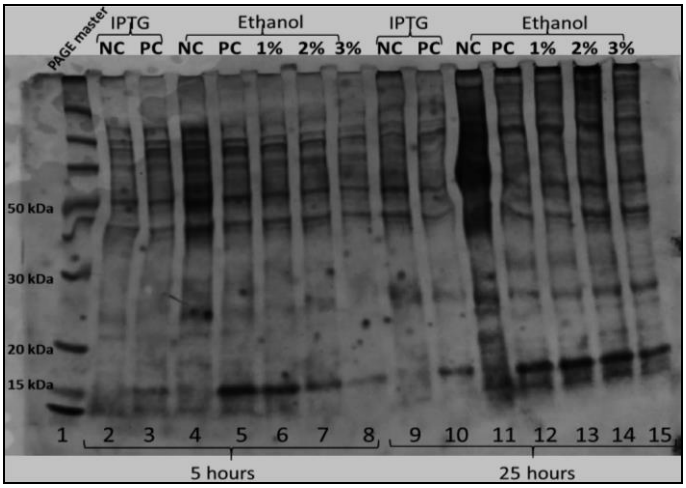


Figure 4.10 Time dependent comparative analysis of protein expression using 1%, 2%, 3% ethanol in LB broth, to analyse whole cell proteins.

Proteins from whole cells were resolved on 12% commercial SDS-PAGE gel, stained with Coomassie blue, and observed under the imager. NC and PC represent negative control and positive control, respectively. Gel representation from left to right, PAGE ladder: lane 1; samples collected after 5 hours (lanes 2, 3, 4, 5, 6, 7, and 8) and samples collected after 24 hours (lanes 9, 10, 11, 12, 13, 14, and 15).

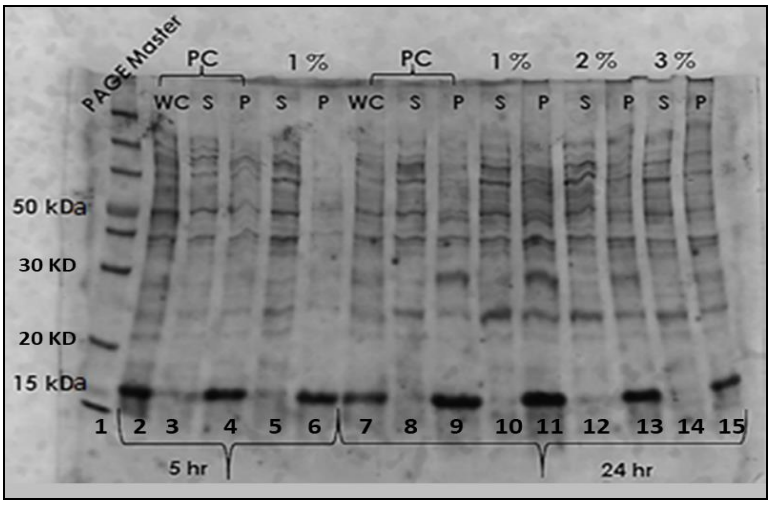


Figure 4.11 Time dependent comparative analysis of protein expression using 1%, 2%, 3% ethanol in LB broth, to analyse recombinant protein.

Proteins obtained from whole cells and after sonication were run on 12% commercial SDS-PAGE gel and stained with Coomassie blue. The gel was observed under the imager. soluble (S: Supernatant) and insoluble (P: Pellet) portions of the samples Here, PC was used alone.

4.5.2 Temperature dependent comparative analysis

To observe for any further solubility of the protein or allow soluble recombinant protein production at lower temperatures, the experiment was run at 22°C as per Section 2.7. However, the bacterial proteins remained to be highly expressed at 37°C in the pellet (Figure 4.12 A, lane 8) rather than the supernatant (Figure 4.12 B, lane 9).

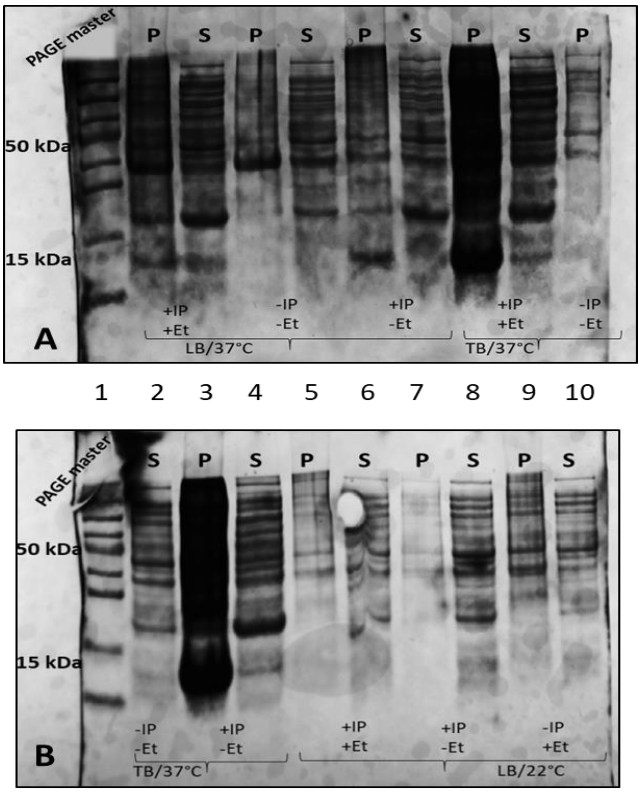


Figure 4.12 Temperature and media dependent comparative analysis of protein solubility using 1%, 2%, 3% ethanol in LB and TB broth.

Samples were resolved on 12% hand casted SDS-PAGE gels and stained with Coomassie blue. A) Gel image showing comparison between protein produced in LB and TB at 37°C each. B) Gel image showing comparison between protein produced in TB and LB at 37°C and 22°C, respectively Samples indicated as +IP +Et are test samples under study. IP and Et represents IPTG and 1% Ethanol respectively, while (+) indicates addition and (–) indicates absence of particular component. +IP -Et is negative control and -IP +Et is positive control.

In conclusion, the best optimised conditions for production of recombinant proteins were seen when *E. coli* Rosetta was grown in Terrific broth at 37°C, with dual induction of 1mM IPTG and 1 % ethanol after 24 hours. The next step was to produce large quantity of proteins to enable purification.

4.6 Large-scale recombinant protein production

Large scale protein production was followed as per Section 2.8 and 2.8.1. *E. coli* pET28 (+) PLysS cells were grown in 250 ml flask containing 100ml of TB broth supplemented with antibiotics and in presence of 1 mM IPTG and 1% ethanol at 37°C for 24 hours. All conditions remained same as per the best optimized parameters. The cells were spun, and the pellet was resuspended in Lysis Buffer and EDTA-free Protease inhibitor. After the pellet was completely dissolved, the sample was sonicated, centrifuged, denatured, and electrophoresed. For better resolution of the bands, the pellet sample was diluted (1:10, 1:100, 1:1000) from the observations made from previous gels which showed too much of proteins during the PAGE run. Also, samples from small-scale production were retained at -20°C and resolved on the same gel (lanes 2, 3, 4, 5, 6, 7, and 8) along with the processed large-scale samples (lanes 9, 10, 11, 12, and 13) to check for consistency between results.

In Figure 4.13, lanes 5 and 6, show consistency with scaled up recombinant protein production via more intense 15 kDa bands (lanes 10 and 11) that came from the insoluble fraction.

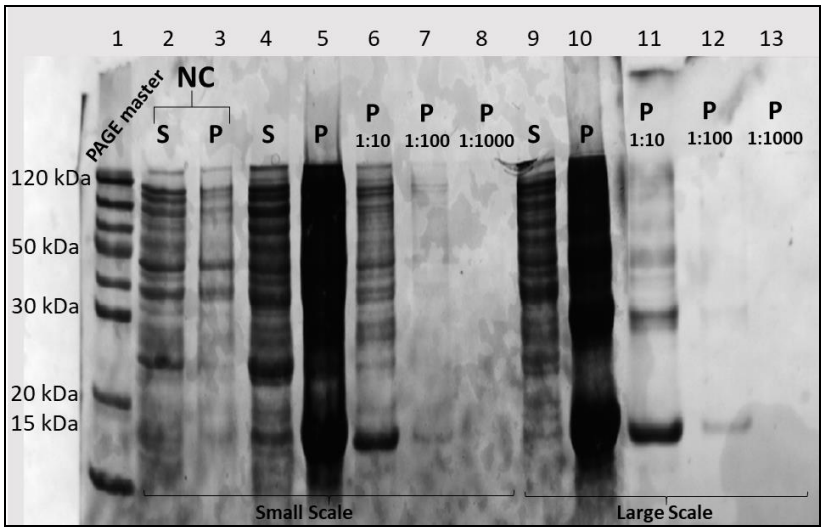


Figure 4.13 Comparison between small scale and large-scale production of recombinant proteins in TB broth having 1% ethanol and 1mM IPTG.

Samples were resolved on 12% hand casted SDS-PAGE gels and stained with Coomassie blue. Gel representation from left to right, Lane 1: PAGE master ladder, Lanes 2-8: small-scale samples, Lanes 9-13: large-scale samples. NC represents negative control (no IPTG and ethanol).

Overall, a consistent band was seen at 15 kDa for both small and large-scale protein production. Also, a better separation of band was seen that eliminated most of the other proteins, especially after 1:10 dilution (lanes 6 and 11).

4.7 Purification of His-tagged VMO1 recombinant protein

The reserved insoluble protein lysate (Figure 4.13, Lane 10) was treated with DNase. Once solubilized in Binding Buffer it was loaded onto a His-resin column (Section 2.10). Protein purification methods were investigated without imidazole (non-denaturing conditions) and with imidazole (native conditions).

4.7.1 Purification of recombinant protein lysate without imidazole

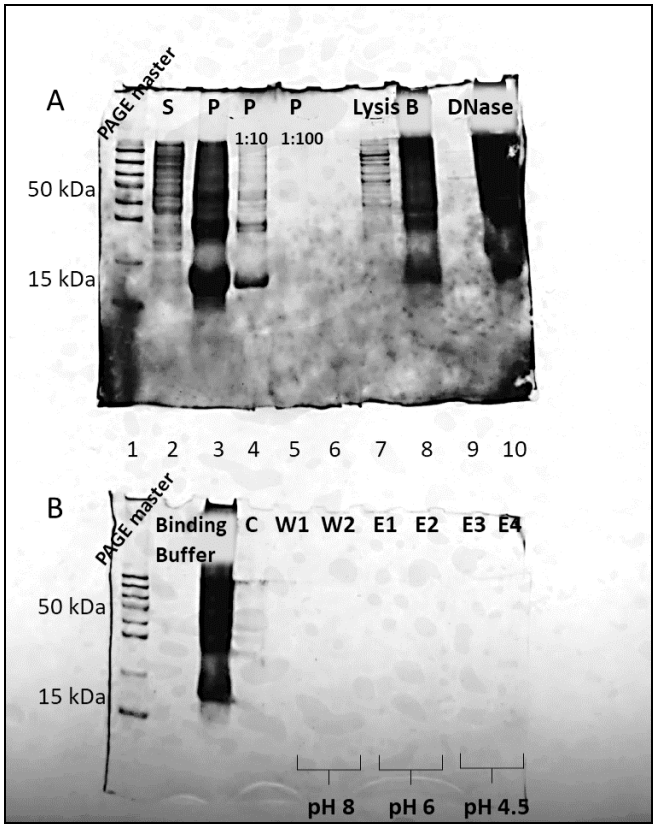


Figure 4.14 Coomassie Blue staining of His-VMO1 purification under native conditions.

Samples were resolved on 12% hand casted SDS-PAGE gels and stained with Coomassie blue. A) Protein separation of sample when treated with Lysis buffer and DNase. B) Protein separation of sample when treated with Binding buffer, and followed with His purification buffer as follows; Lane 4: sample passed through affinity column, Lanes 5 and 6: Wash buffer pH 8, Lane 7 and 8: Elution buffer pH 6.0, Lanes 9 and 10: Elution buffer pH 4.5 respectively.

Fate of protein lysate when treated with Lysis Buffer and DNase, yielded the 15 kDa protein (Figure 4.14 A, lanes 8 and 10) followed by Binding buffer (Figure 4.14 B, lanes 3). However, on following purification steps (Section 2.10.1) there was no evidence of bands (Figure 4.14 B, lanes 5, 6, 7, 8, 9 and 10). This suggests that the protein remains unpurified in absence of imidazole.

4.7.2 Purification of recombinant protein lysate with imidazole

pH modifications were made with purification buffers when Imidazole was added to allow better solubility of the recombinant proteins (Section 2.10.3). There was evidence of protein bands at 15kDa after purification steps. This suggests that the protein purifies in the presence of imidazole due to its correct refolding. Majority of the protein was seen eluting using Elution buffer, pH 6 as expected (Figure 4.15, lanes 7 and 8) followed by traces at pH 5 (Figure 4.15, lane 9).

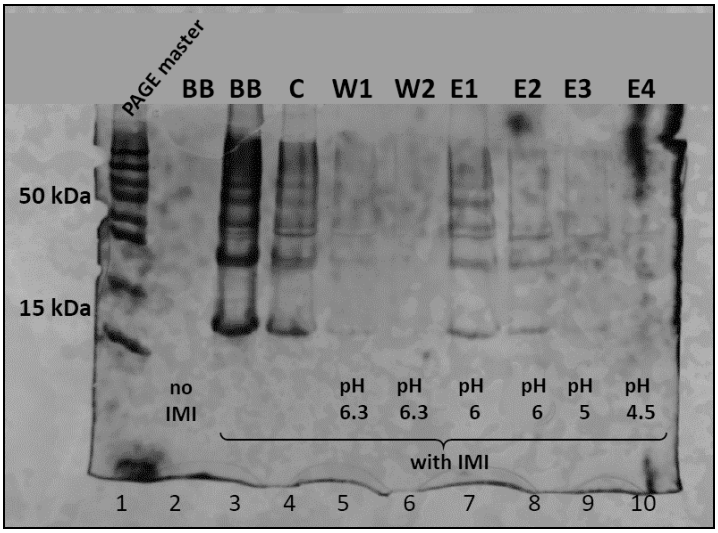


Figure 4.15 Coomassie Blue staining of His-VMO1 purification with Imidazole and urea.

Samples were resolved on 12% hand casted SDS-PAGE gels and stained with Coomassie blue. Gel representation from left to right, Lane 1: PAGE master ladder, Lane 2: sample treated with Binding buffer without IMI, Lane 3: sample treated with Binding buffer with IMI, Lane 4: sample passed through affinity column, Lanes 5 and 6: samples treated with Wash buffer having IMI, pH 6.3, Lane 7 and 8: sample treated with Elution buffer having IMI, pH 6.0, Lanes 9 and 10: sample treated with Elution buffer having IMI, pH 4.5 respectively. IMI represents Imidazole.

Once it was confirmed that protein purification works under denaturing conditions, western blot detection was performed. It was not only important that the protein is purified but also be recognized by His-tag and VMO1 antibody to further confirm that a true recombinant protein was expressed and is folded correctly.

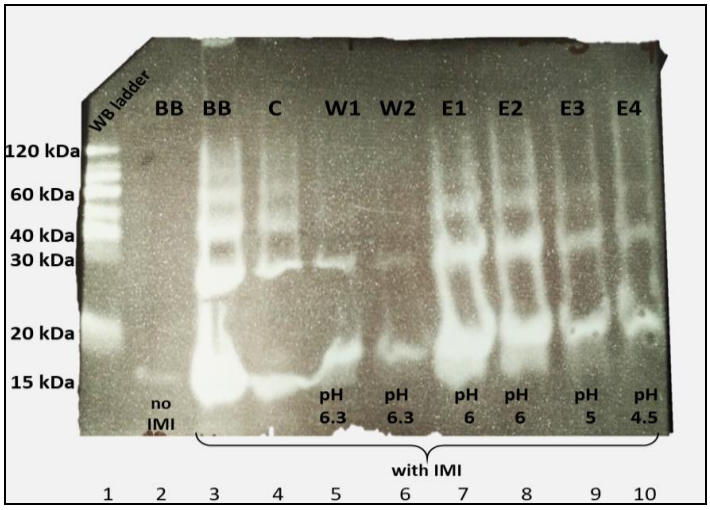


Figure 4.16 Western blot detection of His-VMO1 purification with Imidazole and urea using VMO1 (1:1000) antibody.

The blot is an Inverted image developed in the imager to detect visible protein bands. Representation from left to right, Lane 1: Western blot ladder, Lane 2: sample treated with Binding buffer without IMI, Lane 3: sample treated with Binding buffer with IMI, Lane 4: sample passed through affinity column, Lanes 5 and 6: samples treated with Wash buffer having IMI, pH 6.3, Lane 7 and 8: sample treated with Elution buffer having IMI, pH 6.0, Lanes 9 and 10: sample treated with Elution buffer having IMI, pH 4.5 respectively. IMI represents Imidazole.

Both the antibodies were able to detect a range of protein bands from 60kDa to 15kDa (Figures 4.16 and 4.17). The possible reason for this scenario is that VMO1 protein is "sticky" in nature and there is high possibility of it getting attached to other proteins. Evidence and detection of band at ~30kDa raises another plausibility that VMO1 is a protein dimer as observed for chicken VMO1 in PDB (Figure 1.7 C). The 3 bands were consistent compared to our earlier westerns i.e. 15, 30 and 40 kDa (Figure 4.16, lanes 7,8, 9, and 10 and Figure 4.17, lanes 7 and 8). Because it was difficult to see bands after membrane stripping in the imager and auto fluorescence of PVDF membrane, a manual chemiluminescent setting was adjusted that increased the exposure time of UV

light to 10-15 mins. Hence the bands now appear white, which otherwise show high background signal. This could have also affected visualising western ladder in Figure 4.17 (lane 1). Also, there appears an air bubble between lanes 3-5 (Figures 4.16 and 4.17).

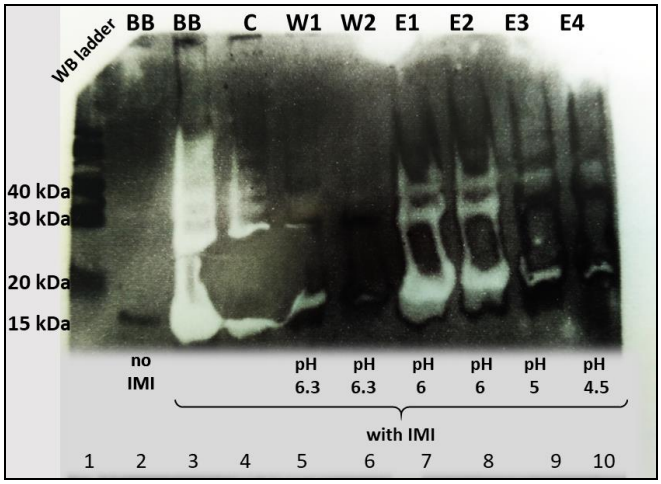


Figure 4.17 Western blot detection of His-VMO1 purification with Imidazole and urea using His-tag (1:1000) antibody.

The blot is an Inverted image developed in the imager to detect visible protein bands. Representation from left to right, Lane 1: Western blot ladder, Lane 2: sample treated with Binding buffer without IMI, Lane 3: sample treated with Binding buffer with IMI, Lane 4: sample passed through affinity column, Lanes 5 and 6: samples treated with Wash buffer having IMI, pH 6.3, Lane 7 and 8: sample treated with Elution buffer having IMI, pH 6.0, Lanes 9 and 10: sample treated with Elution buffer having IMI, pH 4.5 respectively. IMI represents Imidazole.

4.8 Immunohistological evaluation using validated VMO1 antibody

Longitudinal Sections (LS) of paraffin embedded tissues were sliced from P240 mouse ear, P5 mouse cochlea and L5 rat trachea on to Gelatin-coated slides for H&E staining and IHC. Cellular integrity was found to be intact in the sectioned tissues, with good anatomical distinction of the RM and OOC. A nuclear, blue counterstain (DAPI stain) helped visualization of nuclei in the tissue when excited at 358 nm during fluorescent microscopy.

4.8.1 H and E staining with DAPI

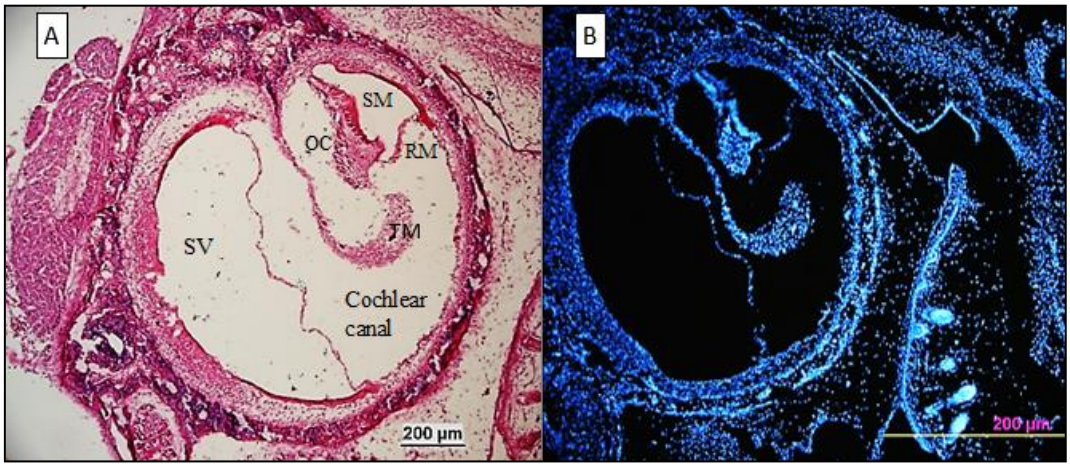


Figure 4.18 Anatomy of the mouse inner ear (P240).

A) Haematoxylin and Eosin staining of L.S of mouse ear, P5. Parts labelled as follows, SV: Scala vestibular; SM: Scala media; OC: organ of Corti; RM: Reissner's membrane; TM: tectorial membrane. B) DAPI staining of L.S of mouse cochlea, P240. Blue fluorescence emitted by DAPI under U.V (358nm) shows structural location of nuclei.

IHC was performed as per the protocol (Section 2.11.4, Table 2.8) using two different primary antibody concentrations (1:200, and 1:400) with 1:2000 goat anti-Rabbit IgG (H+L) Alex fluro488 secondary antibody solution. The slides were then viewed using phase contrast. Phase contrast helps unstained objects (e.g. flat cells) to fluoresce and visualize cellular components by immunofluorescence.

Immovable bright green dots (excitation wavelength of 490nm) were considered as positive signals when green filter was used. However, an orange filter (647nm) helped to distinguish between true signals. Hence, it was also used as a control.

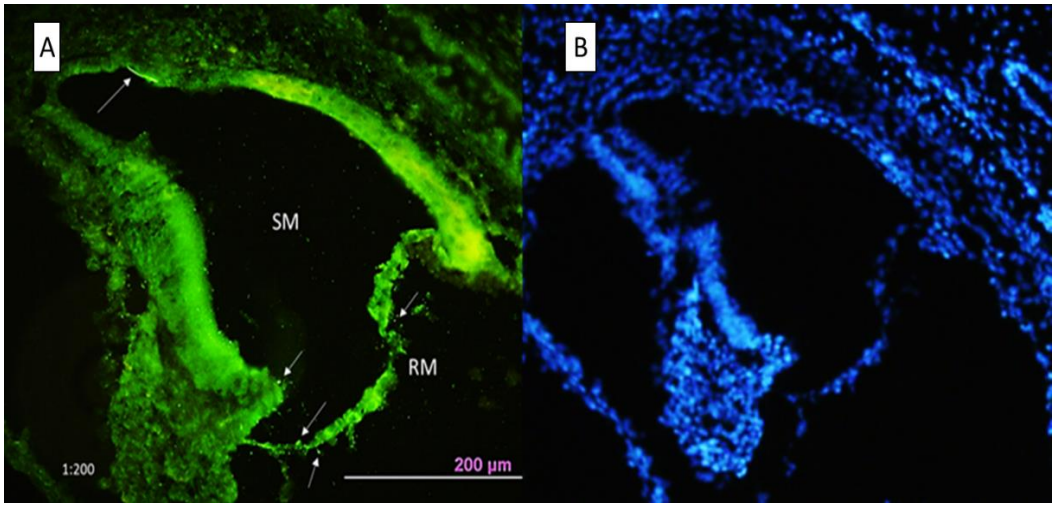
Antigen revival was performed initially to unmask the epitopes or antigenic sites with heat induced epitope revival using Sodium citrate buffer, pH 6.0. But, this method disintegrated many cochlear sections hence an alternative method was approached. Here, the slides were set in the buffer bath O/N at 60°C with gentle shaking. The mouse inner ear sections were however fully intact employing the microwave method. The cochleas' complex shape, along with delicate constitution of the OOC, the bony enclosure of the inner ear poses a challenge for histological and analysis (Montgomery & Cox, 2016).

Signals were seen within the tissues (e.g., Figure 4.23), the peripheries (e.g., Figure 4.19) as well as in the extracellular regions of tissue (e.g., Figure 4.22).

In conclusion it can be said that VMO1 antibody recognized protein localization in the tissue specimen but a further optimisation of control slides needs to be done in the future that will further help to distinguish true signals from the background. This can be achieved by using a higher diluted secondary antibody e.g., 1:3000 or 1:5000 or 1:500 or 1:1000 diluted primary antibody. The test and control slides can be compared

4.8.2 IHC of mouse inner ear

IHC on mouse ear sections showed signals mostly in the SM (Figure 4.19 A and 4.19 C), around RM, and OOC (Figure 4.19 A'). Since no control slide was available, true signal was only detected by comparing the fluorescence from green and orange filters (Personal communication with Barry O'Brien, 2018).



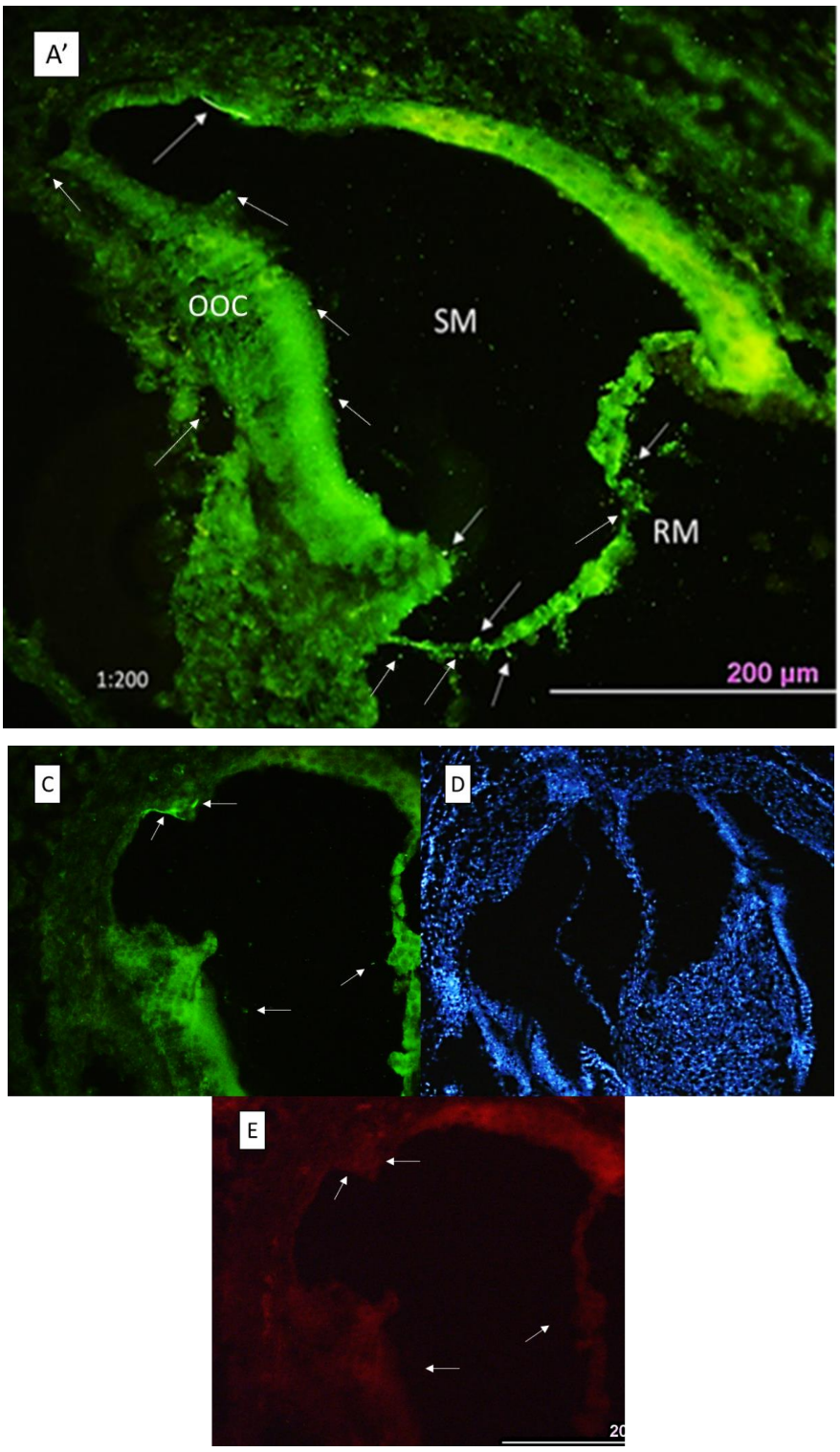


Figure 4.19 Fluorescent microscopy of IHC on mouse ear (P240) with 1:200 VMO1 antibody.

Figures A, B, A', C, D, and E represents tissues that were treated with 1:200 diluted VMO1 antibody. Green, blue, and orange images were developed under green fluorescence, DAPI stain (U.V), and orange fluorescence, respectively. The arrow shows presence of signals under green light that does not get fluoresced in presence of orange light. Figure A' is a higher magnification of A.

Figure 4.20 B and 4.21 B showed high background signals clumped together, when their respective control slides (Figure 4.20 E and 4.21 E). This indicates too much of secondary antibody and background signal being detected and thus required more washing steps to be employed during IHC.

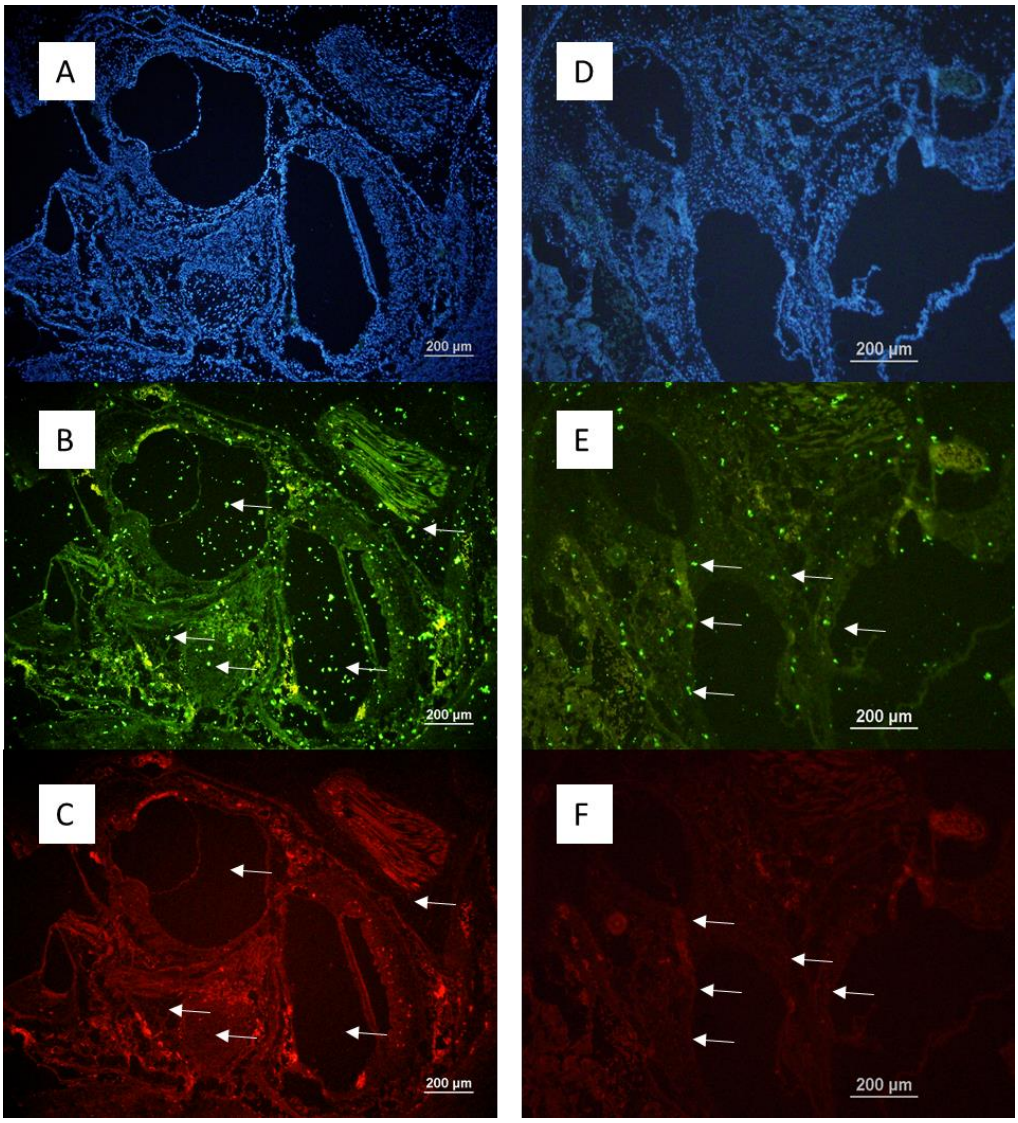


Figure 4.20 Fluorescent microscopy of IHC on mouse ear (P5) with 1:200 VMO1 antibody and control.

Figures A, B, C, represents tissues that were treated with 1:200 diluted VMO1 antibody, while D, E, and F represents control slides. Blue, green, and Orange images were developed under DAPI stain (U.V), green fluorescence, and orange fluorescence, respectively. The arrows show presence of signals under green light that do not get fluoresced in the presence of orange light.

With further washings employed with PBS (pH 7.4); a reduced background can be observed in Figure 4.21 B when compared to its control slide E.

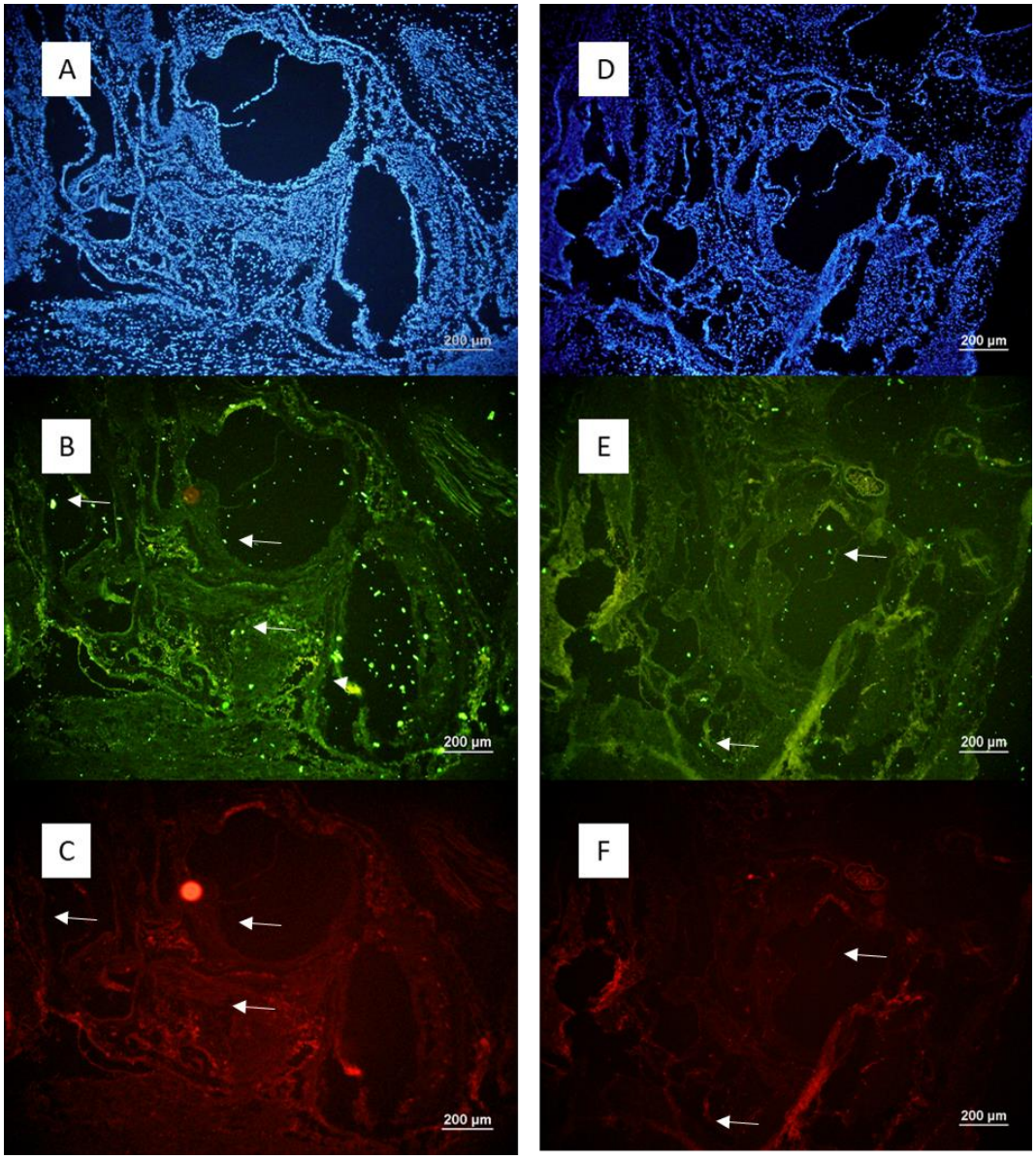


Figure 4.21 Fluorescent microscopy of IHC on mouse ear (P5) with 1:400 VMO1 antibody and control.

Figures A, B, C, represents tissues that were treated with 1:400 diluted VMO1 antibody, while D, E, and F represents control slides. Blue, green, and orange images were developed under DAPI stain (U.V), green fluorescence, and orange fluorescence, respectively. The arrows show presence of signals under green light that do not get fluoresced in the presence of orange light.

4.8.3 IHC of mouse cochlea

When IHC was performed on cochlear sections, few more additional washings were introduced. This further reduced the background as seen in Figure 4.22 B and 4.23 B, when compared with their respective control slides (Figure 4.22 D and 4.23 D).

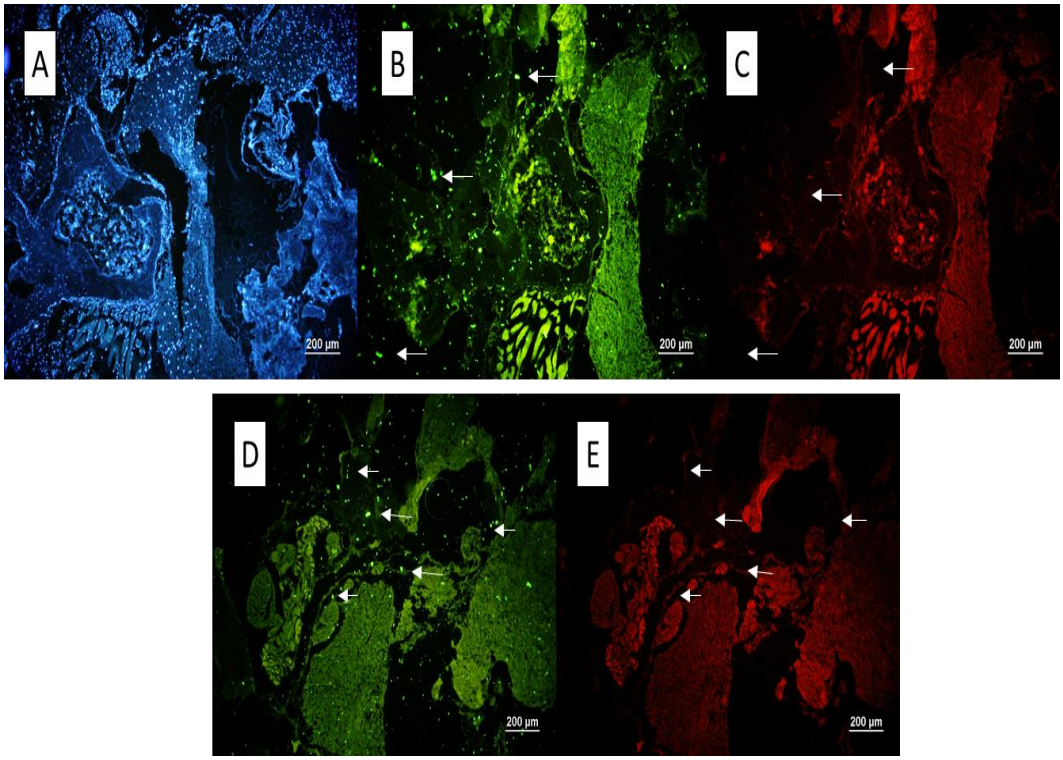


Figure 4.22 Fluorescent microscopy of IHC on mouse cochlea (P5) with 1:200 VMO1 antibody and control.

Figures A, B, C, represents tissues that were treated with 1:200 diluted VMO1 antibody, while D, and E represents control slides. Blue, green, and orange images were developed under DAPI stain (U.V), green fluorescence, and orange fluorescence, respectively. The arrows show presence of signals under green light that do not get fluoresced in the presence of orange light.

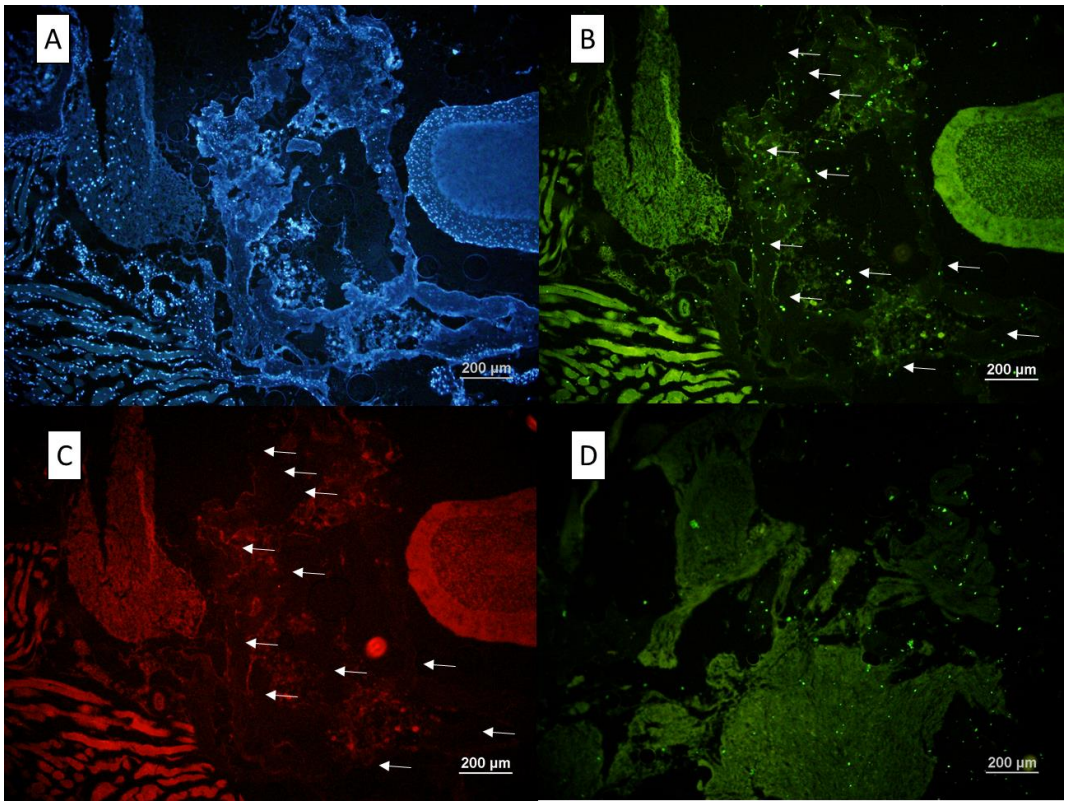


Figure 4.23 Fluorescent microscopy of IHC on mouse cochlea (P5) having 1:400 VMO1 antibody and control.

Figures A, B, C, represents tissues that were treated with 1:400 diluted VMO1 antibody, while D, represents control slide. Blue, green, and orange images were developed under DAPI stain (U.V), green fluorescence, and orange fluorescence, respectively. The arrows show presence of signals under green light that do not get fluoresced in the presence of orange light.

4.8.4 IHC of rat lung

For the rat lung sections, signals were seen inside the alveolar spaces and some in the alveolar walls (Figure 4.24 B and E).

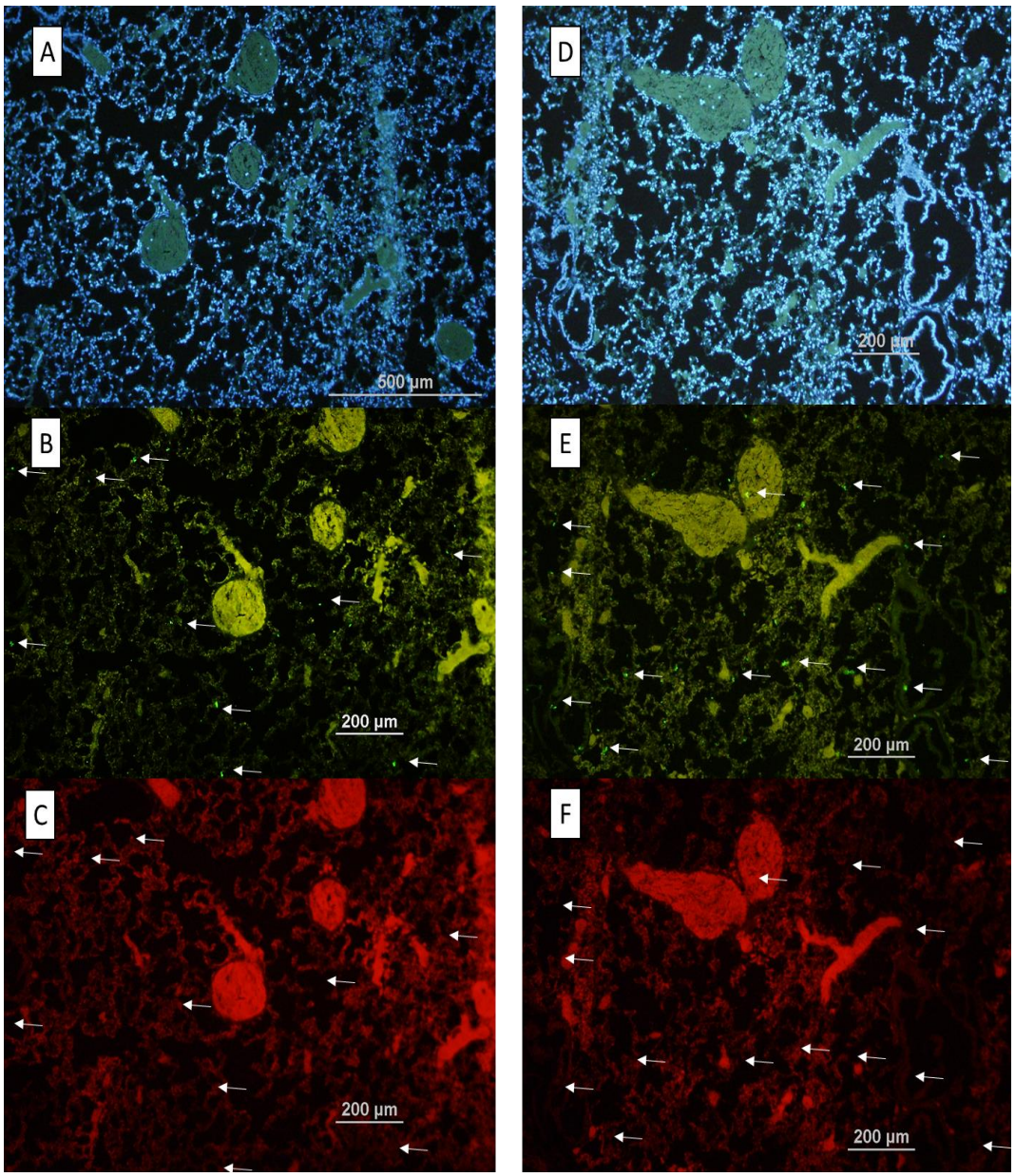


Figure 4.24 Fluorescent microscopy of IHC on Rat lung (L5) tissues having 1:200 and 1:400 VMO1 antibody with control.

Figures A, B represent tissues that were treated with 1:200 diluted VMO1 antibody, while D and E represent tissues that were treated with 1:400 diluted VMO1 antibody. Figures C and F represent their respective control slides. Blue, green, and orange images were developed under DAPI stain (U.V), green fluorescence, and orange fluorescence, respectively. The arrows show presence of signals under green light that do not get fluoresced in the presence of orange light. C and F represent control slides.

CHAPTER FIVE

DISCUSSION

The main aim of this research was to investigate the expression and localization of Human VMO1 protein using molecular and cellular assays. This was achieved using four objectives. First, we studied the structural and functional properties of *VMO1* gene through bioinformatics. Secondly, we induced and expressed a recombinant VMO1 protein having His-tag using a traditional method and an alternative new method that used ethanol. Third, we purified the induced recombinant protein using a His Nickel affinity column and protein purification buffers. Finally, we tested the specificity of commercial antibodies to enable recognition of the purified recombinant protein by western blot methodologies and IHC. Thus, this chapter will discuss the outcomes of the four objectives and the recommended 11 future directions that should be carried out to further understand the functionalities of VMO1.

5.1 Objective 1: Comparative analysis of *VMO1* gene

After analysing *VMO1* reference sequences (both nucleotide and amino acid) in Chapter 3, it was found that this gene was conserved across many organisms including Rhesus monkey, chicken, mouse, and zebra fish (Table 3.2). More specifically, 176 organisms have orthologs to the human *VMO1* gene. Structurally, four alternative *VMO1* mRNA transcript having three exons in humans have been identified and they map to chromosome 17: Isoform 1 (785 nt), Isoform 2 (821 nt), Isoform 3 (765 nt), and Isoform 4 (669 nt). In mouse (672 nt) and chicken (2626 nt) it is present as only one isoform, on chromosome 11 and chromosome 1, respectively.

The importance of using these reference organisms is that, chicken was the first where *VMO1* was discovered and thus, well-studied with respect to its biological function in the VM and its extracellular expression (Kido *et al.*, 1995; Mann, 2008). Secondly, the mouse serves as a good model to study human genetics since its genome has 85% similar while most of its genes are 99% identical (National

Human Genome Research Institute, 2018). Mouse also has a similar anatomical structure as that of a human ear. Also, an estimated 90% of rat genes have orthologs to the mouse and human genomes (Mullins & Mullins, 2004). Having said that, *VMO1* transcript in rat is highly expressed in the lung, kidney, and brain but not in the mouse; where most of its identifications have come from RNA sequencing data sets. Rat tissues were thus subjected to histochemical identification in this project. The phylogenetic tree developed in Geneious show that our consensus plasmid is closely related to the clade formed by human VMO1 sequences. Mouse and rat appear to be sister taxa's (Figure 3.19).

Mouse Vmo1 protein is 201 amino acids in length with a signal peptide from 1-21 and polypeptide chain from 22-201. The sequence has four disulphide linkages where cysteine residues forms a covalent bond to retain the tertiary configuration of a protein. This in turn makes the protein resistant under harsh external environment like thermal or biological encounters (proteolytic cleavage) (Burns *et al.*, 2016). The signal peptide directs the entire protein towards proteolytic cleavage, and to be ultimately released from the endoplasmic reticulum to serve a particular function. Thus, VMO1 is predicted to be a secreted protein.

When the mouse Vmo1 protein sequence was analysed using UniProt it revealed regions of local similarities with 249 other organisms with varying identities. Here, rat had the highest amino acid identity with 96%, followed by Human isoform 1 with 77.8%, and chicken with 52.7% identity. Fifty three percent similarity was also found between chicken *VMO1* and, mouse and rat. This data is in agreement with Lee *et al.*, 2015.

When mouse *Vmo1* was blasted against the four human *VMO1* isoforms, Isoform 1 showed 77.8% identity with eight conserved cysteine residues with a match length of 202 amino acids. Isoform 2 showed 79.5% identity with seven conserved cysteine residues with a match length of 151 amino acids. Isoform 3 showed 79.1% identity with six conserved cysteine residues with a match length of 102 amino acids. Finally, Isoform 4 showed 78.5% identity with six conserved cysteine residues with a match length of 114 amino acids. Additionally, there are 10 conserved cysteine residues between mouse and chicken *VMO1* that gives a match length of 183 amino acids. This data suggests that the *VMO1* gene is highly

conserved between humans and mouse. Majority of these sequences (Figure: 3.16) have hydrophobic (those having alanine, glycine, valine, proline, leucine, isoleucine, methionine, phenylalanine, and tryptophan) and polar amino acids (lysine, arginine, histidine, glutamate, and aspartate). Hydrophobic amino acids tend to shield against water molecules, termed as the hydrophobic effect; which makes the protein fold stable. The strength of this effect is also temperature dependent where it offers lower protein stability at low temperatures (cold denaturation) (Van Dijk *et al.*, 2015).

5.2 Objectives 2 and 4: Recombinant protein production and antibody detection.

Frozen *E. coli* Rosetta™ (DE3) pLysS cells having the *VMO1* construct was streaked onto LB agar plate in order to extract and sequence the plasmid DNA using SH3R and SH3F primers. The resulting electropherograms (Figure 3.7) were trimmed accordingly to provide high quality reads and made into a consensus sequence containing 423 nucleotides. This sequence was then blasted in NCBI against the four human *VMO1* isoforms, with a direct hit of 100% identity to *VMO1* isoform 3 mRNA. The plasmid translates to 140 amino acids with a calculated molecular mass of 15 kDa (Figure 3.8). This recombinant protein was in frame with the N terminus His-tag with a single stop codon at the end. An ORF search, found seven potential ORFs within the consensus sequence; four in the plus strand and three in the minus strand. Additionally, a 3D structure generated online had one of the conserved Greek key motifs (Figure 3.23) and a characteristic DNA fold (Figure 3.22). Less information on the DNA fold however limited its understanding and nature.

Potential for three bands seen on western blots could be the sequences having start codons in different ORF frames (Figure 3.9). Predicted MW to distinguish between the three frames can thus be attributed to 15, 30, 40 kDa in Geneious.

Rosetta™ (DE3) pLysS cells harbouring *VMO1* isoform 1 in pET28b (+) vector was grown using 1 mM IPTG at 24 hours. The whole cell lysate and insoluble fractions were resolved on 12% PAGE gels that yielded a 15 kDa band. Along

with this other bands of 20, 30, 40, 45 kDa were observed in Sarah Hardie's research. The 6X His antibody however detected a 20 kDa band, while VMO1 antibody (both from GeneTex) detected three bands of ~20, 35 and 50 kDa (Hardie, 2015). This advocates the sticky nature of VMO1 protein.

ProtParam tool (web.expasy.org) helped with computing many physical and chemical parameters based on the amino acid sequence of our recombinant protein. It revealed molecular weight of 15071.08 Da, and a theoretical pI of 8.62. Majority of the amino acids where composed of Glycine (15.7%), Leucine (9.3%), Alanine (7.1%), and Arginine (7.9%). The instability index (II) computed by this program was 15.75, which classified this protein as being stable. These findings are suggestive that the recombinant protein is mildly basic in nature.

Theoretically the steps for obtaining recombinant proteins are simple and involves acquiring the gene of interest using PCR, cloning the amplicon into an expression vector, transforming into a host cell, inducing protein expression, and finally purifying and characterising the biological properties. Suitability of protein induction had given rise to three important questions: (1) what organism to use for a particular gene of interest? (2) which vector to use? (3) and would be the most appropriate host?

Recent optimisations in genetic engineering has led to production of various alterations in *E. coli* that has improved protein yield and solubility. An example of such, is overexpression of heterologous tRNAs which encourages mammalian favoured codon usage (Galloway *et al.*, 2003). Other advantages of using *E. coli* is its fast growth kinetics in glucose-salts media producing stationary phase, high cell densities saturated starter culture in few hours (Rosano & Ceccarelli, 2014). Rosetta™ strains (BL21 derivatives) are tailored to enhance eukaryotic proteins by providing tRNAs for the following six codons AGA, AGG, AUA, CCC, GGA, and CUA. They also allow universal translation of proteins. Apart from these, a DNA optimization tool (atgme.org) helped us identify other very rare codons (**TGA**, **CGA** and AGG) and rare codons (**TGC**, CTA, **TGT**, CCC, AGA, **CAC**, **CGG**, **TCG**, **CTC**, and **CCT**) indicated in bold.

Rosetta (DE3) pLysS are more suitable to induce protein from target genes by IPTG and exhibits T7 lysozyme. This, suppresses the expression of DNA polymerase (an enzyme that synthesizes DNA) before induction and allows stability in the recombinant cells (www.merckmillipore.com). On the other hand, IPTG induces protein expression by triggering the Lac Operon. Significant enhancement in protein production was observed when many expression parameters were optimized using commercial pET expression vectors (Appendix 2, 1. Figure 4.25) and pLysS plasmid by Novagen (Sivashanmugam *et al.*, 2009). Hence, the vector used was pET-28b (+), having 5368bp size, a derivative of the classical plasmid- pBR322. This vector possesses various characteristics T7 promoter, T7 transcription start, His•Tag coding sequence (270-287), T7•Tag coding sequence (207-239), multiple regions for cloning, T7 terminator, lacI coding sequence and Kanamycin coding sequence.

The liquid cultures of *E. coli* strain, Rosetta™ (DE3) pLysS were grown to the late log phase (OD_{600} = 5.0-1.0) followed by the addition of IPTG. The late log phase culture was an influential factor to obtain high protein yield (Galloway *et al.*, 2008). In general, aliquots were taken out as per Section 2.4 and separated on 12% SDS-PAGE gels; either to stain the proteins via Coomassie blue or perform protein transfer from the pre-run gel onto the PVDF membrane for Ponceau staining and western blot detection. Both whole cell lysates (total proteins) and sonicated cells (recombinant proteins) were considered and compared. Protein separation was observed from 80 kDa to 15 kDa upon Coomassie staining while 60 kDa to 15 kDa upon western blot analysis.

Firstly, we used the traditional method for protein induction (Figure 4.5). To allow high recombinant protein expression, we optimized the growth conditions by using two concentrations of IPTG (0.1 or 1 mM), and different expression times (0, 3, 6, and 24 hours). While assessing whole cell protein extracts of induced versus control samples, four bright bands were consistently observed at 50, 40, 25 and 15 kDa. Amongst which, the 15 kD band was the most prominent, after 6 and 24 hours. The size for the VMO1 recombinant protein was expected to be at 15 kDa based on sequencing and bioinformatics, hence any band around this size was significant.

When the sonicated samples were resolved on the PAGE gel and stained with Coomassie blue, high intensity bands at 15kDa were seen for the pellet samples. This showed that the expressed protein was indeed insoluble and found in inclusion bodies formed inside the bacterial cell. Additionally, two light intensity bands were also seen at 30 and 50 kDa. During western blot detection, the His-tag antibody recognized multiple bands (15, 30, and 40 kDa) of the sonicated samples while only one band (15 kDa) of the whole cell lysates (Figure 4.8). Interestingly, VMO1 antibody recognized the bands at 15, 30, 40, 60 and 80 kDa of the sonicated samples along with same bands of the whole cell lysates (Figure 4.9).

The possible reason of presence of multiple bands can be hypothesized as VMO1 to be present as a dimer at ~30 kDa. Secondly, VMO1 can be a "sticky" protein and could attach to other *E. coli* proteins when expressed in recombinant form, they form monstrously sized aggregates and bind to a plethora of irrelevant proteins from *E. coli,* expressing inappropriate fragments that expose their hydrophobic core regions (Feller & Lewitzky, 2012). In addition, surface hydrophobic interactions can also add to the sticky nature of proteins (Dorh *et al.*, 2015). Thirdly, through bioinformatics it was observed that *E. coli* is not predicted to express full length VMO1-like proteins but a few amino acids from 21-32 (Figure 3.20). VMO1 antibody sequence is unlikely to recognise this due sequence dissimilarities between the two.

Anti-VMO1 antibody (Abcam: ab126510) is a polyclonal antibody raised in rabbit against Human VMO1 peptide (amino acids 25-112) and recognizes all four isoforms with a 99% match to isoform 3 (Figure 3.10).

In conclusion, our data shows high expression of recombinant VMO1 protein using the following optimized conditions: induction with 1mM IPTG, cells grown for 24 hours at 37°C in LB media incorporated with 30 mg/ml Kanamycin and 34 mg/ml Chloramphenicol under good aeration.

To improve upon this data, we also investigated the introducing of ethanol (Chhetri *et al.*, 2015) to improve protein expression and solubility. After the bacterial growth parameters were adjusted as indicated for the traditional method above, varying concentrations of ethanol (1%, 2%, and 3%) were added after

induction with 1mM IPTG (Section 2.7). The samples were taken out after 5 and 24 hours. Once again, whole cell extracts, and sonicated lysates were resolved on 12% denaturing SDS-PAGE gels. Multiple bands (10, 15, 25, 30, 40, 50, 60 and 80 kDa) were seen for whole cell proteins, and once again the 15kDa band was found to be the brightest when the cells were grown in presence of 1% ethanol than compared to 2% or 3% after 24 hours (Figure 4.10).

To investigate whether any improvements were made with solubility of the insoluble nature of the recombinant protein, the samples were sonicated again and reolved for PAGE analysis (Figure 4.11). The pellet was resolved into eight bands (15, 25, 30, 40, 50, 60, 80 and 120 kDa) while the supernatant separated into seven faint bands (20, 25, 40, 50, 60, 80 and 120 kDa). Compared to the traditional method, protein expression was further enhanced by using two component inductions: IPTG and ethanol. However, the recombinant protein still remained insoluble after western blot detection by the same antibodies.

To further improve the solubility, the bacterial cells were then grown at 22°C (Figure 4.12) (Vera *et al.*, 2007). Also, another media called Terrific broth (TB) was introduced. TB is an enriched medium and has increased concentrations of tryptone, yeast extract, glycerol as a carbon source and Potassium phosphate unlike LB. Hence improves yield of recombinant *E. coli* strains (www.thermofisher.com, 2018; Novagen. pET System Manual, 2017). Coomassie blue staining revealed a more intense band at 15 kDa while using TB than LB at 37°C (Figure 4.12 B, lane 3). Also, no band was seen for the test sample around 15 kDa at 22°C (Figure 4.12 B, lanes 6 and 7).

In conclusion, we were able to induce, express and detect a 15 kDa His-tagged, insoluble recombinant protein by subjecting the bacterial cells to grow in TB media supplemented with 30 mg/ml Kanamycin, 34 mg/ml Chloramphenicol, 1mM IPTG, and 1% ethanol at 37°C with shaking (200 rpm) conditions. Additionally, His-tag and VMO1 antibodies also detected the required 15 kDa band. With these optimized conditions, the experiment was scaled up successfully (Section 2.8) from small-scale set up (5 mL) to a large-scale synthesis (100 mL) so as to obtain sufficient protein for downstream purification experiments (Figure 4.13).

5.3 Objectives 3 and 4: Recombinant protein purification using immobilized metal ion affinity chromatography (IMAC) and antibody detection

In the event of recovering biologically active products, employing an appropriate purification strategy becomes important. HIS-Select® Nickel Affinity Gel (Sigma-Aldrich, USA) was used as a purification column. Histidine (His) exhibits strongest interaction with an immobilized transition metal ion matrix. Thus, peptides bearing poly-His-tags are retained effectively in IMAC columns. After washing the matrix, proteins can then be eluted by adjusting the pH of buffers or addition of salts under native or denaturing conditions. Purification of poly-histidine affinity tagged proteins through IMAC has resulted in 100-fold enrichment in a single step giving 95% purity in high yields (Bornhorst & Falke, 2000).

Once the bacterial cells were harvested from large scale production, the pellet was surrendered to pre-purification strategies. Starting with its resuspension in Lysis Buffer and protease inhibitors for cell lysis. The cells were then sonicated and centrifuged. It is to be noted, that different sonication settings were implied compared to small scale induction of recombinant proteins. Furthermore, the pellet was then treated with DNase I to reduce viscosity and finally resuspended in Binding buffer. All the samples derived through centrifugation were stored for SDS-PAGE gel analysis. After this step, the protein slurry was transferred into the His-Ni column.

A trial scale purification using a His-Ni column was conducted under native conditions using Equilibrium and Wash buffers having no imidazole. All the pre-purification steps and His purification for trail scale were done as per Section 2.10.1. PAGE gel separation revealed a consistent protein band at 15kDa for all pre-purified samples. However, no bands were seen throughout the His purification steps (Figure 4.14). This showed us either the concentration of protein was too low, or it could be still attached to the resin column or not folded correctly; that most likely made the protein to remain impurified.

To address these issues, we used denaturing conditions by introducing imidazole (Section 2.10.3) to our buffers. Urea forms hydrogen bonds with the N-H and C-O groups of the peptide and solubilizes proteins thermodynamically by unfolding it, while Imidazole helps in the removal of nonspecific bounded contaminants (Lim *et al.*, 2009; Bennion & Daggett, 2003; Oganesyan *et al.*, 2004). The bacterial cell pellet was resuspended again in the Binding buffer and incubated for 1 hour at room temperature (22°C) to allow complete solubilization. The Equilibrium buffer and Wash buffers were now pH 8.0 and pH 6.0, respectively. Elution buffers were prepared at three different pH values; 6.0, 5.0 and 4.5. Following PAGE, Coomassie blue staining showed increased proteins that were recovered at pH 6.0 after all the modifications at the expected band of 15kDa. In addition, a better resolved gel was seen as well with the presence of 15 kDa band throughout each purification step (Figure 4.15). There VMO1 and His-tag antibody was able to recognize these bands (Figures 4.16 and 4.17). Also, the three consistent bands (15, 30, 40 kDa) were seen throughout the denaturation-purification steps as seen before during protein expression.

Apart from that, there were traces of protein being extracted in the washing step. One of the possibility is due elution of protein at pH 6.3. This shows that we need to further optimize the Wash buffer. Or there is too much of unbound protein that still remains in the column and this can be easily optimized by increasing the amount of resin in the column. Another possibility is nonspecific binding of untagged proteins, which is a rare phenomenon (Bornhorst & Falke, 2000). This has been a major disadvantage when tagged protein fails to express at high level. In such situation, use of other affinity tags become suitable. For example, biotinylation-accepting domain affinity tags, and GST affinity tags. (Nilsson *et al.*, 1996; Panagiotidis & Silverstein, 1995). A trial scale purification can be re-run with the optimized Wash buffer and perform a large-scale purification.

In conclusion, with the trial scale purification, it can be inferred that the recombinant proteins were purified where other His-tagged proteins. The next step to this would be; to first excise the approximate sized bands (15, 30, and 40 kDa) from the PAGE gels and send to mass spectrometry analysis. Or prior to loading protein onto nickel column, the protein mixture could be placed into a size exclusion column. We can use a molecular weight cut off (MWCO) to exclude

sizes, for e.g. 30 kDa and follow Western blot detection to further authenticate the VMO1 antibody from Abcam. It would also be interesting to subject this antibody for detection on proteins extracted from mouse ears, mammalian cell lines, and VM lysates where true VMO1 proteins have been identified.

5.4 Objective 4: VMO1 antibody for IHC

Rat lung (L5), mouse hemi-dissected head (P5) or finely dissected cochlea (P5) was sectioned onto Gelatin-coated slides. To begin IHC it was important that the tissue under study had proper biological integration. Haematoxylin and Eosin stained slides helped to visualise the morphology of the inner ear and navigate tissue position. DAPI staining also helped visualise all the nuclear material present inside the cells (Figure 4.18).

The commercial suppliers of VMO1 antibody recommend an antigen revival step. Antigen masking for mouse inner ear was conducted using the microwave method since the inner ear was protected within the bony structure, the tissue was not disrupted during boiling and remained on the slide. However, for the mouse cochlea, antigen masking was performed under mild shaking at 60°C overnight. The microwave method washed away most of the tissue on the slide, hence the reduced temperature and longer incubation time was found to be more suitable for protecting the cochlea's integrity.

Using the VMO1 antibody, bright green fluorescent signals (apple green colour) in the form of dots were observed across Scala media, and small signals adjacent to organ of Corti and Reissner's membrane in the mouse sections (Figures 4.19 and 4.20). VMO1 antibody was first used in 1:200 dilution, here signals generated background noise in most parts of the tissue specimen (Figures 4.20 and 4.22). Hence, a 1:400 dilution was applied and this reduced the background noise and improved IHC detection (Figure 4.23). The control slide (available for cochlea) showed too much of background noise and a further dilution of secondary antibody should be considered for future studies. In the absence of a comparative control slide, another technique to visually distinguish between true signals was to

investigate for presence and absence of a particular signal (dot) in different wavelengths that produce green and orange fluorescence.

The signal localisation in the ear tissues also indicates that VMO1 is a secreted protein. To further strengthen this, it is recommended to include a comparative tissue section that has no primary antibody exposure. This would help to clearly demonstrate the levels of background signal.

In addition, we screened the rat lung (L5) for evidence of VMO1. The rat lung was available for testing, and mRNA expression data exists from a transcriptome project where VMO1 expression was observed in 11 organs (brain, thymus, lung, heart, adrenal gland, muscle, kidney, liver, spleen, and uterus or testis) and across four developmental stages (2-, 6-, 21- and 104-weeks-old) (NCBI, 2018). At the present, no antibody expression data is available for the mouse or rat on any online databases.

Positive signals were present in rat lower left lung tissue, especially in the alveolar space (Figure 4.24). However, whether the secreted signal protein was released from the epithelial cells (surfactant proteins) or cytokines can be studied further ahead by extracting these components using sucrose density gradient or bronchoalveolar lavage fluid respectively, from lung lysates. A polyclonal antibody against Interleukins could be directed for its identification or by using ELISA (enzyme-linked immunosorbent assay) (Mittal & Sanyal, 2011).

Our data is in agreement with the information available on The Expression Atlas (Mouse Genome Informatics, 2017; EMBL-EBI, 2018). Upon further examination of these databases, *Vmo1* in the mouse was found to be expressed in eighteen tissue samples including adrenal gland, cerebral cortex, kidney, and lung. The level detected was in below cut off values, low or at a moderate level. High expression was found in lungs, olfactory epithelium, and vomeronasal organ. Across a total of nineteen experiments, the cell types that expressed Vmo1 were documented in astrocytes, cardiomyocyte, dendritic cell, embryonic fibroblasts, embryonic stem cell, inner ear hair cell, lung macrophage, lung dendritic cell, skin, Langerhans cells, and Sox+ supporting cells. Three experiments documented expression at developmental stage in Post-natal mouse at days 1, 4, 22, and 29

with medium expressions. Twelve experiments had differential expression data on cell type, disease, genotype, and phenotype based on log_2-fold change in *Vmo1*. Of these, two were studied for diseases in acute graft versus host disease versus normal strain and for clear cell sarcoma. Here, the log change was observed from -1.6 to -3.8 for the former and 1.7 to 6.6-fold for the latter.

5.5 Conclusion and future prospects

The first aim of our research was to induce and express a recombinant VMO1 protein and understand the function of this gene using bioinformatics. The next aim was to purify the recombinant protein and confirm the antibody's recognition capabilities for the protein in the mouse ear and cochlea tissues. Thus, helping us to understand the molecular function in the RM and OOC.

The results obtained in this thesis provided further insight into the conservation of the *VMO1* gene with respect to its structure and function, and if the expressed protein was secreted and where. Based on the comparative genomics analysis and anatomy, the mouse is most suitable for studying the role of human VMO1 in the auditory system with results showing that the *VMO1* gene is highly conserved across mammalian species.

The best conditions for protein induction were obtained by culturing the Rosetta bacterial cells in Terrific broth supplemented with 1mM IPTG and 1% ethanol, at 37°C for 24 hours. A 15 kDa insoluble His-tag protein was expressed in *E. coli*. Elution with Imidazole and urea was allowed for protein refolding. Thus, it is recommended to optimize the trial scale buffers and perform a large-scale purification.

Also, the three different bands consecutively seen can be excised from the gel and subjected to mass spectrometry (MS) to understand proteomics with regards to any spliced isoforms or post-translational modifications (PTMs). Primarily there are three important applications of MS which deals with understanding of the protein expression, protein interactions, and exploring sites of protein modification (e.g. phosphorylation, acetylation) (Han *et al.*, 2008).

Since the function of this gene and protein is still unknown it would be recommended to test the purified protein on medically important microbial isolates e.g. *Staphylococcus aureus*. Chicken VMO1 was found to be active against bacterial pathogen *Listeria monocytogenes* but not *Salmonella enterica* at a concentration of 100 µg/mL through its heparin-binding domains (Guyot *et al.*, 2016). Enzymatic hydrolysis yielded functional VMO1 peptides (Lee *et al.*, 2017) hence the VMO1 protein could be subjected to similar treatment. The transferase activity can be assessed following the methods from Raftery & Rand-Meir. (1968) and Kido *et al.* (1995). Furthermore, VMO1 and lysozyme's interaction could be studied to better understand the functionality between the two molecules using the protocols described by Wang *et al.* (2014).

Mammals do not have lysozyme. Therefore, methods like Pull-down assay, yeast two hybrid (Y2H) screen or phage library display can be used to find interacting partners of VMO1. A Pull-down assay (ThermoFisher Scientific, 2018) uses a "bait" protein instead of an antibody. In this case, the bait could be the purified pET28 (+) VMO1 protein immobilized onto affinity ligands such as Nickel/cobalt chelate complexes to identify proteins from a cell lysate sample. The protein-protein complex can then be eluted and resolved on SDS-PAGE and detected by western blotting. In Y2H and phage library, detection of interacting proteins take place in living yeast cells or phages, and subsequently screened for full-length ORFs or by analysing phage-displayed protein libraries (Brückner *et al.*, 2009; Sidhu *et al.*, 2003).

Following purification of the recombinant VMO1 protein, crystals could be developed to allow the determination of tertiary structure by X ray crystallography. It would be interesting to see how the structural difference of the human protein crystals are with respect to the chicken egg and then comment on the Greek key motif predicted in Pymol. The protein crystals produce X-ray diffraction patterns to view their composition. The diffraction pattern is directly correlated with the internal order of the crystal. This could be done by using cryo-crystallography, high-intensity synchrotron X-ray sources or phasing methods (McPherson & Gavira, 2014). To obtain crystals, the protein macromolecule is purified to homogeneity by bringing the macromolecule to a supersaturated state with an approximate concentration of 250 mg/mL.

Next, when the sample is precipitated to promote nucleation of the protein crystals it results in the growth of large three-dimensional crystals. This could be achieved by vapor diffusion (water vapor diffusion and dehydration of protein solution) or batch crystallization (using precipitant) (Dessau & Modis, 2011; McPherson & Gavira, 2014).

In addition, it would be interesting to compare and contrast the expression of VMO1 in mammalian cell line (e.g. *HeLa* cell line). One of the important aim would be to review and purchase the commercial mammalian vector that contains mouse and/or human *VMO1* or to transfect *VMO1* gene into a mammalian vector and then introduced into the cell line using chemical transfection. This would help in understanding morphological features of the cells i.e. study growth rate, cell viability, identify organelles using under the confocal microscope by stained cell markers or any dramatic effect (over expression). In addition, IHC can be undertaken using VMO1 and FLAG-tag/ GFP antibodies to investigate protein localisation within the cell. The cell lines can also be subjected to CRISPRi knockdown of *VMO1* and assess phenotypic changes or a knockout animal model can be designed.

One of the other recommendations is also doing a genome wide association study (GWAS) with individuals with hearing loss or use animal models and look for if any, presence of single nucleotide polymorphisms (SNPs). For example, a GWAS was conducted for noise-induced HL in 100 inbred strains of mouse (Hybrid Mouse Diversity Panel). Their Auditory brainstem response (ABR) thresholds were recorded (40 dB at the 4-kHz frequency), and the mRNA was extracted from cochlea's and gene expressions was measured. Out of 14 candidate genes expressed in the cochlea, 5 genes (*Rassf4, Zfp422, Olfr212, Olfr215, and March8*) were identified with rs37517079, being a peak SNP for noise susceptibility. This was later mapped to chromosome 6 (Lavinsky *et al.*, 2016). In another study, 352 Finnish Saami individuals aged between 50 and 75 years were subjected to GWAS and 83,381 SNPs were included in the statistics. Genomic DNA was extracted and Genotyping was performed with the Affymetrix 100 K chip using BRLMM algorithm (for genotype calling and estimates allele signals (Teo *et al.*, 2008)). One of the top SNP identified was rs457717. It was found to be localised

in an intron of the GTPase-activating-like protein (IQGAP2), which expressed inside the cochlea (Van Laer *et al.*, 2010).

Many assays are available to evaluate the phenotypic traits of deafness-related mutations in the mouse to study human deafness. These include, microscopy (bright field, scanning electron, and transmission, fluorescent) and paint-fill analysis (Bissonnette & Fekete, 1996). Apart from these, a physiological test called as patch clamp assays (measure membrane potentials or electro-motility in a single cell) can be utilised (Liberman *et al.*, 2002). In situ hybridization can also be used to study expression patterns of specific mRNAs in wild type and mutant inner ears (Chatterjee & Lufkin, 2011).

APPENDIX ONE

BUFFERS AND SOLUTIONS

Antigen revival buffer – sodium citrate buffer pH 6

2.94g Tri-sodium citrate (dehydrate)

Make up to 900 mL with mQH2O and autoclave

Mix to dissolve. Adjust pH to 6.0 with 1N NaCl

Add 0.5 ml of Tween 20 and mix well

Make upto 1000mL with milli. Q H2O

10% Ammonium persulfate (APS)

0.1 g Ammonium persulfate

1 mL ddH2O

Store at -20°C.

8X Binding Buffer (4M NaCl, 160mM Tris-HCl, 40mM imidazole, pH7.9)

11.69g NaCl

8 ml 1M Tris-HCl

0.14g Imidazole

Make up to 50mL with deionised dH_20

1X Binding Buffer with 6 M Urea

6.25 mL 8X Binding Buffer (40 mM imidazole, 4 M NaCI, 160 mM Tris-HCI, pH 7.9)

18g Urea

Make up to 45 mL with deionized dH_20. Stir until dissolved.

Adjust pH to 7.9 and make up to a final volume of 50 mL with deionised dH_20

Blocking solution

5g Low fat milk powder

50mL 1X TBS-T (pH 7.6)

8X Charge Buffer (400 mM $NiSO_4$)

5.24g $NiSO_4$

Make up to 50mL with deionised dH_20

Chloramphenicol

34 mg Chloramphenicol

1 mL Ethanol

Filter sterilise, store at -20°C

Cracking Buffer

3.6 g Urea

Make up to 10 mL with 2X SDS Loading Buffer

DNAse Reaction buffer

50 µL 10X DNAse Reaction Buffer (50 mM Tris-HCl (pH 7.5), 10mM CaCl2, 50%% glycerol, ThermoScientific™, USA)

445 µL Deionised dH_20

Add DNase to make a final volume of 500 µL with 1X reaction buffer.

0.5M EDTA pH 8.0

93.05g EDTA

Make up to 500mL with mQH2O and autoclave

Eosin (1% solution)

10g Eosin Y

2.0mL Acetic acid (5% aqueous)

0.1g Phloxine B

Make up with 1L mQH2O

4X Elute Buffer (4M imidazole, 2M NaCl, 80mM Tris-HCl, pH7.9)

13.62g Imidazole

5.84g NaCl

4.4ml 1M Tris-HCl

Make up to 50mL with deionised dH_20

1X Elute Buffer with 6 M Urea

12.5 mL 4X Elute Buffer (4M imidazole, 2M NaCI, 80 mM Tris-HCI, pH 7.9)

18 g Urea

Make up to 45 mL with deionised dH_20. Stir until dissolved.

Adjust pH to 7.9 and make up to a final volume of 50 mL with deionised dH_20

Elution Buffer pH 4.5

0.1M Sodium phosphate

8 M Urea

Dissolve in ddH2O. Adjust pH to 4.5 with appropriate acids or bases.

Elution Buffer pH 6.0

0.1M Sodium phosphate

8 M Urea

Dissolve in ddH2O. Adjust pH to 4.5 with appropriate acids or bases.

Equilibration and Wash buffer pH 8

50 mM sodium phosphate

0.3 M sodium chloride and

250 mM imidazole (when required)

Add 2.839g of Disodium phosphate in 50 mL DD water

Equilibrium Buffer (Sigma Aldrich P6611)

0.1 M sodium phosphate, pH 8.0, with 8 M urea

Wash Buffer (Sigma Aldrich P6611)

0.1 M sodium phosphate, pH 6.3, with 8 M urea

Elution Buffer (Sigma Aldrich P6611)

0.1 M sodium phosphate, pH 4.5–6.0, with 8 M urea

20% Glucose

2 g D Glucose

10 mL milli. Q H2O

Filter sterilise. Aliquot into 2 mL tubes, store at -20°C

Haematoxylin

4.0g	Haematoxylin
0.4g	Sodium iodate
35.2g	Aluminium sulphate
250mL	Ethylene glycol
40mL	Glacial acetic acid

Make up to 1L with mQH2O

100mM IPTG

2.38 g	IPTG
100 mL	ddH2O

Filter sterilise and aliquot into 1 mL tubes. Store at -20°C.

Kanamycin

30 mg	Kanamycin
1 mL	mQH2O

Filter sterilise, store at -20°C.

LB broth

25.0 g	LB

Add to 1 L of distilled H2O and autoclave.

LB agar plates

25 g	LB
15 g	Agar

Make up to 1 L with mQH2O. Autoclave. Cool to ~50°C. Add antibiotics if necessary and pour. Once set, refrigerate upside down.

6X Loading Dye

3 mL	Glycerol
25 mg	Bromophenol Blue
20 μL	Xylene Cyanole

Make up to 10 mL with sterile mQH2O.

Lysis Buffer

1 mL	1 M Tris-HCl pH 7.5
6.25 mL	4 M NaCl
5 mL	Glycerol
50 µL	TritonX-100

Make up to 50 mL with ddH2O

1X Phosphate buffered saline (PBS), pH 7.4

8 g	NaCl
0.25 g	KCl
1.44 g	KH2PO4
0.24 g	Na2HPO4

Make to 1 L with milli. Q H2O

1X Phosphate buffered saline Tween 20 (PBS-T), pH 7.2

8 g	NaCl
0.2 g	KCl
1.44 g	KH2PO4
0.24 g	Na2HPO4

2 ml of Tween-20 (0.2% final conc)

Make to 1 L with milli. Q H2O

4% Paraformaldehyde (PFA)

4.0g	PFA
10.0µl	10M NaOH

Make up to 50mL with mQH2O and heat in 65°C water bath to dissolve

10mL	10X PBS (pH 7.4)

Make up to 100mL with sterile milli. Q H2O

1X Ponceau S stain

0.5 g	Ponceau S
5 mL	Acetic acid

Make to 500 mL with milli. Q H2O

Ponceau S Destaining solution (1% acetic acid)

5 mL Acetic acid

Make up to 500 mL with sterile milli. Q H2O

2X Protease Inhibitor (EDTA free)

One Roche cOmplete ULTRA Tablets, to be added to 5mL of Milli.Q H2O. Vortex until dissolved.

0.4 M Sodium phosphate

2.839 g Disodium phosphate (anhydrous)

50 mL Sterile DDH2O

Use at final concentration of 0.1 M.

12% Resolving gel

10.2 mL DDH2O

7.5 mL 1.5 M Tris-HCl, pH 8.8

300 µL 10% SDS

12.0 mL 30% Bis-Acrylamide. Makes up to 30 ml

150 µL 10% APS

15 µL TEMED

Scott's tap water

1g Sodium hydrogen carbonate

10g Magnesium sulphate

500 mL distilled water

10% SDS

1 g SDS

Dissolve in 10 mL of milli. Q H2O

2X SDS Loading Buffer

To make 10 mL:

1 mL 1 M Tris.HCl pH 6.8

3 mL 100% Glycerol

4 mL 10% SDS

10 mg Bromophenol Blue

Make up to 10 mL with sterile milli. Q H2O

Prior to use, add 20 μL of β-Mercaptoethanol to 1 mL of 2X SDS Loading Buffer

5% Stacking gel

8.5 mL DDH2O

1.6 mL 1.5 M Tris-HCl, pH 6.8

0.125 mL 10% SDS

2.125 mL 30% Bis-Acrylamide. Makes up to 12.35 ml

0.063 mL 10% APS

0.0063 mL TEMED

1M Tris-HCl pH 8.0

121.1 g Tris base

Make up to 800 mL of milli. Q H2O and adjust to correct pH. Autoclave.

50X TAE - Tris-acetate EDTA buffer

242g Tris base dissolved in 800 mL milli. Q H2O

57.1mL Glacial acetic acid

100mL 0.5M EDTA (pH 8.0)

Make up to 1L with milli. Q H2O

1X TAE running buffer

20mL 50X TAE

980mL Milli Q H2O

1X TBS - Tris buffered saline, pH 7.6

50mL 1M Tris

30mL 5M Sodium chloride

Make up to 1L with milli. Q H2O and autoclave

1X TBS-T - Tris buffered saline + Tween-20, pH 8

999mL TBS (pH 7.6)

1mL Tween-20

TE buffer – Tris EDTA pH 8.0

10mL	1X Tris-HCl
2mL	0.5M EDTA

Terrific Broth

47g	Terrific broth base
1000 mL	DDH2O

Add 4ml glycerol and autoclave.

0.5% Triton-X 100

250µl	Triton-X 100

Make up to 50mL with 1X PBS (pH 7.4)

8M Urea

4.8048 g	Urea
10 mL	Sterile DDH2O

Wash Buffer

0.1M	Sodium phosphate
8 M	Urea

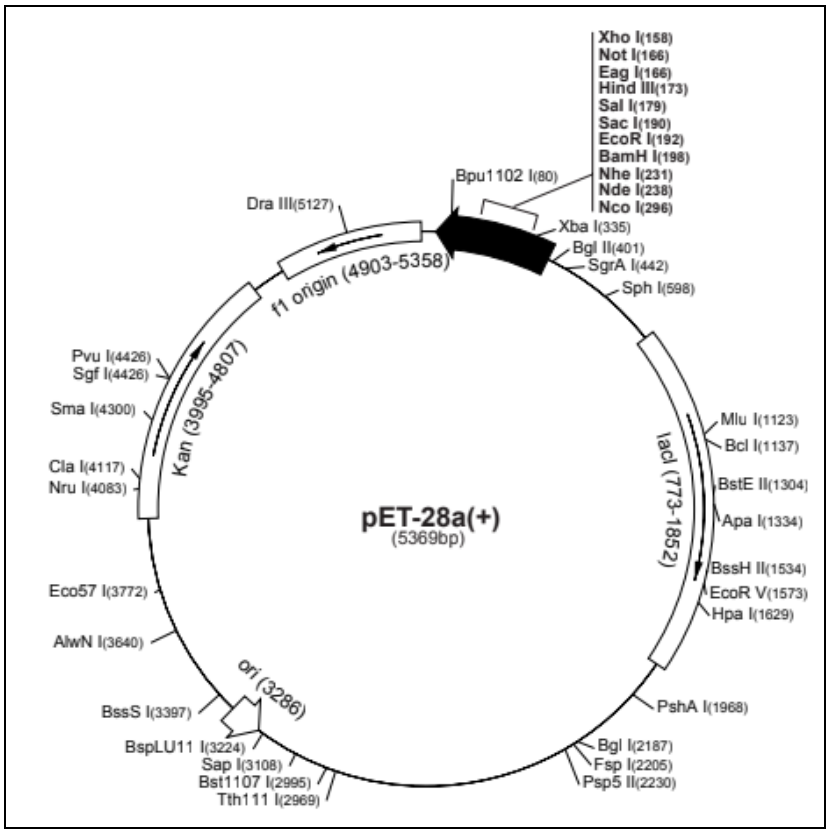
Dissolve in DDH2O. Adjust pH to 6.3 with appropriate acids or bases.

APPENDIX TWO

1: Vector Map

pET28b (+) Vector

An expression vector designed to produce recombinant proteins. The human *VMO1* gene, isoform 3 was cloned into the vector between the EcoRI and BamHI restriction enzyme sites.



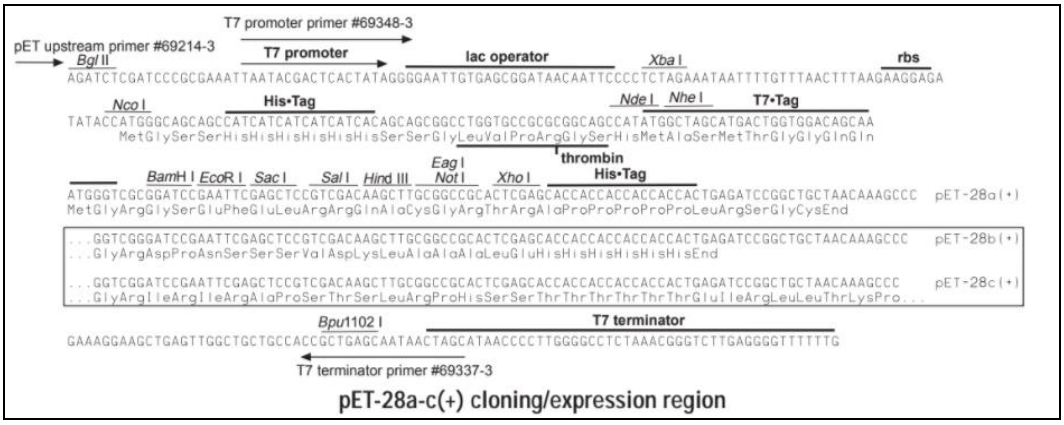


Figure 4.25: pET28b (+) vector map. Top Figure shows multiple cloning site, restriction sites, sequencing primer sites, and kanamycin antibiotic resistance. The bottom Figure shows the position of T7 promoter and T7 terminator (Novagen, 2017)

2: BLAST and Database searching

The following databases were used to investigate genomics

B Chicken Genome –
http://www.ncbi.nlm.nih.gov/projects/genome/guide/chicken/
http://genome.ucsc.edu/cgibin/
hgGateway?clade=vertebrate&org=Chicken&db=0&hgsid=67521880

E Ensembl - joint project between EMBL - EBI and the Welcome Trust Sanger Institute
http://www.ensembl.org/index.html

F EMBL - European Molecular Biology Laboratory
http://www.ebi.ac.uk/embl/

G GenBank – National institute of health genetic sequence database
http://www.ncbi.nlm.nih.gov/

K Mouse Genome –
http://www.ncbi.nlm.nih.gov/projects/genome/guide/mouse/
http://genome.ucsc.edu/cgibin/hgGateway?clade=vertebrate&org=Mouse&db=0&hgsid=67521880
http://www.informatics.jax.org/marker/MGI:2685587

L NCBI - National centre for biotechnology information
http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome
https://www.ncbi.nlm.nih.gov/gene/?term=vmo1

Molecular weight prediction

The theoretical molecular weight of proteins was predicted using the Swiss Institute of Bioinformatics Resource Portal (ExPASy)
https://web.expasy.org/protparam/

Rare codon prediction

http://atgme.org/

ORF finder

https://www.ncbi.nlm.nih.gov/orffinder/

References

abcam. (2018). *Goat Anti-Rabbit IgG H&L (Alexa Fluor® 488) (ab150077)*. Retrieved 22 March, 2018, from http://www.abcam.com/goat-rabbit-igg-hl-alexa-fluor-488-ab150077-references.html.

Abdou, A. M., Kim, M., & Sato, K. (2013). Functional proteins and peptides of hen's egg origin. *In Bioactive Food Peptides in Health and Disease. InTech.*, 115-144.

addgene. (2018a). *Analyze Sequence: pET-28 b (+)*. Retrieved 17 March, 2018, from https://www.addgene.org/browse/sequence_vdb/2566/.

addgene. (2018b). *Polymerase Chain Reaction (PCR)*. Retrieved 10 March, 2018, from https://www.addgene.org/protocols/pcr/.

Ahamad, S. R., Raish, M., Ahmad, A., & Shakeel, F. (2017). Potential health benefits and metabolomics of camel milk by GC-MS and ICP-MS. *Biological trace element research, 175(2)*, 322-330.

Alberti, P. W. (2001). The anatomy and physiology of the ear and hearing. *Occupational exposure to noise: Evaluation, prevention, and control.*, 53-62.

Alföldi, J., Di Palma, F., Grabherr, M., Williams, C., Kong, L., Mauceli, E., & Ray, D. A. (2011). The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature, 477(7366), 587-591*.

Alters, S. (2000). *Biology: understanding life*. (3rd Edition, google books. ed.). Sudbury, MA: Jones and Bartlett.

Alves, G., Pereira, D. A., Sandim, V., Ornellas, A. A., Escher, N., & Melle, C. (2013). Urine screening by Seldi-Tof, followed by biomarker identification, in a Brazilian cohort of patients with renal cell carcinoma (RCC). *International braz j urol, 39(2)*, 228-239.

Beitz, E., Zenner, H. P., & Schultz, J. E. (2003). Aquaporin-mediated fluid regulation in the inner ear. *Cellular and molecular neurobiology, 23(3)*, 315-329.

Bennion, B. J., & Daggett, V. (2003). The molecular basis for the chemical denaturation of proteins by urea. *Proceedings of the National Academy of Sciences, 100(9)*, 5142-5147.

BIO-RAD. (2017). *Quick StartTM Bradford Protein Assay Instruction Manual*. Retrieved 04 April, 2017, from http://www.bio-rad.com/webroot/web/pdf/lsr/literature/4110065A.pdf.

Bissonnette, J. P., & Fekete, D. M. (1996). Standard atlas of the gross anatomy of the developing inner ear of the chicken. *Journal of Comparative Neurology, 368 (4)*, 620–630.

Bornhorst, J. A., & Falke, J. J. (2000). [16] Purification of proteins using polyhistidine affinity tags. *In Methods in enzymology, Academic Press., 326*, 245-254.

Bos, E. J., Doerga, P., Breugem, C. C., & van Zuijlen, P. P. (2016). The burned ear; possibilities and challenges in framework reconstruction and coverage. *Burns, 42(7)*, 1387-1395.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry, 72(1-2)*, 248-254.

Brückner, A., Polge, C., Lentze, N., Auerbach, D., & Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. . *International journal of molecular sciences, 10(6)*, 2763-2788.

Burns, A., Olszowy, P., & Ciborowski, P. (2016). *Biomolecules.* Proteomic Profiling and Analytical Chemistry ((Second Edition) ed.). Elsevier.

Camp, G. V., ;, & Smith, R. (2018). *The Hereditary Hearing loss*. Retrieved 24 February, 2018, from http://hereditaryhearingloss.org/.

Campo, P., Morata, T. C., & Hong, O. (2013). Chemical exposure and hearing loss. *Disease-a-month, 59(4)*, 119-138.

Cegielska-Radziejewska, R., Lesnierowski, G., & Kijowski, J. (2008). Properties and application of egg white lysozyme and its modified preparations-a review. *Polish Journal of Food and Nutrition Sciences, 58(1)*, 5-10.

Cernilogar, F. M., Fabbri, F., Andrenacci, D., Taddei, C., & Gargiulo, G. (2001). Drosophila vitelline membrane cross-linking requires the fs (1) Nasrat, fs (1) polehole and chorion genes activities. *Development genes and evolution, 211(12)*, 573-580.

Chang, K. W. (2015). Genetics of Hearing Loss—Nonsyndromic. *Otolaryngologic Clinics of North America, 48(6)*, 1063-1072.

Chatterjee, S., & Lufkin, T. (2011). The sound of silence: mouse models for hearing loss. *Genetics research international, 2011*, 1-9.

Chen, Z., Shamsi, F. A., Li, K., Huang, Q., Al-Rajhi, A. A., Chaudhry, I. A., & Wu, K. (2011). Comparison of camel tear proteins between summer and winter. *Molecular vision, 17*, 323-331.

Chhetri, G., Kalita, P., & Tripathi, T. (2015). An efficient protocol to enhance recombinant protein expression using ethanol in Escherichia coli. *MethodsX, 2*, 385-391.

Clark, J. G. (1981). Uses and abuses of hearing loss classification. *Asha, 23(7)*, 493-500.

Cohen, B. E., Durstenfeld, A., & Roehm, P. C. (2014). Viral causes of hearing loss: a review for hearing health professionals. *Trends in hearing, 18*, 1-17.

Cordeiro, C. M., & Hincke, M. T. (2015). Quantitative proteomics analysis of eggshell membrane proteins during chick embryonic development. *Journal of proteomics, 130*, 11-25.

Dallos, P., Popper, A. N., & Fay, R. R. (1996). *The cochlea.*

Dessau, M. A., & Modis, Y. (2011). Protein crystallization for X-ray crystallography. Journal of visualized experiments. *Journal of Visualized Experiments, 47*, 1-6.

Dhooge, I. J. (2003). Risk factors for the development of otitis media. *Current allergy and asthma reports, 3(4)*(321-325).

Ding, D., Liu, H., Qi, W., Jiang, H., Li, Y., Wu, X., & Salvi, R. (2016). Ototoxic effects and mechanisms of loop diuretics. *Journal of Otology, 11(4)*, 145-156.

Dorh, N., Zhu, S., Dhungana, K. B., Pati, R., Luo, F. T., Liu, H., & Tiwari, A. (2015). BODIPY-Based fluorescent probes for sensing protein surface-hydrophobicity. *Scientific reports, 5, 18337.*, 1-10.

Eckhard, A., Müller, M., Salt, A., Smolders, J., Rask-Andersen, H., & Löwenheim, H. (2014). Water permeability of the mammalian cochlea: functional features of an aquaporin-facilitated water shunt at the perilymph–endolymph barrier. *Pflügers Archiv-European Journal of Physiology, 466(10)*, 1963-1985.

Egilmez, O. K., & Kalcioglu, M. T. (2016). Genetics of nonsyndromic congenital hearing loss. *Scientifica 2016*, 1-9.

EMBL-EBI. (2018). *ENSMUSG00000020830 (Vmo1) Mus musculus vitelline membrane outer layer 1 homolog (chicken)*. Expression Atlas. Retrieved 12 March, 2018, from https://www.ebi.ac.uk/gxa/genes/ENSMUSG00000020830?bs=%7B%22mus%20musculus%22%3A%5B%22ORGANISM_PART%22%2C%22CELL_TYPE%22%2C%22DEVELOPMENTAL_STAGE%22%2C%22STRAIN%22%2C%22SAMPLING_TIME_POINT%22%5D%7D&ds=%7B%22kingdom%22%3A%5B%22animals%22%5D%7D#baseline.

Etournay, R., Lepelletier, L., de Monvel, J. B., Michel, V., Cayet, N., Leibovici, M., & Petit, C. (2010). Cochlear outer hair cells undergo an apical circumference remodeling constrained by the hair bundle shape. *Development, 137(8)*, 1373-1383.

Exeter, D. J., Wu, B., Lee, A. C., & Searchfield, G. D. (2015). The projected burden of hearing loss in New Zealand (2011-2061) and the implications for the hearing health workforce. *The New Zealand Medial Journal, 128*(1419), 12-21.

Feller, S. M., & Lewitzky, M. (2012). Very ‘sticky’proteins–not too sticky after all? *Cell Communication and Signaling, 10(1), 15.*, 1-2.

Fischer, A. H., Jacobson, K. A., Rose, J., & Zeller, R. (2008). Hematoxylin and eosin staining of tissue and cell sections. *Cold Spring Harbor Protocols, 2008(5)*(pdb-prot4986.).

Flanagan, J. L., & Willcox, M. D. P. (2009). Role of lactoferrin in the tear film. *Biochimie, 91(1)*, 35-43.

Forrester-Gauntlett, B. K. E. (2013). *Developmental gene expression profile of Vmo1 in the mouse auditory system.* thesis, University of Waikato, New Zealand.

Galibert, F., Azzouzi, N., Quignon, P., & Chaudieu, G. (2016). The genetics of canine olfaction. . *Journal of Veterinary Behavior: Clinical Applications and Research, 16*, 86-93.

Galloway, C. A., Sowden, M. P., & Smith, H. C. (2003). Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. *Biotechniques, 34(3)*, 524-6.

Galloway, C. A., Sowden, M. P., & Smith, H. C. (2008). Increasing the yield of soluble recombinant protein expressed in *E. coli* by induction during late log phase. *Biotechniques, 34(3)*, 524-6.

Goldman, A., Harper, S., & Speicher, D. W. (2016). Detection of proteins on blot membranes. *Current protocols in protein science, 86*, 1-14.

Guérin-Dubiard, C., Pasco, M., Mollé, D., Désert, C., Croguennec, T., & Nau, F. (2006). Proteomic analysis of hen egg white. *Journal of Agricultural and Food Chemistry, 54(11)*, 3901-3910.

Guérin-Dubiard, C., & Nau, F. (2007). Minor proteins. *Bioactive Egg Compounds*, 93-98.

Guyot, N., Labas, V., Harichaux, G., Chessé, M., Poirier, J. C., Nys, Y., & Réhault-Godbert, S. (2016). Proteomic analysis of egg white heparin-binding proteins: towards the identification of natural antibacterial molecules. *Scientific reports, 6, 27974.*, 1-11.

Han, X., Aslanian, A., & Yates III, J. R. (2008). Mass spectrometry for proteomics. *Current opinion in chemical biology, 12(5)*, 483-490.

Harada, Y. (2012). Atlas of the Ear: By Scanning Electron Microscopy.

Hardie, S. E. (2015). *Expression and Purification of Recombinant VMO1 Protein.* thesis, The University of Waikato, Hamilton, NZ.

Hervé-Grépinet, V., Réhault-Godbert, S., Labas, V., Magallon, T., Derache, C., Lavergne, M., & Nys, Y. (2010). Purification and characterization of avian β-defensin 11, an antimicrobial peptide of the hen egg. *Antimicrobial agents and chemotherapy, 54(10)*, 4401-4409.

Hibino, H., Nin, F., Tsuzuki, C., & Kurachi, Y. (2010). How is the highly positive endocochlear potential formed? The specific architecture of the stria

vascularis and the roles of the ion-transport apparatus. *Pflügers Archiv-European Journal of Physiology, 459(4)*, 521-533.

Hutchinson, E. G., & Thornton, J. M. (1993). The Greek key motif: extraction, classification and analysis. *Protein Engineering, Design and Selection, 6(3)*, 233-245.

Jing, Z., Yan–chun, J., & Min–Xin, G. (2012). Mitochondrial tRNA mutations associated with deafness. *Journal of Otology, 7(1)*, 36-44.

Kido, S., Morimoto, A., Kim, F., & Doi, Y. (1992). Isolation of a novel protein from the outer layer of the vitelline membrane. *Biochemical journal, 286(1)*, 17-22.

Kido, S., Doi, Y., Kim, F., Morishita, E., Narita, H., Kanaya, S., & Ooi, T. (1995). Characterization of vitelline membrane outer layer protein I, VMO-I: amino acid sequence and structural stability. *The Journal of Biochemistry, 117(6)*, 1183-1191.

Lab, S. C. (2018). *Electrophoresis*. Retrieved 17 March, 2018, from http://www.bioinformatics.nl/molbi/SimpleCloningLab/electrophoresis.htm.

Lavinsky, J., Ge, M., Crow, A. L., Pan, C., Wang, J., Salehi, P., & Friedman, R. A. (2016). The genetic architecture of noise-induced hearing loss: evidence for a gene-by-environment interaction. *G3: Genes, Genomes, Genetics, 6(10)*, 3219-3228.

Lee, D., Bamdad, F., Khey, K., & Sunwoo, H. H. (2017). Antioxidant and anti-inflammatory properties of chicken egg vitelline membrane hydrolysates. *Poultry science, 96(9)*, 3510-3516.

Lee, S. I., Ji, M. R., Jang, Y. J., Jeon, M. H., Kim, J. S., Park, J. K., & Byun, S. J. (2015). Characterization and miRNA-mediated posttranscriptional regulation of vitelline membrane outer layer protein I in the adult chicken oviduct. *In Vitro Cellular & Developmental Biology-Animal, 51(3)*, 222-229.

Lemaire, L., & Kessel, M. (1997). Gastrulation and homeobox genes in chick embryos. *Mechanisms of development, 67(1)*, 3-16.

Li, C., Tan, X. F., Lim, T. K., Lin, Q., & Gong, Z. (2016). Comprehensive and quantitative proteomic analyses of zebrafish plasma reveals conserved protein profiles between genders and between zebrafish and human. *Scientific reports, 6, 24329.*, 1-15.

Li, Q., Li, W., Li, X., Liu, L., Zhang, Y., Guo, Y., & Zheng, J. (2017). The Distribution Characteristics and Applications for Maternal Cells on Chicken Egg Vitelline Membrane. *Scientific Reports, 7.*, 1-11.

Liberman, M. C., Gao, J., He, D. Z., Wu, X., Jia, S., & Zuo, J. (2002). Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature, 419(6904)*, 300-304.

Lim, D. J. (1986). Functional structure of the organ of Corti: a review. *Hearing research, 22(1-3)*, 117-146.

Lim, W., & Song, G. (2015). Differential expression of vitelline membrane outer layer protein 1: hormonal regulation of expression in the oviduct and in ovarian carcinomas from laying hens. *Molecular and cellular endocrinology, 399*, 250-258.

Lim, W. K., Rösgen, J., & Englander, S. W. (2009). Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. *Proceedings of the National Academy of Sciences, 106(8)*, 2595-2600.

Lu, P., Chen, X., Liu, X., Yu, L., Kang, Y., Xie, Q., & Wei, X. (2008). Dry eye syndrome in elderly Tibetans at high altitude: a population-based study in China. *Cornea, 27(5)*, 545-551.

Luscombe, N. M., Greenbaum, D., & Gerstein, M. (2001). What is bioinformatics? An introduction and overview. *Yearbook of Medical Informatics, 1(83-100), 2*.

Mahmood, T., & Yang, P. C. (2012). Western blot: technique, theory, and trouble shooting. *North American journal of medical sciences, 4(9)*, 429–434.

Mann, K. (2008). Proteomic analysis of the chicken egg vitelline membrane. *Proteomics, 8(11)*, 2322-2332.

Marcus, D. C. (2011). *Acoustic transduction.* In Cell Physiology Source Book (Fourth Edition)

Masinick, S. A., Montgomery, C. P., Montgomery, P. C., & Hazlett, L. D. (1997). Secretory IgA inhibits Pseudomonas aeruginosa binding to cornea and protects against keratitis. *Investigative ophthalmology & visual science, 38(5)*, 910-918.

Matlock, B. (2018). *Assessment of Nucleic Acid Purity*. Retrieved 17-03-2018, from https://tools.thermofisher.com/content/sfs/brochures/TN52646-E-0215M-NucleicAcid.pdf.

McNally, E. H. (1943). The origin and structure of the vitelline membrane of the domestic fowl's egg. *Poultry Science, 22(1)*, 40-43.

McPherson, A., & Gavira, J. A. (2014). Introduction to protein crystallization. *Acta Crystallographica Section F: Structural Biology Communications, 70(1)*, 2-20.

Mellor, D. J., & Diesch, T. J. (2007). Birth and hatching: key events in the onset of awareness in the lamb and chick. *New Zealand Veterinary Journal, 55(2)*, 51-60.

Meyer, X., Delevoye, E., & Chopard, B. (2017). Study of fluid flow within the hearing organ. *arXiv preprint arXiv:1709.06792*.

Mittal, N., & Sanyal, S. N. (2011). In vivo effect of surfactant on inflammatory cytokines during endotoxin-induced lung injury in rodents. *Journal of immunotoxicology, 8(4)*, 274-283.

Montgomery, S. C., & Cox, B. C. (2016). Whole mount dissection and immunofluorescence of the adult mouse cochlea. *Journal of visualized experiments: JoVE, 107*, 1-9.

Moran Jr, E. T. (2007). Nutrition of the developing embryo and hatchling. *Poultry Science, 86(5)*, 1043-1049.

Morris, L. M., DeGagne, J. M., Kempton, J. B., Hausman, F., & Trune, D. R. (2012). Mouse middle ear ion homeostasis channels and intercellular junctions. *PloS one, 7(6)*(e39004), 1-9.

Mouse Genome Informatics. (2017). *Vmo1*. Retrieved 12 April, 2017, from http://www.informatics.jax.org/marker/MGI:2685587.

Mullins, L. J., & Mullins, J. J. (2004). Insights from the rat genome sequence. *Genome biology, 5(5)*(221).

National Human Genome Research Institute. (2018). *Why Mouse Matters*. Retrieved 22 March, 2018, from https://www.genome.gov/10001345/importance-of-mouse-genome/.

NCBI. (2018). *Vmo1 vitelline membrane outer layer 1 homolog [Rattus norvegicus (Norway rat)]*. Retrieved 10 March, 2018, from https://www.ncbi.nlm.nih.gov/gene/360553.

NCBI AceView. (2018a). *Homo sapiens complex locus VMO1, encoding vitelline membrane outer layer 1 homolog (chicken).* Retrieved 13 March, 2017, from https://www.ncbi.nlm.nih.gov/ieb/research/acembly/av.cgi?db=human&term=VMO1&submit=Go.

NCBI AceView. (2018b). *Mus musculus gene Vmo1, encoding vitelline membrane outer layer 1 homolog (chicken).* Retrieved 15 March, 2018, from https://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?db=mouse&q=Vmo1.

NCBI Gene. (2018). *VMO1 vitelline membrane outer layer 1 homolog [Homo sapiens (human)]*. Retrieved 15 March, 2018, from https://www.ncbi.nlm.nih.gov/gene/284013/?report=expression.

NIH, N. I. o. H. (Compiler) (2011): National Institute on Deafness and Other Communication Disorders

Accessed 22-02-2018 from https://www.nidcd.nih.gov/health/statistics/what-numbers-mean-epidemiological-perspective-hearing.

Nilsson, J., Larsson, M., Ståhl, S., Nygren, P. Å., & Uhlén, M. (1996). Multiple affinity domains for the detection, purification and immobilization of

recombinant proteins. *Journal of Molecular Recognition, 9(5 - 6)*, 585-594.

Novagen. (2017). *pET-28a-c(+) Vectors*. Retrieved 04 April, 2017, from https://www.staff.ncl.ac.uk/p.dean/pET28.pdf

Novagen. (2017) pET System Manual. 10th Edition Rev.B 0403. Retrieved. 04 April, 2017, from http://lifeserv.bgu.ac.il/wp/zarivach/wp-content/uploads/2017/11/Novagen-pET-system-manual-1.pdf

Oganesyan, N., Kim, S. H., & Kim, R. (2004). On-column chemical refolding of proteins. *PharmaGenomics, 4(7)*, 22-25.

Ohashi, Y., Dogru, M., & Tsubota, K. (2006). Laboratory findings in tear fluid analysis. *Clinica chimica acta,, 369(1)*, 17-28.

Olender, T., Keydar, I., Pinto, J. M., Tatarskyy, P., Alkelai, A., Chien, M. S., & Lancet, D. (2016). The human olfactory transcriptome. *BMC genomics, 17(1), 619.*, 1-18.

Oliveira, P. F. D., Oliveira, C. S., Andrade, J. S., Santos, T. F. D. C., & Oliveira-Barreto, A. C. D. (2016). Cancer treatment in determination of hearing loss. *Brazilian journal of otorhinolaryngology, 82(1)*, 65-69.

Ottinger, M. A., & Bakst, M. R. (1995). Endocrinology of the avian reproductive system. *Journal of Avian Medicine and Surgery*, 242-250.

Panagiotidis, C. A., & Silverstein, S. J. (1995). pALEX, a dual-tag prokaryotic expression vector for the purification of full-length proteins. *Genetics research international, 164(1)*, 45-47.

Park, H. J., Cho, H. J., Baek, J. I., Ben-Yosef, T., Kwon, T. J., Griffith, A. J., & Kim, U. K. (2010). Evidence for a founder mutation causing DFNA5 hearing loss in East Asians. *Journal of human genetics, 55(1)*, 59-62.

Pascucci, T., Perrino, J., Mahowald, A. P., & Waring, G. L. (1996). Eggshell Assembly inDrosophila: Processing and Localization of Vitelline Membrane and Chorion Proteins. *Developmental biology, 177(2)*, 590-598.

Patel, A., & Groppo, E. (2010). Management of temporal bone trauma. . *Craniomaxillofacial trauma & reconstruction, 3(2)*, 105-113.

Peters, L. M., Anderson, D. W., Griffith, A. J., Grundfast, K. M., San Agustin, T. B., Madeo, A. C., & Morell, R. J. (2002). Mutation of a transcription factor, TFCP2L3, causes progressive autosomal dominant hearing loss, DFNA28. *Human molecular genetics, 11(23)*, 2877-2885.

Peters, L. M., Belyantseva, I. A., Lagziel, A., Battey, J. F., Friedman, T. B., & Morell, R. J. (2007). Signatures from tissue-specific MPSS libraries identify transcripts preferentially expressed in the mouse inner ear. *Genomics, 89(2)*, 197-206.

PyMOL™. *The PyMOL Molecular Graphics System*. from https://www.schrodinger.com/pymol.

Raftery, M. A., & Rand-Meir, T. (1968). Distinguishing between possible mechanistic pathways during lysozyme-catalyzed cleavage of glycosidic bonds. *Biochemistry, 7(9)*, 3281-3289.

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: advances and challenges. *Frontiers in microbiology, 5, 172*.

Russell, I. J. (1987). *The physiology of the organ of Corti.* (Vol. 43(4)). British medical bulletin.

Schierenberg, E., & Junkersdorf, B. (1992). Schierenberg, E., & Junkersdorf, B. (1992). The role of eggshell and underlying vitelline membrane for normal pattern formation in the early C. elegans embryo. Roux's archives of developmental biology, 202(1), 10-16. *The role of eggshell and underlying vitelline membrane for normal pattern formation in the early C. elegans embryo., Roux's archives of developmental biology*(202(1)), 10-16.

Shalit, E., & Avraham., K. B. (2008). Genetics of hearing loss. *Auditory Trauma, Protection, and Repair. Springer US*, 9-47.

Shamsi, F. A., Chen, Z., Liang, J., Li, K., Al-Rajhi, A. A., Chaudhry, I. A., & Wu, K. (2011). Analysis and comparison of proteomic profiles of tear fluid from human, cow, sheep, and camel eyes. *Investigative ophthalmology & visual science, 52(12)*, 9156-9165.

Shearer AE, Hildebrand MS, & RJH., S. (1999). *Hereditary Hearing Loss and Deafness Overview.* University of Washington, Seattle: GeneReviews®

Sheng, G. (2014). Day - 1 chick development. *Developmental Dynamics, 243(3)*, 357-367.

Shimizu, T., Vassylyev, D. G., Kido, S., & Morikawa, K. (1994). Crystal structure of vitelline membrane outer layer protein I (VMO-I): a folding motif with homologous Greek key structures related by an internal three-fold symmetry. *The EMBO journal, 13(5)*, 1003-1010.

Sidhu, S. S., Fairbrother, W. J., & Deshayes, K. (2003). Exploring protein–protein interactions with phage display. *Chembiochem, 4(1)*, 14-25.

SIGMA-ALDRICH. (2018). *Poly-Prep Slides*. Retrieved 22 March, 2018, from https://www.sigmaaldrich.com/catalog/product/sigma/p0425?lang=en&region=NZ.

Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J., & Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using Escherichia coli. *Protein Science, 18(5)*, 936-948.

Song, M. H., Lee, K. Y., Choi, J. Y., Bok, J., & Kim, U. K. (2012). Nonsyndromic X-linked hearing loss. *Frontiers in bioscience (Elite edition), 4*, 924-933.

Stein, K. K., & Golden, A. (2015). The C. elegans eggshell. *WormBook: the online review of C. elegans biology, 1.*, 1-35.

Subha, S. T., & Raman, R. (2006). Role of impacted cerumen in hearing loss. *Ear, nose & throat journal, 85(10)*, 650-653.

Systems, R. D. (2018). *Protocol for the Preparation of Gelatin-coated Slides for Histological Tissue Sections*. Retrieved 22 March, 2018, from https://www.rndsystems.com/resources/protocols/protocol-preparation-gelatin-coated-slides-histological-tissue-sections.

Teo, Y. Y., Small, K. S., Clark, T. G., & Kwiatkowski, D. P. (2008). BRLMM: An improved genotype calling method for the GeneChip human mapping 500K array set. *ANNALS OF HUMAN GENETICS, 72*, 368-374.

The American Speech-Language-Hearing Association (ASHA) (2018). *Types of Hearing Loss*. Retrieved 23 March, 2018, from https://www.asha.org/public/hearing/types-of-hearing-loss/.

The National foundation for the Deaf. (2017). *Social and economic costs of hearing loss in New Zealand*. Retrieved 25 October, 2017, from https://www2.deloitte.com/content/dam/Deloitte/au/Documents/Economics/deloitte-au-economics-social-economic-cost-hearing-loss-new-zealand-021216.pdf.

ThermoBlog. (2018). [Blog]. Retrieved 27 March, 2018, from https://blog.thermoworks.com/2017/03/easter-hard-boiled-eggs/.

ThermoFisher Scientific. (2018). *Pull-Down Assays*. Retrieved 22 March, 2017, from https://www.thermofisher.com/nz/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/pull-down-assays.html.

Thomas Scientific. (2018). *Positive Charged Slides*. Retrieved 22 March, 2018, from https://www.thomassci.com/scientific-supplies/Positive-Charged-Slides.

Tong, Q., Romanini, C. E., Exadaktylos, V., Bahr, C., Berckmans, D., Bergoug, H., & Demmers, T. (2013). Embryonic development and the physiological factors that coordinate hatching in domestic chickens. *Poultry science, 92(3)*, 620-628.

Touchman, J. (2010). Comparative Genomics. *Nature Education Knowledge, 3(10), 13.*

Van Dijk, E., Hoogeveen, A., & Abeln, S. (2015). The hydrophobic temperature dependence of amino acids directly calculated from protein structures. *PLoS computational biology, 11(5), e1004277.*, 1-17.

Van Laer, L., Huyghe, J. R., Hannula, S., Van Eyken, E., Stephan, D. A., Mäki-Torkko, E., & Huentelman, M. J. (2010). A genome-wide association study for age-related hearing impairment in the Saami. *European Journal of Human Genetics, 18(6)*, 685–693.

Vera, A., González - Montalbán, N., Arís, A., & Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and bioengineering, 96(6)*, 1101-1106.

Wang, Z., Chen, Z., Yang, Q., Jiang, Y., Lin, L., Liu, X., & Wu, K. (2014). Vitelline Membrane Outer Layer 1 Homolog Interacts With Lysozyme C and Promotes the Stabilization of Tear FilmVMO1 Promotes Stable Tear Film. *Investigative ophthalmology & visual science, 55(10)*, 6722-6727.

Watson, C., Paxinos, G., & Puelles, L. E. (2012). *The mouse nervous system.* Academic Press.

WHO, W. H. O. (2015).

WHO World Health Organisation. (2017). Deafness and hearing loss. Retrieved 03 September, 2017.

Wu, T., Manogaran, A. L., Beauchamp, J. M., & Waring, G. L. (2010). Drosophila vitelline membrane assembly: a critical role for an evolutionarily conserved cysteine in the "VM domain" of sV23. *Developmental biology, 347(2)*.

www.leicabiosystems.com. (2018). Retrieved 09-03-2018

www.thepoultrysite.com. (2018). Retrieved 27-02-2018.

www.thermofisher.com. (2018). Retrieved 09-03-2018.

Xing, G., Chen, Z., & Cao, X. (2007). Mitochondrial rRNA and tRNA and hearing function. *Cell research, 17(3)*, 227-239.

Zhang, C., & Kim, S. H. (2000). A comprehensive analysis of the Greek key motifs in protein β - barrels and β - sandwiches. *Proteins: Structure, Function, and Bioinformatics, 40(3)*, 409-419.

Zhou, L., Huang, L. Q., Beuerman, R. W., Grigg, M. E., Li, S. F. Y., Chew, F. T., & Tan, D. (2004). Proteomic analysis of human tears: defensin expression after ocular surface surgery. *Journal of proteome research, 3(3)*, 410-416.

Zhou, L., Zhao, S. Z., Koh, S. K., Chen, L., Vaz, C., Tanavde, V., & Beuerman, R. W. (2012). In-depth analysis of the human tear proteome. *Journal of proteomics, 75(13)*, 3877-3885.

Hardie, S. E. (2015). *Expression and Purification of Recombinant VMO1 Protein.* thesis, The University of Waikato, Hamilton, NZ.

Simple Cloning Lab. (2018). *Agarose gel electrophoresis*. Electrophoresis. Retrieved 17 March, 2018, from http://www.bioinformatics.nl/molbi/SimpleCloningLab/electrophoresis.htm.

World Health Organisation. (2017). Deafness and hearing loss. Retrieved 03 September, 2017.