Expression and purification of immunity-related GTPase protein family M (IRGM) for structural studies

Sarah Sabboobeh

Department of Biochemistry

Groupe de Recherche Axé sur la Structure des Protéines (GRASP)

McGill University

Montreal, QC

September 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science under the Faculty of Medicine

“You know, Prince Zuko, destiny is a funny thing. You never know how things are going to work out. But if you keep an open mind and an open heart, I promise you will find your own destiny someday.

“

--- General Iroh, from Avatar: The Last Airbender

Contents

List of Tables and Figures

Preface

This thesis is based on original work and details the research carried out solely by its author, Sarah Sabboobeh, under the supervision of Dr. Bhushan Nagar.

List of Abbreviations

Ala	alanine
AMPK	adenosine monophosphate kinase
ATG	autophagy-related protein
βME	β-mercaptoethanol
CD	Crohn's disease
Cys	cysteine
DNA	deoxyribonucleic acid
DPG	diphosphatidylglycerol
DTT	dithiothreitol
ER	endoplasmic reticulum
GAP	GTPase activating protein
GBP	guanylate binding protein
GDP	guanosine diphosphate
GEF	guanine nucleotide exchange factor
GFP	green fluorescent protein
Glu	glutamic acid/glutamate
Gly	glycine
GppNHp	5'-guanylyl imidodiphsphate
GST	glutathione S-transferase
GTP	guanosine triphosphate
His	histidine
IBD	irritable bowel disease
IFN-γ	interferon- gamma
IIGP1	interferon-gamma-inducible GTPase protein-1
IPTG	isopropyl-β-D-thiogalactoside
IRG	Immunity-related GTPase
IRGM	Immunity-related GTPase Family M
LC3	mictrotubule-associated protein 1A/1B-light chain 3
LPS	lipopolysaccharide
MDP	muramyl dipeptide

MPG	mycobacterial phagosome
Mtb	*Mycobacterium tuberculosis*
mRNA	messenger RNA
mTOR	mammalian target of rapamycin
NOD2	nucleotide-binding oligomerization domain-containing protein 2
NOS	nitric oxide synthase
ORF	open reading frame
PAMP	pathogen-associated molecular pattern
PBS	phosphate buffered saline
PMSF	phenylmethylsulfonyl fluoride
PRR	pattern recognition receptor
PtdIns	phosphoinositides
RNA	ribonucleic acid
ROS	reactive oxygen species
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SHIP1	SH2-containing inositol-5'-phosphatase 1
siRNA	small interfering RNA
SNP	single nucleotide polymorphism
Strep	streptadivin
TEV	Tobacco Etch virus
TLR	Toll-like receptor
Ub	ubiquitin
ULK1	unc-51 like autophagy activating kinase 1

Abstract

Immunity-related GTPase Family M protein (IRGM) is involved in regulating the process of cellular autophagy. Recent studies in genetics have shown that underexpression of IRGM is associated with the incidence of Crohn's disease and increased susceptibility to infection by *Mycobacterium tuberculosis*. The protein's mechanism of action is thought to occur by recognition of mitochondrial lipid, and inducing mitochondrial fission to instigate autophagy. The function of IRGM has been shown to be pathogen-specific since there has been no clear role of its function under conditions without infection.

This thesis focuses on biochemical analysis of human and murine *IRGM* proteins. I report the different conditions under which a pure stable form of the *IRGM* protein can be produced for structural determination with X-ray crystallography. The optimal expression system for these proteins was in Sf9 insect cells: murine Irgm1 constructs and IRGM(d) isoform express far better than their IRGM(a) counterpart. Generally, low salt buffers (150 mM NaCl) are most favorable for both the human and murine proteins. Murine Irgm1 requires the presence of 0.1% Tween-20 for stability. Additionally, it is possible to obtain the pure form of the construct IRGM(23-181) from bacterial cells that can be refolded in 2M urea. Finally, I report the presence of numerous degradation products associated with both the human and mouse proteins, illustrating solutions that were attempted to address this problem. These solutions include using different protease inhibitors as well as utilizing a C-terminal tag that can serve as a protective entity against proteases.

Résumé

La protéine M de la famille des GTPases liées à l'immunité (IRGM) est impliquée dans la régulation du processus cellulaire d'autophagie. Récemment, des études génétiques ont révélé le lien entre la sous-expression d'IRGM et l'incidence de la maladie de Crohn ainsi que la susceptibilité à l'infection par *Mycobacterium tuberculosis.* Le mode d'action de cette protéine passerait par la reconnaissance de lipides mitochondriaux et l'initiation de la fission mitochondriale afin de débuter l'autophagie. La fonction d'IRGM a aussi été démontrée être spécifique en réponse aux agents infectieux, et aucun rôle n'y est clairement attribué en l'absence d'infection.

Cette thèse se concentre sur l'analyse biochimique de la protéine IRGM de la souris et de l'humain. Je présente les différentes conditions sous lesquelles une forme stable et pure d'IRGM peut être produite pour la détermination structurale par cristallographie aux rayons X. Le système d'expression optimal était en cellules d'insecte Sf9. Irgm1 de la souris et l'isoforme IRGM(d) étaient produites bien mieux que leur homologue IRGM(a). Généralement, les solutions tampon à faible concentration de sel (150 mM NaCl) sont plus appropriées pour ces protéines. Irgm1 de la souris requiert la présence de 0,1% de Tween-20 pour la stabilité. De plus, il est possible d'obtenir une forme pure de la construction d'IRGM(23-181) produite en *E. coli*, qui peut être repliée dans une solution de 2M d'urée. Enfin, je signale la présence de nombreux produits de dégradation associés à ces protéines de la souris et de l'humain, illustrant quelques méthodes que j'ai employées pour parer à ce problème, telles l'utilisation d'inhibiteurs de protéases ou d'une étiquette C-terminale.

Acknowledgements

I would first like to thank my supervisor, Dr. Bhushan Nagar for giving me this great opportunity to participate in a high-end research laboratory that will ultimately help advance my scientific career. More importantly, I would like to acknowledge his guidance throughout my research project, his support and advice on my writing and presentations. For someone who wishes to pursue a career in journalism and writing, I cannot be more grateful for the helpful literary pointers Dr. Nagar has offered me throughout the course of my studies. I would like to point out that there are rarely leaders in science who exude kindness, compassion and understanding towards their students and staff but Dr. Nagar has been an inspiration and a leadership role model in that regard.

The advancement of this project and my professional improvement in terms of lab skills and scientific thinking could not have been possible were it not for our lab manager and research assistant, Katalin Illes Kocsis as well as Jon Labriola - both of whom have been mentors and great role models right from the start. I would like to acknowledge Alexei Gorelik and Ahmad Gebai for maintaining the underworkings of our lab, for ensuring buffers and stocks are always available, for being there to answer questions and more importantly, for the fun and mischievous times during late nights at the lab. Big thanks to Christina Dimech for unrelenting moral support. Also, to Yazan Abbas for guidance and help on the AKTA purifiers.

I would like to acknowledge members of my Research Advisory Committee Drs. Jason Young and Martin Thomas Schmeing as well as the generous support from the Chemical Biology training initiative and GRASP.

Additionally, members from the Berghuis lab, Barry Sleno and Dimitro Rodinov for advice on molecular biology techniques, maintaining equipment on our floor and for the encouragement and warmth that really motivated me to toil through the slow and unrewarding days.

There is no one to thank most for all of this other than my mother who stood by me through the good times and the bad times. And also my father for his insights and inputs.

Thank you to my wonderful friend and colleague Nikita Desai from Dr. Michel Tremblay's lab at the Goodman Cancer Centre. Finally, it goes without saying, a heartfelt thanks to my wonderful sisters, friends and those whose love touched my life and kept me going.

Prologue: The Innate Immunity

With millions and millions of pathogens around and in contact with us, it is quite miraculous that we do not contract an infection every day. Our bodies have biologically evolved numerous mechanisms to protect us from harmful microorganisms. We are exposed to pathogens through the food we consume, the air we breathe and the objects we touch with our skin: ingestion, inhalation and direct contact, respectively. We have thus evolved exterior "shields" to protect us from the multiple possible ways of infection: skin to shield our blood vessels, nerve endings and organs, enforced by a fat layer to strengthen the barrier; protective mucosal membranes lining our gut and lungs; acidic fluids expelled from tear ducts, saliva and sweat glands to destroy any pathogen lurking on the surface, seeking entry.

Pathogens that succeed in overcoming these primary defenses are courageously met with our body's first interior line of defense: the innate immune system. White blood cells, namely macrophages, circulating in the blood stream come into contact with specific signals marking the invading pathogen and they become consequently activated. In their activated state, macrophages begin to engulf the pathogens to destroy them while simultaneously signaling for enforcements in the form of neutrophils and dendritic cells. This cascade of activation, engulfment of pathogen and recruitment of more white blood cells creates an inflammatory response that alerts the body to an infection. Dendritic cells then engulf and deliver remnants of the pathogen to the T-cells of our adaptive immune system. The latter develops with age as we become exposed to more pathogens. The adaptive immune system retains memory of each specific pathogen so that further infection is prevented in the future. But when our bodies encounter an unfamiliar entity, it is our innate immunity that springs to life to save us within just a few hours of the infection – a critical time for the body since pathogens can multiply at rapid rates.

Mutations in the genes encoding the protein machinery involved in innate immunity can have drastic results that manifest in diseases such autoimmune disorders; susceptibility to virulent pathogens – notoriously tuberculosis that affects around 9 million people globally, with 1.5 million deaths worldwide; certain types of cancer; an entire family of disorders that fall under irritable bowel disease. It has thus become of vital importance for us to gain heavy insight into the inner workings of the immune system - starting from its genes and corresponding proteins and finally to its interaction as a whole system and how it ties in to our adaptive immunity.

Chapter 1: The Role of Immunity Related GTPases (IRG's) in the Innate Immunity and their Implication in Disease

1.1 Introduction: Autophagy, Phagocytosis and the Innate Immunity

The innate immune system represents our first line of internal defense against foreign particles. These particles can be inanimate or living organisms and may be virulent pathogens from bacterial, viral and protozoan origin. Our immune system has devised many mechanisms that operate in organized hierarchal steps that ultimately serve to eliminate the invader.

Upon entry, the pathogen is met with and recognized by macrophages. Pathogen-associated molecular patterns (PAMPs), a tier of small molecules specific to pathogens, are recognized by pathogen recognition receptors (PRRs) such as Toll-Like Receptors (TLRs) and NOD-like receptors (NLRs) on the macrophage's surface and in its cytoplasm, respectively. Recognition triggers a series of downstream events that lead to the activation of proteins involved in the inflammatory response and that ultimately lead to downstream events that rid the

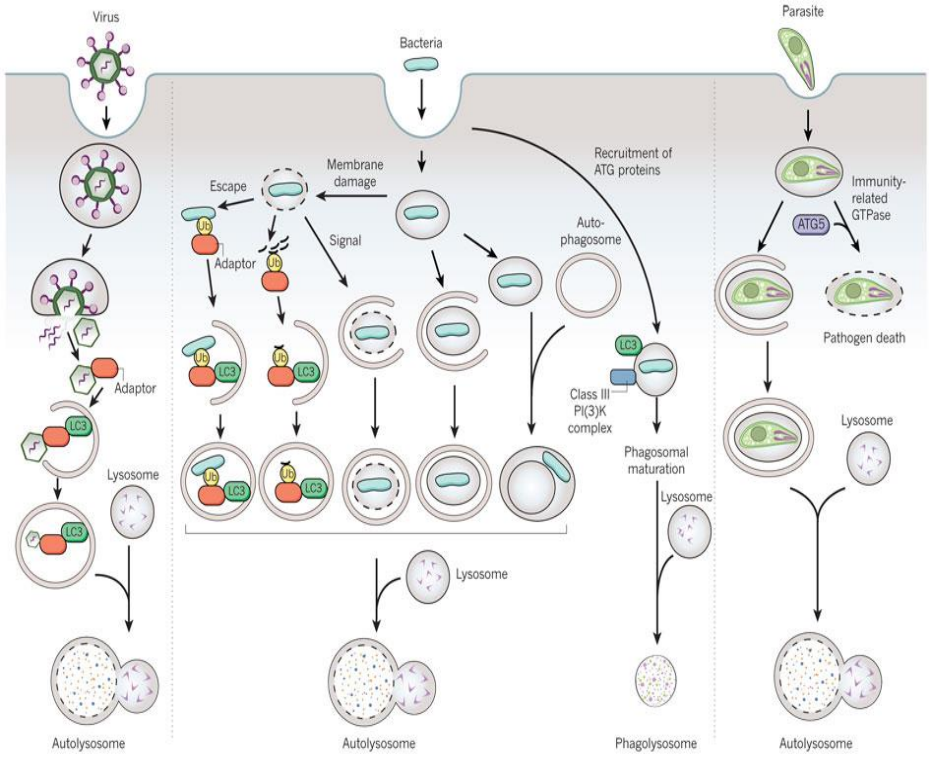

Figure 1: Pathogen-specific xenophagy pathways (Levine et al. 2011)

body of the microorganism. These downstream actions include phagocytosis and autophagy, where the cell first engulfs the pathogen (phagocytosis) and then delivers the engulfed pathogen to the late endosomes and autolysosomes for digestion by proteases (autophagy, or under conditions of infection, "xenophagy"), see Figure 1.

Autophagy is generally a regulatory process whereby cells commit to degrading their own selves or some of their components and organelles. Autophagy functions to rid cells of large entities such as whole organelles, engulfed pathogens and protein aggregates that are too large to be broken down by the proteasome and was initially thought to have evolved as a coping mechanism for starvation conditions in simple and complex organisms. However, in complex eukaryotic organisms, autophagy has evolved to work in conjunction with the immune and inflammatory response: The scope of this thesis will mainly encompass the role of autophagy in infection and inflammation.

Upon infection, TLRs on the cell surface bind to pathogenic PAMPs such as bacterial LPS – a molecule that is a membrane constituent of Gram-negative bacteria. Binding and activation of TLR's activate the protein complex NF-κB and the MyD88 pathway, thus eliciting an inflammatory response (Xu et al., 2007). NF-κB initiates transcription of proinflammatory cytokines and interferons such as IFNα/β and type II interferons IFN-γ.

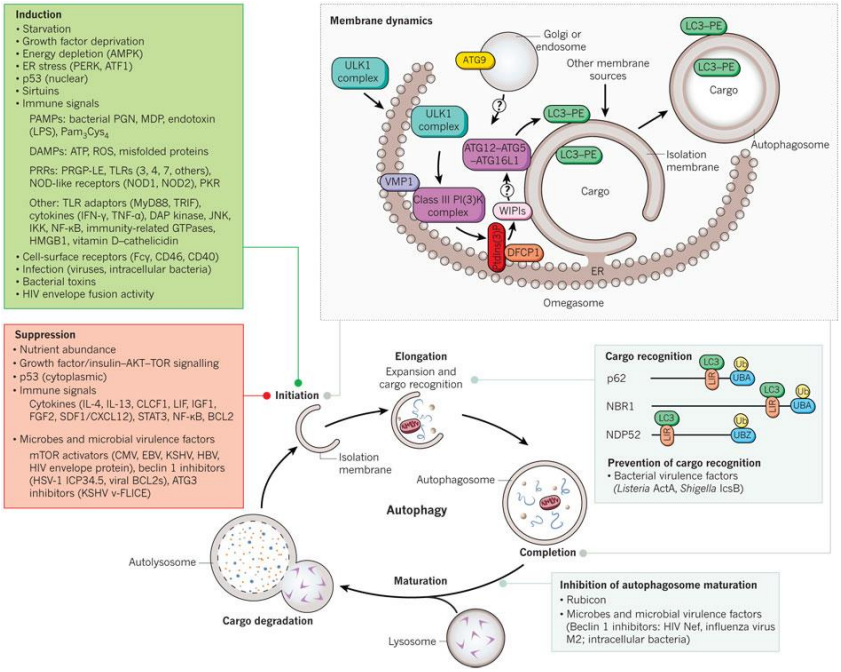

Figure 2: Autophagy pathway and autophagosome formation (Levine et al. 2011)

Additionally, infection inhibits the mTOR pathway leading to the activation and translocation of the ULK1 complex to the ER, where it associates with VMP1. Simultaneously, class III PI(3)K which contains Beclin1, VPS43 and other ATG autophagy factors also translocate to the ER to associate with VMP1, consequently becoming activated by the ULK1 complex. Activated PI(3)K complex begins to produce PtdIns(3)P on the membrane of the ER, which help recruit double FYVE-containing protein 1 (DFCP1) and WD-repeat domain phosphoinositide interacting proteins WIPIs. DFCP1 and WIPIs recruit autophagy factors ATG12-ATG5 as well as ATG16L1 (Itakura et al., 2010).

In the meantime, IFN-γ facilitates the conjugation of the autophagic marker LC3-I with phosphatidylethanolamine (PE), through JAK1/2 and p38 MAPK, independent of STAT1. The presence of the assembled ATG12-ATG5 complex is also vital for LC3 conversion. LC3-PE (LC3-II) is then recruited to phagosomes and nascent autophagic vesicles that are budding from the ER, the "omegasomes". LC3-II on the autophagosomal membrane is responsible for the elongation and maturation of the vesicles and ultimately dissociation from the ER (Matsuzawa et al., 2012).

Furthermore, IFN-γ, through activation of p38 MAPK, facilitates regulation and binding of P38IP to Atg9. Atg9 is a transmembrane protein that shuttles between the Golgi and the ER, forming pre-autophagosomal vesicles that can provide preliminary membranous material for the formation of the autophagosome. Atg9 is also phosphorylated and activated by Atg1 (Papinski et al., 2014).

Another role for the IFN-γ is to upregulate the expression of the interferon-inducible p47 GTPases that are downstream effectors of autophagy. The p47 GTPases play a vital role in eradicating an invading pathogen through the pathway of autophagy. However, their exact mechanism of action is not very well understood, although a few recent studies have presented data and models that provide some insight into their function. These studies are mainly on the murine Irgm1 and its human homologue IRGM, which will be discussed further in upcoming sections of this chapter.

In conclusion, the processes and pathways described above converge and intersect to form the autophagosome which then fuses with the lysosome to create autolysosomes. In case of an infection, the phagosome containing the pathogen might either fuse with the autophagosome or be enveloped by the autophagosome and then mature into the autophagolysosome – the ultimate desired product which will house the pathogen in the necessary conditions to eradicate it.

The autolysosome comprises an acidic environment of pH 4.8- 5.5 in which digestive enzymes such as hydrolases, proteases and nucleases can function optimally to digest pathogenic DNA, proteins and other components. Additionally, macrophages employ oxygen-dependent eradicating mechanisms such as the production of NO via the nitrogen oxide synthetase 2 (NOS2) pathway and the production of reactive oxygen species (ROS) in the autophagolysosome

and the phagosome (Ng et al., 2004). Both NO and ROS are potently toxic compounds for pathogens.

Virulent pathogens like *Mycobacteria tuberculosis* are able to repair nitroxidative damage, thereby persisting in the infection and taking up residence in the autophagosome. Other pathogens like the *Salmonella* family interfere with cellular trafficking to prevent fusion of phagosome to lysosome, thus evading degradation. The presence of these persistent vacuolar pathogens gives rise to the need for specialized proteins that can drive an alternative immune pathway – a role recently discovered to belong to the p47 GTPase family.

1.2 Interferon-Inducible p47 GTPases

Virulent bacterial and protozoan species can succeed in disrupting several molecular pathways elicited by the host to contain infection and eradicate the invader. Upon doing so, the pathogen is free to reside in the phagosomal vacuole and/or hijack the cellular machinery to escape into the cytoplasm, as is the case with the protozoan *Trypanosoma cruzi*. However, according to recent research (Collazo et al., 2001 & Singh et al., 2006), protein members of the interferon-inducible p47 GTPase family are loaded onto phagosomal vesicles and possess the ability to recognize specific pathogens and then initiate an eradication through the activation of the autophagic process - such is the case with the *IRGM* proteins investigated in this thesis. However, the exact mechanism of recognition and initiation of autophagy that these proteins carry out is not yet fully understood.

Upon infection, IFN-γ, a cytokine that is a product of the NF-κB inflammatory response, binds to its receptors on the cell surface in macrophages and other cells. Binding of IFN-γ activates JAK proteins tethered to the receptor, allowing JAK to activate STAT by phosphorylation. Upon activation, STAT dimerizes and translocates to the nucleus where it activates transcription of genes of proteins that are involved in the immune response (Darwich et al., 2009). Among such proteins are the aforementioned nitrogen oxide synthetase 2 NOS2, phagocyte oxidase, phox (that gives rise to ROS), GBP (a class of proteins that mediate antiviral activity) and finally, the p47 GTPases (See Fig. 3). The role of p47 GTPases was not very clear until recent studies showed that individuals who have aberrant loading of p47 GTPases on phagosomal vesicles, or possess mutations in p47 GTPases are highly susceptible to infection by *Mtb* and *Toxoplasma gondii.*

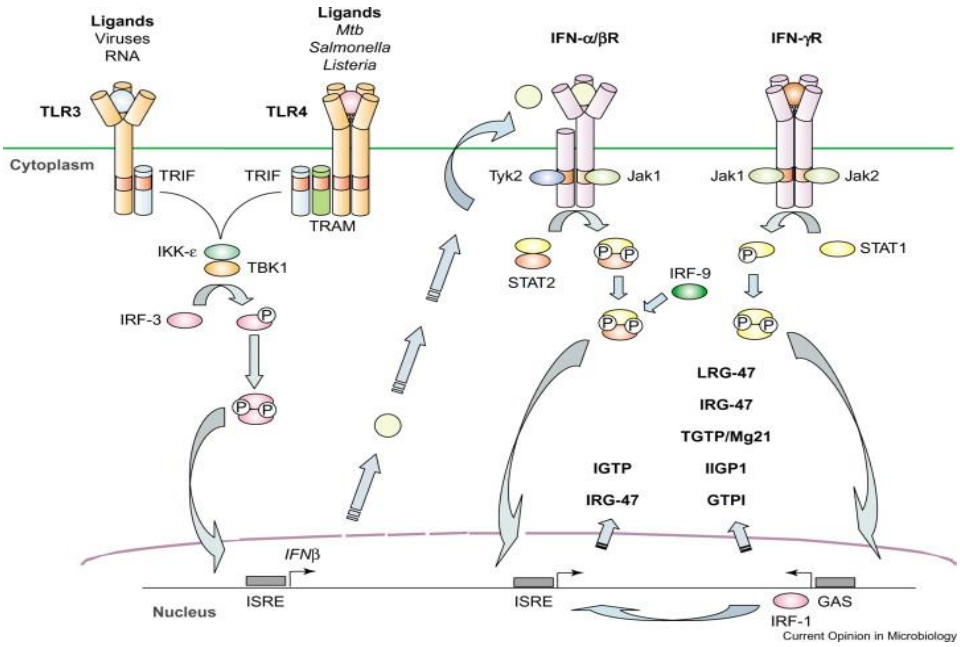

Figure 3: IFN-gamma initiates an inflammatory response by activating JAK/STAT pathway and leading to the transcription of over a thousand proteins involved in containing an infection. Among these proteins are NOS, phox and the p47 GTPases (Macmicking, 2005)

The p47 GTPaseses are ~47 kDa in size and were first characterized in mouse (*Mus musculus).* Interestingly, there are no p47 GTPase homologs in yeast – the first system to give us clues on the process of autophagy as well as the existence of autophagy factors and proteins (Atgs). The p47 GTpases are found to be present in several different species of mammals, birds and fish.

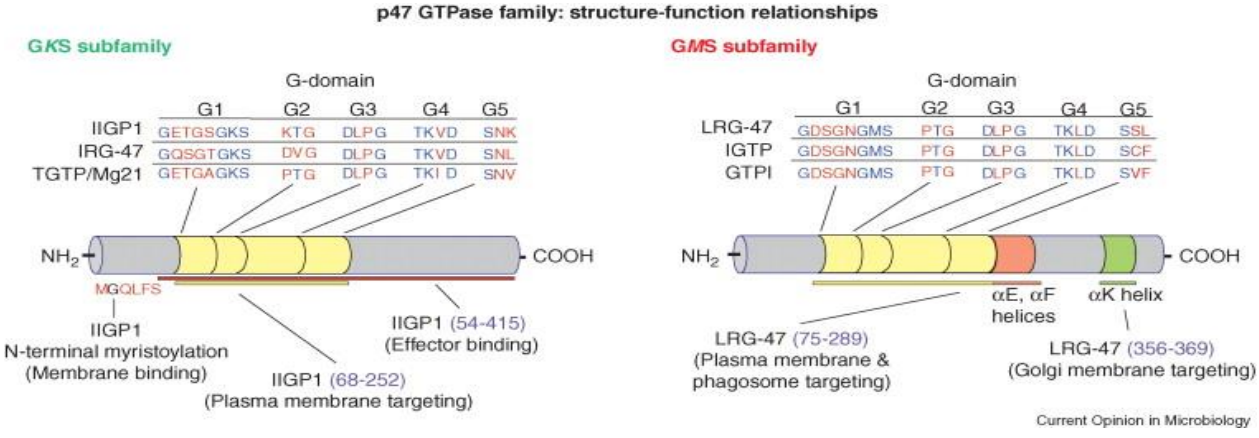

Figure 4: Domain organization of the p47 GTPase family (MacMicking, 2005)

So far, 22 genes have been discovered in mouse that code for the p47 GTPases but only 6 have been characterized: IRG-47, LRG-47 (Irgm1), TGTP/Mg21, IGTP, IIGP1 and GTPI. They share 26-63% sequence identity and they are divided into two classes based on the composition of their GTPase domain, mainly in the G1 motif as illustrated in Fig 4. Table 1 discusses the anatomy of the G-domain of the p47 GTPases and the evolutionary function of the motifs within the domain.

Table 1: Outlining the anatomy of the p47 GTPase family G-domain (MacMicking, 2005)

Domain	**Function**	**Motif**
G1	Hydroxylates the gamma Pi of GTP to help stabilize P-loop	G(x4)GK(S/T)
G2 and G5		Loosely conserved among both subfamilies
G3		(D/N)(x2)G
G4	Makes contact with the GDP or GTP	(N/T)(K/Q)xD

Additionally, p47 GTPases have been shown to possess membrane-binding properties and specific roles in infection, conferring pathogen-specificity. Table 2 summarizes the localization, roles and properties of the 6 characterized p47 GTPases.

Table 2: Summary of the 6 characterized murine p47 GTPases (Ghosh et al, 2004; Martens, et al., 2005; MacMicking, 2005; Henry, et al. 2014)

p47 GTPase	Subfamily	Localization	GTP/GDP binding ratio	Membrane Binding Signals	Role in infection
IIGP1	GKS	Cytosol, ER	GDP > GTP	N-term myristoylation on Gly2	Resistance against *Plasmodium berghei* *Toxoplasma gondii*
IGTP	GMS	ER (calnexin-positive regions), Lipid Bodies (LB)	90-95% GTP-bound	No data.	Loaded onto parasitophorous vacuole containing *Toxoplasma gondi* and *Leishmania major* LRG-47 is specifically resistant to and is loaded on phagosome containing *Salmonella typherium; Listeria monocytogenes; Mycobacteria avium; Mycobacteria tuberculosis.* Accelerates the acidification of phagosome.
LRG-47/Irgm1	GMS	*Cis*- and *medial* -Golgi	Shows GTPase activity but no ratio determined.	Amphipathic αK helix and palmitoylation on Cys371,373,374, 375 close to C-terminal	
TGTP/Mg21	GKS	Cytosol, membrane	Shows GTPase activity but no ratio determined.	No data.	Specific resistance to vesicular stomatitis virus
IRG-47	GKS	Cytosol	No data.	No data.	Chronic resistance against *Toxoplasma gondii*
GTPI	GMS	Golgi	No data.	No data.	Unknown

1.2.1 Structural Mechanisms of GTPases: the Ras Model

Proteins belonging to the class of GTPases, including the p47 IRGM GTPases discussed here, possess a G-domain that is capable of binding GDP in its inactive state and allowing it to be exchanged with GTP for activation. GTPases performing a regulatory role in cells belong to

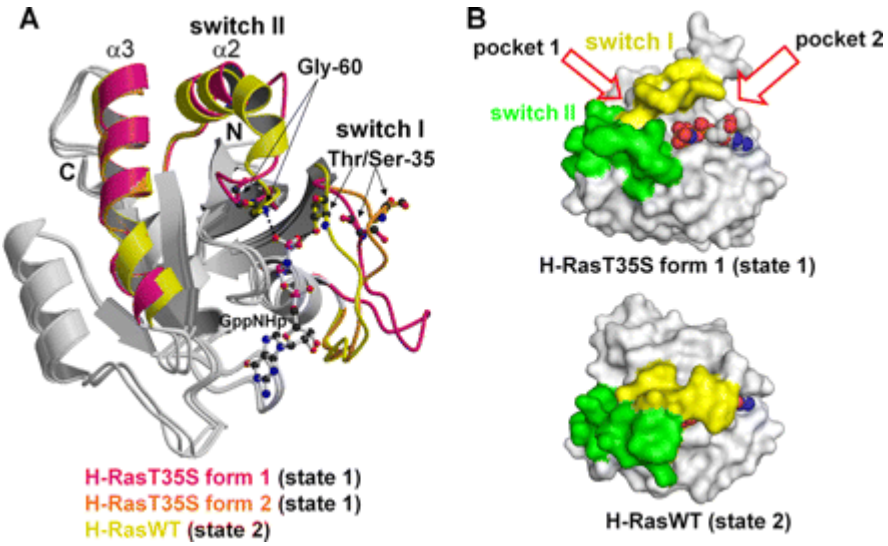

Figure 5: Structural mechanism of switch I and switch II states of H-Ras (Shima et al, 2010).

Ras-GTPase superfamily. The mechanism for the function of a GTPase is illustrated in Figure 5. Cycling between the active and inactive states (GTP-bound and GDP-bound, respectively) drives surrounding chemical reactions. GTPase Activating Proteins, GAPs, act as accessory proteins that facilitate the hydrolysis of GTP by stabilizing the binding pocket further. On the other hand, Guanosine Exchange Factors, GEFs, facilitate the exchange of GDP to GTP to activate the GTPase. The enzymatic mechanism for GTP hydrolysis has been worked out in detail for Ras GTPase family (Maegly et al, 1996; Du et al., 2000; Shima et al 2010). Briefly, a GTPase hydrolyzes GTP in an Sn2 mechanism. A water molecule attacks the P on the gamma phosphate of GTP and produces inorganic phosphate (Pi) and energy. The GTP hydrolysis and the water molecule are stabilized within the binding pocket of the GTPase. The mechanism of GTP hydrolysis is dissociative: negative charges shift towards the beta phosphate (from the beta-gamma bridge) upon nucleophilic attack of the water molecule and consequently, the positive charges on the backbone of the binding pocket (Val14, Gly15, Lys16 and the Mg^{2+} cofactor) help stabilize these negative charges. Simultaneously, Gln61 orients the water molecule in the switch II region (residues 59-97) so that the nucleophilic attack can take place and the transition state is stabilized. Gly13 plays a big role in the P-loop to further stabilize interactions by forming a strong hydrogen bond to the beta phosphate of GDP. Additionally, Thr35 in the Switch I region (residues 32-38) and Gly60 in the switch II region (G3 motif) stabilize the gamma phosphate of GTP.

The binding of cofactor Mg^{2+} induces a conformational change in the switch I region, causing it to adopt an open conformation to receive a GTP molecule. Loss of a side chain methyl group on Thr35 allows switch I to dissociate from the gamma phosphate, causing a

conformational change that permits Gly60 to dissociate, thereby releasing Pi and retaining the GDP molecule that can be exchanged for GTP to restart the cycle.

1.2.2 Structural Characterization of the p47 GTPase IIGP1

In 2004, Ghosh et al. solved the only crystal structure of a p47 GTPase to-date: the murine Interferon-Inducible p47 Resistance GTPase (IIGP1). This protein shares 16% and 28% sequence identity with IRGM and murine Irgm1, respectively. The structure of IIGP1 may thus provide some insight into the function of the *IRGM* proteins. IIGP1 crystallizes as a dimer and its structure has been solved with different ligands: nucleotide-free, GDP- and GppNHp- bound (at 2.3, 2.0 and 2.7Å, respectively).

The structure distinctively shows the two domains of IIGP1: a G domain and a helical domain. The G domain has 6 beta strands and 6 alpha helices ((residues 67-252), similar to the G domain of a Ras-GTPases. The helical domain consists of 10 alpha helices, with αA, B and C (residues 14-65) forming the N-terminal region prior to the G domain. Helices αF-L (residues 265-413) make up the C-terminal region following the G domain. The helix αE serves as a linker between the G-domain and αF.

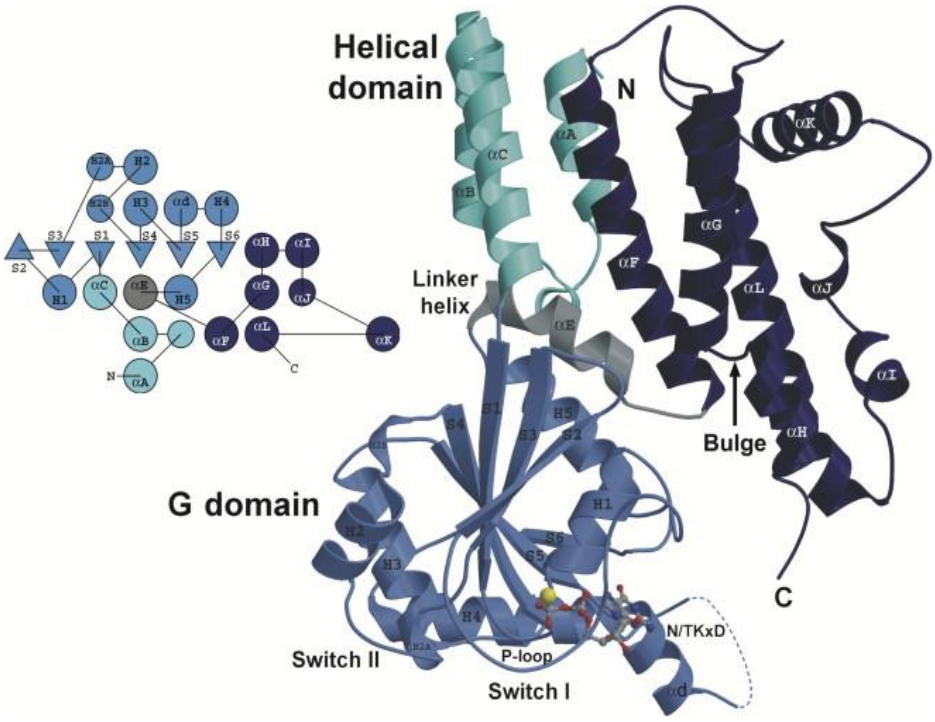

Figure 6: 3D structure of IIGP1 monomer (Ghosh et al., 2004)

IIGP1 has a stretch of negatively-charged amino acid residues on one side of its surface which encompass αH-K (Glu324, Glu335, Glu337, Asp339, Glu342, Asp355, Glu356, Glu357

and Glu361). In addition, IIGP1 becomes myristoylated at Gly2, which help anchor it onto the ER membrane.

The group also found that the mutant IIGP1 K43A (GDP-bound) and M173A (GppNHp-bound) abolished IIGP1's oligomerization ability, thus crystallizing as monomers. Within the mutants, the α- and β-phosphates of the nucleotides form a different set of hydrogen binds to the Switch I region and the Mg^{2+} cofactor. The Mg^{2+} in both mutants is loosely bound in an anionic pocket formed by Asp126(G3), Ser38 and the β- and γ-phosphates.

In the IIGP1wt-GDP, however, an equivalent to the Thr35 in Ras that coordinates the gamma phosphate was absent. Asp126 was observed to partly perform the role of Thr35 in Mg^{2+} coordination. Furthermore, IIGP1 coordinates the γ-phosphate through Thr78 in the P loop, which corresponds to Gly12 in Ras. Thr108 in IIGP1wt-GppNHp is hypothesized to play a catalytic role in the activation or coordination of the water molecule, similar to Gln61 in Ras.

1.3 Murine Irgm1/LRG-47

1.3.1 Localization and Membrane Association

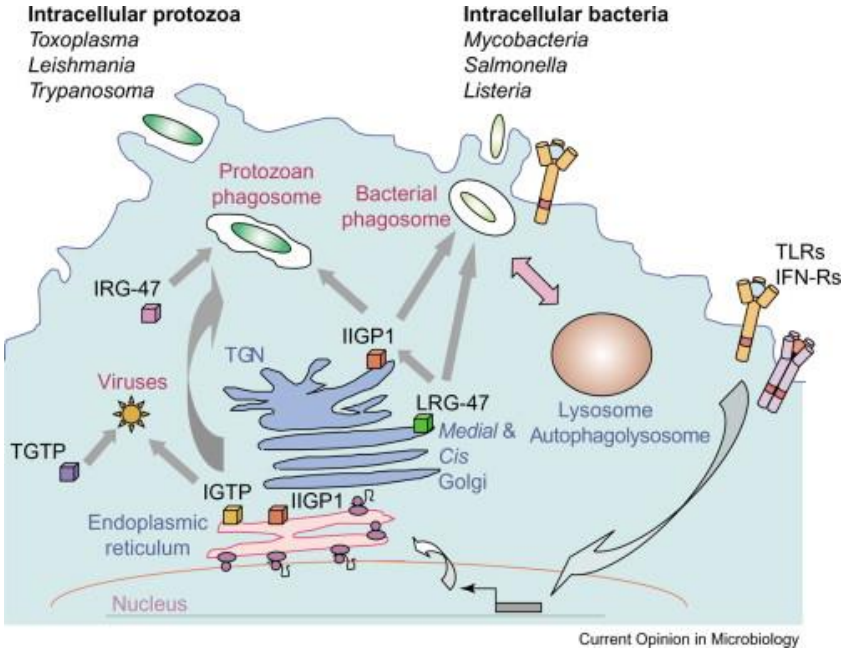

Figure 7: Different mouse IRGs and their mode of action in response to specific pathogens. Irgm1 was more commonly known as LRG-47 at the time of the publication of this figure. The arrows in this figure illustrate the trafficking of their designated IRG to its destination, where it will perform its function. In this example, LRG-47/Irgm1 is trafficked from the Golgi to a pathogen-specific phagosome that harbors a bacteria such as *Mtb* or a protozoan. (MacMicking, 2005)

In recent years, Henry et al. (2014) and Tiwari et al. (2009) provided evidence that Irgm1 attaches to Golgi membranes via its amphipathic αK helix (residues 356-369 – Henry et al, 2014) and via lipid group insertions (palmitoyl groups found on Cys371, 373,374, 375 at the C-terminal). Mutating Cys palmitoylation sites to Ala and disrupting the amphipathicity of the αK through the insertion of negatively-charged Glu, are both needed to completely abolish membrane binding. Otherwise, Irgm1 retains its binding properties to the membrane, although not to full extent. Speculatively, the mechanism of dissociation from the membrane is thought to occur via a conformational change of the αK helix in relation to the membrane. Normally, the hydrophobic region of the αK helix is tucked into the G-domain, rendering the protein inactive and in a closed conformation. Disruption causes the helix to shift, causing its positively charged residues to move away from the membrane. The disintegration of the electrostatic forces between the positively charged residues and the negatively charged membrane surface is what abolishes binding of Irgm1 to the membrane.

Palmitoylation was confirmed through radiolabeling of [^{3}H]-palmitate first and performing progressive mutagenesis on all Cys residues to uncover the exact residues that become palmitoylated. This experiment also illustrated that palmitoylation is dynamic, as some of these cysteines may lose the palmitoylation and regain it later.

Furthermore, staining with Tom20 (a mitochondrial marker) showed that Irgm1 localizes to mitochondrial membranes where it is hypothesized to promote mitochondrial fission in preparation of autophagy.

1.3.2 Function in Phagocytosis and link to Tuberculosis

Irgm1 mouse knockouts display a very pronounced susceptibility to infection with *Mtb.* Additionally, Irgm1 macrophage knockouts show aberrant trafficking of the phagosome towards the autophagolysosome pathway. Irgm1 is induced by IFN-γ, upon registering an infection at the TLR4 level (See Fig. 7). However, not much is understood about the molecular mechanism of activation and function of Irgm1.

To shed more light on the mechanism, Tiwari et al. (2009) have recently discovered a role for Irgm1 in instigating autophagy upon infection via the PI3K pathway. Screening for lipid-binding, it was shown that Irgm1 associates with phosphatidylinositol-3,4-bisphosphate (PtdIns(3,4)P_2), phosphatidylinositol-3,4,5-triphosphate (PtdIns(3,4,5,)P_3), DPG and weakly, phosphatidic acid. These lipids carry net negative charges whose empirical value depends on pH and interaction with the protein. The range of these charged phosphate entities runs from -2 to -7 and form attractive binding partners for clusters of positively charged basic amino acid residues on a protein surface.

Irgm1 is located on the cytosolic side of the Golgi where it binds to DPG, however not to PtdIns. The binding occurs through two features on the Irgm1 protein: 1) the amphipathic αK helix (350-374 – Tiwari et al., 2009) that has an isoelectric point of 9.0, and is rich in basic amino acid restudies (6 Arg and 1 lys) and 2) palmitoylation on Cys371, 373,374,375. When cells are infected by *Mtb,* Irgm1 travels along pseudo pods to the region of phagocytosis of the pathogen - the mycobacterial phagosome (MPG).

Tiwari et al. show that PtdIns(3,4)P_2 and PtdIns(3,4,5)P_3/PI3P serve to recruit Irgm1 to the phagosome. The lipids are synthesized by class I PI(3)K and are found on the surface of MPG's. DPG is detected on MGP in later stages as the phagosome matures and it is hypothesized that it accumulates on the MPG from the pathogen, as opposed to the host. The presence of DPG, therefore, acts in a way to secure Irgm1 on the MPG. Recruitment to the phagosomal membrane is mainly governed by the αK helix of Irgm1. Additionally, the class I

PI(3)K isoforms that synthesize PtdIns(3,4)P_2 and PIP3, that specifically target Irgm1 to the MPG membrane, were shown to be heterodimers consisting of Pik3ca-Pik3r1 or Pik3ca-Pik3r2. These, along with SHIP1 enrich the MPG membrane with PtdIns(3,4)P_2 and PIP3. The class I PI(3)K factors bind directly to Irgm1 (except SHIP1). Binding to the PI3K complex and cofactors enhances the activity of PI3K in converting PtdIns(3,4)P_2 to PIP3. In turn, PI3K and the act of lipid binding increase GTPase activity of Irgm1 in an unknown mechanism. These events create a positive feedback loop reminiscent the GTPase, p21/Ras. In conclusion, the binding to PI3K and membrane lipids may activate the allosteric switch of the GTPase.

The increase in GTP-hydrolysis of Irgm1 and allows it to bind to the fusogenic SNARE proteins. Screening for effectors that Irgm1 binds to showed association with Snapin and Tmed10. Snapin localizes to the MPG regardless of the presence of Irgm1. Ultimately, activation of Snapin and the SNARE complex serves to merge the phagosome to the autophagolysosome (Pan et al., 2009).

1.4 Human IRGM: Genetics, Function and link to Crohn's and IBD

1.4.1 Evolutionary and Genetic Studies

Unlike other mammals that possess from 3-21 IRG genes, humans have evolved with only two IRGs: IRGC and a truncated IRGM. The truncation in the IRGM gene occurred about 40-50 million years ago within the anthropoid lineage of evolution. Further studies in evolution attribute this truncation event to an integrated $AluS_c$ repeat within the gene's ORF in the anthropoid ancestor, rendering the IRGM locus nonfunctional since

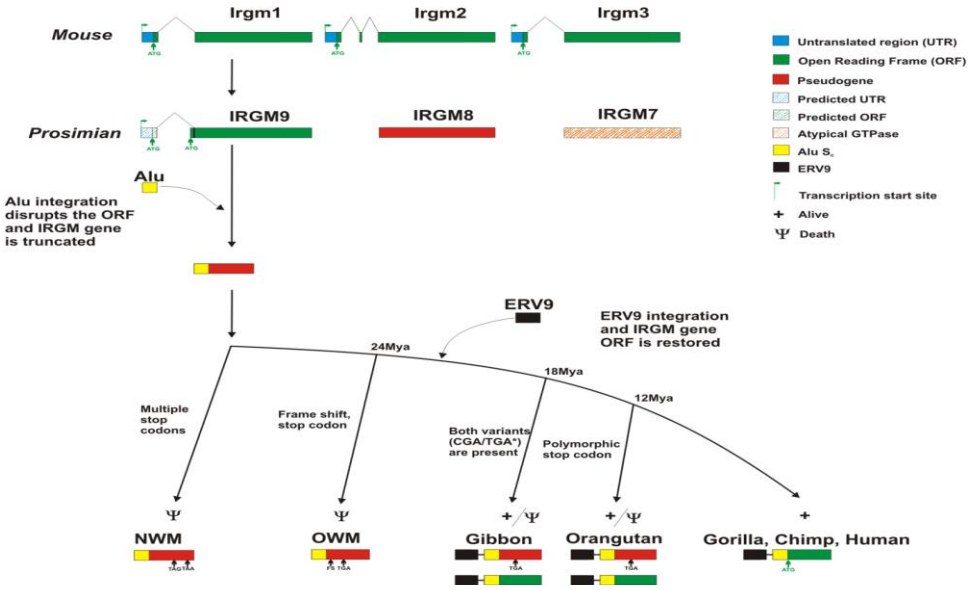

Figure 8: Timeline of the evolution, death and resurrection of the human IRGM gene across generations (Bekpen et al, 2009).

the divergence of the New World and Old World monkey lineages (35—40 million years ago) See Fig. 8.

The human *IRGM* gene is found on locus 5q33. In mouse, the IRG genes cluster map on chromosomes 11 and 18 which correspond to the human chromosome 5.

Some millennia later following the truncation of the IRGM gene (~20 million years ago), a retroviral transposable element, ERV9, was discovered to have been integrated upstream of the IRGM gene in the common ancestor of humans and African great apes. The integration of ERV9 provided a 5' UTR exon and an alternative splice site that may have restored function to the IRGM gene.

Resurrected as it may be, whole genome association studies identified a synonymous exonic variant of IRGM that was prevalent amongst patients with CD. The variant is characterized by an SNP at 313C>T (CTG>TTG, leucine) and occurs in perfect linkage disequilibrium with a 20-kb deletion polymorphism in the LTR-rich region upstream of the *IRGM*. Research from McCarrol et al (2008) demonstrated the existence of a relationship between aberrant expression of IRGM and the deletion polymorphism, causing phenotypes that manifest in loss of function of autophagy and accordingly, resistance to infection. Thus, the c.313T variant of IRGM has been deemed a risk-associated allele for CD.

Since the SNP present within the risk-associated allele was synonymous (i.e. it did not alter the protein sequence of IRGM), Brest et al. (2011) investigated the mechanism of gene regulation and transcription for the *IRGM* gene. Their research found that the C>T base change itself alters the regulation of this gene by means of miRNA-mediated regulation. Since miRNAs regulate their targets based on base complementarity at the mRNA level, it was found that mi-RNA196 regulates the processing of the *IRGM* mRNA through binding to the RISC complex. In the c.313T risk-allele, there is decreased binding of mi-RNA196 to the mRNA since the SNP occurs within the seed region necessary for RISC binding. This aberrant processing of the mRNA results in decreased expression of IRGM which interferes with the process of autophagy, thus causing persisting symptoms in CD and susceptibility to bacteria.

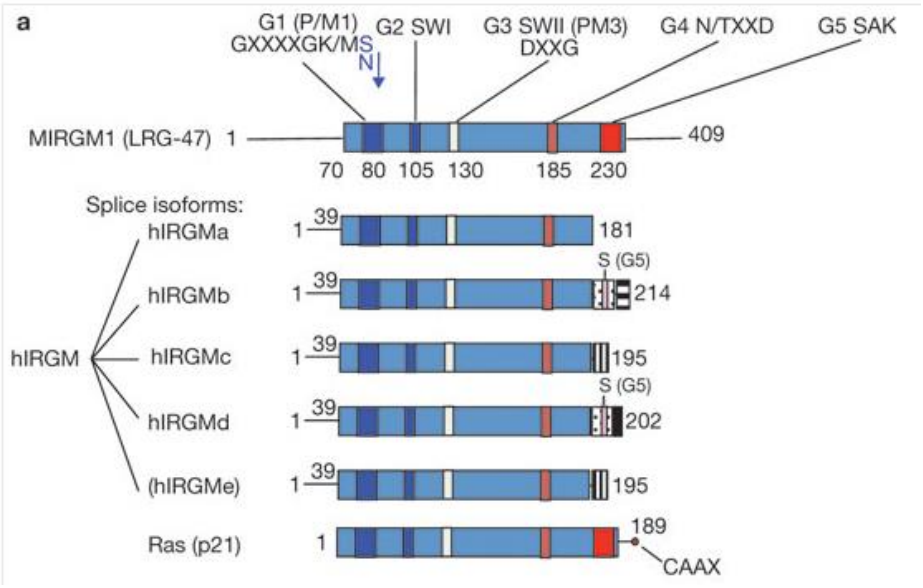

Figure 9: IRGM isoforms compared to the Ras GTPase. IRGMb and IRGMd only have the complete G domain with G5 at the C-terminal. The other isoforms are truncated (Singh et al., 2010).

In other studies by Bekpen et al (2009), the IRGM gene transcript was found to possess five different 3' splicing isoforms, extending more than 50kb 3' of the long coding exon. The isoforms are termed IRGM(a) – IRGM(e). The two isoforms concerned in this thesis are the truncated IRGM(a) and its longer counterpart IRGM(d). Figure 9 shows some features of the IRGM isoforms – IRGM(b) and IRGM(d) possess the complete G domain of a GTPase, with the G5 motif at their C-terminal.

1.4.2 Implication of IRGM in Disease through Autophagy

Clues as to how underexpression of the IRGM protein can cause susceptibility to infection by *Mtb* were brought to light in 2006 by Singh et al. Immunoblotting showed that knockdown of IRGM through siRNA in the macrophage cell line U937 prevented the conversion of LC3-I into its active form, LC3-II – a vital step in the formation of the autophagosome. Fluorescence microscopy and staining showed a decrease in LC3 puncta in cells treated with siRNA-IRGM.

More importantly, cells infected with the mycobacterial strain, *Mycobacteria Bovis*, showed reduced acidification of the BCG phagosome and higher bacterial survival rates when treated with siRNA-IRGM. The studies have also demonstrated that the action of IRGM is independent of IFN-γ. Unlike its mouse orthologue, Irgm1, the human IRGM is not directly induced by IFNγ, largely because it lacks the transcriptional element required for gene transcription induction by the IFN. Moreover, when autophagy was induced with rapamycin – independent of IFN-γ – and cells were treated with siRNA-IRGM, LC3-II formation was also inhibited, indicating that the function of IRGM is indeed independent of IFN-γ. Later studies by Singh et al. (2010) showed that IRGM localizes to the mitochondria to induce mitochondrial fission in the autophagy response. IRGM did not appear to localize to phagosomes or other compartments within the cells. Specifically, the isoform IRGM(d) has the most affinity to bind to mitochondrial cardiolipin in the inner mitochondrial membrane or mitochondrial matrix to induce mitochondrial depolarization. Cardiolipin is a very important lipid in mitochondrial function. It facilitates protein import; stabilizes respiratory chain supercomplexes; is involved in driving the process of apoptosis; and consequently has been implicated with Barth syndrome and cancer. Mutation of the Ser47 to Asn in IRGM(d) interferes with binding to cardiolipin, thus preventing

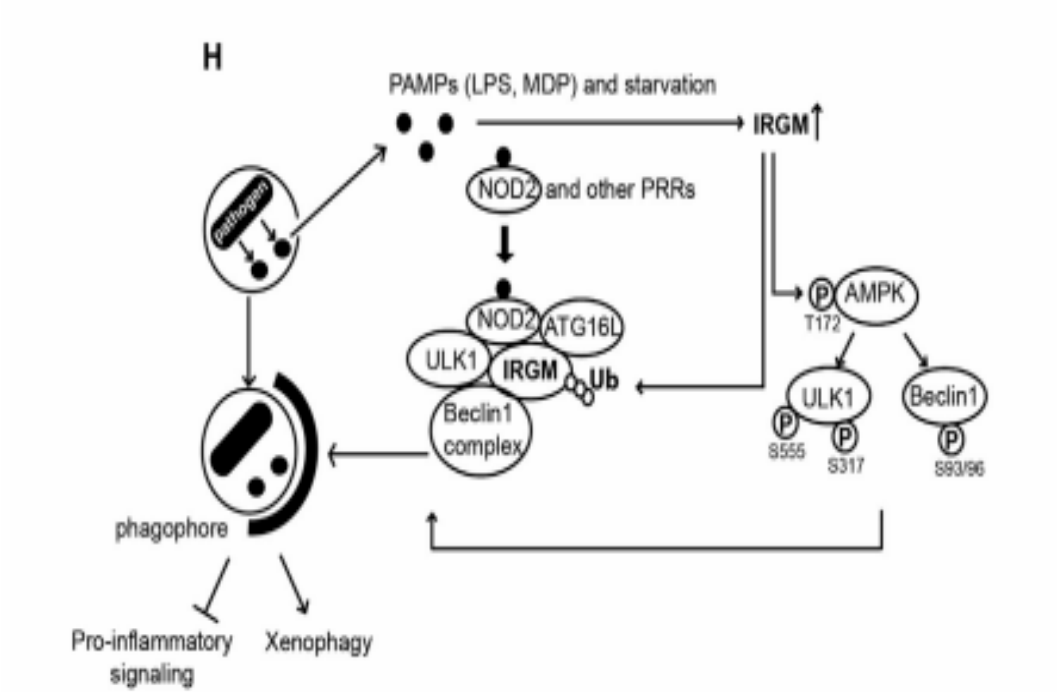

Figure 10: Model of IRGM protein interactions in autophagy upon encountering a PAMP from an invading pathogen. (Chauhan et al., 2015)

the occurrence of the appropriate morphological changes in the mitochondria that are necessary for autophagy.

Novel data (Chauhan et al 2015) placed IRGM at the hub of the autophagy-activating complex, giving it a vital role in the immune response to infection.

IRGM increases the levels of AMPK – a protein that directly phosphorylates the autophagy protein complexes ULK1 and Beclin1.IRGM consequently enhances phosphorylation of ULK1 and Beclin 1 at their respective activating phosphorylation sites. Knockdown of IRGM reduces levels of ULK1, ATG14L and AMBRA1 (a protein associated with Beclin1). Beclin 1 and the autophagy factors ATG5-ATG12 remain unaffected. However, the ATG5-ATG12 physical organization was affected. Specifically, ATG5 organization is governed by ATG16L1 – an autophagy factor that was found to bind directly to IRGM. ATG16L1 involved in LC3 conversion and autophagosome formation. Therefore, IRGM stabilizes autophagy key regulators. Under basal conditions, no interaction with ULK1 or ATG16L1 was detected. Only under infection of autophagy-inducing conditions were interactions detected.

Western blots of pulldown assays with purified proteins showed a direct protein-protein interaction between IRGM and NOD2. The direct interaction between IRGM and NOD2 was further confirmed through a proximity ligation assay (PLA). IRGM interacts with NOD2's CARD domains specifically, with enhanced interaction upon deletion of NOD2's LRR region which supposedly folds into the CARD domains to render the protein in a closed inactive conformation. NOD2 enhances polyubiquitination of IRGM at Lys63. Mutation of all Lys residues in IRGM abolishes its ubiquitination potential completely, which greatly reduces its binding to ULK1, Beclin1 and other autophagic regulators except ATG16L1. IRGM gets polyubquitinated at K63.

Finally, polyubiquintinated IRGM inversely controls NOD2 levels by targeting it for degradation. This forms a negative feedback loop that tightly regulates the system, preventing any aberrant inflammation. These findings illustrate more clearly why genetic defects in NOD2, ATG16L1 and IRGM lead to susceptibility to *Mtb* and implication in diseases characterized with chronic inflammation (e.g. IBD and CD)

1.5 Rationale of Study

The IRGM protein has been implicated in Crohn's Disease – a disease that greatly reduces the quality of life for the individuals affected. Aberration in IRGM expression also renders individuals susceptible to infection with *Mtb*. Moreover, recent studies have shown that IRGM levels are heightened in patients with gastric cancer.

The aim of this study is to advance research and scientific knowledge in the field of structural biology in order to be able to obtain a pure and stable form of the IRGM protein fit for analysis by X-ray crystallography. Due to its central role in disease, the structure of this protein can greatly aid in targeted drug design for diseases related to autophagy such as IBD, CD and susceptibility to virulent pathogens.

Chapter 2: Study and Characterization of Human and Mouse IRGM Proteins

2.1 Introduction

Aberrant expression of the human IRGM protein has been heavily implicated in the malfunction of autophagy, consequently leading to chronic inflammatory diseases such as Crohn's Disease, as well as increased susceptibility to virulent pathogens (e.g. *M. tuberculosis).*

My project focuses on obtaining a pure form of the human IRGM protein as well as its mouse orthologue, Irgm1. Pure protein is the precursor step in setting up good-quality protein crystals that can be analyzed by X-ray crystallography to obtain a 3D structure of the protein. The structure of the IRGM protein is of great importance in understanding its mechanism of action at an atomic level, especially the mechanism of its unique GTPase domain that functions in an immunity-related manner. Additionally, a stable and pure form of the protein can be used for GTPase assays and binding assays to confirm its direct interaction with other proteins in the autophagy hub (e.g. NOD2). The information from its structure and functional assays can ultimately pave the way for potential therapeutic opportunities and drug design.

In this chapter, I discuss the methods used to prepare a purified form of the protein: different protein constructs; expression systems; different purification streams. Every construct was purified differently to maximize chances of finding optimal conditions for protein stability. While the stage of crystallography was not attained due to experimental limitations, I will discuss the challenging aspects of this project, providing hypothesis and solutions that can be considered in future investigation of the IRGM protein.

2.2 Methods and Results

2.2.1 Cloning and Expression in Bacteria

Constructs were subcloned into the BamHI and NotI restriction sites of the bacterial expression vectors pProEx HTb (selection by ampicillin) and/or pSMT3 (selection by kanamycin). The vectors were then transformed into an *E. coli* expression strain to produce fusion proteins with an N-terminal 6xHis and/or N-terminal His6-SUMO-tag, respectively. The tags can be cleaved off by Ulp1. (See Table 3 for a list of bacterial expression strains used).

A starter culture containing the appropriate antibiotics (50mg/L of Ampicillin and/or 30mg/L of Kanamycin) was left to grow overnight, at 37 ^{o}C while shaking at 220 rpm. The following day, flasks containing 1L of LB plus the appropriate antibiotics were inoculated the with the starter culture to achieve a 1:50 ratio of starter culture to growth culture. Cells were grown to an OD_{600} = 0.6-0.8 (about 2-3 hours) and then induced with 1mM IPTG. Time of induction was 4-5 hours or overnight in the case of Arctic Express. (See Table 4 for the different induction temperatures used).

Cells were then harvested by centrifugation at 4,500 rpm and 4^{o}C for 20 minutes. Pellets were scraped and stored at -80 ^{o}C until the time of purification.

Table 3: Bacterial expression strains used and their associated features

Strain	Resistance	Genotype	Application
BL21(DE3)	None	*Lon* and *ompT* protease deficient; F– ompT hsdSB(rB – mB –) gal dcm (DE3)	Protease deficient to prevent degradation of recombinant proteins; high yield protein expression.
Rosetta2(DE3)	Chloramphenicol (34 μg/ml)	BL21(DE3)-derivative; F– ompT hsdSB(rB – mB –) gal dcm (DE3) pRARE23 (CamR)	Expresses rare tRNAs that facilitate expression of genes that encode rare *E.coli* codons.
Arctic Express RIL (DE3)	Gentamycin	BL21 XL-10 Gold-dervative; E. coli B F– ompT hsdS(rB – mB –) dcm+ Tetr gal λ(DE3) endA Hte [cpn10 cpn60 Gentr] [argU ileY leuW Strr]	Lower temperature cultivation and encoding of chaperones allo for proper proper folding of protein and enhanced solubility.
C43(DE3)	None	BL21(DE3)-derivative; F *– ompT hsdSB (rB- mB-) gal dcm* (DE3)	Effective in expressing toxic and membrane proteins.

Table 4: Results of the protein expression trials in different bacterial expression strains, under different conditions.

Construct	Vectors	Strain	Induction Temperature/ ºC	Expression results
IRGM23-181	pProEx Htb	BL21(DE3)	30	Insoluble
		Rosetta2(DE3)	30	
IRGM31-181	pProEX Htb pSMT3	Rosetta2(DE3)	37, 20	Insoluble
		C43	37	
		Arctic Express RIL (DE3)	16	
IRGM(3-181)	pProEx Htb pSMT3	BL21 (DE3)	37	Traces of soluble protein, but insufficient
		Rosetta2 (DE3)	37, 30, 25, 20, 18	
		Arctic Express RIL (DE3)	16	
		C43	25, 20	No Expression
IRGM(d)	pProEx Htb pSMT3	Rosetta2(DE3)	37, 30	Insoluble
		BL21(DE3)	37, 30	
		Arctic Express RIL(DE3)	16	
Irgm1	pProEx pSMT3	Rosetta2 (DE3)	37, 30	Insoluble
		BL21(DE3)	37, 30	
		Arctic Express RIL (DE3)	16	
Palm-αK/ Irgm1(ins.362,367E; C371, 373, 374, 375A)	pProEx Htb	BL21(DE3)	37, 30	Insoluble
		Rosetta2(DE3)	37, 30	
		Arctic Express RIL (DE3)	16	

Table 4 shows the results from small scale and large scale purifications that were carried out to determine the protein solubility and its extent of expression in a given bacterial cell line (data not shown).

2.2.2 Cloning and Expression in Insect Cells

Constructs were subcloned into the BamHI and NotI restriction sites of the baculovirus vectors His-strep-GFP-TEV- pFastBac, C-terminal-GFP-pFastBac and His-GST-pFastBac (IRGMd only), to generate recombinant proteins carrying the tag His-strep-GFP-TEV, C-terminal TEV-GFP-His and His-GST-TEV respectively.

Work from here on until cell harvest was performed by our lab manager and research assistant Katalin Illes Kocsis, according to a specific protocol. Her work encompassed the generation of the baculovirus and the following transfection of Sf9 insect cells with the virus. At the last step, cells were grown at 27^{o}C at 85 RPM and then harvested by centrifugation at 1,000 x g for 10 minutes. Pellets were then resuspended with 1X PBS and stored at -80^{o}C until the time of purification. Table 5 shows expression results of the different constructs in insect cells.

Table 5: Results of protein expression in insect cell lines

Construct	**Tag**	**Cell Type**	**Expression and appearance**
IRGM23-181	His-Strep-GFP	Sf9	Very low
IRGM(3-181)	His-Strep-GFP	High Five	Acceptable, diffuse and soluble
	C-terminal GFP-His	Sf9	Very low, punctate
IRGM(d)	His-Strep-GFP	Sf9	Good; diffuse and soluble
	C-terminal GFP-His		
	His-GST		N/A
Irgm1	His-Strep-GFP	High Five	Good, punctate and localizing to vesicles
Palm-αK/ Irgm1(ins.362,367E; C371, 373, 374, 375A)	His-Strep-GFP	Sf9	Good; some diffuse and some punctate and localizing to vesicles
	C-terminal GFP-His		
αK/ Irgm1(ins.362, 367E)	His-Strep-GFP	Sf9	Good; punctate

2.2.3 Protein Purifications

Protein Purification Protocol from insect cells

Cells were thawed on ice and the volume was made up to 30 ml with Sonication Buffer. Cells were then disrupted by sonication and the volume was made up to 50 ml with Buffer A (refer forward to tables). In case of use of detergent, the 50 ml of cells were allowed to sit on ice for about twenty minutes, after inverting the tube five times. After that, the lysate was separated into two ultracentrifuge tubes and spun in the ultracentrifuge for 1 hour at 4^{o}C and 41,000 RPM (100,000 x g, 70Ti 09U345 rotor). The supernatant containing the soluble protein was then aspirated, retained and the pellet was discarded. Both the color of supernatant and the pellet were recorded to observe the distribution of GFP to track the recombinant protein.

Meanwhile, HisPur Ni^{2+}-NTA Agarose resin (Thermo Scientific; pre-packed column or loose beads in batch mode) was equilibrated with Buffer A in preparation for purification by affinity of the 6xHis tag on the protein to the Ni^{2+}-NTA resin. The supernatant containing the soluble His_6 -tagged protein was applied onto the resin for binding (~ 1 hour for both a pre-packed column or in batch mode). After binding, the protein was washed with Buffer A and eluted off the Ni^{2+}-NTA with 20% of Buffer B (refer forward to tables). The protein was then concentrated by centrifugation at 4,000 rpm and 4^{o}C in an Amicon Ultra-15 Centrifugal Filter Unit (EMD Millipore) with a molecular cut-off of 30kDa. Following concentration to about 250-300 µl, the protein was applied onto a 0.5ml Streptadivin gravity column and washed with 0.5- 1 ml Buffer C (refer forward to tables). The protein was then eluted with 1X Biotin in Buffer C to a volume of 1.5 ml.

If gel filtration was the next step, the protein was concentrated to 0.5 ml and then applied onto the gel filtration column. Fractions of 0.5 ml were collected during gel filtration, concentrated and analyzed on 12.5% SDS PAGE and/or Western Blot. See the following section for details regarding exact buffer concentrations, binding times, volume of Ni-NTA resin and other conditions used in the protein purification.

2.2.3.1 Protein Purification Results of *IRGM* Human Protein Constructs

IRGM(3-181)

Table 6: Purifications of IRGM(3-181)

Parameters	Purification 1	Purification 2	Purification 3
Cell Line	High Five	High Five	High Five
Cell Volume	400 ml	200 ml	200 ml
Harvest Buffer	500 mM NaCl; 50 mM TRIS; 10% Glycerol; 20 mM Imidazole ; 5 mM $MgCl_2$; 3 mM DTT	1X PBS	1X PBS
Sonication Buffer	Same as harvest buffer	150 mM NaCl; 3 mM $MgCl_2$; 2 mM DTT; 20 mM imidazole; 200 μM PMSF	150 mM NaCl; 3mM $MgCl_2$; 2mM DTT; 20mM imidazole; Roche Cocktail Protease inhibitors
Batch vs. Column	Batch; 2 ml beads	Batch; 1 ml	Batch; 2 ml
Binding Time	2 hours	1.5 hours	1.5 hours
Buffer A	500 mM NaCl; 50 mM TRIS; 10% Glycerol; 20 mM Imidazole ; 1 mM $MgCl_2$; 10μM βME	150 mM NaCl; 3 mM $MgCl_2$; 2 mM DTT; 20 mM imidazole; 200 μM PMSF	150 mM NaCl; 3 mM $MgCl_2$; 2 mM DTT; 20 mM imidazole; Roche Cocktail Protease inhibitors
Volume of Ni Wash	5xC.V. x1 (for each tube) Total: 20 ml	5x C.V. x3 Total: 15 ml	5x C.V. x6 Total: 30 ml
Buffer B	500 mM NaCl; 50 mM TRIS; 10% Glycerol; 200 mM Imidazole ; 5 mM $MgCl_2$; 1mM DTT	150 mM NaCl; 3 mM $MgCl_2$; 2 mM DTT; 250 mM imidazole; 200 μM PMSF	150 mM NaCl; 3 mM $MgCl_2$; 2 mM DTT; 250 mM imidazole;
Volume of Ni Eluent	2.5x C.V. Total: 20 ml	3x C.V. x2 Total: 6 ml	3C.V. x2 Total: 6 ml
Concentrator cut-off	10 kDa	10kDa	10 kDa
Buffer C	500 mM NaCl; 50 mM TRIS; 10% Glycerol; 10μM βME	150 mM NaCl; 3 mM $MgCl_2$; 2 mM DTT	150 mM NaCl; 3 mM $MgCl_2$; 2 mM DTT
Cleavage O/N with TEV?	No.	No.	No.
Gel Filtration?	No.	No.	No.

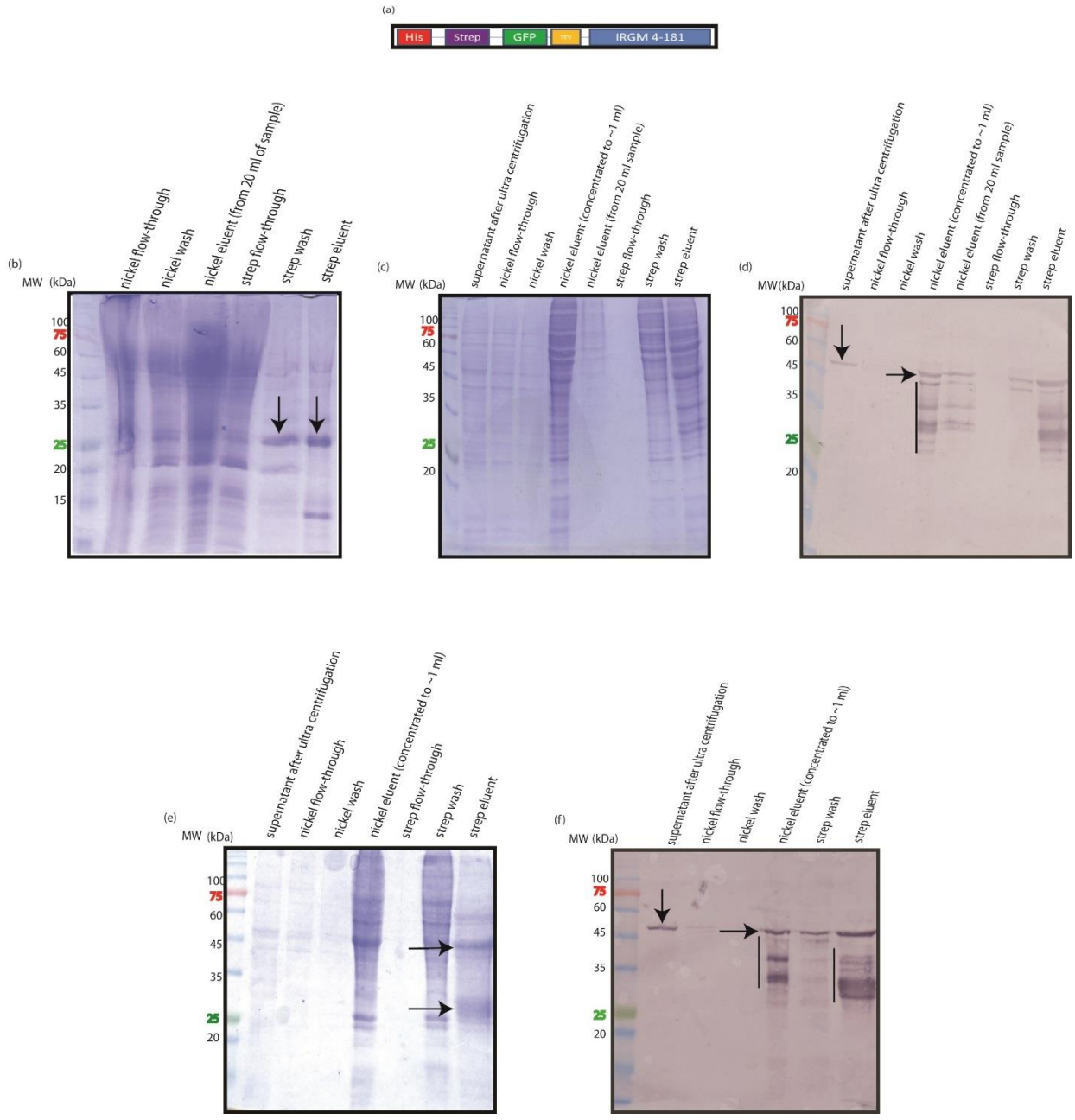

Figure 11: (a) Cartoon representation of the IRGM(3-181) construct. (b) SDS PAGE gel corresponding to Purification 1 in table 6. (c) SDS PAGE gel corresponding to Purification 2 in table 6. (d) Western Blot corresponding to the SDS PAGE gel in panel (c) of this figure. (e) SDS PAGE gel corresponding to Purification 3 in table 6. (f) Western Blot corresponding to the SDS PAGE gel in panel (e). The arrows indicated on several gels and WBs point to prominent protein bands on the gel around 30 kDa and 50 kDa. The vertical black lines seen in the WBs highlight the spectrum of what is possibly degradation products throughout the purifications.

Table 6 details the different conditions attempted for the purification of the IRGM(3-181) construct (52.4 kDa). The first purification follows a standard protocol, suggested by Ni-NTA resin supplier, QiaGen (The QiaEpressionist Manual, 2003) and contains high-salt buffer, with 500 mM NaCl and 1 mM $MgCl_2$. The latter was added merely out of experience from working with GTPases, as they require the Mg^{2+} co-factor for stabilization. The full IRGM construct with the His-strep-GFP tag, however, does not appear clearly in the SDS PAGE for that purification (Figure 11 (b)). The gel only shows a prominent band around 30 kDa after the strep elution, which could indicate a large degradation product that resulted in the dissociation of the tag from the protein.

The second purification was carried out under low salt conditions (150 mM NaCl) and slightly higher $MgCl_2$ (3 mM, compared to only 1 mM in the first purification). Additionally, 200 µM of the serine protease inhibitor PMSF was added in order to avoid degradation products. The SDS PAGE gel shown in Figure11(c) for the second purification does not give useful information about the purification, since we see no prominent band indicating the protein of interest. Thus, a Western blot of the purification was produced and the arrows in Figure 11(d) clearly show a constant band at around 50kDa that corresponds to the size of the full construct. Interestingly, the blot also shows the presence of a spectrum of degradation products (possibly of a C-terminal nature) that seem to increase in intensity throughout the purification; we can see that there is only one band corresponding to the protein at the start but then becomes accompanied by other smaller bands throughout the purification.

To address the degradation problem, the 3rd purification contained a tablet of cocktail protease inhibitors (Roche, see Materials section). The results of the 3rd purification are shown in the SDS PAGE gel in Figure 11(e) and its corresponding WB in Figure 11(f). The gel shows great improvement since the band of interest is clearly seen in the strep eluent (around 50kDa, as indicated by the arrow) but also shows the reappearance of the 30kDa prominent band that might correspond to the tag. The WB shows a reduction in degradation compared to the one in panel (d) and the band corresponding to the full length construct remains fairly strong throughout the purification.

The protein sample from these purifications was ultimately concentrated to ~ 500 µl but its yield proved difficult to determine on the NanuVu since it contained many impurities and displayed a very weak and haphazard signal on the machine. Due to its low purity, the protein

sample was not subjected to size-exclusion chromatography. It is also interesting to note that the green color of GFP was gradually being lost throughout the purification (for all three purifications) and the binding to the Ni-NTA resin beads was not visibly adequate, since most of the green color from GFP was left behind in the flow-through, while the beads remained largely blue in color. It was thus decided to attempt purifying other constructs with a truncated N-terminal to try and minimize degradation products, since it is possible that there is a weak or misfolded region between the N-terminus of the protein and the engineered tag and which is very susceptible to proteases. Additionally, constructs from the mouse orthologue, Irgm1, were explored (see next few sections).

IRGM(23-181)

The His-IRGM23-181 construct (20.1 kDa) was purified from Rosetta2(DE3) cells under denaturing conditions and was then refolded with 2M urea. Cells were thawed in cold water and the volume was made up to 50 ml with NiA buffer (300 mM NaCl, 50 mM TRIS, 20 mM imidazole, 8 M urea and 1 tablet Roche Cocktail Protease Inhibitors, pH 8). Cells were then lysed by stirring in a 500-ml beaker at medium speed and room temperature for an hour. After that, the lysate was separated into two centrifuge tubes and spun in the centrifuge for 30 minutes at 4^{o}C and 22,500 RPM (JA 25.5 rotor). The supernatant containing the soluble protein was then aspirated, retained and the pellet was discarded.

Meanwhile, HisPur Ni^{2+}-NTA Agarose resin (Thermo Scientific; pre-packed column or loose beads in batch mode) was equilibrated with buffer NiA in preparation for purification by affinity of the 6xHis tag on the protein to the Ni^{2+}-NTA resin. The supernatant containing the soluble His_6-tagged protein was applied onto the resin and left on the shaker for 1 hour at room temperature. After binding, the protein was washed twice with 10 ml buffer NiA (for a total of 20 ml) and eluted off the Ni^{2+}-NTA in two 10-ml NiB buffer (300 mM NaCl, 50 mM TRIS, 250 mM imidazole, 8 M urea, pH 8) for a total of 20 ml. The Ni-NTA beads were then cleaned with buffer containing 1M imidazole.

The method chosen for refolding was dialysis over a urea gradient. Following elution, the protein was pipetted into a dialysis tubing (3kDa cut-off) and placed in a 1L beaker containing 1L dialysis buffer D1 (150 mM NaCl, 50 mM Tris, 6M urea, 10% glyercol, 1 mM PMSF, pH 7.6). Dialysis was left to take place overnight, at room temperature with gentle stirring. After 24

hours, buffer D1 was replaced by buffer D2 (150 mM NaCl, 50 mM Tris, 4M urea, 10% glyercol, 1 mM PMSF, pH 7.6) and dialysis was continued at room temperature with gentle stirring for 48 hours. Buffer D2 was then replaced with the final dialysis buffer, D3 (150 mM NaCl, 50 mM Tris, 2M urea, 10% glyercol, 1 mM PMSF, pH 7.6) and dialysis continued with gentle stirring overnight at 4ºC.

Following the last step of dialysis, the protein was concentrated by centrifugation at 4,000 rpm and 4ºC in an Amicon Ultra-15 Centrifugal Filter Unit (EMD Millipore) with a molecular cut-off of 3kDa, to 500 µl. Next, the protein was applied onto the AKTA Pure FPLC machine for the final gel filtration step, in buffer GF (100 mM NaCl; 25 mM Tris; 2M urea; 5% glycerol; 1 mM DTT). Fractions of 0.5 ml were collected during gel filtration, concentrated and analyzed on 12.5% SDS PAGE and/or Western Blot.

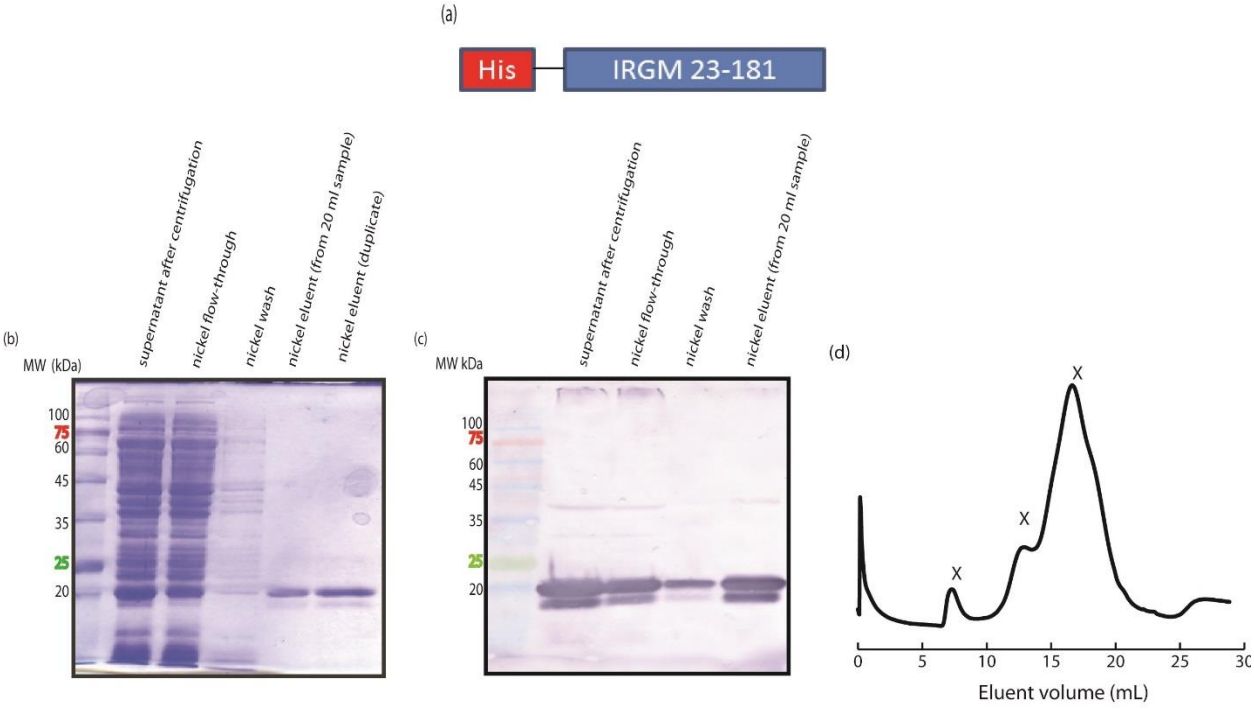

Figure 12: (a) Cartoon representation of the His-IRGM23-181 construct. (b) SDS-PAGE gel corresponding to the purification described in this section. (c) Western Blot corresponding to the SDS gel in panel (b). (d) Chromtagogram of the gel filtration step in this purification.

Figure 12 shows the SDS-PAGE and the corresponding Western Blot of this purification (panels (b) and (c)). The protein appears very clearly in the WB at 20 kDa. Figure 12(d) shows the chromatogram from applying the concentrated protein (following the last step of dialysis), and we observe a large yet asymmterical peak around 17 ml. An SDS-PAGE gel of the peaks

marked “X” in figure 12(d) showed only one band at 25kDa exactly coming from the void volume (~8ml), while the other peaks only showed empty lanes on both SDS-PAGE and a Western Blot. The Western Blot and SDS-PAGE were repeated several times but no protein could be detected.

IRGM(d)

Table 7: Purifications of IRGM(d)

Parameters	Purification 1	Purification 2 (GFP-His)	Purification 3
Cell Line	Sf9	Sf9	Sf9
Cell Volume	400 ml	400 ml	400 ml
Harvest Buffer	1X PBS	1X PBS; Roche cocktail protease inhibitors	1X PBS; Roche cocktail protease inhibitors
Sonication Buffer	1X PBS; 20 mM imidazole; 1 mM PMSF; 10 mM βME	1X PBS 1 mM PMSF Roche cocktail inhibitors 10 mM βME	1X PBS 1 mM PMSF Roche cocktail inhibitors 10 mM βME
Batch vs. Column	1 ml pre-packed Ni-NTA column on Prime	1 ml pre-packed Ni-NTA column on Prime	1 ml pre-packed Ni-NTA column on Prime
Binding Time	Approx. 45 minutes	Approx. 45 minutes	Approx. 45 minutes
Buffer A	1X PBS; 20 mM imidazole; 1 mM PMSF; 10 mM βME	1X PBS; 20 mM imidazole; 1 mM PMSF; 10 mM βME	1X PBS; 20 mM imidazole; 1 mM PMSF; 10 mM βME
Volume of Ni Wash	35 ml	35 ml	25 ml
Buffer B	1X PBS; 200 mM imidazole; 1 mM PMSF; 10 mM βME	1X PBS; 200 mM imidazole; 1 mM PMSF; 10 mM βME	1X PBS; 200 mM imidazole; 1 mM PMSF; 10 mM βME
Volume of Ni Eluent	20 ml	20 ml	15 ml
Concentrator cut-off	30 kDa	30 kDa	10 kDa
Buffer C	1X PBS	N/A	1X PBS; 1 mM PMSF; 1 mM DTT
Cleavage O/N with TEV?	No.	No.	No.
Gel Filtration?	No.	No.	No.

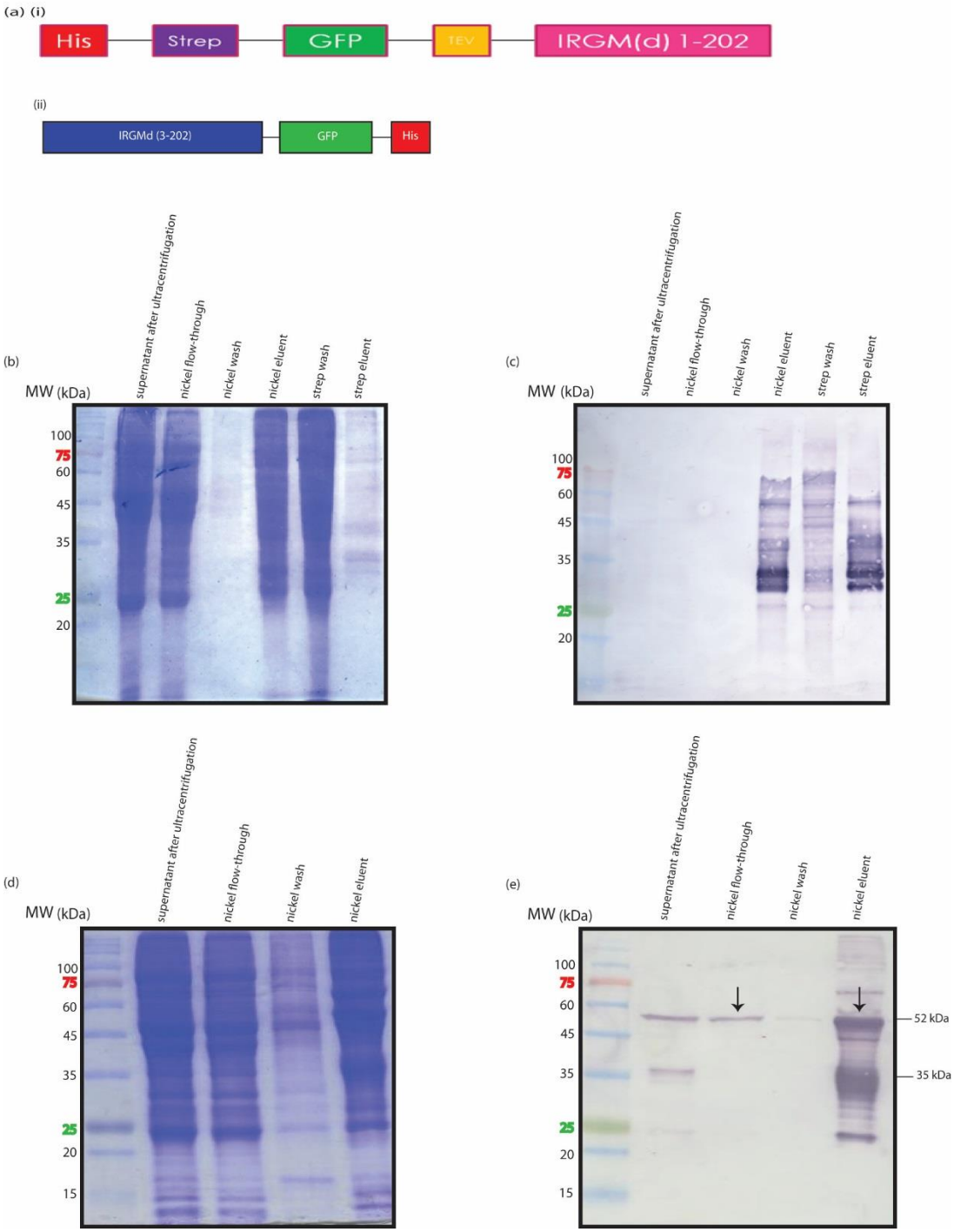

Figure 13: (a) Cartoon representation of the His-strep-GFP-IRGMd construct(i) and IRGMd-GFP-His construct (ii). (b) SDS PAGE gel corresponding to Purification 1 in table 7. (c) Western Blot corresponding to the SDS PAGE gel in panel (b). (d) SDS PAGE gel corresponding to Purification 2 in table 7. (e) Western Blot corresponding to the SDS PAGE gel in panel (d) of this figure. The arrows indicated the WB in panel (e) points to the protein construct around 50 kDa.

IRGM(d) is an isoform of the IRGM(1-181) protein first described by Bekpen et al. (2005) and later claimed to be purified and assessed for lipid binding activity by Singh et al. (2006).

Table 7 herein details several purifications of the IRGM(d) construct. The first purification concerns the His-Strep-GFP-IRGM(d) (55.1 kDa, figure 13(a)-i) construct and was performed as described at the beginning of the Methods section. Figure 13(b) and 13(c) show the SDS-PAGE and Western Blot corresponding to the purification. From these, we can see that the purification produced many degradation products which made the isolation of the protein very difficult – a re-occurring theme with the IRGM proteins. For this reason, a C-terminal tag was added to the protein in an attempt to act as a protective entity from degradation.

Purification 2 concerns the IRGM(d)-GFP-His construct (figure 13(a)-ii) that carries this C-terminal tag. Figure 13(d) and 13(e) show the SDS-PAGE gel and the Western Blot corresponding to this purification. We can see from these results that there is a remarkable improvement upon incorporation of the C-terminal tag. However, some degradation products still remain (prominent band at 35 kDa). Since this tag only carries affinity for the nickel column, it might be a prudent step in the future to engineer a C-terminal tag with multiple affinities for different columns so as to eliminate as much as possible of the degradation products.

The final purification concerned the construct His-GST-IRGM(d). This construct was described by Chauhan et al. (2015) where the protein was shown on an SDS-PAGE in the supplementary figures of the publication. Since the protocol was not described in the paper, the protein was purified in the same manner as previous constructs with the addition of applying it onto a GST column. The data for this purification are not shown, since the protein was washed off of the nickel column and could not recovered.

2.2.3.2 *IRGM* Mouse Protein Constructs

Irgm1

Table 8: Purifications of Irgm1

Parameters	Purification 1	Purification	Purification 3
Cell Line	High Five	High Five	High Five
Cell Volume	200 ml	200 ml	400 ml
Harvest Buffer	1X PBS	1X PBS	1X PBS
Sonication Buffer	1X PBS; 20 mM imidazole; 10 mM βME	1X PBS; 20 mM imidazole; 10 mM βME; 1 mM PMSF	1X PBS; 20 mM imidazole; 10 mM βME; 1 mM PMSF
Batch vs. Column	Batch 1 ml Ni-NTA resin	Batch 1 ml Ni-NTA resin	Batch 2 ml Ni-NTA resin
Binding Time	1 hour	1 hour	1 hour
Buffer A	1X PBS; 0.5% Triton X-100; 20 mM imidazole; 1 mM PMSF; 10 mM βME	1X PBS; 0.5% Triton X-100; 20 mM imidazole; 1 mM PMSF; 10 mM βME	1X PBS; 0.5% Triton X-100; 20 mM imidazole; 1 mM PMSF; 10 mM βME
Volume of Ni Wash	10 C.V. x3 30 ml	10 C.V. x3 30 ml	10 C.V. x5 50 ml
Buffer B	1X PBS; 250 mM imidazole; 10 mM βME	1X PBS; 250 mM imidazole; 10 mM βME	1X PBS; 250 mM imidazole; 10 mM βME
Volume of Ni Eluent	5 C.V. x3 15 ml	5 C.V. x3 15 ml	5 C.V. x4 20 ml
Concentrator cut-off	30 kDa	30 kDa	30 kDa
Buffer C	1X PBS	1X PBS	1X PBS
Cleavage O/N with TEV?	Yes.	Yes.	No.
Gel Filtration?	Yes. 1 X PBS; 1 mM DTT	Yes. 1X PBS; 1 mM DTT	Yes. 1X PBS; 1 mM DTT.

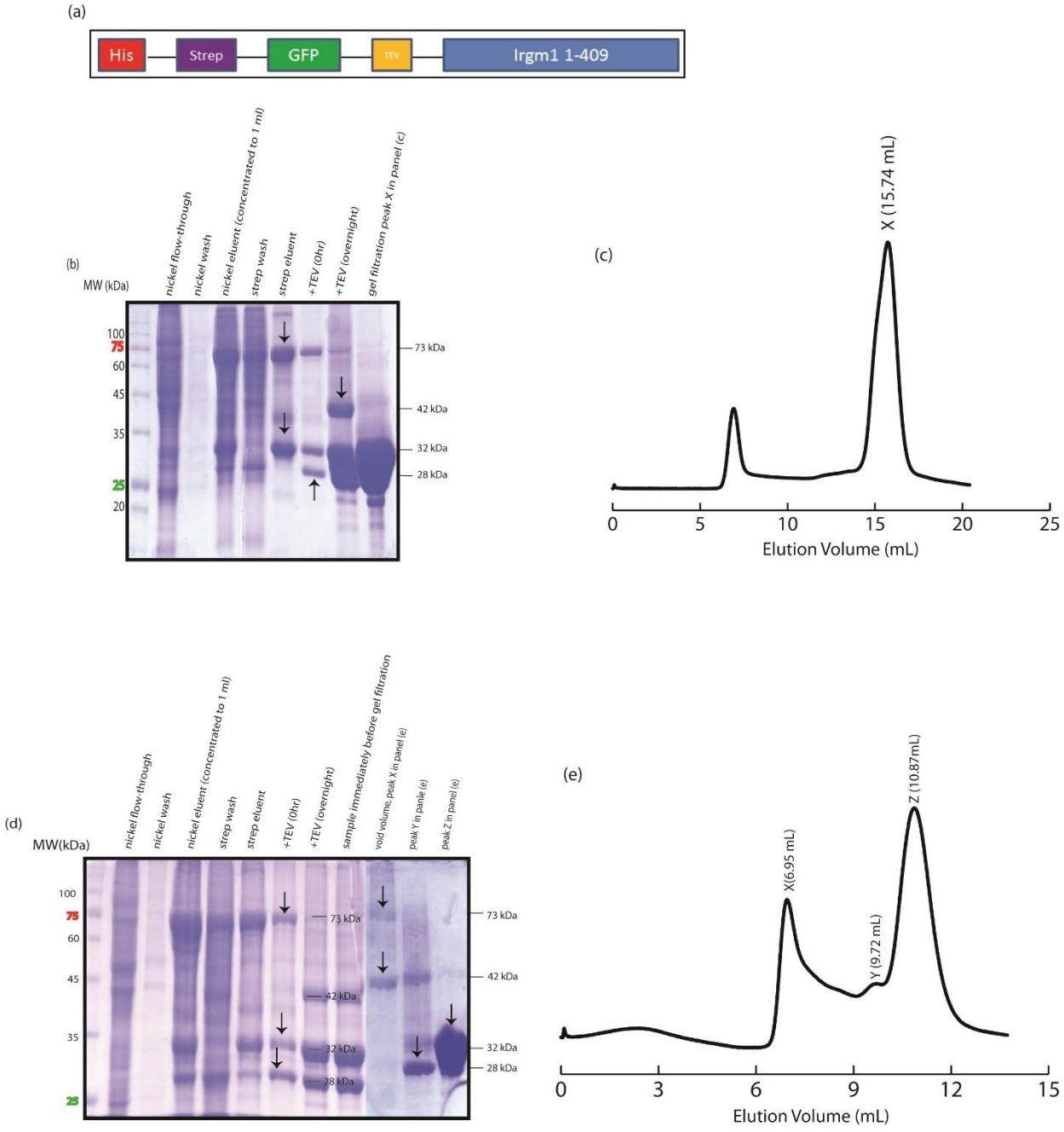

Figure 14: (a) Cartoon representation of the His-Strep-GFP-Irgm1 construct. (b) SDS PAGE gel corresponding to Purification 1 in table 8. (b). (c) chromatogram corresponding to the gel filtration step from purification 1 in table 8. (d) SDS PAGE gel corresponding to Purification 2 in table 7. (e) Chromatogram corresponding to the gel filtration step from purification 2 in table 8. The arrows indicated point to prominent protein bands at 73kDa (full-length construct); 42 kDa (cleaved full-length protein); 32 kDa (GFP tag); 25kDa (TEV).

Table 8 shows the first 3 purifications attempted for the His-Strep-GFP-Irgm1 construct (79.1 kDa, figure 14(a)). The main buffer chosen for these purifications was 1X PBS, following an experimental protocol presented by Ghosh et al (2006) who were successful in purifying the interferon-inducible IIGP1 (see Chapter 1, section 1.2.2). Since Irgm1 strongly associates with Golgi membranes (Henry et al., 2014 and Tiwari et al., 2009), it was a sound strategy to use a detergent to be able to extract the protein from the membranous portion of the cellular lysate. The choice of using Triton X-100 was merely out of convenience and practicality since it is readily available. A small-scale purification was first conducted to see whether the protein does indeed precipitate in the cellular pellet following centrifugation without the presence of detergent. Indeed, the protein required detergent to be solubilized and present in the supernatant following cell lysis and ultracentrifugation (data not shown).

The first purification follows a standard purification stream with standard buffer concentrations and procedures. In this purification, however, the serine protease inhibitor was added after sonication. Following the strep elution step, the protein was concentrated and incubated with TEV for overnight cleavage at 4^{o}C. Gel filtration was then performed the following day with the absence of detergent. Figure 14(b) shows the SDS-PAGE gel of the purification, along with the corresponding chromatogram from the size-exclusion chromatography performed (Figure 14(c)). The chromatogram in figure 13c shows two prominent peaks, the first and smaller one indicates the void volume and the more prominent one (at 15.74 ml) has been calculated to correspond to the His-Strep-GFP tag that was previously cleaved. Indeed, the last lane in the SDS-PAGE gel in figure 14(b) confirms the prominent presence of the tag, at ~32 kDa. The SDS-PAGE gel in figure14(b) also clearly shows the TEV cleavage: the protein band at ~73 kDa that corresponds to the Irgm1 protein construct does indeed disappear after cleavage with TEV to produce a band at ~32 kDa (GFP tag) and one ~42 kDa (Irgm1: 46 kDa). We can also see that the band corresponding to the tag was present in the nickel and strep eluent, prior to cleavage with TEV, indicating a degradation product that probably occurs as a result of a weak linker between the tag and the protein.

To address the degradation problem, purification 2 (see Table 8) shows that 1 mM PMSF was added to prior to sonication – a step that was not performed in the previous purification. From then on, purification 2 is identical to its precedent in buffer concentrations and experimental method. Figures 14(d) and 14(e) show the SDS-PAGE gel and gel filtration

chromatogram corresponding to purification 2 (S75). From the chromatogram, we see that the spectrum is similar to the one produced in purification 1, with the added feature of a shoulder peak at 9.72 ml. this shoulder peak chases after the main GFP peak that corresponds to the cleaved tag. The fractions from the shoulder peak (indicated as peak Y in figure 14(e)) were concentrated and a sample was run analyzed with SDS-PAGE (figure 14(d)). the lane indicated as “peak Y” shows the presence of the cleaved Irgm1 protein with traces of the tag and some degradation product. Unfortunately, the protein could not be recovered after the gel filtration step, for further analysis. In general, the SDS-PAGE of purification 2 is similar to that of purification 1and continue to show the GFP tag on its own in the nickel and strep eluents, prior to TEV cleavage.

To gain more insight into this phenomenon, the purification was repeated with the same conditions as with the previous two purifications, with the exception of eliminating cleavage with TEV and applying the uncleaved protein onto the size-exclusion column. The SDS-PAGE gel from purification 3 is not shown in this section, but figure 15(a) shows the SDS-PAGE analysis of several fractions from the gel filtration step, spanning from fraction 29 to 35. Figure 15(b) shows the chromatogram with the fractions 29 and fraction 35 indicated (11.50 ml and 17.00 ml, respectively). From the SDS-PAGE, we see that the prominent inhabitant of these fractions is the GFP tag. The shoulder peak that shows at fraction 29 (11.50 ml) corresponds to the shoulder peak from the previous purification, with the difference in elution volume being due to the S200 column being used in this instance as opposed to the S75 in purification 2. The SDS gel, however, does not show a clear nor prominent band corresponding to the Irgm1 protein. It was therefore concluded that the protein largely ended up in the void volume (SDS-PAGE analysis of the void volume is not shown).

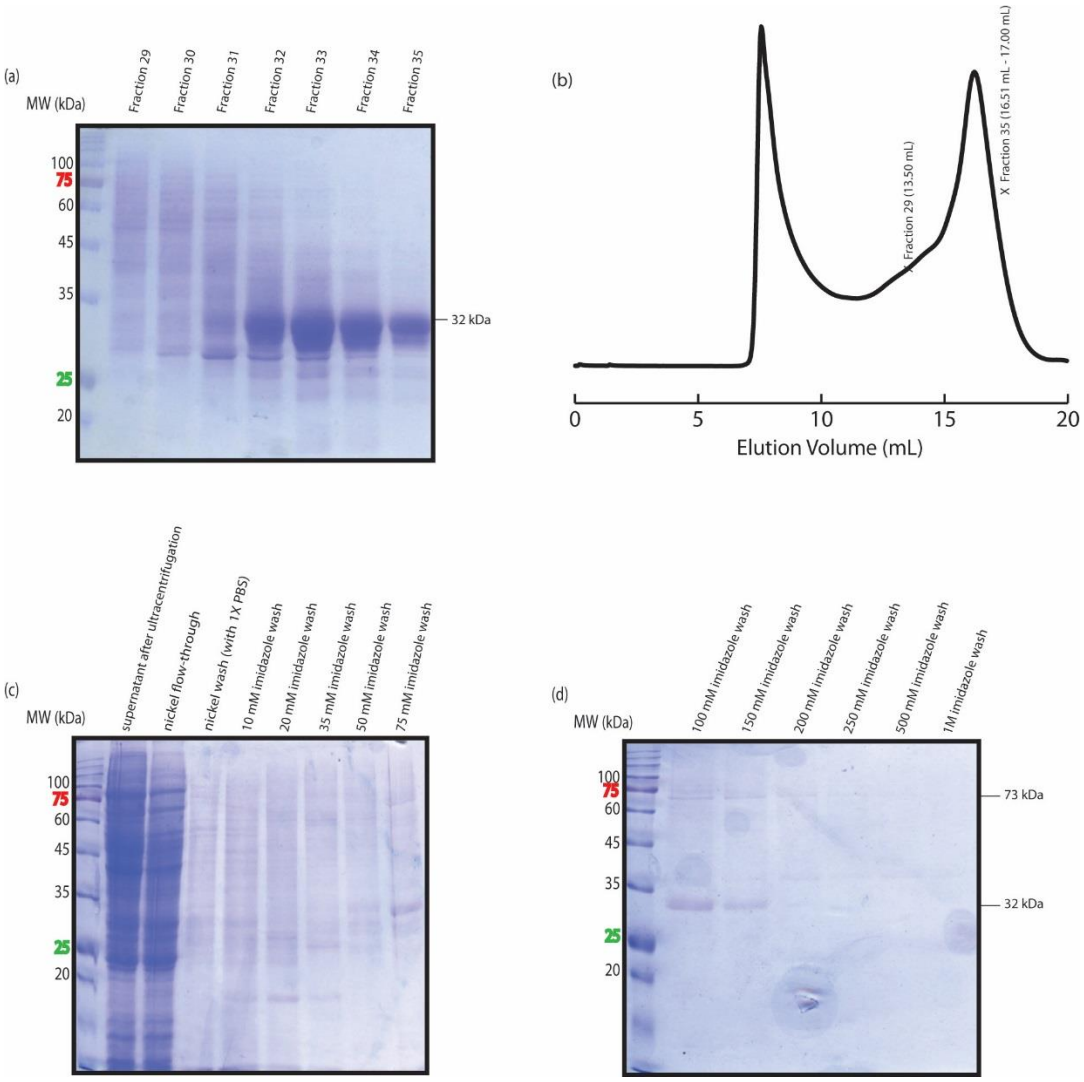

Figure 15: (a) SDS PAGE gel corresponding to Purification 3 in table 8. (b) Chromatogram corresponding to the gel filtration step from purification 3 in table 8. (c) SDS PAGE gel corresponding to Purification 4 in table 9 and shows the first half of the imidazole gradient (10 mM – 75 mM). (d) SDS PAGE gel corresponding to Purification 4 in table 9 and shows the second half of the imidazole gradient (100 mM – 1 M).

Table 9: Further purifications of Irgm1

Parameters	Purification 4	Purification 5	Purification 6
Cell Line	High Five	High Five	High Five
Cell Volume	400 ml	400 ml	400 ml
Harvest Buffer	1X PBS	1X PBS	1X PBS; cocktail protease inhibitors
Sonication Buffer	1X PBS	50 mM NaCl; 50 mM TRIS; 0.1% Tween-20; 1 mM PMSF; 10 mM βME	1X PBS
Batch vs. Column	Pre-packed Ni-NTA 1 ml column on Prime	5 ml pre-packed Q column, followed by 1 ml pre-packed Ni-NTA column on Purifier	Pre-packed Ni-NTA 1 ml column on Prime
Binding Time	Approx. 45 minutes	2 hours/ Q col. 45 min./Ni-NTA	1 hour
Buffer A	1X PBS; 0.1% Tween-20; 1 mM PMSF; 10 mM βME	150 mM NaCl; 50 mM TRIS; 0.1% Tween-20; 1 mM PMSF; 10 mM βME; 20 mM imidazole	1X PBS; 0.1% Tween-20; 1 mM PMSF; 10 mM βME; Roche cocktail protease inhibitors
Volume of Ni Wash	Imidazole gradient	20 ml	30 ml
Buffer B	1X PBS; 0.1% Tween-20; 1mM PMSF; 10 mM βME; 1M imidazole	150 mM NaCl; 50 mM TRIS; 0.1% Tween-20; 1 mM PMSF; 10 mM βME; 1M imidazole	1X PBS; 0.1% Tween-20; 1 mM PMSF; 10 mM βME; 150 mM imidazole
Volume of Ni Eluent	Imidazole gradient; final eluent volume 20 ml	15 ml	12 ml
Concentrator cut-off	30 kDa	30 kDa	30 kDa
Buffer C	N/A	150 mM NaCl; 50 mM TRIS	1X PBS; 0.1% Tween-20
Cleavage O/N with TEV?	No.	No.	No.
Gel Filtration?	No.	Yes. 1X PBS; 0.1% Tween20; 1 mM DTT	Yes. 1X PBS; 0.1% Tween20; 1 mM DTT

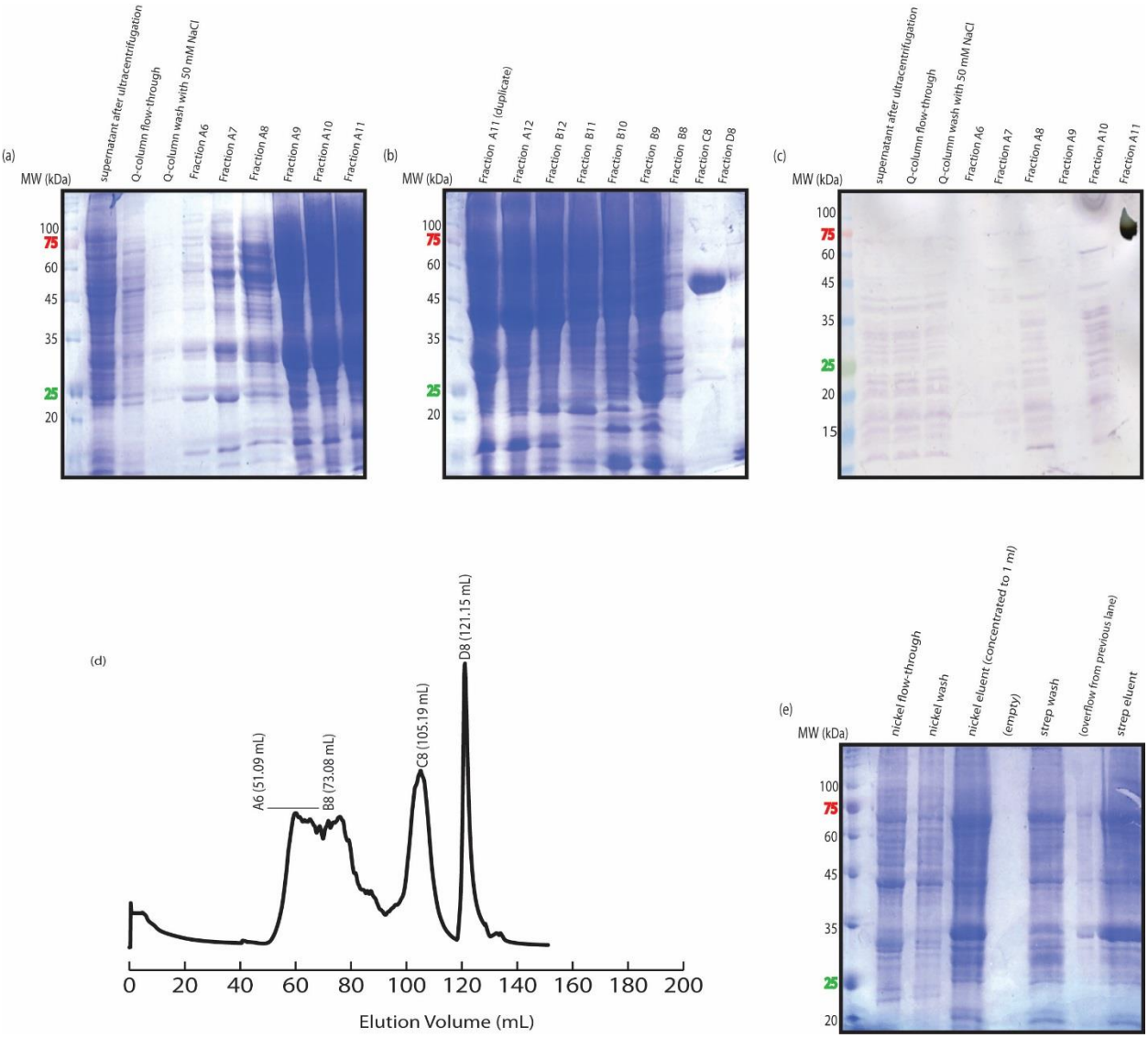

Figure 16: (a) SDS PAGE gel corresponding to Purification 5 in table 9 and shows fractions A6-A11 corresponding to the chromatogram displayed in panel (d) of this figure. (b) SDS PAGE gel corresponding to fractions A11-B8, as well as C8 and D8, shown in the chromatogram displayed in panel. (c) Western Blot corresponding to the SDS-PAGE gel in panel (a) of this figure. (d) Chromatogram of the anion exchange step in purification 5. (e) SDS PAGE gel corresponding to the nickel affinity step in purification 5.

Following purification 3, it was decided to opt for a strategy that might minimize the contaminating proteins and/or degradation products during the nickel affinity step. Additionally, the detergent was changed to 0.1% Tween-20 due to its lower cmc and lack of UV absorbance – an important feature if detergent were to be applied onto the gel filtration column. Accordingly, purification 4 (described in table 9) shows that Buffer A contained no imidazole but a step-wise gradient ranging from 10 mM to 1M imidazole was performed. Figures 15(c) and 15(d) show the

SDS-PAGE gels from the purification, with each lane corresponding to an imidazole concentration within the range indicated. At 100-200 mM imidazole (figure 15(d)), we begin to see the protein eluting, with a band at ~73kDa. The gel also shows another band at ~32 kDa, indicating the presence of the His-Strep-GFP tag. It seems from figures 15(c) and 15(d) that the tag and the full-length construct are difficult to separate on the nickel column, so this strategy failed to separate the construct from its degradation products. It was then suggested that the protein be purified by adding the lysate first onto an anion-exchange column, where the full length construct would bind to the column at low salt conditions (pI: 6.55).

Table 9 describes the details of purification 5, featuring the anion exchange Q column. The purification was performed by applying the lysate onto a 5 ml pre-packed anion exchange column, on the AKTA Purifier, in the same manner used in previous purifications for applying the lysate onto a pre-packed nickel column. A continuous salt gradient was performed, ranging from 50 mM NaCl to 1M NaCl (with detergent present). It is worth noting that due to the dense lysate produced from the insect cells (presence of DNA), the loading onto the anion exchange column took place in two hours – a time long enough to allow for protein degradation. The protein was then eluted off of the anion-exchange column and applied into a nickel column. Figure 16 shows the SDS-PAGE gels from this purification, with panels (a), (b) and (c) showing the fractions from the anion exchange column; panel (d) shows the chromatogram; panel (e) shows the SDS PAGE of the subsequent nickel and strep affinity steps. Overall, the SDS-PAGE gels do not indicate any improvement of clearing the protein from contaminants nor degradation products. In fact, we continue to see the prominent His-Strep-GFP band in the strep eluent in figure 16(e).

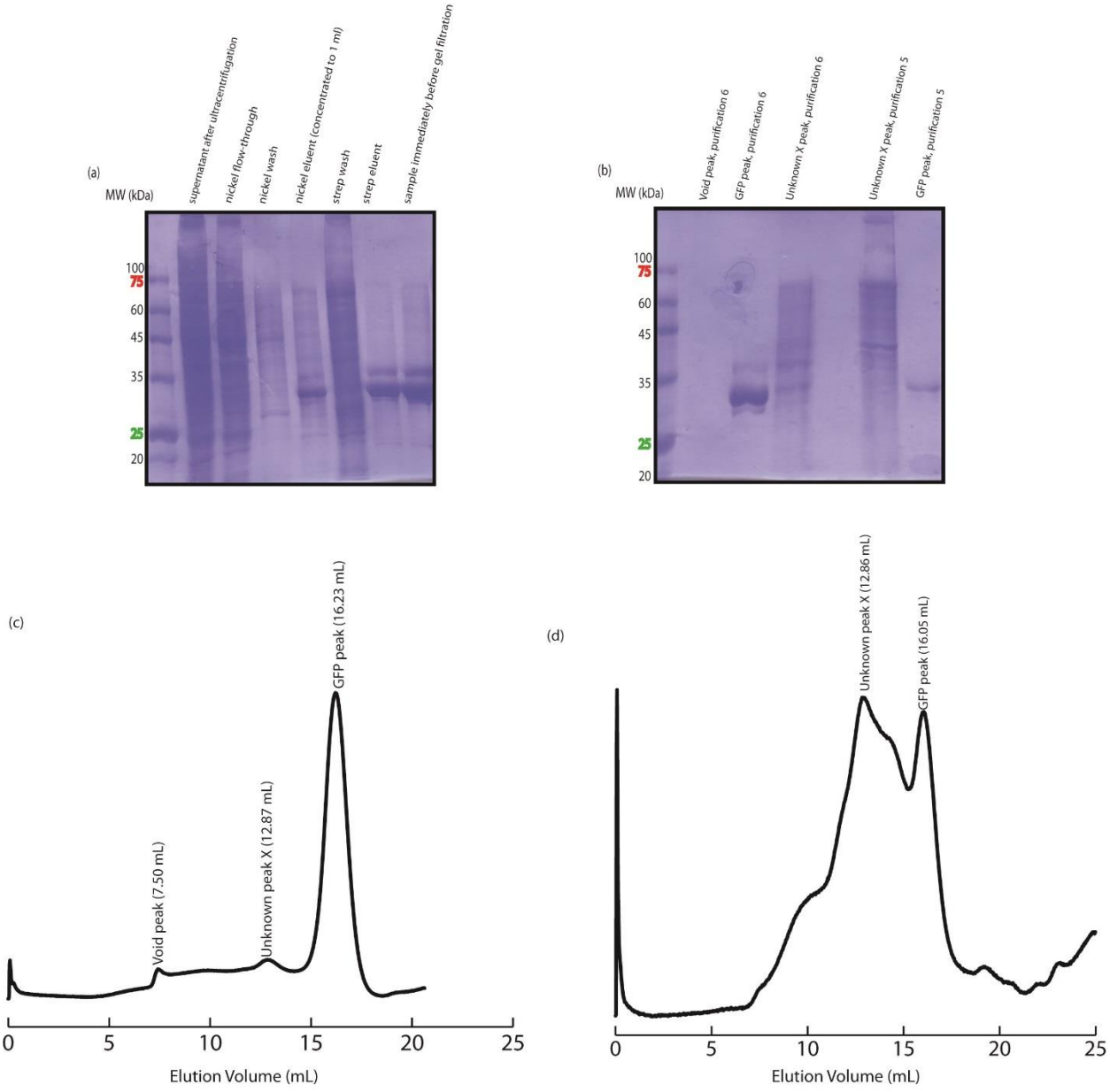

Figure 17: (a) SDS PAGE gel corresponding to Purification 6 in table 9. (b) The first 3 lanes on this SDS PAGE gel correspond to the gel filtration step in purification 6. The last two lanes correspond to the gel filtration step of the protein purified from purification 5 described in table 9. (c) Chromatogram of the gel filtration step in purification 6. (d) Chromatogram of the gel filtration step from purification 5.

Finally, purification 6 (also described in table 9), features the use of cocktail protease inhibitors consistently throughout the purification. The purification was performed following the same protocol from the first purification, with the difference of using Tween20 instead of Triton X-100 and the exception of not cleaving the protein with TEV. The uncleaved protein was then

subjected to size-exclusion chromatography. Figures 17(a), (b) and (c) show the SDS-PAGE gels and chromatogram for this purification. The chromatogram in figure 17(c) shows three distinct peaks: the void volume peak, unknown peak at 12.87 ml, and possibly a GFP peak at 16.23 ml. indeed, the first half of the SDS gel in figure 17(b) shows that the peak at 16.23 ml belongs to the GFP tag. The shoulder peak at 12.87 ml shows some degradation products but no clear presence of the Irgm1 protein. The void volume also appears inconclusive. Simultaneously, another size-exclusion chromatography was performed on the protein sample from purification 5 (see table 9) and its corresponding chromatogram appears in figure 17(d). The graph shows several peaks, with two prominent ones at 12.86 ml and 16.05 ml (suspected GFP peak) – again, the former appearing as a shoulder peak to the latter. Both these experiments were performed on the S200 column. The latter half of the SDS gel in figure 17(b) shows samples from the unknown shoulder peak at 12.86 ml and the suspected GFP peak from purification 5. Interestingly, the unknown peak shows a band ~73 kDa and another one around ~42 kDa, which might indicate the presence of the protein construct. However, due to its very low concentration, the protein could not be recovered appropriately and its yield was too low to be determined properly.

With the failure of several of these attempts in recovering a pure form of the Irgm1 protein, it was decided to move on to other constructs. One construct that had prominent expression in Sf9 insect cells and showed potential promise upon its purification was the mutant Palm -αK/ Irgm1(ins. 362, 367E; C371,373,374,375A) – a mutant described by Henry et al. (2014). Dr. Gregory Taylor kindly provided the cDNA necessary for the cloning and production of this specific construct that has already been shown to dissociate largely from membranes due to its mutated palmitoylation sites (C371,373,374,375A) and disrupted amphipathic nature of its alphaK helix that anchors it to the Golgi membranes (ins. 362, 367E).

Palm -αK/ Irgm1(ins. 362, 367E; C371,373,374,375A)

Table 10: Purifications of Palm-alphaK

Parameters	Purification 1	Purification 2	Purification 3
Cell Line	Sf9 (P3)	Sf9	Sf9
Cell Volume	200 ml	400 ml	400 ml
Harvest Buffer	1X PBS	1X PBS	1X PBS; Roche cocktail protease inhibitors
Sonication Buffer	1X PBS 1 mM PMSF 10 mM βME	1X PBS 1 mM PMSF Roche cocktail inhibitors 10 mM βME	1X PBS 1 mM PMSF Roche cocktail inhibitors 10 mM βME
Batch vs. Column	1ml Pre-packed Ni-NTA column on Prime	1ml Pre-packed Ni-NTA column on Purifier	1ml Pre-packed Ni-NTA column on Prime
Binding Time	Approx. 45 min	Approx. 45 min.	Approx. 1 hour
Buffer A	1X PBS 1 mM PMSF 10 mM βME 20 mM imidazole	1X PBS 1 mM PMSF Roche cocktail inhibitors 10 mM βME 30 mM imidazole	1X PBS 1 mM PMSF Roche cocktail inhibitors 10 mM βME 30 mM imidazole
Volume of Ni Wash	35 ml	35 ml	35 ml
Buffer B	1X PBS 1 mM PMSF 10 mM βME 200 mM imidazole	1X PBS 1 mM PMSF 10 mM βME 200 mM imidazole	1X PBS 1 mM PMSF 10 mM βME 200 mM imidazole
Volume of Ni Eluent	20 ml	12 ml	22 ml
Concentrator cut-off	30 kDa	30 kDa	30 kDa
Buffer C	1X PBS	1X PBS	N/A
Cleavage O/N with TEV?	No.	No.	No.

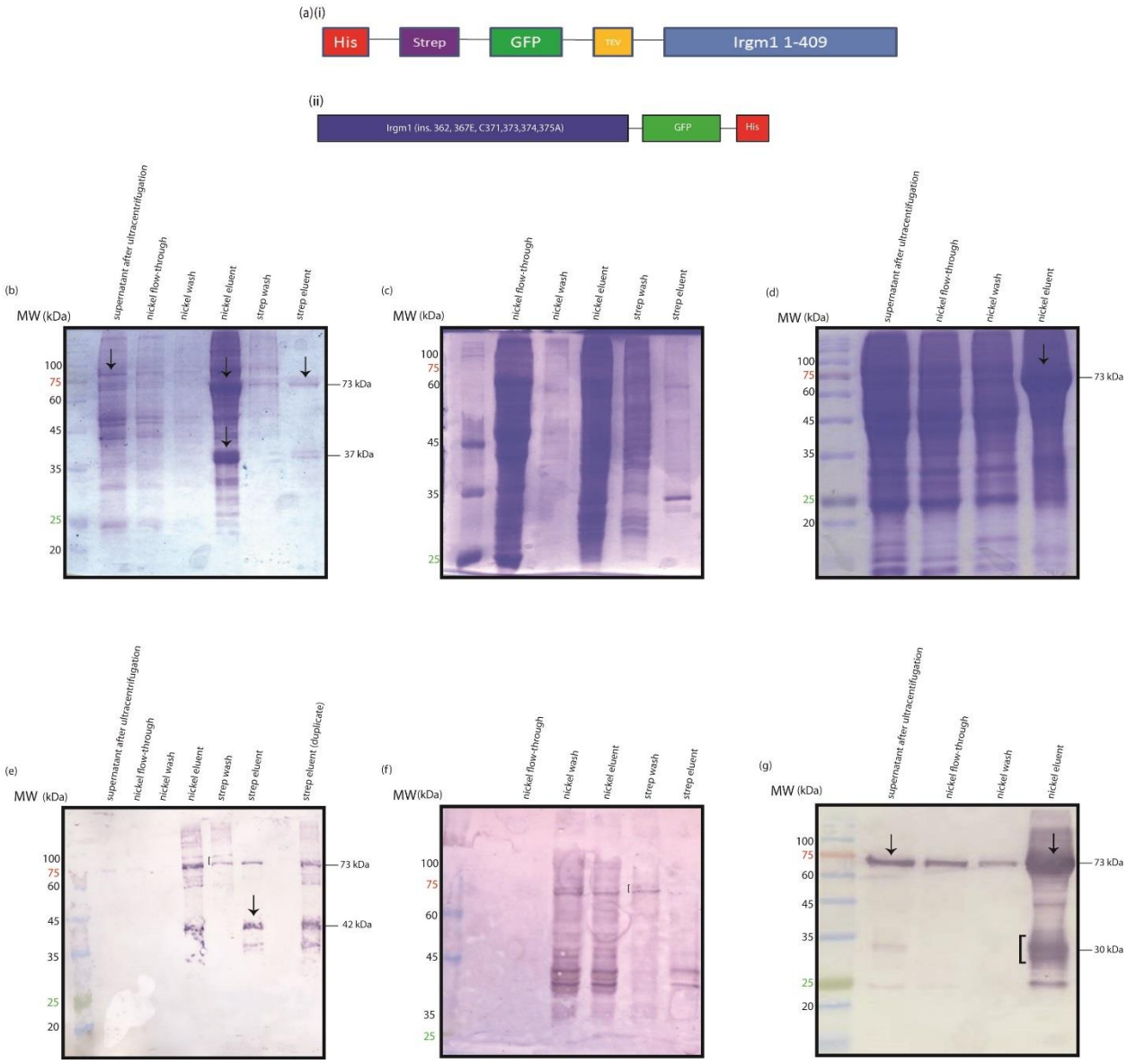

Figure 18: (a)-i Cartoon representation of the His-Strep-GFP-PalmαK construct. (a)-ii Cartoon representation of the PalmαK-GFP-His construct. (b) SDS PAGE gel corresponding to Purification 1 in table 10. (c) SDS PAGE gel corresponding to Purification 2 in table 10. (d) SDS PAGE gel corresponding to Purification 3 in table 10. (e) Western Blot corresponding to the SDS PAGE gel in panel (b) of this figure. (f) Western Blot corresponding to the SDS PAGE gel in panel (c) of this figure. (g) Western Blot corresponding to the SDS PAGE gel in panel (d) of this figure. The arrows indicated on several gels and WBs point to prominent protein bands on the gel around 42 kDa and 73 kDa. The vertical black lines seen in the WBs highlight the spectrum of what is possibly degradation products throughout the purifications.

Table 10 highlights the details concerning three purifications of the Palm-αK Irgm1 construct. Purifications 1 and 2 concern the His-Strep-GFP-PalmαK construct (79.3 kDa, figure 18(a)-(i)), whereas purification 3 concerns the PalmαK-GFP-His construct (Figure 18(a)-(ii)). The latter was designed in order to tackle the C-terminal degradation problem occurring during the purifications of the IRGM proteins, with reasoning that a C-terminal tag may help protect the protein from C-terminal degradation as seen in the purifications of Irgm1 and in this section.

The first two purifications were conducted in the same manner as the first purification of the Irgm1 construct (see previous section), in terms of buffer concentrations and protocol. The advantage of using the PalmαK construct was that it was feasible to purify it without the use of detergent, as opposed to its counterpart, the full-length Irgm1 construct. Indeed, the reason for this is largely due to the fact that the insertions disrupting the amphipathicity of the αK helix (ins. 362, 367 E) of the protein, plus the point mutations (C371, 373, 374, 375) that prevent palmitoylation have the advantage of abolishing the tight binding ability of Irgm1 to the membrane, thus making it more soluble.

The difference between the first two purifications of His-Strep-GFP-PalmαK construct (Table 10) lay in the addition of cocktail protease inhibitors in purification 2 to minimize degradation products that were largely dominant on the SDS-PAGE gel and Western Blot of the first purification (figure 18(b) & (e)). From figure 18(c) & (f), which depict the SDS-PAGE gel and the Western Blot from the 2nd purification, we observe that the cocktail protease inhibitors had minimal effects in reducing the degradation products. In fact, the SDS gel in figure 18(b) shows improved purity of the construct (band at ~73 kDa), but the protein yield was too low and could not be recovered for gel filtration analysis.

A final attempt to purify a version of this protein was carried out with the PalmαK-GFP-His construct, with only preliminary analysis conducted due to time limitations. Figure 18(d) & (g) show the SDS-PAGE gel and the Western Blot of the 3rd purification and a clear band at ~73 kDa corresponding to the protein construct can clearly be seen throughout the purification. However, some degradation or contaminants still persist. From the results of the Western Blot in Figure18(g), this purification shows the most promise and could potentially be expanded upon in terms of refined purification techniques that can produce protein of high yield and quality to be analyzed by X-ray crystallography. These purification techniques might be utilization of an anion exchange column, cleavage of the tag and performing a second nickel step and then

finally, size-exclusion chromatography. Additional tags and constructs may be explored but overall, the PalmαK construct shows the most promise among its counterparts. However, a GTPase assay or another functional assay needs to be employed in order to test the protein for correct folding and hence functionality, especially since this is the mutated version of the protein which does not perform its physiological function of binding to membranes.

2.3 Materials

All chemicals were purchased from BioShop and Sigma-Aldrich.

BioBasic Miniprep Kit

BLUeye Prestained Protein Ladder, Tris-Glycerine 4-20%

Competent cell glycerol stocks were purchased from Novagen and Agilent Technologies

Enzymes and associated buffers for cloning were purchased from New England BioLabs (NEB).

pProEx HTb (Invitrogen, Carlsbad)

pSMT3 vector provided by Mossessova, E. and Lima, (C.D. (2000) Ulp1-SUMO. *Mol. Cell* 5, 865-67)

Qiagen PCR purification and Gel Extraction Kits

SMOBIO ExcelBand 3-color Pre-stained Protein Marker Broad Range PM2700

Equiment:

ÄKTAprime GE Healthcare

ÄKTApure GE Healthcare

ÄKTApurifier GE Healthcare

Allegra X-15R Centrifuge Beckman Coulter

Amicon Ultra-15 Centrifugal Filter Units EMD Millipore

Avanti J-26XP Centrifuge Beckman Coulter

Branson Digital Sonifier 102C (CE)

EmulsiFlex-C3 Avestin

HisPur Ni-NTA Resin Thermo Scientific

MaxQ 8000 Incubated Stackable Shaker Thermo Scientific

Microcentrifuge 5415R Eppendorf

MJ Mini Thermal Cycler BIO-RAD

Multitron Standard, incubation shaker Infors HT

NanoVue Plus GE Healthcare

PowerPac Basic Power Supply BIO-RAD

Sorvall RT3 Centrifuge Thermo Scientific

Superdex 200 Increase 10/300 GL GE Healthcare

2.4 Discussion

The literature on the *IRGM* proteins, reviewed in the first chapter of this thesis shows a remarkable gap in our understanding of these proteins. Preliminary studies by Tiwari et al. (2009) and Henry et al. (2014) shed some light on the membrane binding capabilities of the murine Irgm1, as well as its role in autophagy. More recent studies by the Deretic group (Chauhan et al., 2015) have put forth a model for how the human IRGM functions within the autophagic pathway and how it could potentially act as a scaffolding protein that binds to NOD2 and other proteins to drive the autophagic response. The claims in this paper have not been investigated further and one might propose the question of whether a 20 kDa protein might indeed be acting as a scaffold due to its small size. A speculation would be that it is oligomerizing, in which case it must have another domain or surface attribute that allows it to do so.

The purpose of the study described in this thesis was to contribute further to the literature on these IRGM proteins and gain more understanding into its mechanism of action within the cell, especially under infection conditions. The 3D structure of the protein would be able to answer some questions such as:

- How does the GTPase domain of human IRGM function?
- How structurally related are the human IRGM and murine Irgm1, given that the human IRGM is a truncated version of the mouse orthologue?
- Do other domains (other than the GTPase domain) for both murine and human versions of the IRGM protein exist?
- How is the murine Irgm1 structurally binding to membranes? How does a conformational change cause it to migrate to phagosomal membranes and activate its G-domain?
- How does the human IRGM bind to NOD2?
- Are there other functions for these proteins?

Naturally, a combination of cellular biology, biochemical analysis and structural biology is essential to be able to answer all of these questions and advance our knowledge on these mysterious proteins.

From a structural biology point of view, working with the IRGM proteins was a real challenge: the proteins are largely unstable and degrade rapidly even with the presence of

cocktail protease inhibitors, and applying sensitive conditions (such as maintaining temperature at 4^{o}C at all times). An approach that has shown considerable potential to work for the human IRGM was the denaturation and re-folding method, which requires a massive amount of bacterial cell since it is expected that considerable amounts of the protein will be lost during dialysis (Section 2.2.3.2 IRGM23-181).

On the other hand, the murine Irgm1 protein adds another fold of complexity to the problem: it has membrane-binding properties. During preliminary expression trials in insect cells, the protein, tagged with GFP, showed a punctate distribution and an affinity for cellular membranes (pictures not shown). Indeed, the use of detergent was absolutely required to extract the full-length wildtype protein from cellular lysates, and gradual dilution of the detergent causes the protein to aggregate. Although the PalmαK construct shows approximately the same pattern under the microscope, it does not require detergent during the purifications and shows the most promise among all IRGM constructs attempted. The only issue is, if crystals were indeed obtained and the structure solved, how would the disruption in the αK helix pan out, structurally? Would it change the protein conformation considerably?

In retrospect, there seems to be a missing factor that might help stabilize these proteins in solution and *in vivo*. At first, it was suspected that adding this factor was Mg^{2+} magnesium (see purifications of IRGM(3-181), section 2.2.3.1), but addition of this cofactor did not improve the situation at all. As mentioned earlier, the Deretic group (Chauhan et al.2015) indeed show that the IRGM protein is versatile and binds to many different proteins, so one might speculate that this protein requires binding partners or even oligomerization for stability. However, pull-down assays aren't always definitive of whether proteins interact with each other directly. In this case, it would also be interesting to confirm the direct protein-protein interaction between human IRGM and NOD2, especially with NOD2's CARD domains. This can be done on a size exclusion column using the pure form of both proteins and/or co-expression in bacterial/insect cells.

Finally, a definitive GTPase or another functional assay needs to be performed on any IRGM protein construct that has a claim to a pure protein. In fact, it might be prudent to conduct these tests before one plunges into solving the crystal structure of the protein.

Researchers wishing to investigate this protein further might now have some helpful information as to what to avoid while attempting to express and purify the *IRGM* proteins.

To summarize:

- Regarding the human IRGM, it fares better to work with the isoform IRGM(d) as opposed to IRGM(a).
- Regarding the murine Irgm1, it fares better to employ detergents such as Tween-20 or DDM that have relatively low cmc, low micelle molecular weight and are mild.
- For all constructs, it might be better to incorporate a tag at the c-terminus or employ a double tag that may protect from degradation further.
- The PalmαK/Irgm1(ins. 362, 367E; C371,373,374,375A) construct does not require detergent, can be easily dissociated from membranes and is convenient to work with for preliminary analysis.

2.5 Acknowledgements

This work could not have been possible or complete without the diligence and massive contribution of Katalin Illes Kocsis who provided me with the pFastBac cloning empty vectors and primers, and performed all the work related to the expression and production of the recombinant proteins in insect cell lines. Her insight and experience in cellular and molecular biology along with her knowledge of molecular pathways were absolutely indispensable for the advancement of the project. Additionally, constant guidance and supervision from Dr. Bhushan Nagar allowed the project to move as smoothly as possible, despite the challenges and difficulties it presented.

Many thanks also to Jon Labriola who freely offered guidance on operating the gel filtration machinery and ideas on how to troubleshoot expression in bacteria and to screen for suitable detergents.

Finally, I would like to extend thanks to Dr. Gregory Taylor, Associate Professor of Molecular Genetics and Microbiology at Duke University, who was very helpful by correspondence and also provided us with the Palm-αK and the αK cDNA.

Epilogue

Overall, it is important to continue researching the IRGM proteins as they play a pivotal role in autophagy and ridding the body of virulent pathogens such as *Mtb, T. gondii and Litseria monocytogens.* Their involvement in disease is not only restricted to protection from pathogens; disruption due to genetic fallacies in the *IRGM* gene interfere with the process of autophagy, resulting in chronic diseases such as Crohn's, IBD and ultimately gastric and colon cancer.

This project began with the aim of obtaining a crystal structure of the IRGM protein in an effort to understand its mechanism of action in autophagy along with its specific role in disease. However, the difficulties and challenges involved with the expression and purification of these proteins as illustrated in this thesis created a major obstacle in reaching this aim. On the other hand, the conditions that were tried and sampled, described herein and in the literature, has increased the overall chances of finding optimal conditions for obtaining a pure soluble protein that can be and further analyzed through X-ray crystallography. The conditions that were sampled included: salt concentrations, protease inhibitors, use of detergent, expression systems.

In this day and age where antibiotic resistance has become prevalent and chronic illnesses progress to complex forms of cancer, it is important to do our best to understand the biological underworkings of our basic homeostasis processes. Eventually, the more we know, the better we understand, and the more we are capable of developing therapeutic strategies that have a high probability and more positive outcome in curing – or at least largely alleviating – illnesses resulting from aberrations in these proteins.

References

Admiraal, S. J., & Herschlag, D. (1995). Mapping the transition state for ATP hydrolysis: implications for enzymatic catalysis. *Chemistry & biology*, *2*(11), 729-739.

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

Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002. Innate Immunity. Available from: http://www.ncbi.nlm.nih.gov/books/NBK26846/

Allin, C., & Gerwert, K. (2001). Ras catalyzes GTP hydrolysis by shifting negative charges from γ-to β-phosphate as revealed by time-resolved FTIR difference spectroscopy. *Biochemistry*, *40*(10), 3037-3046.

Bekpen, C., Hunn, J. P., Rohde, C., Parvanova, I., Guethlein, L., Dunn, D. M., ... & Howard, J. C. (2005). The interferon-inducible p47 (IRG) GTPases in vertebrates: loss of the cell autonomous resistance mechanism in the human lineage. *Genome biology*, *6*(11), R92.

Bekpen, C., Marques-Bonet, T., Alkan, C., Antonacci, F., Leogrande, M. B., Ventura, M., ... & Eichler, E. E. (2009). Death and resurrection of the human IRGM gene. *PLoS genetics*, *5*(3), e1000403.

Bilevich, K. A., Bubnov, N. N., Okhlobystin, O. Y., & Radzhabov, N. G. (1971). Mechanism of nucleophilic substitution. *Russian Chemical Bulletin*, *20*(8), 1704-1706.

Boehm, U., Guethlein, L., Klamp, T., Ozbek, K., Schaub, A., Fütterer, A., ... & Howard, J. C. (1998). Two families of GTPases dominate the complex cellular response to IFN-γ. *The Journal of Immunology*, *161*(12), 6715-6723.

Boehm, U., Klamp, T., Groot, M., & Howard, J. C. (1997). Cellular responses to interferon-γ. *Annual review of immunology*, *15*(1), 749-795

Bos, J. L., Rehmann, H., & Wittinghofer, A. (2007). GEFs and GAPs: critical elements in the control of small G proteins. *Cell*, *129*(5), 865-877.

Bougneres, L., Helft, J., Tiwari, S., Vargas, P., Chang, B. H. J., Chan, L., ... & Guermonprez, P. (2009). A role for lipid bodies in the cross-presentation of phagocytosed antigens by MHC class I in dendritic cells. *Immunity*, *31*(2), 232-244.

Brest, P., Lapaquette, P., Souidi, M., Lebrigand, K., Cesaro, A., Vouret-Craviari, V., ... & Hofman, P. (2011). A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nature genetics*, *43*(3), 242-245.

Brest, P., Lapaquette, P., Souidi, M., Lebrigand, K., Cesaro, A., Vouret-Craviari, V., ... & Hofman, P. (2011). A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nature genetics*, *43*(3), 242-245.

Cai, Q., Lu, L., Tian, J. H., Zhu, Y. B., Qiao, H., & Sheng, Z. H. (2010). Snapin-regulated late endosomal transport is critical for efficient autophagy-lysosomal function in neurons. *Neuron*, *68*(1), 73-86.

Carlow, D. A., Teh, S. J. & Teh, H. S. Specific antiviral activity demonstrated by TGTP, a member of a new family of interferon-induced GTPases. *J. Immunol.* **161**, 2348–2355 (1998).

Carlow, D. A., Teh, S. J., & Teh, H. S. (1998). Specific antiviral activity demonstrated by TGTP, a member of a new family of interferon-induced GTPases. *The Journal of Immunology*, *161*(5), 2348-2355.

Chauhan, S., Mandell, M. A., & Deretic, V. (2015). IRGM Governs the Core Autophagy Machinery to Conduct Antimicrobial Defense. *Molecular cell*.

Chen, X., Bi, Y., Wang, T., Li, P., Yan, X., Hou, S., ... & Bi, L. (2015). Lysosomal Targeting with Stable and Sensitive Fluorescent Probes (Superior LysoProbes): Applications for Lysosome Labeling and Tracking during Apoptosis. *Scientific reports*, *5*.

Chheda, M. G., Ashery, U., Thakur, P., Rettig, J., & Sheng, Z. H. (2001). Phosphorylation of Snapin by PKA modulates its interaction with the SNARE complex. *Nature cell biology*, *3*(4), 331-338.

Christoforidis, S., Miaczynska, M., Ashman, K., Wilm, M., Zhao, L., Yip, S. C., ... & Zerial, M. (1999). Phosphatidylinositol-3-OH kinases are Rab5 effectors.*Nature cell biology*, *1*(4), 249-252.

Collazo, C. M., Yap, G. S., Sempowski, G. D., Lusby, K. C., Tessarollo, L., Woude, G. F. V., ... & Taylor, G. A. (2001). Inactivation of LRG-47 and IRG-47 reveals a family of interferon γ–inducible genes with essential, pathogen-specific roles in resistance to infection. *The Journal of experimental medicine*,*194*(2), 181-188.

Darwich, L., Coma, G., Peña, R., Bellido, R., Blanco, E. J., Este, J. A., ... & Bofill, M. (2009). Secretion of interferon-γ by human macrophages demonstrated at the single-cell level after costimulation with interleukin (IL)-12 plus IL-18. *Immunology*, *126*(3), 386-393.

Darwin, K. H., Ehrt, S., Gutierrez-Ramos, J. C., Weich, N., & Nathan, C. F. (2003). The proteasome of Mycobacterium tuberculosis is required for resistance to nitric oxide. *Science*, *302*(5652), 1963-1966.

Deretic, V. (2008). *Autophagosome and phagosome* (pp. 1-10). Humana Press.

DeVay, R. M., Dominguez-Ramirez, L., Lackner, L. L., Hoppins, S., Stahlberg, H., & Nunnari, J. (2009). Coassembly of Mgm1 isoforms requires cardiolipin and mediates mitochondrial inner membrane fusion. *The Journal of cell biology*,*186*(6), 793-803.

Du, X., Frei, H., & Kim, S. H. (2000). The mechanism of GTP hydrolysis by Ras probed by Fourier transform infrared spectroscopy. *Journal of Biological Chemistry*, *275*(12), 8492-8500.

Du, X., Poltorak, A., Silva, M., & Beutler, B. (1999). Analysis of Tlr4-mediated LPS signal transduction in macrophages by mutational modification of the receptor. *Blood Cell*
Ehrt, S. *et al*. Reprogramming the macrophage transcriptome in response to interferon-ɣ and *mycobacterium tuberculosis*: signaling roles for nitric oxide sythase-2 and phagosome oxidase. *J. Exp. Med.* **194**, 1123–1140 (2001).

ever, T. E., Glynias, J. & Merrick, W. C. GTP-binding domain: three consensus sequence elements with distinct spacing. *Proc. Natl Acad. Sci. USA* **84**, 1814–1818 (1987).

Ferguson, K. M., Kavran, J. M., Sankaran, V. G., Fournier, E., Isakoff, S. J., Skolnik, E. Y., & Lemmon, M. A. (2000). Structural basis for discrimination of 3-phosphoinositides by pleckstrin homology domains. *Molecular cell*, *6*(2), 373-384.

Fruh, K., Karlson, L. & Yang, Y. In ɣ-*Interferon in Antiviral defense* (ed. Karupiah, G.) 39 (Springer, Heidelberg, Germany, 1997).

Gebert, N., Joshi, A. S., Kutik, S., Becker, T., McKenzie, M., Guan, X. L., ... & Pfanner, N. (2009). Mitochondrial cardiolipin involved in outer-membrane protein biogenesis: implications for Barth syndrome. *Current Biology*, *19*(24), 2133-2139.

Gebert, N., Ryan, M. T., Pfanner, N., Wiedemann, N., & Stojanovski, D. (2011).

Mitochondrial protein import machineries and lipids: A functional connection. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, *1808*(3), 1002-1011.

Ghosh, A., Uthaiah, R., Howard, J., Herrmann, C., & Wolf, E. (2004). Crystal structure of IIGP1: a paradigm for interferon-inducible p47 resistance GTPases.*Molecular cell*, *15*(5), 727-739.

Gibbs, J. B., Marshall, M. S., Scolnick, E. M., Dixon, R. A. F. & Vogel, U. S.Modulation of guanine nucleotides bound to Ras in NIH3T3 cells by oncogenes, growth factors, and GTPase activating protein (GAP). *J. Biol. Chem.* **265**, 20437–20442 (1990).

Gilly, M. I. C. H. A. E. L., & Wall, R. A. N. D. O. L. P. H. (1992). The IRG-47 gene is IFN-gamma induced in B cells and encodes a protein with GTP-binding motifs. *The Journal of Immunology*, *148*(10), 3275-3281.

Gräslund, S., Nordlund, P., Weigelt, J., Bray, J., Gileadi, O., Knapp, S., ... & Zhang, F. (2008). Protein production and purification. *Nature methods*, *5*(2), 135-146.

Grégoire, I. P., Richetta, C., Meynie f-3l-Schicklin, L., Borel, S., Pradezynski, F., Diaz, O., ... & Faure, M. (2011). IRGM is a common target of RNA viruses that subvert the autophagy network. *PLoS Pathog*, *7*(12), e1002422.

Hakoshima, T., Shimizu, T., & Maesaki, R. (2003). Structural basis of the Rho GTPase signaling. *Journal of biochemistry*, *134*(3), 327-331.

Hawkins, P. T., Anderson, K. E., Davidson, K., & Stephens, L. R. (2006). Signalling through Class I PI3Ks in mammalian cells. *Biochemical Society Transactions*, *34*(5), 647-662.

Henry, S. C., Schmidt, E. A., Fessler, M. B., & Taylor, G. A. (2014). Palmitoylation of the immunity related GTPase, Irgm1: impact on membrane localization and ability to promote mitochondrial fission. *PloS one*, *9*(4), e95021.

Hoffmann, J. A., Kafatos, F. C., Janeway, C. A., & Ezekowitz, R. A. B. (1999). Phylogenetic perspectives in innate immunity. *Science*, *284*(5418), 1313-1318.

Hoshino, K., Takeuchi, O., Kawai, T., Sanjo, H., Ogawa, T., Takeda, Y., ... & Akira, S. (1999). Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *The Journal of Immunology*, *162*(7), 3749-3752.

Howard, J. C., Hunn, J. P., & Steinfeldt, T. (2011). The IRG protein-based resistance mechanism in mice and its relation to virulence in Toxoplasma gondii. *Current opinion in microbiology*, *14*(4), 414-421.

http://www.sivabio.50webs.com/lysosomes.htm

Hundeshagen, P., Hamacher-Brady, A., Eils, R., & Brady, N. R. (2011). Concurrent detection of autolysosome formation and lysosomal degradation by flow cytometry in a high-content screen for inducers of autophagy. *BMC biology*, *9*(1), 38.

Itakuraa, E., & Mizushima, N. (2010). Characterization of autophagosome formation site by a hierarchical analysis of mammalian Atg proteins. *Autophagy*, *6*(6), 764-776

Kamen, L. A., Levinsohn, J., & Swanson, J. A. (2007). Differential association of phosphatidylinositol 3-kinase, SHIP-1, and PTEN with forming phagosomes.*Molecular biology of the cell*, *18*(7), 2463-2472.

Khaminets, A., Hunn, J. P., Könen-Waisman, S., Zhao, Y. O., Preukschat, D., Coers, J., ... & Howard, J. C. (2010). Coordinated loading of IRG resistance GTPases on to the Toxoplasma gondii parasitophorous vacuole. *Cellular microbiology*, *12*(7), 939-961.

Klionsky, D. J., Eskelinen, E. L., & Deretic, V. (2014). Autophagosomes, phagosomes, autolysosomes, phagolysosomes, autophagolysosomes… Wait, I'm confused. *Autophagy*, *10*(4), 549-551.

Levine, B., Mizushima, N., & Virgin, H. (2011). Autophagy in Immunity and Inflammation. Nature, 469, 323-335. doi:10.1038.

Liu, B., Gulati, A. S., Cantillana, V., Henry, S. C., Schmidt, E. A., Daniell, X., ... & Taylor, G. A. (2013). Irgm1-deficient mice exhibit Paneth cell abnormalities and increased susceptibility to acute intestinal inflammation.*American Journal of Physiology-Gastrointestinal and Liver Physiology*, *305*(8), G573-G584.

MacMicking, J. D. (2005). Immune control of phagosomal bacteria by p47 GTPases. *Current opinion in microbiology*, *8*(1), 74-82.

Maegley, K. A., Admiraal, S. J., & Herschlag, D. (1996). Ras-catalyzed hydrolysis of GTP: a new perspective from model studies. *Proceedings of the National Academy of Sciences*, *93*(16), 8160-8166.

Martens, S., & Howard, J. (2006). The interferon-inducible GTPases. *Annu. Rev. Cell Dev. Biol.*, *22*, 559-589.

Martens, S., Sabel, K., Lange, R., Uthaiah, R., Wolf, E., & Howard, J. C. (2004). Mechanisms regulating the positioning of mouse p47 resistance GTPases LRG-47 and IIGP1 on cellular membranes: retargeting to plasma membrane induced by phagocytosis. *The Journal of Immunology*, *173*(4), 2594-2606.

Matsuzawa, T., Kim, B. H., Shenoy, A. R., Kamitani, S., Miyake, M., & MacMicking, J. D. (2012). IFN-γ elicits macrophage autophagy via the p38 MAPK signaling pathway. *The Journal of Immunology*, *189*(2), 813-818.

McCarroll, S. A., Huett, A., Kuballa, P., Chilewski, S. D., Landry, A., Goyette, P., ... & Xavier, R. J. (2008). Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nature genetics*, *40*(9), 1107-1112.

McLaughlin, S. , Wang, J. , Gambhir, A. & Murray, D. PIP_2 and proteins: interactions, organization, and information flow. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 151–175
McLaughlin, S., & Murray, D. (2005). Plasma membrane phosphoinositide organization by protein electrostatics. *Nature*, *438*(7068), 605-611.

Medzhitov, R., & Janeway, C. A. (1997). Innate immunity: the virtues of a nonclonal system of recognition. *cell*, *91*(3), 295-298.

Melzer, T., Duffy, A., Weiss, L. M., & Halonen, S. K. (2008). The gamma interferon (IFN-γ)-inducible GTP-binding protein IGTP is necessary for Toxoplasma vacuolar disruption and induces parasite egression in IFN-γ-stimulated astrocytes. *Infection and immunity*, *76*(11), 4883-4894.

Miyairi, I., Tatireddigari, V. R. A., Mahdi, O. S., Rose, L. A., Belland, R. J., Lu, L., ... & Byrne, G. I. (2007). The p47 GTPases Iigp2 and Irgb10 regulate innate immunity and inflammation to murine Chlamydia psittaci infection. *The Journal of Immunology*, *179*(3), 1814-1824.

Ng, V. H., Cox, J. S., Sousa, A. O., MacMicking, J. D., & McKinney, J. D. (2004). Role of KatG catalase-peroxidase in mycobacterial pathogenesis: countering the phagocyte oxidative burst. *Molecular microbiology*, *52*(5), 1291-1302.

Okamoto, M., Oshikawa, T., Tano, T., Ohe, G., Furuichi, S., Nishikawa, H., ... & Sato, M. (2003). Involvement of Toll-like receptor 4 signaling in interferon-γ production and antitumor effect by streptococcal agent OK-432. *Journal of the National Cancer Institute*, *95*(4), 316-326..

Pan, P. Y., Tian, J. H., & Sheng, Z. H. (2009). Snapin facilitates the synchronization of synaptic vesicle fusion. *Neuron*, *61*(3), 412-424.

Papinski, D., Schuschnig, M., Reiter, W., Wilhelm, L., Barnes, C. A., Maiolica, A., ... & Kraft, C. (2014). Early steps in autophagy depend on direct phosphorylation of Atg9 by the Atg1 kinase. *Molecular cell*, *53*(3), 471-483.

Parkes, M., Barrett, J. C., Prescott, N. J., Tremelling, M., Anderson, C. A., Fisher, S. A., ... & Mathew, C. G. (2007). Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nature genetics*, *39*(7), 830-832.

Parvanova, I. A. (2005). *Analysis of the role of the p47 GTPase IIGP1 in resistance against intracellular pathogens* (Doctoral dissertation, Universität zu Köln).

Pfeiffer, K., Gohil, V., Stuart, R. A., Hunte, C., Brandt, U., Greenberg, M. L., & Schägger, H. (2003). Cardiolipin stabilizes respiratory chain supercomplexes.*Journal of Biological Chemistry*, *278*(52), 52873-52880.

Prescott, N. J., Dominy, K. M., Kubo, M., Lewis, C. M., Fisher, S. A., Redon, R., ... & Mathew, C. G. (2010). Independent and population-specific association of risk variants at the IRGM locus with Crohn's disease. *Human molecular genetics*, *19*(9), 1828-1839.

Prive, G. G., Milburn, M. V., Tong, L., de Vos, A. M., Yamaizumi, Z., Nishimura, S., & Kim, S. H. (1992). X-ray crystal structures of transforming p21 ras mutants suggest a transition-state stabilization mechanism for GTP hydrolysis. *Proceedings of the National Academy of Sciences*, *89*(8), 3649-3653.

Roux, C. M., Rolán, H. G., Santos, R. L., Beremand, P. D., Thomas, T. L., Adams, L. G., & Tsolis, R. M. (2007). Brucella requires a functional Type IV secretion system to elicit innate immune responses in mice. *Cellular microbiology*, *9*(7), 1851-1869.

Saitoh, T., Fujita, N., Hayashi, T., Takahara, K., Satoh, T., Lee, H., ... & Akira, S. (2009). Atg9a controls dsDNA-driven dynamic translocation of STING and the innate immune response. *Proceedings of the National Academy of Sciences*, *106*(49), 20842-20846.

Schoenborn, J. R., & Wilson, C. B. (2007). Regulation of interferon-γ during innate and adaptive immune responses. *Advances in immunology*, *96*, 41-101.

Schweins, T., Geyer, M., Scheffzek, K., Warshel, A., Kalbitzer, H. R., & Wittinghofer, A. (1995). Substrate-assisted catalysis as a mechanism for GTP hydrolysis of p21ras and other GTP-binding proteins. *Nature Structural & Molecular Biology*, *2*(1), 36-44.

Shima, F., Ijiri, Y., Muraoka, S., Liao, J., Ye, M., Araki, M., ... & Kataoka, T. (2010). Structural basis for conformational dynamics of GTP-bound Ras protein. *Journal of Biological Chemistry*, *285*(29), 22696-22705.

Shotland, Y., Krämer, H., & Groisman, E. A. (2003). The Salmonella SpiC protein targets the mammalian Hook3 protein function to alter cellular trafficking. *Molecular microbiology*, *49*(6), 1565-1576.

Singh, S. B., Davis, A. S., Taylor, G. A., & Deretic, V. (2006). Human IRGM induces autophagy to eliminate intracellular mycobacteria. *Science*, *313*(5792), 1438-1441.

Singh, S. B., Ornatowski, W., Vergne, I., Naylor, J., Delgado, M., Roberts, E., ... & Deretic, V. (2010). Human IRGM regulates autophagy and cell-autonomous immunity functions through mitochondria. *Nature cell biology*,*12*(12), 1154-1165.

Sondek, J., Lambright, D. G., Noel, J. P., Hamm, H. E., & Sigler, P. B. (1994). GTPase mechanism of Gproteins from the 1.7-Å crystal structure of transducin α-GDP AIF− 4. *Nature*, *372*(6503), 276-279.

Stamnes, M. A., Craighead, M. W., Hoe, M. H., Lampen, N., Geromanos, S., Tempst, P., & RoTHmAN, J. E. (1995). An integral membrane component of coatomer-coated transport vesicles defines a family of proteins involved in budding. *Proceedings of the National Academy of Sciences*, *92*(17), 8011-8015.

Takeda, K., & Akira, S. (2005). Toll-like receptors in innate immunity.*International immunology*, *17*(1), 1-14.
Taylor GA, Feng CG, Sher A. p47 GTPases: regulators of immunity to intracellular pathogens. Nat Rev Immunol. 2004;4:100–109.

Taylor, G. A. *et al*. The inducibly expressed GTPase (IGTP) localizes to the endoplasmic reticulum independently of GTP binding. *J. Biol. Chem.* **272**, 10639–10645 (1997).
Taylor, G. A., Feng, C. G., & Sher, A. (2004). p47 GTPases: regulators of immunity to intracellular pathogens. *Nature Reviews Immunology*, *4*(2), 100-109.

Taylor, G. A., Stauber, R., Rulong, S., Hudson, E., Pei, V., Pavlakis, G. N., ... & Woude, G. F. V. (1997). The inducibly expressed GTPase localizes to the endoplasmic reticulum, independently of GTP binding. *Journal of Biological Chemistry*, *272*(16), 10639-10645.

Thatcher, G. R., & Kluger, R. (1989). Mechanism and catalysis of nucleophilic substitution in phosphate esters. *Adv. Phys. Org. Chem*, *25*, 99-265.

Tiwari, S., Choi, H. P., Matsuzawa, T., Pypaert, M., & MacMicking, J. D. (2009). Targeting of the GTPase Irgm1 to the phagosomal membrane via PtdIns (3, 4) P2 and PtdIns (3, 4, 5) P3 promotes immunity to mycobacteria.*Nature immunology*, *10*(8), 907-917.

Toner M, Vaio G, McLaughlin A, McLaughlin S. 1988. Adsorption of cations to phosphatidylinositol 4,5-bisphosphate. *Biochemistry*27:7435–43.

Uthaiah, R. C., Praefcke, G. J., Howard, J. C., & Herrmann, C. (2003). IIGP1, an interferon-γ-inducible 47-kDa GTPase of the mouse, showing cooperative enzymatic activity and GTP-dependent multimerization. *Journal of Biological Chemistry*, *278*(31), 29336-29343.

Virgin, H. W., & Levine, B. (2009). Autophagy genes in immunity. *Nature immunology*, *10*(5), 461-470.

Wang J, Arbuzova A, Hangyas-Mihalyne G, McLaughlin S. 2001. The effector domain of myristoylated alanine-rich C kinase substrate binds strongly to phosphatidylinositol 4,5-bisphosphate.*J. Biol. Chem.* 276:5012–19.

Wei, Y., Zhang, Y., Derewenda, U., Liu, X., Minor, W., Nakamoto, R.K., Somlyo, A.V., Somlyo, A.P., and Derewenda, Z.S. (1997) Crystal structure of RhoA-GDP and its functional implications. Nat. Struct. Biol. 4, 699–703.

Xu, Y. I., Jagannath, C., Liu, X. D., Sharafkhaneh, A., Kolodziejska, K. E., & Eissa, N. T. (2007). Toll-like receptor 4 is a sensor for autophagy associated with innate immunity. *Immunity*, *27*(1), 135-144.

Zhao, Y. O., Könen-Waisman, S., Taylor, G. A., Martens, S., & Howard, J. C. (2010). Localisation and mislocalisation of the interferon-inducible immunity-related GTPase, Irgm1 (LRG-47) in mouse cells. *PloS one*, *5*(1), e8648.

Zhu, G., Liu, J., Terzyan, S., Zhai, P., Li, G., & Zhang, X. C. (2003). High resolution crystal structures of human Rab5a and five mutants with substitutions in the catalytically important phosphate-binding loop. *Journal of Biological Chemistry*, *278*(4), 2452-2460.