XB130: *in silico* and *in vivo* studies of a novel signal adaptor protein

By

Matthew Rubacha

A thesis submitted in conformity with the requirements for the degree of Master of Science
Department of Physiology
University of Toronto

© Copyright by Matthew Rubacha 2009
Abstract

Matthew Rubacha
Master of Science, 2009
Department of Physiology, University of Toronto

XB130: *in silico* and *in vivo* studies of a novel signal adaptor protein

XB130 is a relatively unstudied novel signal adaptor protein. In the first phase of this study, an in silico search for proteins related to XB130 was conducted. Two other proteins (AFAP and AFAP1L1) were found to have a significant similarity to XB130 and were compared in detail. After an analysis of these three proteins, it was proposed that they are members of a novel protein family, termed the “AFAP family of signal adaptor proteins”.

XB130 has previously been found to regulate cell cycle progression, death, and migration in lung epithelial cells. It was therefore hypothesized that XB130 is protective in acute lung injury (ALI) and important for facilitating repair after injury. XB130 was found to be differentially regulated in ALI depending on the initial insult. Engineering XB130 transgenic mice to further characterize the role of XB130 in lung injury/regeneration revealed that this protein could be essential for early embryo development.
Acknowledgements

Firstly I would like to thank my supervisor, Dr. Mingyao Liu. His knowledgeable support has been instrumental to the completion of this thesis. It is difficult to say enough with respect to my experiences under his supervision. My interactions with Dr. Liu over the past years are a continued source of motivation to become a better scientist myself.

The members of my graduate committee, Drs. Shaf Keshavjee and Haibo Zhang, have also been instrumental towards the completion of this project. I am grateful for our discussions which constantly challenged my work and made me more knowledgeable in the process. It was their guidance that also forced me to keep to a schedule and complete this project on time.

I was fortunate to have the opportunity to work with the technician for whom the protein XB130 was named. I learned early in my graduate degree that Xiao-Hui Bai is a much respected scientist and convincing her of my theories, results, or conclusions was the key to convincing Dr Liu and the rest of the lab! She has been very supportive of my project: without her help I may never have been able to perform some of the work that I have included in this thesis.

Other lab members have also been helpful in the production of this thesis. They have all contributed in some way to my work in the lab whether on a technical, intellectual, or social level. They include Roli Bawa, Dr. Marcelo Cypel, Pascal Duchesneau, Dr. Bing Han, Monika Lodyga, Dr. Jeya Nadesalingam, Dr. Takeshi Oyaizu, Dr. Atsushi Shiozaki, Dr. Dirk Wagnetz Helan Xiao, and Dr. Jon Yeung as well as other members of the Latner Thoracic Surgery Laboratory. This group of people has made my stay in the lab truly productive and enjoyable.
# Table of Contents

Abstract .................................................................................................................. ii  
Acknowledgements ................................................................................................. iii  
Table of Contents ..................................................................................................... iv  
List of Abbreviations ............................................................................................... vii  
List of Tables ........................................................................................................... x  
List of Figures .......................................................................................................... xi

Chapter 1. Introduction ....................................................................................... 1  
A. Thesis Overview ............................................................................................................. 1  
B. A brief introduction to signal transduction ......................................................... 2  
C. Signal Adaptor Proteins ....................................................................................... 3  
   1. General overview ........................................................................................................ 3  
   2. Functions in signal transduction ............................................................................. 5  
      a) Cooperative domains of adaptor proteins ......................................................... 5  
      b) Adaptor proteins regulate the oligomerization of signaling complexes .......... 6  
      c) Multiple interaction domains lead to diversity in signaling ......................... 6  
      d) Sorting/detecting and coordinating multiple events .................................... 7  
      e) Conformational change in adaptor proteins regulates function .................. 8  
      f) Functions in embryo and early development ................................................. 9  
   3. AFAP, a signal adaptor protein ......................................................................... 10  
      a) Overview of AFAP .......................................................................................... 11  
      b) Sequence and domains ................................................................................... 11  
      c) Function .......................................................................................................... 14  
         (1) Cytoskeletal organization ........................................................................ 14  
         (2) Role of AFAP in Brain ............................................................................. 16  
         (3) Role of AFAP in Cancer ....................................................................... 17  
   4. XB130, another signal adaptor protein .......................................................... 18  
      a) In signal transduction ................................................................................... 18  
      b) Function .......................................................................................................... 21  
         (1) Cell cycle regulation .............................................................................. 21  
         (2) Cell death regulation ............................................................................. 22  
         (3) Cell migration regulation .................................................................... 22  
      c) In cancer ......................................................................................................... 23  
         (1) Thyroid cancer observations ................................................................. 23  
         (2) In animal models .................................................................................... 24  
      d) XB130 Summary ............................................................................................ 24  

Chapter 2. Objectives and Aims ......................................................................... 26  

Chapter 3. Characterization of XB130 In Sicilo ............................................... 26  
A. Section Abstract ...................................................................................................... 26  
B. Background and Objective ................................................................................... 27  
C. Introduction .......................................................................................................... 28
Chapter 4. Characterization of XB130 in vivo
A. Section Abstract........................................................................................................ 48
B. Introduction ................................................................................................................ 49
   1. Acute Lung Injury.................................................................................................... 49
   2. Prospective role of XB130 in lung injury and repair........................................... 49
   3. Models of lung injury............................................................................................ 50
      a) Indirect lung injury: intestinal ischemia/reperfusion ....................................... 50
      b) Direct lung injury: LPS induced lung injury .................................................... 51
      c) Direct lung injury: Lung transplantation ......................................................... 51
C. Methods ....................................................................................................................... 52
   1. Animal Care ........................................................................................................... 52
   2. Lung injury study .................................................................................................. 52
   3. RT-PCR .................................................................................................................. 54
D. Results .......................................................................................................................... 56
   1. XB130 in mice ....................................................................................................... 56
      a) Tissue distribution in normal mice ..................................................................... 56
      b) Lung Injury ......................................................................................................... 59
         (1) Indirect lung injury: Intestinal ischemia/reperfusion ................................. 59
         (2) LPS induced lung injury ................................................................................ 61
         (3) Human lung transplantation ......................................................................... 64
E. Discussion ...................................................................................................................... 65
   1. Mouse and human XB130 distribution ................................................................. 65
   2. Lung Injury ............................................................................................................. 66
a) Intestinal Ischemia/Reperfusion ................................................................. 66
b) LPS induced lung injury ........................................................................... 67
c) Human Lung Transplantation ................................................................. 67

Chapter 5. XB130 in Embryological Development .................................... 69
A. Section Abstract ......................................................................................... 69
B. Introduction ............................................................................................... 70
   1. Transgenic mice ...................................................................................... 70
      a) Introduction to transgenic mouse generation ...................................... 70
      b) Basic strategies for targeted genetic manipulation ......................... 71
         (1) Complete KO .................................................................................. 72
         (2) Conditional and inducible KO ......................................................... 73
C. Methods ..................................................................................................... 74
D. Results ....................................................................................................... 81
   b) Genotyping ............................................................................................ 81
      (1) PCR product sequence compared to database ............................... 81
      (2) Population analysis ......................................................................... 83
      (3) Embryo dissections ......................................................................... 83
         (a) Embryo Genotyping .................................................................. 86
         (b) Western ......................................................................................... 87
E. Discussion .................................................................................................. 88
   1. XB130 is essential for embryonic development .................................... 88
      a) Population distribution ..................................................................... 88
      b) Potential role of XB130 in embryo development ............................ 89

Chapter 6. Future directions ...................................................................... 91
A. How are XB130 knockout embryos dying? ............................................. 91
B. Conditional XB130 knockout mice ......................................................... 91

Chapter 7. References .............................................................................. 92
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFAP</td>
<td>Actin Filament Associated Protein</td>
</tr>
<tr>
<td>AFAP-120</td>
<td>Actin Filament Associated Protein – 120 kDa</td>
</tr>
<tr>
<td>AFAP1L1</td>
<td>Actin Filament Associated Protein 1 like 1</td>
</tr>
<tr>
<td>AKAP150</td>
<td>A kinase anchoring protein 150kDa</td>
</tr>
<tr>
<td>AMPA</td>
<td>α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate</td>
</tr>
<tr>
<td>BD</td>
<td>Brain Death</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>BLOSUM62</td>
<td>Blocks substitution matrix 62</td>
</tr>
<tr>
<td>Btk</td>
<td>Bruton’s tyrpsione kinase</td>
</tr>
<tr>
<td>Cas</td>
<td>Crk associated substrate</td>
</tr>
<tr>
<td>c-Cbl</td>
<td>Cellular casitas B-lineage Lymphoma</td>
</tr>
<tr>
<td>CD</td>
<td>Cardiac Death</td>
</tr>
<tr>
<td>CDK/CDK6</td>
<td>Cyclin-dependant kinase 4/6</td>
</tr>
<tr>
<td>CIT</td>
<td>Cold ischemia time</td>
</tr>
<tr>
<td>co-IP</td>
<td>Co-immunoprecipitation</td>
</tr>
<tr>
<td>DDBJ</td>
<td>DNA Data bank of Japan</td>
</tr>
<tr>
<td>EBI</td>
<td>European Bioinformatics institute</td>
</tr>
<tr>
<td>EGF</td>
<td>epidermal growth factor</td>
</tr>
<tr>
<td>EMBL</td>
<td>European molecular biology laboratory</td>
</tr>
<tr>
<td>Flp</td>
<td>Flipase</td>
</tr>
<tr>
<td>Gab1/2</td>
<td>Grb2-associated protein 1/2</td>
</tr>
<tr>
<td>GAP</td>
<td>GTPase activating protein</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>GEF</td>
<td>Guanine exchange factor</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescent protein</td>
</tr>
<tr>
<td>GluR1</td>
<td>Glutamate receptor 1</td>
</tr>
<tr>
<td>Grb2</td>
<td>Growth factor receptor bound protein2</td>
</tr>
<tr>
<td>GSK3β</td>
<td>Glycogen synthase kinase 3 β</td>
</tr>
<tr>
<td>GTPase</td>
<td>Guanosine triphosphate hydrolase</td>
</tr>
<tr>
<td>HEK293</td>
<td>Human embryonic Kidney 293</td>
</tr>
<tr>
<td>HSV-tk</td>
<td>Herpes Simplex Virus Thymidine Kinase</td>
</tr>
<tr>
<td>Ick</td>
<td>Intestinal cell kinase</td>
</tr>
<tr>
<td>ICM</td>
<td>Inner cell mass</td>
</tr>
<tr>
<td>KCNQ2</td>
<td>Potassium voltage-gated channel, KQT-like subfamily, member 2</td>
</tr>
<tr>
<td>LAT</td>
<td>Linker for activation of T cells</td>
</tr>
<tr>
<td>MyD88</td>
<td>Myeloid differentiation primary response gene</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>PI3K</td>
<td>Phosphoinositide 3-kinase</td>
</tr>
<tr>
<td>PIP2</td>
<td>Phosphatidylinositol 4,5-bisphosphate or PtdIns(4,5)P₂</td>
</tr>
<tr>
<td>PSI-BLAST</td>
<td>Position-specific iterated basic local alignment search tool</td>
</tr>
<tr>
<td>PtdIns</td>
<td>Phosphatidylinositol</td>
</tr>
<tr>
<td>PTEN</td>
<td>Phosphatase and tensin homolog</td>
</tr>
<tr>
<td>RET/PTC</td>
<td>Rearranged in Transformation/Papillary Thyroid Carcinomas</td>
</tr>
<tr>
<td>SAP97 MAGUK</td>
<td>Synaptic associated protein 97 Membrane-Associated Guanylate Kinase</td>
</tr>
<tr>
<td>Shp2</td>
<td>Src Homology Phosphatase 2</td>
</tr>
<tr>
<td>Sos</td>
<td>Son of Sevenless</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SRE</td>
<td>Serum response element</td>
</tr>
<tr>
<td>IL-8</td>
<td>Interleukin-8</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris-Acetate EDTA buffer</td>
</tr>
<tr>
<td>TIRAP</td>
<td>Toll-Interleukin 1 Receptor Domain-Containing Adapter Protein</td>
</tr>
<tr>
<td>TKB1</td>
<td>Tyrosine kinase derivative of the BL21 strain (E. coli)</td>
</tr>
<tr>
<td>TPC1</td>
<td>Two pore segment channel 1</td>
</tr>
<tr>
<td>TRAF6</td>
<td>TNF receptor associated factor 6</td>
</tr>
<tr>
<td>WRO</td>
<td>Human thyroid cancer cells</td>
</tr>
<tr>
<td>XB130</td>
<td>Xiao-hui Bai 130 KDa</td>
</tr>
<tr>
<td>ZAP-70</td>
<td>Zeta-chain-associated protein kinase 70KDa</td>
</tr>
</tbody>
</table>
List of Tables

Table 1 Proteins related to XB130 ........................................................................................................ 33
Table 2 Human XB130-related proteins .............................................................................................. 36
Table 3 Reverse transcriptase protocol ................................................................................................. 55
Table 4 RT-PCR primers ....................................................................................................................... 55
Table 5 Modified PCR reaction mixture for genotyping ....................................................................... 78
Table 6 PCR reaction protocol and primers ........................................................................................... 78
Table 7 Genotype of XB130 transgenic mice at 3 weeks of age ................................................................. 83
Table 8 Genotype of XB130 transgenic embryos .................................................................................. 83
List of Figures

Figure 1 Summary of Bioinformatics methods.......................................................... 31
Figure 2 Evolutionary Comparison of XB130-related proteins................................. 35
Figure 3 Comparison of functional domains found in AFAP, AFAP1L1 and XB130 .......... 38
Figure 4 Expression of AFAP, AFAP1L1, and XB130 in human tissue ..................... 40
Figure 5 XB130 expression in C57/B6 mice............................................................. 58
Figure 6 Expression of XB130 in the airway after intestinal ischemia-reperfusion......... 60
Figure 7 Quantification of XB130 staining after direct lung injury........................... 62
Figure 8 Quantification of XB130 mRNA expression after direct lung injury .............. 63
Figure 9 Quantification of XB130 mRNA expression after human lung transplantation .. 64
Figure 10 Mutation of the murine XB130 gene....................................................... 76
Figure 11 Strategy for genotyping by PCR ............................................................... 79
Figure 12 Confirmation of specificity of genotyping by PCR..................................... 82
Figure 13 Representation of normal and abnormal E13.5 embryos............................. 84
Figure 14 Structure of normal E13.5 embryo.......................................................... 84
Figure 15 Size comparison of E7.5 and E9.5 embryos............................................. 85
Figure 16 Representative genotyping results from E9.5 embryos............................... 86
Figure 17 Western blot from E10.5 Embryos............................................................ 87
Chapter 1. Introduction

A. Thesis Overview

Recently, XB130 has been found to play an important role in regulating cell death, cell cycle progression, and cell migration in two different types of lung specific cell lines (1). Currently, there is only one lab in the world investigating this protein. The main focus of my study was to characterize the role of XB130 in vivo. The first stage of this study was an in silico search for structural, functional, and expressional information related to XB130. It was hoped that this information could be used to increase our in vivo knowledge of XB130 and give insight for further study. Concurrently with the data mining search and the known in vitro function of XB130 in lung specific cell lines, XB130 expression was investigated in multiple models of lung injury using C57/B6 mice. This portion of the study was to be used as a pilot study to find clues for further in vivo characterization of XB130 using knockout mice as a study tool. During the generation of these mice, a homozygous knockout animal could not be identified in the breeding colony (mice tested were minimum 3 weeks old) leading to the hypothesis that XB130 is critical for embryological development.

This thesis is divided into three distinct sections which represent the progression of the study and the direction that the project took as more data were discovered. It will begin with a general introduction to the entire thesis. This section will introduce important concepts in signal transduction and adaptor protein function including an overview of AFAP, a protein highly related to XB130, before focusing on the structure and function of XB130 itself. The first section contains a detailed report of the in silico information on XB130 found through data mining. Following the in silico data is an introduction to the experimental section. This includes a brief introduction to the models of lung injury which were used to investigate the in vivo role...
of XB130 as well as the rationale for the involvement of this novel signal adaptor protein in lung injury/regeneration. The third and final section begins with an introduction to transgenic mouse development and mouse embryo development. The introduction also touches on the role of PI3K signaling, the canonical XB130 signaling pathway, in early embryo development. This section focuses on observations of the XB130 transgenic mice and the attempt to identify a homozygous knockout embryo. The thesis ends with a discussion of the results and an indication of several promising areas for further study.

**B. A brief introduction to signal transduction**

Signal transduction between cells, within cells, and from the extracellular to intracellular environment is critical for life. Proteins are responsible for mediating these signals through their intrinsic dynamic properties. Recently, there has been a large influx of information regarding interaction networks based on genetic and physical interaction data (2). Smock and Gierasch hypothesized that there are eight different ways for proteins to interact and transmit signals based on their “energy landscape”, adding yet another layer of complexity to cell signaling as we know it (3).

The appropriate response of a cell to an incoming signal requires the integration of multiple signaling pathways. The activation of a cell surface receptor is transmitted through the cell via post-translational modifications, recruitment of binding partners to specific subcellular compartments, and/or protein-protein interactions. In the field of signal transduction, the mechanism by which the specificity of a signal cascade and cross-talk between cascades is achieved is constantly under investigation. It is believed that adaptor (also known as adapter) proteins are one of the contributing factors to this specificity and are an integral part of signal transduction (4).
C. **Signal Adaptor Proteins**

1. **General overview**

Signal adaptors are proteins composed mainly of protein-protein interaction motifs, capable of linking cell surface receptors to downstream signaling targets while regulating the specificity of signal cascades (5). Studies on several different families of signal adaptors suggest that these proteins are not only used to temporally and spatially couple successive signaling proteins but are highly dynamic interactors capable of directing and regulating downstream signals. Signaling adaptors have protein binding domains that are able to specifically recognize and bind to an activated receptor, and go on to recruit one or several downstream effectors. Conformational change or specific combinations of binding partners allows adaptor proteins to regulate the specificity of signaling and in turn, the downstream effect (5).

Adaptor proteins contain several protein-protein interaction domains which function to link binding partners together, creating a signaling complex. Through various combinations of protein interaction domains, adaptor proteins can link specific proteins together in response to cellular signals providing specificity and proper subcellular localization of signaling partners. In this manner, adaptor proteins are important for regulating the specificity of downstream cell signaling and regulating proteins which are important for individual signaling cascades in a spatial and temporal manner (4).

Typically, adaptors contain functional domains which can selectively recognize activated cellular surface receptors or intracellular signaling molecules and use one or more of their other domains to couple the receptor with their downstream effector(s) (6). A simple example of this phenomenon is the SH2/SH3 adaptor Grb2. Grb2 contains an SH2 (Src homology 2) domain which recognizes pYXN regions on activated receptor tyrosine kinases (RTK). Two SH3 (Src
homology 3) domains are then able to recruit effector proteins, such as the Ras guanine nucleotide exchange factor Sos or docking proteins such as Gab1/2. Recruited proteins become phosphorylated and interact with other SH2 containing, downstream effectors such as PI3K or Shp2. In this manner, Grb2 is capable of facilitating the downstream signal cascade that links an activated RTK to Ras-MAPK or PI3K signaling pathways ultimately having an effect on cellular function such as cell proliferation, migration, and differentiation in the case of the aforementioned pathways (7).

Structural analysis of Grb2 demonstrated that the interaction domains of this adaptor function almost independently of each other (8). This observation implies a very simple “linker” or organizer function for adaptor proteins as one binding motif can interact with upstream signaling proteins while another binds to downstream effectors, propagating a signaling cascade. As more adaptor proteins are studied, the complexity of their function is becoming more evident and the “linker”/organizer model for adaptor proteins appear to be the exception rather than the rule (9). Adaptor proteins themselves can be dynamically regulated, adding an extra layer of complexity to the signaling cascade (5). It has been demonstrated that the function of adaptor proteins can even be regulated by intracellular signals such as DNA damage (10). The association of signaling proteins with adaptors can act cooperatively to enhance signaling or even antagonistically to gate a signal cascade.

The complexity of signal adaptor proteins doesn’t end there. A single adaptor protein can be associated with different proteins depending on the signaling pathway functioning at any given time. In this way they serve to link different signaling proteins to different downstream effectors depending on the cell type being investigated or even in different subcellular compartments. A classic example of this is intracellular protein trafficking of signaling proteins. This important cellular function is dependent on the diversity and specificity of adaptors.
involved in endocytosis and sorting in the endosomes (11). The various capabilities of adaptors demonstrate that the interaction subunits of these proteins are capable of much more than the simple “beads on a string” model with which they were originally labeled (5).

2. Functions in signal transduction

a) Cooperative domains of adaptor proteins

Certain domains of adaptor proteins have the unique ability to cooperatively interact with each other to increase their specificity in recognizing the appropriate binding partner, or to form a unique interaction network capable of binding other proteins. An example of cooperative interaction among domains of a single adaptor protein is the ZAP-70 tyrosine kinase. This protein has two SH2 domains which bind obliquely to biphosphorylated motifs. This interaction is required in part because the N-terminal SH2 domain is not able to recognize phosphorylated tyrosine residues and this function is therefore supplied by the C-terminal SH2 domain (12).

Another interesting case of cooperative interaction comes from the c-Cbl protein. This adaptor is responsible for ubiquitination of RTKs by connecting an autophosphorylated receptor to E2 protein-ubiquitin ligases (13). In this case three distinct adaptor regions (an N-terminal four-helix bundle, an EF hand, and an SH2 domain) contact each other to generate a single protein-protein interaction domain. This novel interaction site, not unlike conventional SH2 domains, specifically recognizes phosphotyrosine sites, however actually binds to the target protein at a site no fewer than six residues on the N-terminal side of the phosphorylated tyrosine residue. Although this phenomenon is still being investigated, it is believed that the four-helix bundle is mainly responsible for extending the interaction domain of the c-Cbl protein (14).
b) Adaptor proteins regulate the oligomerization of signaling complexes

Grb2 and other SH2/SH3 adaptor proteins are able to oligomerize signaling clusters, a process important for cell signaling. Through activation of the T-cell antigen receptor, the SH2 domain of Grb2 is able to recruit the membrane associated LAT docking protein through its three pYXN motifs (15). At the same time, the SH3 domains are able to recruit other signaling molecules such as Sos1 and Cbl. It has been shown that the Sos1 adaptor is able to bind two Grb2 adaptor proteins, which allows for the further recruitment of LAT to the signaling complex (16). By this mechanism, adaptor proteins are able to cluster signaling complexes through the oligomerization of signaling components which leads to localized and potent stimulation of signaling pathways (5).

c) Multiple interaction domains lead to diversity in signaling

Adaptor proteins often have multiple protein interaction domains which allow them to interact with many different proteins. It is therefore possible that adaptors use only a portion of these binding domains at any given time. The binding of a protein to a particular domain could depend on the cell type or subcellular localization, allowing the adaptor to exert a number of unique biological effects tailored to the situation. Hoshi et al. have proposed just such a function for the murine A-kinase anchoring protein (AKAP150) in regulating ion channel function in neurons. AKAP150 is localized to the GluR1 subunit of the AMPA-type glutamate receptor via SAP97 MAGUK. In this position, AKAP150 functions to facilitate the phosphorylation status of Glu1, and by extension, the stability/rundown status of the ion current. This is accomplished by
linking Glu1 to either protein kinase A (PKA) or PP2B phosphatase, depending on the required phosphorylation status (17). Interestingly, when AKAP150 is recruited to M-type potassium channels (via a different function subunit which interacts with phosphoinositides at the lipid membrane) it can interact with protein kinase C (PKC) to facilitate the phosphorylation of the KCNQ2 subunit of the ion channel. This causes suppression of M currents in response to muscarinic agonists (17). In these two examples it is clear that AKAP150 is able to use different combinations of its functional subunits to interact with two different types of ion channels through interaction with either another adaptor protein (SAP97 MAGUK) or the lipid membrane and regulate their current by delivering PKA, PKC, or PP2B to a specific ion channel subunit. The observation that AKAP150 is able to interact with a variety of distinct ligands to control ion channel function implies diversity and specificity in function of adaptor proteins with situational dependence and physiological implications.

**d) Sorting/detecting and coordinating multiple events**

In certain instances, cell signaling requires two distinct events to occur simultaneously to initiate a signaling cascade. Adaptor proteins have the ability to link signaling proteins, or activated receptors from separate pathways to consolidate a signaling cascade (18). In this manner, an adaptor protein can act as a coincidence detector to facilitate the activation of downstream signaling proteins, only in the coincident activation of two upstream pathways.

A well known example of an adaptor acting as a coincidence detector is in TLR signaling, of which TLR4 is the best characterized. TLR4 and its co-receptor MD2 are activated when cells are stimulated with lipopolysaccharide (LPS). After activation, MyD88 dependant and MyD88 independent pathways are activated. The MyD88 dependant pathway involves activation of IRAK and TRAF6 to activate NF-κB and AP-1 transcription factors (19). The
MyD88 independent pathway activates TBK1 leading to IRF3 activation. There is also a certain degree of crosstalk between these pathways which allows the MyD88 independent pathway to elicit a delayed activation of NF-κB and AP-1 (19).

MyD88, however, is not able to bind to TLR4 directly, and requires the recruitment of TIRAP, which functions as a “sorting adaptor”, to recruit MyD88 to certain membrane bound TLRs as well as aid in the assembly of the signaling complex. In this respect, MyD88 is considered a true “signal adaptor” as it links an activated receptor to its downstream effectors (20). This distinction between sorting and signaling adaptors is believed to be the reason why no TLRs have been found to utilize TIRAP in the absence of MyD88, specifically; a sorting adaptor has no function in the absence of a signaling adaptor (20).

TIRAP itself contains a phosphatidylinositol 3, 4 bisphosphate (PIP2) binding region which allows it to associate with the plasma membrane. Generation of PIP2 by integrin activation or other signaling pathways (20, 21) leads to the recruitment of TIRAP to the plasma membrane, where it can then associate with activated TLR4 and recruit MyD88. In this respect, TIRAP acts as a coincidence detector, assembling the MyD88 dependant signaling cascade only when PIP2 is generated at the plasma membrane in areas near activated TLR4.

e) Conformational change in adaptor proteins regulates function

Adaptor proteins play an important role in the signal cascade which regulates cell morphology and migration. One of the commonly studied pathways involves Src or FAK phosphorylation of 15 YXXP motifs on the docking protein Cas (p130Cas). Once phosphorylated, these motifs are able to bind the N-terminal SH2 domain of the CrkII adaptor
protein. CrkII also contains two SH3 domains, the first of which binds GEFs to act on Rap1 and Rac GTPases for control over cellular adhesion and cytoskeletal organization (5, 6).

The unique property of the Cas and CrkII adaptor proteins is that their phosphorylation status, which regulates their function, can be regulated through conformational changes that result from cell stretching or sheer force. Cas is recruited to focal adhesion sites by an N-terminal SH3 domain and by a C-terminal Src family kinase binding region. Sawada et al. have shown that when human embryonic kidney (HEK293) cells were subjected to stretch forces, a conformational change in the Cas adaptor lead to the exposure of the YXXP motifs, allowing for their “extension dependent phosphorylation” and enhanced downstream signaling (22). This mechanism of signaling implies that Cas may be a sensor for the conversion of mechanical forces into a biochemical signal. In a related mechanism, actin filament associated protein (AFAP), another adaptor protein, has been found to function as a mechanosensor under conditions of cell stretching which will be discussed later (23).

**f) Functions in embryo and early development**

The Grb2 adaptor has also been found to be important in early embryo development. Cheng et al. have demonstrated that in Grb2 -/- embryos differentiation into the trophoblast lineage while other cells are designated to the inner cell mass (ICM) lineage, (a change required for preimplantation development) occurs normally. Further development of the ICM into the egg cylinder requires the differentiation of a portion of the ICM into an endodermal lineage. It is this differentiation that is blocked in Grb2 null embryos, effectively arresting embryological development past embryonic day 5 (E5) (24).

Coupling of RTK signals to the Ras-MAPK pathway, which induces expression of Gata6, is critical for the formation of the primitive endoderm and therefore development of the
extraembryonic endoderm layers of the visceral and parietal yolksack. This signaling pathway has also been found to suppress expression of Nanog, an epiblast specific protein, in primitive endoderm progenitor cells. Grb2-/- embryos which fail to express Gata6 arrest development at E4.5 due to a lack of primitive endoderm development (25).

These observations indicate that the Grb2 adaptor protein plays a critical role in a specific stage of embryological differentiation as prior levels of differentiation occur normally and cell proliferation is unaltered. It should be noted however, that certain endodermal differentiation markers were observed in Grb 2 -/- embryos implying that there is more than one pathway involved in this type of differentiation highlighting the complexity of embryological development (24).

3. **AFAP, a signal adaptor protein**

As described, it is clear that adaptor proteins are important mediators in the signal transduction process. They can have complex functions and operate in a variety of situations. One interesting function for signal adaptor proteins that was previously alluded to is in cytoskeletal mediated mechanotransduction. Previously it was thought that most of the force sensing function in cells is at the cytoplasm membrane, such as calcium movement and ion channel gating (26, 27). Work from the Latner Thoracic Surgery Research Laboratories at the University of Toronto has indicated an important role for signal adaptor proteins in the sensing of forces imparted to lung tissue through mechanical ventilation. They have identified Actin Filament Associated Protein -110 (AFAP) as being an integral component in converting the physical expansion force of ventilation into biochemical signals through activation of c-Src (23). While further characterizing AFAP, the same group identified a novel, related signal adaptor protein and named it XB130 (1). There have been relatively few studies done on XB130 and
only one laboratory is working to characterize it. The focus of this thesis will be a systematic
investigation into the role of XB130 in vivo, however, in order to fully understand the project,
one must have some knowledge of AFAP and the work performed to date on XB130.

a) Overview of AFAP

AFAP was first discovered through research directed at discovering the mechanism by
which activated Src could induce morphological transformations in cells. AFAP was found to
form a stable complex with Src through SH2 or SH3 binding domains. The observation that
AFAP was associated with stress fibers in chicken embryonic fibroblast (CEF) cells and rosette-
like structures in v-Src transformed CEF cells made by Kanner et al. indicated that AFAP may
be involved in mediating the association of Src and actin filaments (28, 29).

The avain form of AFAP was first cloned by Flynn et al. who reported the sequence (30).
Soon after, the mammalian form of AFAP was sequenced and shown to have an 87% homology
with the avian form (31). It was also found by Flynn et al. that there is a splice variant of AFAP,
known as AFAP120. This version contains a 258bp novel insert and can only be found in brain
tissue (32).

b) Sequence and domains

AFAP is a complex signal adaptor protein in that it contains an N-terminal SH3 domain
binding motif followed by a WW binding motif, then two PH domains which flank multiple
internal Ser/Thr phosphorylation sites which in turn are flanked by two SH2 domain binding
motifs finally followed by a C-terminal leucine zipper/actin binding domain (Fig 3) (33). With
multiple protein interaction domains, it is easy to hypothesize that AFAP could be involved in
many different biological functions which are linked to the actin cytoskeleton.
The SH3 domain binding motif of AFAP has been demonstrated to be important for Src family kinase binding. Indeed a single amino acid mutation of Pro$^{71} \rightarrow$ Ala$^{71}$ (AFAP$^{71A}$) was able to abrogate this interaction and was found to inhibit cell stretch induced c-Src activation (23, 33). It was observed that co-expression of AFAP$^{71A}$ with constitutively active Src (Src$^{527F}$) failed to form a stable complex, indicating that the SH3 domain binding motif of AFAP is required for effective interaction between these two proteins. Lyn, Fyn and Yes are Src protein tyrosine kinase family members (34). It has been demonstrated that SH3 domains of Lyn and Fyn, but not of Yes are able to bind with AFAP. This differential binding to AFAP between Src family kinases implies that there is a degree of specificity to the signaling cascade or mechanism of Src protein tyrosine kinase based signal transduction (33).

The SH2 domain binding motifs of AFAP have also been found to be important for interaction with Src. Mutation of five tyrosine phosphorylation sites (Y$^{93}$, Y$^{94}$, Y$^{125}$, Y$^{451}$, Y$^{453}$) completely inhibited the ability of Src to bind or phosphorylate AFAP. It is possible that the SH2 domain binding motifs of AFAP can facilitate the interaction of AFAP with multiple Src signaling proteins. Han et al. have proposed a mechanism whereby the initial interaction between c-Src and the SH3 domain binding motif of AFAP further facilitates the interaction between c-Src and AFAP via the SH2 domain of c-Src (23, 33).

AFAP contains two PH domains that are distinct from each other in sequence and function. While the N-terminal PH domain shares a high sequence homology with the PH domains of β-spectrin and Dynamin, the C-terminal PH domain is more similar to the Btk PH domain. The N-terminal PH domain has been found to interact with both the classical family of PKCs and the novel PKCλ/ι, however the C-terminal PH domain is unable to bind any of the PKC family members. Both the N- and C-terminal PH domains contain a highly conserved tryptophan residue as well as several basic residues which are required for interaction with
Ins(3,4,5)P3 and PtdIns(3,4,5)P3 (33), however AFAP has not been shown to associate with the plasma membrane making it more likely that these domains are used for interacting with other signaling molecules. It has also been hypothesized that both PH domains contain a consensus sequence for binding WD40 repeats, however, it has been suggested that many of the amino acid residues in this sequence are actually involved in determining secondary structure and are not exposed on the protein surface and therefore not available for binding other proteins (33, 35).

Leucine zipper motifs are currently being investigated for their ability to self associate for the generation of nanostructures with many different sizes and shapes (36). Previously, Qian et al. proposed that AFAP has the ability to self associate via its leucine zipper domain (37). Expression of a C-terminus-GST fusion protein (includes the leucine zipper) allowed for affinity absorption of AFAP, however the interaction is not believed to be a classical leucine zipper interaction (37). Interestingly, co-expression of Src$^{527F}$ (a constitutively active form of c-Src) with the fusion protein eliminates the observed affinity absorption, implying that Src$^{527F}$ has the ability to alter the C-terminal region of AFAP effectively eliminating the ability to self-associate presumably through Src dependent phosphorylation of AFAP. Specific deletion of the leucine zipper domain causes a rearrangement of actin filaments into rosette-like structures where AFAP is colocalized (37). This deletion also removed the ability of AFAP to interact with the c-terminus-GST fusion protein, providing evidence for leucine zipper based self assembly (33). This effect is similar to the results observed when Src$^{527F}$ is overexpressed (37). The observation of AFAP self-association in conjunction with its ability to bind actin filaments indicates that AFAP could function as an actin cross-linking protein and that cell signals which target the self-association of AFAP could alter cross-linking functions and ultimately actin filament integrity (33).
Similar to the way that α-actinin can induce loose networks of actin at low concentration and tight actin bundles at high concentration (38), AFAP has been found to crosslink actin in a dose-dependent manner. At low concentration, AFAP could only induce loosely cross-linked actin filaments as observed by co-sedimentation and a lack of visible microscopic bundles. At high concentration, however, AFAP was able to induce the generation of large, branched actin filaments which could be visualized under a microscope (33).

There is an alternatively processed version of AFAP, AFAP120, which contains a novel insert (NINS) region. This isoform of AFAP is 258bps longer and 10 kDa larger by western blot. AFAP120 mRNA could not be detected in chick embryonic fibroblasts, and was only found to be expressed in brain tissue. The NINS region contains a proline rich section which could facilitate the interaction between other brain specific proteins and actin filaments (32).

More recently it has been found that AFAP120 is associated with neurite elongation in a tyrosine phosphorylation dependent manner in differentiating cerebellar granule cells. AFAP120 was found to be enriched in the growth cone of growing neurons and neuron extension was enhanced when AFAP120 was phosphorylated. It is proposed that AFAP120 coordinates Src signaling with changes in the actin filament structure in the elongating growth cone to facilitate growth cone mobility and extension (39).

c) Function

(1) Cytoskeletal organization

The observation that deletion of the leucine zipper domain of AFAP (AFAPΔzip) induced a similar phenotype to Src527F transfected cells (marked by cytoskeletal rearrangement which includes colocalization of AFAP to actin-rich rosette structures, as well as the observation of cell motility structures) implies that the mutant form of AFAP was able to alter actin filament
organization (37, 40). Since then, it has been thought that AFAP may act as an upstream activator of c-Src to facilitate these changes (33).

Baisden et al. have observed that when transfected with AFAP\textsuperscript{Δlip}, cells displayed hyperphosphorylation of AFAP\textsuperscript{Δlip} on its Ser, Thr, and Tyr residues. Total cellular p-Tyr was also increased. This observation coincided with an increase in phosphorylation cSrc (pY\textsuperscript{416}, an indicator of activated Src), however, transfecting cells with a double mutant of AFAP (AFAP\textsuperscript{Δlip} + AFAP\textsuperscript{71A} = AFAP\textsuperscript{71A/Δlip}), which was unable to interact with the Src SH3 domain, abolished these observations (28). Dominant positive Rho is able to block the effects of AFAP\textsuperscript{Δlip} implying that the changes in actin filaments maybe a result of Src-stimulated Rho GTPase activity (28). This data shows that mutant forms of AFAP are capable of modulating actin filament organization by activating Src family kinases (33). It is therefore possible that changes in AFAP confirmation may enable c-Src activation and modulations in actin filament structure (33).

Tyrosine phosphorylation of AFAP allows for the formation of a stable complex between Src and AFAP (23). It has previously been noted that Src\textsuperscript{527F} alters the ability of AFAP to self-associate, however Baisden et al. have observed that the effect of Src\textsuperscript{527F} occurs independently of tyrosine phosphorylation (37). It appears that hyperphosphorylation of AFAP by PKC is the major mechanism by which changes in confirmation, self-association, and actin cross-linking properties of AFAP are transmitted. Once phosphorylated, AFAP also has the ability to be a direct activator of Src. These observations indicate that the modulatory role of AFAP on actin filaments can act through both direct and indirect mechanisms (33).

PKC\(\alpha\) activation has been shown to induce the formation of cell mobility structures including lamellipodia and filopodia (41). This is characterized by a reduction in stress filaments and an increase in actin networks or bundles (42). After phosphorylation of AFAP by PKC, the
ability of AFAP to cross-link actin filaments is increased and larger aggregates of actin are observed. PKC phosphorylation also reduces the ability of AFAP to multimerize in vitro, whereas in vivo, it is found to reposition AFAP from stress fibers to cell motility structures (43). Based on these observations, Baisden et al. propose two ways in which AFAP may transmit the effect of PKCα signaling. Firstly, phosphorylation of AFAP by PKCα could induce a conformational change which allows for more efficient binding to the SH3 domain of Src by AFAP, leading to Src activation. Activated Src can then go on to facilitate changes in actin filament integrity. Secondly, PKCα phosphorylation of AFAP reduces AFAP self-association which allows AFAP to migrate to cell motility structures including lamellipodia and filopodia, where it can link actin filaments and assist in the formation of the aforementioned structures (33).

(2) Role of AFAP in Brain

AFAP120 has been found to be highly expressed in the brain of mouse embryos and pups. At E16, AFAP120 expression was found to be very high in the olfactory blub, cortex, forebrain, cerebellum, and peripheral sensory structures. Compared to E16, expression of AFAP120 at postnatal day 3 (P3) was markedly reduced and found in the olfactory bulb, cortex, and cerebellum (39). Isoforms of PKC and Src family kinases signaling proteins associated with AFAP function are developmentally regulated in the brain and expressed in the adult olfactory bulb. Several papers have demonstrated the expression of these proteins in different neural structures and indicated their role in neuronal development, function, and regeneration (44, 45). As previously indicated AFAP is an SH3 domain binding partner for the Src family kinase members Fyn and Lyn, and is also able to interact with PKCα (28). Different mutants have shown AFAP’s ability to regulate actin filaments as a c-Src activating protein or downstream of
PKCα signaling. Because the dynamic nature of actin filaments in the brain is important for growth cone extension and neural path finding, it can be inferred that AFAP120 could play a role in regulating signals involved in reorganizing actin structures and in brain development (39).

In the adult brain, AFAP120 was found to be strongly expressed in olfactory receptor neurons as well as neuroproliferative zones in the margins of the lateral ventricle where neural stems cells are believed to reside. These observations imply that AFAP120 could be involved in the activation of Src family kinases and actin cross linking in developing neurons, two processes important for neurite extension and path finding which are critical functions for the normal activity of neural progenitor cells or regenerating neurons (32, 39). Expression of AFAP120 in the Purkinje cells of the developing cerebellum has led researchers to believe that it plays a transient role in regulating actin filament integrity in this region. It is believed that the high expression of AFAP120 in peripheral sensory structures which decreases in adulthood is reflective of a similar role for this protein as is hypothesized in the cerebellum (39).

(3) Role of AFAP in Cancer

There have been several experiments providing evidence for a role of AFAP in tumorigenesis. Transfection of fibroblasts with AFAPΔzip induces changes similar to v-Src transformed fibroblasts which include stress fiber disruption and podosome formation, changes which are consistent with malignant transformation (46). It has also been demonstrated that AFAP is required for the assembly of actin into invasive structures such as those found in PKCα-induced podosomes (40, 47). Finally, AFAP is known to be a c-Src activator regulated by PKC phosphorylation (28). Aberrantly expressed or activated c-Src and/or PKC are able to interact with a wide variety of cell signaling pathways. Some of these pathways are capable of modifying cancer cell adhesion, migration, proliferation, invasion, and apoptosis (48, 49).
Recently, Zhang et al. demonstrated that AFAP plays a role in the development and progression of prostate adenocarcinoma (46). It is therefore not hard to imagine that AFAP could play a role in the development and metastasis of multiple types of cancer.

4. **XB130, another signal adaptor protein**

While investigating the role of AFAP in mechanotransduction, a novel signal adaptor protein was found to have a high sequence identity to AFAP. It was first cloned in the Latner Thoracic Surgery Research Laboratories where it was named XB130 (1). Currently, the aforementioned lab has made several interesting discoveries related to the structure and function of this protein.

**a) In signal transduction**

In the original cloning process for XB130, it was found that this novel signal adaptor protein, in a similar fashion to AFAP, contains several Src SH2/SH3 domain binding motifs (1). By performing in vitro co-immunoprecipitation studies as well as incubating XB130 with various GST-fusion proteins, containing the SH2 or SH3 domain from Src, GAP, and Nck, it was found that XB130 could only interact with the SH2/SH3 domains of Src and not GAP or Nck (1). This is in contrast to the SH2 and SH3 domain binding motifs of AFAP which are able to bind the SH2 and SH3 regions from all of Src, GAP, and Nck (1), implying that XB130 has different specificity for its binding partners.

It was further determined that the specific interaction between XB130 and c-Src is able to induce c-Src activation. When XB130 was co-expressed with c-Src in COS-7 cells, there was an increase in cellular tyrosine phosphorylation that was not found when c-Src was replaced by a kinase dead form of Src. XB130 was found by western blotting to be able to induce this change.
by facilitating autophosphorylation of c-Src on Tyr^{416}, a process known to be required for c-Src activation. It was also found that phosphorylation of Src at Tyr^{527}, a known site for Src deactivation, was not affected by XB130 over expression (1).

  c-Src activation is involved in regulating the transcription of many genes (50-52). One mechanism is through activation of the SRE and AP-1 transcription factors, leading to the transcriptional regulation of genes critical for mitogenic, transforming, and death/survival signals (53, 54). When XB130 was co-expressed with both c-Src and a luciferase reporter for either the SRE or AP-1 transcription factors in HEK293 cells, the luciferase signal was significantly increased relative to the untransfected controls or cells transfected with c-Src or XB130 plus luciferase reporter alone. Co-expression of XB130 with c-Src drove an IL-8 promoter based luciferase reporter, which contains AP-1 and NF-κB among other transcription factor binding elements, in BEAS-2B cells, a lung epithelial cell line (1). Although the mechanistic link between XB130 and transcriptional regulation are yet to be resolved, these observations indicate that the interaction between XB130 and c-Src has physiological significance.

  The PI3K/Akt pathway is important for regulating cell cycle progression and survival. XB130 contains a YxxM motif which begins at Y^{54} and is capable of binding to either the N- or C-terminal SH2 domain of the p85α regulatory subunit of PI3K (55). It is hypothesized that binding of XB130 to this regulatory subunit induces a conformational change allowing for the activation of the kinase subunit of PI3K. When XB130 was knocked down with small interference RNA (siRNA) in TPC1 (a human thyroid cancer cell line) cells, the phosphorylation of Akt, a downstream signaling protein of the PI3K signaling pathway, was markedly reduced (55). It should be noted that the phosphorylation status of MAP kinase signaling proteins were unchanged under this condition (55).
To further investigate the role of XB130 in cell cycle regulation and survival, Shiozaki et al. examined downstream signaling proteins from the Akt signaling pathway in both A549 (a lung adenocarcinoma cell line) and WRO (a human thyroid cancer cell line) cells. Akt phosphorylates the cyclin kinase inhibitors p27 and p21. Upon phosphorylation, these proteins undergo nuclear translocation where they can induce G_{1} phase arrest (56, 57). After treatment of WRO and A549 cells with XB130 siRNA to downregulate XB130, phosphorylation of both p27 and p21 was reduced compared to untransfected controls. Furthermore, treatment also increased total p21 levels while having no effect on total p27 or p53 expression levels (Shiozaki et al., unpublished observations).

Akt is also able to regulate cell proliferation through GSK3β. Phosphorylation of GSK3β by Akt inhibits cyclin D1 degradation while enhancing CDK4/CDK6 complex formation, ultimately promoting S phase progression (58). By treating WRO and A549 cells with XB130 siRNA, phosphorylation of GSK3β was decreased. This observation in conjunction with the p21 and p27 data indicates a possible mechanism by which XB130 exerts control over cell cycle progression (Shiozaki et al., unpublished observations).

Akt also delivers anti-apoptotic signals to both the intrinsic and extrinsic cell death pathways (59). Knocking down XB130 with siRNA in WRO cells caused an increase in cleavage activation of both caspase-8 and caspase-9. It should be noted however that when XB130 was knocked down in A549 or WRO cells, a decrease in procaspase-8 and an increase in cleaved caspase-9 was observed. Based on the effects of XB130 on both the intrinsic and extrinsic cell death pathways, it is believed that XB130 is not a specific regulator of either pathway, but rather functions to counterbalance death signals against survival signals via the PI3K/Akt signaling pathway (Shiozaki et al., unpublished observations).
b) Function

(1) Cell cycle regulation

XB130 has been found to mediate cell cycle progression in two distinct cell types. When the human thyroid follicular carcinoma (WRO) cells or human lung adenocarcinoma (A549) cells were treated with siRNA against XB130, both cells lines were arrested in their progression from G1 to S phase of division. When A549 cells were transfected with siRNA directed against XB130, then stimulated with EGF, there was an accumulation of cells in the G1 phase of arrested cell division and a reduced number of cells in the G2/M phases relative to non-transfected cells (1).

The expression of Ki-67 or PCNA (markers of cell proliferation) were reduced after XB130 siRNA treatment. While further investigating the expression of XB130 in various cancer cell lines, it was found that A549, WRO, NPA, and TPC1, but not ARO or MRO cell lines express high levels of XB130. The cell lines expressing XB130 had shorter doubling times compared to those cell lines without detectable XB130 expression. This implies that XB130 plays some role in controlling the proliferation of cancer cells (Shiozaki et al., unpublished observations).

The mechanism of action for XB130 on controlling cell cycle appears to involve the PI3K signaling pathway as it was found that knocking down XB130 was able to reduce the phosphorylation in response to EGF stimulation of Src, Akt, and GSK3β at the Tyr416, Ser473, and Ser9 respectively (1).
(2) Cell death regulation

The role of XB130 in spontaneous, intrinsic (staurosporine induced) and extrinsic (FasAb induced) cell death has been analyzed in WRO and A549 cells. It was found that XB130 siRNA was able to increase the spontaneous cell death of WRO but not A549 cells. When these cell types were transfected with XB130 siRNA and subjected to intrinsic and extrinsic cell death stimuli, only WRO cells showed an increase in both types of cell death (Shiozaki et al., unpublished observations), implying that XB130 plays an important role in the apoptosis/cell survival pathways of this cell type.

It was difficult to induce apoptosis in A549 cells under conditions that have been shown to kill other cell types. However, when stimulated with intrinsic cell death stimuli in a serum free condition, cells transfected with XB130 siRNA had an increased rate of apoptosis relative to untransfected cells (Shiozaki et al., unpublished observations).

Anoikis is a type of apoptotic pathway which is activated when anchorage dependant cells become detached from their surrounding matrix. Dysregulation of this cell death pathway is important in the development of metastases from primary tumors (60). When the thyroid cancer cell line, TPC1, was subjected to anoikis challenge, cells were found to undergo more anoikis when XB130 was down regulated. Lodyga et al. suggest the role of XB130 in anoikis dependent cell death may be an important clue for determining how the PI3K signaling pathway facilitates irregular cell proliferation and survival in cells containing the RET/PTC rearrangement (55).

(3) Cell migration regulation

PI3K signaling is believed to be important for cell migration (61), leading to the investigation of the role of XB130, a modulator of PI3K signaling (55), in cell migration. In a
cell culture model of wound healing using normal airway bronchiole epithelial (BEAS-2B) cells, XB130 plays an important role in cell migration. Stably transfecting these cells with siRNA against XB130 lead to slower migration and repair into a simulated wound, relative to untransfected controls ((55), unpublished observations). Conversely, when cells were stably transfected with vector expression GFP-XB130 fusion protein, cell migration into the simulated wound was found to increase over cells transfected with GFP alone (Nadeslingam et al., unpublished observations). These studies have demonstrated that XB130 plays an important role in regulating cell migration, however how it accomplishes this function is yet unknown.

c) In cancer

(1) Thyroid cancer observations

Lodyga et al. have determined, by northern blot in human tissues, that XB130 is abundantly expressed in the spleen and thyroid. Through immunohistochemistry, staining for XB130 with a specific monoclonal antibody showed that protein expression was mainly expressed in the thyroid follicular cells. XB130 expression was further found in human papillary thyroid carcinomas. Approximately half of the samples stained showed normal levels of XB130 while the other half showed a considerably weaker staining pattern (55).

In TPC1 cells, which contain the RET/PTC rearrangement, a genetic disorder important for the development of papillary thyroid carcinoma, XB130 was found to be a RET/PTC substrate. Cotransfection of RET/PTC with XB130 into HEK293 cells showed RET/PTC induced phosphorylation of Y54 on XB130, and that this phosphorylation induced binding of XB130 to the p85α subunit of PI3K. This also lead to the activation of the PI3K/Akt signaling pathway (55). These observations indicate that XB130 could be a mediator of the
RET/PTC/PI3K/Akt signaling pathway which is important for regulating cell cycle and survival in papillary thyroid carcinoma (62).

(2) In animal models

Shiozaki et al. have developed an animal model for studying subcutaneous tumor growth using normal WRO cells, or WRO cells that have been stably transfected with XB130 small hairpin RNA (shRNA) to knock down expression of XB130. When injected into nude mice, it was found that both normal and stably transfected cells were able to form tumors. The XB130 shRNA cells, however, formed tumors that were much smaller than those formed from control vector transfected cells. This implies that XB130 is able to promote proliferation of thyroid tumors, and in its absence, tumors are not able to grow normally (Shiozaki et al., unpublished observations).

d) XB130 Summary

XB130 was first cloned at the Latner Thoracic Surgery Research Laboratory because of its close similarity to AFAP, a protein being studied by this group for its role in mechanotransduction. Since then, XB130 has been found to be capable of binding and activating Src, leading to the activation of SRE and AP-1 transcription elements (1). Although a very interesting finding, it was subsequently discovered that Src activation may not be the main pathway by which XB130 functions in vitro. Further experimentation demonstrated that XB130 is able to bind to the p85 regulatory subunit of PI3K, facilitating a conformational change and activation of this kinase (55).

Investigations into the role of XB130 in the PI3K signaling pathway lead to exciting observations in vitro. It was found that XB130 could regulate cell cycle progression, cell death,
and cell migration ((1) and unpublished observations). These observations lead to experiments
directed at determining the role of XB130 in cancer. Lodyga et al. demonstrated that XB130
plays an important role in cancer cells containing the RET/PTC rearrangement, especially in the
anoikis cell death pathway, implying a potential role for XB130 in the development of
metastases. Since XB130 has also been found to be important for regulating the growth of
primary tumors, the role of XB130 in cancer development and progression remains an exciting
potential avenue for discovery ((55) and unpublished observations).

Absent from the introduction to XB130 is any study investigating the role of this protein in
vivo. I hope to build on the in vitro studies of others and be the first to characterize XB130 in
vivo.
Chapter 2. Objectives and Aims

The main focus of my thesis study was to improve our understanding of the function of XB130 in vivo. Towards that end, the study began with a systematic data mining search for information on or related to XB130. It was hoped that information gained from this process could be applied to further direct our study of XB130 in vivo. My investigation of XB130 in vivo began with a pilot lung injury study in C57/B6 mice and samples collected from human lung transplantation. The objective of this study was to determine if XB130 is involved in the pathogenesis of acute lung injury and to determine models for elucidating this function. This study provided rationale for the development of XB130 knockout mice to further characterize the role of XB130 in vivo. XB130 heterozygous mice were developed on a 129/OLA and C57/B6 background and backcrossed onto a C57/B6 background for 10 generations. During and after the backcrossing process, no homozygous knockout mice could be identified. This observation meant that the pilot study could not be turned into a full investigation, however implied the very exciting hypothesis that XB130 knockout mice were not surviving the embryo development process. Embryos from XB130 heterozygous breeding pairs were then analyzed to investigate this hypothesis.

Chapter 3. Characterization of XB130 In Sicilo

A. Section Abstract

XB130 is a novel signal adaptor protein first cloned in our laboratory. Because of its similarity to actin filament associated protein 1 (AFAP), it was reported in GenBank as AFAP1L2 (AFAP 1-like 2). The objective of the present study was to perform a genome wide
search of proteins which are similar to XB130 and AFAP, and to gain systemic knowledge of these proteins for their expression, distribution, and potential functions. GenBank database analysis has lead to the discovery of a third, related protein currently identified as actin filament associated protein 1-like 1 (AFAP1L1). Sequence alignment showed that all three of these proteins have approximately 25% identical amino acids and share many conserved functional domains including N-terminal SH2/SH3 binding domains, two consecutive pleckstrin homology domains; however differ in their unique C-terminal regions. Here we propose that these three proteins are members of a novel protein family which we have named the AFAP family of adaptor proteins. The known functions of these proteins are diverse. AFAP, the first found and best characterized, has many functions, including Src activation, stress fiber formation, axonal guidance, and mechanotransduction. XB130 was discovered more recently and has been found to play an important regulatory role in cell cycle progression, cell death, and cell migration. Unfortunately, there have been no studies to investigate the function of AFAP1L1. Information from several databases has indicated that the AFAP family can be found in many different species including the Zebrafish (*Danio Rerio*), mouse (*Mus Musculus*), and humans (*Homo Sapiens*). It has also been indicated that the expression pattern in humans is variable within the family. Taken together, we conclude that this new adaptor protein family may play similar or distinct roles in signal transduction related to multiple cellular functions depending on their expression and interaction with specific binding partners.

**B. Background and Objective**

While investigating AFAP and its role in mechanotransduction, Xu et al. cloned a novel gene that encodes XB130, a novel signal adaptor protein (1). Analysis of XB130 demonstrated that it contained many of the functional motifs found in AFAP and had a high sequence
homology with a few important differences. The objective of the present study was to perform a genome wide search of proteins which are similar to XB130 and AFAP, and to gain systematic knowledge of these proteins for their expression, distribution, and potential functions. GenBank database analysis has lead to the discovery of a third, related protein currently identified as actin filament associated protein 1-like 1 (AFAP1L1). Here we propose that these three proteins are members of a novel protein family which we have named the AFAP family of adaptor proteins.

C. Introduction

1. Genebank History and development

Genbank is a database for all publicly available nucleotide sequences and the proteins they encode. The data base was originally created and maintained at the Los Alamos National Laboratory; however in the early 1990s this responsibility was transferred to NCBI (National Center for Biotechnology Information). At that time, sequence data was manually typed into the database and annotations were added based on published literature. Currently, almost all sequences are deposited into the database directly by the lab which finds the sequence, eliminating the massive amount of manual labor required for data entry. This change has been attributed to the requirement of most journals to have a sequence first deposited into a public database prior to publication so that the accession number can be cited and that other researchers can find the information once the article has been published. This type of submission accounts for approximately 20,000 sequences per month, however the largest source of sequence information (approximately 200,000 per month) comes from bulk submissions generated by sequencing laboratories (63).
Currently, the GenBank database is an international collaboration between the NCBI, the European Molecular Biology Laboratory (EMBL), Data Library from the European Bioinformatics Institute (EBI), and the DNA Data Bank of Japan (DDBJ). Sequence information is routinely received from laboratories around the world for over 100,000 organisms. The database is growing exponentially with a doubling time of approximately 10 months (63).

2. Potential usefulness for researchers

With the wealth of gene information available in this database, it is obvious that it could be a valuable tool to researchers in diverse fields. The maintenance of a centralized resource for genome and proteome information will allow basic science, translation, and clinical labs to mine information about their gene or protein of interest. This knowledge can in turn be used to further direct their research and improve our understanding of biological science. I hope to demonstrate through my own bioinformatics research that by using GenBank, scientists can compile and analyze data from multiple sources to critically assess their gene, protein, or protein family.

D. Methods

Database search

The human form of XB130 (Accession no. Q8N4X5) was used in its entirety as cloned in the Latner Thoracic Surgery Research Laboratories to search for related proteins. The Swissprot database was probed with no phylogenetic restrictions. The default PSI-BLAST threshold for
significance of 0.005 was used, and proteins outside of this significance level were considered to be unrelated. BLAST expect values refer to alignments of the discovered sequence with the human form of XB130 and was calculated using the BLOSUM62 substitution matrix (64).

**Domain prediction**

The prediction of coiled-coil domains was performed using COILS2 software (65). Only sequences with a probability greater than 0.98 to form a coiled-coil region were considered putative coiled-coil regions.

**AFAP family gene expression in different tissues**

The expression pattern of all AFAP family members was performed by searching the Genecard database for the accession numbers of the human proteins found by our original database search. This database uses both the Affymetrix GeneChip HG-U95 set, for normal human tissue and human tissue specific cancer cell lines (66).

**Phylogenic Comparison**

Phylogenetic analysis of the AFAP family of adaptor proteins was performed by using Phylogeny.fr software. The “one-click” feature of this software allows for all known AFAP, AFAP1L1, and XB130 sequences from multiple species to be compared using multiple alignments, phylogenetic branching patterns, and tree-rendering to obtain a robust evolutionary comparison (67).
Figure 1 Summary of Bioinformatics methods

Summary of methods used for identifying and gaining more information on XB130 and related proteins based on database and software information available.
E. Results

1. GenBank search reveals a novel family of adaptor proteins

   a) Database search results

   There were 12 proteins from 7 different species that met the search criteria and are listed in Table 1. Based on their relation to human proteins, they can be classified into three proteins, AFAP, AFAP1L1, or XB130 as indicated by their common names. This classification was further confirmed by evolutionary analysis. When all 12 proteins found were organized into a cladogram, they separated into three distinct groups as shown in Figure 2.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Accession no.</th>
<th>Common Name</th>
<th>No. identical residues/length</th>
<th>No. similar residues/length</th>
<th>Gaps</th>
<th>BLAST expect value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal</td>
<td>Accession no.</td>
<td>Protein</td>
<td>No. identical</td>
<td>No. Similar</td>
<td>No. Different</td>
<td>DB Score</td>
</tr>
<tr>
<td>------------</td>
<td>---------------</td>
<td>---------</td>
<td>---------------</td>
<td>-------------</td>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>Human</td>
<td>Q8N4X5</td>
<td>XB130, AFAP1L2</td>
<td>818/818 (100%)</td>
<td>818/818 (100%)</td>
<td>0/818 (0%)</td>
<td>0.0</td>
</tr>
<tr>
<td>Mouse (Mus Musculus)</td>
<td>Q5DTU0</td>
<td>XB130, AFAP1L2</td>
<td>683/825 (82%)</td>
<td>735/825 (89%)</td>
<td>7/825 (0%)</td>
<td>0.0</td>
</tr>
<tr>
<td>Bovine (Bos Taurus)</td>
<td>Q17R10</td>
<td>XB130, AFAP1L2</td>
<td>695/819 (84%)</td>
<td>735/819 (89%)</td>
<td>3/819 (0%)</td>
<td>0.0</td>
</tr>
<tr>
<td>Frog (Xenopus laevis)</td>
<td>Q6PF55</td>
<td>XB130, AFAP1L2</td>
<td>487/843 (57%)</td>
<td>613/843 (72%)</td>
<td>57/843 (6%)</td>
<td>0.0</td>
</tr>
<tr>
<td>Human (Homo Sapiens)</td>
<td>Q8N556</td>
<td>AFAP</td>
<td>285/717 (40%)</td>
<td>406/717 (57%)</td>
<td>82/717 (11%)</td>
<td>1e^-83</td>
</tr>
<tr>
<td>Chicken (Gallus Gallus)</td>
<td>Q90738</td>
<td>AFAP</td>
<td>275/720 (38%)</td>
<td>402/720 (56%)</td>
<td>78/720 (11%)</td>
<td>5e^-83</td>
</tr>
<tr>
<td>Mouse (Mus Musculus)</td>
<td>Q80YS6</td>
<td>AFAP</td>
<td>273/705 (39%)</td>
<td>395/705 (56%)</td>
<td>76/705 (12%)</td>
<td>4e^-80</td>
</tr>
<tr>
<td>Bovine (Bos Taurus)</td>
<td>A6QQV9</td>
<td>AFAP</td>
<td>278/748 (37%)</td>
<td>400/748 (53%)</td>
<td>84/748 (11%)</td>
<td>1e^-77</td>
</tr>
<tr>
<td>Mouse (Mus Musculus)</td>
<td>Q8BZI0</td>
<td>AFAP</td>
<td>289/807 (36%)</td>
<td>423/807 (52%)</td>
<td>106/807 (13%)</td>
<td>2e^-77</td>
</tr>
<tr>
<td>Rat (Rattus Norvegicus)</td>
<td>Q8VH46</td>
<td>AFAP</td>
<td>276/704 (39%)</td>
<td>393/704 (56%)</td>
<td>74/704 (11%)</td>
<td>3e^-77</td>
</tr>
<tr>
<td>Human (Homo Sapiens)</td>
<td>Q8TED9</td>
<td>AFAP1L1</td>
<td>280/758 (37%)</td>
<td>406/758 (54%)</td>
<td>105/758 (14%)</td>
<td>1e^-74</td>
</tr>
<tr>
<td>Zebrafish (Danio Rerio)</td>
<td>Q4V8Y7</td>
<td>AFAP1L1</td>
<td>268/723 (37%)</td>
<td>390/723 (54%)</td>
<td>86/723 (12%)</td>
<td>5e^-70</td>
</tr>
</tbody>
</table>

Table 1 Proteins related to XB130

Proteins identified from database search for proteins related to XB130. Accession no. is referenced from the Swissprot database. Common name is generated by database to relate to the human protein. No. identical represents the number of identical amino acids after database alignment. No. Similar represents conserved amino acid properties based on a BLOSUM62
matrix. BLAST expect value is the calculated similarity to the probe sequence based on a PSI-BLAST algorithm. The algorithm incorporates a comparison of the identity and similarity of the amino acid sequence, peptide length, and functional domain comparison of the discovered protein to the probe sequence (64).
Figure 2 Evolutionary Comparison of XB130-related proteins

Evolutionary relation between proteins identified through database search. Red numbers represent confidence values in branch points. Scale represents evolutionary divergence based on sequence changes from origin.
b) Bioinformatics information of AFAP family.

Based on the database search results, it was discovered that there are two human proteins related to XB130 which are conserved in many different species. A summary of the information on these proteins can be found in Table 2.

Table 2 Human XB130-related proteins

<table>
<thead>
<tr>
<th>Protein</th>
<th>AFAP</th>
<th>AFAP1L1</th>
<th>AFAP1L2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternate Names</td>
<td>AFAP1, AFAP-110, AI848729, mKIAA3018, 2600003E23Rik, 9630044L16Rik</td>
<td>FLJ36748, MGC14977</td>
<td>Em:AC005383.4, FLJ14564, KIAA1914, XB130</td>
</tr>
<tr>
<td>Length (AA’s)</td>
<td>730</td>
<td>768</td>
<td>818</td>
</tr>
<tr>
<td>Exons (mRNA coding length)</td>
<td>16 (7,497 bps)</td>
<td>19 (4,153 bps)</td>
<td>19 (3,964 bp)</td>
</tr>
<tr>
<td>Chromosome Location</td>
<td>4p16</td>
<td>5q33.1</td>
<td>10q25.3</td>
</tr>
</tbody>
</table>

Summary of the human proteins identified through database search.
(1) Differences and Similarities

Based on domain prediction software and alignment to known functional motifs, new domains were identified on both AFAP and AFAP1L1. Previously, only the two PH domains were thought to exist on AFAP1L1 and the coiled-coil domain of AFAP was not known to exist. A summary of the functional domains found on AFAP, AFAP1L1 and XB130 are shown in Figure 3.

All proteins shown a characteristic proline-rich region, followed by consecutive SH2 and SH3 domain binding motifs, two PH domains, another SH2 domain binding motif, and a coiled-coil domain at the c-terminus. AFAP1L1 and XB130 contain an extra N-terminal SH2 domain binding motif compared to AFAP. The C-terminal regions of all three proteins appear to be unique. XB130 contains a “unique region” while AFAP contains a leucine zipper/actin binding domain. AFAP1L1 does not contain these regions that found on XB130 and AFAP, making it unique in itself from the other two proteins.
Figure 3 Comparison of functional domains found in AFAP, AFAP1L1 and XB130

Representative diagram showing the functional domains of AFAP family members.
(2) Distribution of AFAP family members in human tissue

By using the Genecard database, compressive data regarding the expression pattern of AFAP, AFAP1L1, and XB130 were found and shown in Figure 4. While the expression pattern of AFAP1L1 and XB130 appear to be somewhat similar, they clearly differ from that of AFAP.
Figure 4 Expression of AFAP, AFAP1L1, and XB130 in human tissue

Expression data from Genecard database representing all three members of the AFAP family of adaptor proteins. Open circles indicate that no sample was measured.
F. Discussion

1. Identification of proteins related to XB130

   a) Overview

Through a comprehensive search of the Swissprot database, we have identified the two proteins related to XB130, namely AFAP and AFAP1L1. Including XB130 itself, these three proteins can be found in a variety of species. Evolutionary analysis shows that these proteins are relatively old and must have diverged in the common ancestor of all the species listed. Close comparison of these proteins using domain prediction and alignment software allowed for the prediction of domains that previously were not known to exist and may add insight to the in vitro function of these proteins. In addition, when these data are analyzed in relation to the expression patterns from the Genecard database, it combines to produce a powerful in silico tool which can allow for the prediction of protein function at the organ and systems level.

b) Differences and Similarities

   (1) N-terminal SH2 domain binding motifs

   The N-terminal SH2 domain binding motif is considered to be one of the most important regions of the XB130 protein. This YxxM motif beginning at the 54th tyrosine residue allows XB130 to perform its most studied function. It is thought that the ability to regulate PI3K signaling by binding to the p85α regulatory subunit occurs when XB130 undergoes phosphorylation at Y54, which is within the N-terminal SH2 domain binding motif (55). Through sequence alignment it was found that neither AFAP nor AFAP1L1 contain this N-terminal YxxM motif. This feature separates XB130 from AFAP1L1 and AFAP suggesting that these protein family members may not share this particular function with XB130.
(2) **SH3 Binding Motif**

The proline rich SH3 binding motif was found in XB130 when it was first cloned (1). It has previously been suggested that proline rich SH3 binding motifs could be involved in subcellular localization, enzymatic function, and protein-protein interaction (68, 69); however the function of this region in XB130 has not yet been determined. Through sequence alignment, the SH3 binding motif of AFAP and AFAP1L1 were found to be very similar to that found in XB130. It was found that a single amino acid mutation from proline to alanine at the 71^{st} amino acid position of AFAP (but not proline at the 77^{th} amino acid position) significantly reduced binding between AFAP and c-Src (22, 33). Using site directed mutagenesis to alter proline to alanine in XB130, however, did not reduced binding between XB130 and c-Src (Lodyga et al. unpublished observations). More detailed mutation studies with single and multiple mutations may show the specificity of XB130, AFAP and AFAP1L1 to interact with c-Src and other proteins. It is speculated that this region could be involved in recruiting proteins to a signaling complex, which would be consistent with the hypothesized role of AFAP and XB130 as signal adaptor proteins (1, 28).

(3) **PH domains**

The two PH domains of AFAP, which differ significantly from each other, have previously been suggested to be involved in protein-protein interactions and subcellular localization; however other experiments have suggested that these are actually internal domains that are only involved in generating secondary structure (33, 35). It has been found that the sequence of the PH domains of XB130 and AFAP1L1 are actually very similar to the PH domains of AFAP (unpublished observations). The N-terminal PH domain of AFAP is a substrate and binding partner for PKC\(\alpha\) (43), which has functional implications that have been
studied in the context of podosome formation. The phosphorylation of the N-terminal PH domain of AFAP allows it to bind to the SH3 domain of Src, leading to Src activation and the recruitment of other proteins required for podosome formation (70). Since the N-terminal PH domain of AFAP1L1 and XB130 are very similar to that in AFAP, it is interesting to speculate that AFAP1L1 and/or XB130 could also play a role in Src activation and podosome structure formation.

(4) Coiled-coil region

Coiled-coil domains are unique in that they are a heptad repeat pattern in which the first and fourth positions are hydrophobic, and residues in the fifth and seventh position are usually charged or polar (71). This predictable arrangement allows for the easy prediction of these domains using in silico software analysis. By using a predictive algorithm, it was found that all three proteins identified contain regions with a very high probability of forming a coiled-coil motif. The most common function of this domain is the interaction with other coiled-coil domains for the formation of dimers. It is interesting to note that AFAP is believed to dimerize via its leucine zipper domain (37) which overlaps with the predicted coiled-coil domain. It is therefore possible that the coiled-coil region actually plays a role in the dimerization of AFAP. With this information, it is easy to extrapolate this function to both AFAP1L1 and XB130. Although no dimerization of these proteins via the coiled-coil domain has been studied, it remains a distinct possibility that this could occur, which is consistent with the hypothesized function as signal adaptor proteins.

(5) C-terminal Regions

All three proteins contain C-terminal regions that are different from each other. This includes the leucine-zipper/actin binding motifs of AFAP and the less characterized unique
region of XB130. In AFAP, it is believed that this region is involved in both dimerization and actin interaction (37). In XB130, the C-terminal unique region was found not to be involved in translocating the protein to lamellipodia structures during cell migration ((55), Nadeslingam et al. unpublished observations). Recently, it has been proposed that XB130 may contain a Type 3 actin binding domain; however the significance of this region is still unclear (Nadeslingam et al., unpublished observations). AFAP contains both a Type 1 and Type 3 actin binding domain that partially overlap (40). So far there has been no identification of any type of actin binding domain on AFAP1L1. To date, XB130 has not been shown to associate with F-actin stress fibers under unstimulated conditions. It should be noted, however, that XB130 has been found to translocate to actin rich structures in the lamellipodia, when different types of cells were stimulated with EGF, or phorbol ester ((55), Nadeslingam et al., unpublished observation) or NNK (Nadeslingam et al., unpublished observation). This process requires both the C-terminal and N-terminal regions of XB130, and may be involved in lamellipodia based cell migration (Nadeslingam et al., unpublished observations). The translocation to actin rich lamellipodia structures however, may be facilitated by interactions with other proteins and not through a direct interaction with actin. Currently, our knowledge of the C-terminal regions of AFAP1L1 and XB130 is limited. As our information on functional domains and motifs expands, it is possible that regions similar to the C-terminal regions of AFAP1L1 and XB130 may be found on other proteins, or the function of these regions may be discovered directly through experimentation.

c) Evolutionary Analysis

Based on the evolutionary comparison of AFAP, AFAP1L1, and XB130, it is clear that XB130 is an outgroup of the three proteins compared. It is possible that the divergence of
XB130 from the other two occurred through the insertion of the XB130 unique region, the function of which is still unknown. Without more information, however, it is impossible to definitively decide which protein is phylogenically the oldest or the cause of divergence. It is interesting to note that whatever the cause or mechanism of divergence, it must have happened in the common ancestor of the species listed in Table 1 and Figure 2, making it a very old event.

d) Expression Pattern

The expression pattern of XB130 and AFAP1L1 appeared to be quite similar. In contrast, there were several notable differences between those expression patterns and that found for AFAP. The most notable differences were found in skin and reproductive organs such as ovaries, testis and cervix. It has previously been demonstrated that AFAP expression has been associated with prostate cancer (46).

All three proteins have relatively high expression in the lung. AFAP has previously been shown to be involved in mechanotransduction in airway epithelial cells (23). The results of the lung injury study (especially Figure 6) demonstrate that the expression of XB130 changes in response to different types of lung injury, specifically in epithelial cells. Kidneys also have a strong epithelial cell component, making it possible that XB130 may have a similar role in kidney injury/regeneration as in lung, however the function of AFAP in kidney is unclear.

All three proteins demonstrate high expression in thymus spleen and bone marrow. These organs are known to be involved in hematopoiesis and immune cell maturation, two processes which require extensive cell migration and proliferation (72). Based on these observations, it is possible that XB130, AFAP, and AFAP1L1 are involved in the maturation, proliferation, and/or migration of hematopoietic and immune cells.
Interpretation of database expression data should be made with caution. Although much of the expression data matches known expression patterns, large scale genechip studies may not be as accurate as other techniques such as western blot or RT-PCR. For example, XB130 is known to have strong expression in the human thyroid (55), however this is not reflected in the Genecard database. It is also suspicious that the expression of AFAP1L1 mirrors that of XB130 so closely. The published information on AFAP1L1 is extremely limited and there are currently no ongoing studies aimed at further characterizing this protein, which does not allow for the confirmation of these expression data. With this in mind, the sequence tag used to detect AFAP1L1 may not be as specific as it could be which could lead to non-specific binding of XB130 or AFAP mRNA and vice versa.

e) AFAP, AFAP1L1, and XB130 as the AFAP family of signal adaptor proteins

AFAP and AFAP1L1 have been found to be closely related to XB130 through a database search. Further analysis revealed that all three proteins have very similar functional domains with a few differences in the C-terminal region (Figure 3). It was also found that the expression pattern of these proteins differs in human tissue. This observation in combination with the evolutionary analysis argues that these proteins have non-redundant function in different tissues. Based on this data, it can be concluded that these three proteins are members of an uncharacterized family of proteins which we will call the “AFAP family of signal adaptor proteins” and that this new adaptor protein family may play similar or distinct roles in signal transduction related to multiple cellular functions depending on their expression and interaction with specific binding partners.
a) Lack of information on AFAP1L1

It has previously been shown that AFAP and XB130 play important roles in various cellular functions which have physiological implications (1, 23, 28, 55). The function of AFAP1L1 is considerably less clear. There is very little information published about this protein, and almost all of it comes from high throughput screening (73). By identifying AFAP1L1 as a member of the AFAP family of signal adaptor proteins, it is very possible that it plays a role in cellular functions, which could also have important physiological implications. The research on AFAP and XB130 are far from complete, and AFAP1L1 possesses much potential that should be tapped.
Chapter 4. Characterization of XB130 in vivo

A. Section Abstract

XB130 is a novel signal adaptor protein that was first cloned in our lab. Since its discovery, it has been found to play an important role in regulating cell cycle progression, cell death, and cell migration by binding and regulating phosphoinositide 3-Kinase signaling. This study was aimed at investigating the role of XB130 in vivo through both animal models and human lung injury. Firstly, the change in XB130 protein expression in the lung was measured using a model of murine intestinal ischemia/reperfusion for 0.5 and 4 hrs respectively as an extrapulmonary model of ARDS (acute respiratory distress syndrome). Second, changes in XB130 expression were measured in a mouse model of intrapulmonary ARDS by instilling 5 mg/kg of LPS intratracheally and measuring XB130 via immunofluorescent chemistry and RT-PCR. Finally, we looked at changes in XB130 mRNA via microarray at the end of cold ischemia and 2 hrs after reperfusion in human lung transplant samples from either brain death or cardiac death donors. It was found that XB130 is strongly expressed in airway epithelial cells of the lung and that the type of lung injury is important for determining changes in XB130 expression. In the intestinal ischemia/reperfusion model of extrapulmonary ARDS, it was found that the XB130 signal was strongly increased over control samples, especially in the peri-nuclear cytoplasmic area of airway epithelial cells. Conversely, in the LPS model of intrapulmonary ARDS we found that XB130 showed a trend to a slight decrease, although this change did not reach significance. The decreasing trend in XB130 expression after intrapulmonary lung injury was confirmed in human samples when, independently of the type of donor (brain death or cardiac death), XB130 was significantly down-regulated at 2 hrs after reperfusion relative to the end of cold ischemia. These results demonstrate that depending on the model of injury, XB130...
expression was either increased or decreased. The fact that XB130 shows significant expressional changes over multiple models of lung injury implies that further in vivo study of this novel signal adaptor protein is needed. The function of XB130 in these forms of lung injury is by no means clear; however we hope to continue this study in an XB130 knockout mouse model to further understand this protein and its role in lung biology.

B. Introduction

1. Acute Lung Injury

Acute lung injury (ALI), and its most severe form, acute respiratory distress syndrome (ARDS), is defined by The American-European Consensus Conference on ARDS as severe gas exchange dysfunction and radiographic abnormalities after a predisposing injury in the absence of heart failure (74). In spite of advances made towards treatment of ARDS there is still a very high mortality rate (75). It has recently been shown that cases of ARDS are biochemically distinct based on the initial insult, even though the clinical manifestations appear to be identical (76). This makes injury-specific and individualized treatment strategies an attractive option for the treatment of ARDS (77). In order to accomplish this however, more must be known about the development and mechanisms of ALI and ARDS (76).

2. Prospective role of XB130 in lung injury and repair

The PI3K pathway is regulated by many extrinsic and intrinsic signaling factors. With that in mind, it is impressive that alterations in XB130 expression via siRNA or gene transfection effectively altered cell survival, proliferation, and migration. It is well known that cell death, migration, and proliferation are important processes in the pathogenesis of lung injury and regeneration (78). Since XB130 has been shown to play a role in these cellular processes via
PI3K signaling regulation by binding to its p85α regulatory subunit, it was hypothesized that this novel signal adaptor protein could be involved in the pathogenesis of ALI and ARDS. Since the biochemical development of ALI depends on the initial injury, both intrapulmonary and extrapulmonary models of ARDS were investigated as a pilot study to determine what, if any, role XB130 plays in the pathogenesis of ALI. It is hypothesized that by modulating the activity of the PI3K pathway, XB130 plays an important role in acute lung injury by preventing epithelial cell death and in lung regeneration by promoting epithelial cell proliferation and migration.

3. Models of lung injury

a) Indirect lung injury: intestinal ischemia/reperfusion

The molecular and cellular mechanism of ALI, and of its most severe form ARDS have been extensively studied (77, 79, 80). Tissue damage associated with sepsis and inflammation are able to induce ARDS (81). Intestinal ischemia/reperfusion often resulting from abdominal trauma, surgery, or infection can also lead to ARDS (82). Multiple organ dysfunction associated with ARDS is the leading cause of death in these patients (83). The mechanisms by which an injury or inflammation distant from the lung leads to ARDS, MODS (multiple organ dysfunction syndrome) or death is still unknown (80). It is currently believed that cell death plays an important role in the pathogenesis of ARDS. Apoptosis of alveolar epithelial and endothelial cells can lead to the breakdown of the alveolar barrier which allows for the leak of fluid, proteins, and cellular debris into the alveolar airspace, effectively reducing the efficiency of gas exchange (84). Other types of cell death such as oncosis, caspase-independent cell death and autophagy are also thought to contribute to ARDS and MODS by initiating an inflammatory response (80).
b) Direct lung injury: LPS induced lung injury

Infection is a leading cause of ALI/ARDS (85). Under histological examination, ALI/ARDS is characterized by severe inflammation and neutrophilic alveolitis. Microbial pathogens such as LPS are inflammatory stimuli capable of inducing pulmonary inflammation (86). PI3K signaling has been shown to play an important role in attenuating the LPS-induced inflammatory response (87). It has been demonstrated that Akt signaling reduces LPS induced NF-κB activation through the phosphorylation of GSK3β, suggesting that PI3K signaling is able to reduce inflammation associated with endotoxemia and sepsis (88).

c) Direct lung injury: Lung transplantation

Ischemia/reperfusion induced lung injury is a significant cause of complications and death after transplantation. This injury is typically characterized by general alveolar damage, edema, and reduced efficiency of gas exchange (89). Many different therapeutic interventions have been designed to treat this form of lung injury, however in its most severe form, prolonged mechanical ventilation or death often result 72 hrs after transplantation (90). It has also been found that severe ischemia/reperfusion injury after transplantation is correlated with an increased risk of acute rejection, potentially leading to long term graft dysfunction (91). The mechanism of ischemia/perfusion induced lung injury is still not well understood, and research is constantly ongoing to understand the pathogenesis and to develop treatment (89).
C. Methods

1. Animal Care

Ethics statement

The Animal Care Committee of the Toronto General Hospital Research Institute approved the experimental protocol as well as the housing and breeding of all animals.

Breeding and living conditions

Mice were housed in standard (30.80 cm x 30.80 cm x 14.05 cm) cages which were ventilated with sterilized air. Light/Dark cycle was maintained at 12 hrs/12 hrs. Mice were given food and water ad libitum.

2. Lung injury study

Intestinal ischemia/reperfusion

Male C57BL/6 mice were anesthetized, ventilated, and intestinal ischemia was induced as previously described (80). Briefly, a lower midline laparotomy was performed to expose the peritoneal cavity. The superior mesenteric artery was occluded below the celiac trunk with an arterial microclamp. Pale appearance of the jejunum and ileum confirmed the placement of the clamp and intestinal ischemia. After 30 min, of ischemia the clamp was removed and 0.5 mL of sterile saline was injected intraperitoneally, and the skin was closed with a running suture. Sham-operated animals underwent the same surgical procedure however clamping of the superior mesenteric artery was omitted. Animals were sacrificed 4 hrs after removal of the microclamp in the treatment group and 4.5 hrs after the start of ventilation for the sham group (80).

LPS induced lung injury

Acute lung injury was induced by intratracheal instillation of 6 mg/Kg LPS (Escherichia coli serotype 055:B5, Sigma, St. Louis, MO). Mice were anesthetized with 5% isoflurane and
suspended by their front teeth at a 60° angle. The tongue was pulled out gently to expose the larynx and to prevent swallow reflex. An aliquot (100 µl) of LPS in saline, or 100 µl of saline alone, was instilled into the trachea with a pipette. Lung injury caused by LPS was confirmed by histology (Han et al. unpublished observations). Animals were sacrificed by exsanguination 24 hrs after LPS or saline instillation.

**Human lung transplantation**

Biopsies from human lung transplantation patients were taken at the end of cold ischemia (CIT) or 2 hrs post-reperfusion (2 hrs). Lungs were biopsied at the end of CIT and 2 hrs from 6 Cardiac death (CD) donors while 12 lungs were biopsied at the end of CIT and 2 hrs from brain death (BD) donors. Each cardiac death donor lung was paired with 2 brain death donor lungs matched for age, gender, and cause of death. XB130 expression was measured by microarray using an Affymetrix chip (Affymetrix Inc).

**Western blots**

Western blots were performed as described previously (23, 31). Briefly, tissue was lysed with modified radioimmune precipitation assay buffer (50 mM Tris-HCl; pH 7.5, 150 mM NaCl, 2 mM EGTA, 2 mM EDTA and 1% Triton X-100) containing 10 µg/ml each aprotinin, leupeptin, pepstatin, 1 mM phenylmethylsulfonyl fluoride, 1 mM Na3VO4, and 10 mM NaF. Protein concentration was measured by Bradford assay (Bio-Rad, Munich, Germany). Total lysate was separated by SDS-PAGE and transferred onto nitrocellulose filters (Schleicher & Schuell, Whatman, Middlesex, UK). Filters were then probed with indicated antibodies. Proteins were visualized with an enhanced chemiluminescence detection kit (ECL, Amersham Pharmacia Biotech, Little Chalfont, UK). Band densities were quantified using the ImageJ software (http://rsb.info.nih.gov/ij/).
Immunofluorescent chemistry

Lung tissue sections (5 µm) were double stained for XB130 and cytokeratin. Formalin fixed and paraffin-embedded lung tissue sections were deparaffinized in three changes of xylene (5 minutes each), and rehydrated in a graded alcohol series (5 minutes in each of 100%, 95%, 70%, and 0% EtOH). The sections were then incubated for 20 min in 10 mM citrate buffer (pH 6.0) which had been preheated to 95-100°C. Slides were cooled at room temperature for 20 min in citrate buffer. Nonspecific antibody interaction with the lung tissue was blocked by incubating the slides in 10% normal goat serum for 60 min at room temperature. Primary antibody incubation was overnight at 4°C with the rabbit anti-XB130 polyclonal antibody (1:500 in PBS) and goat anti-Pan-cytokeratin (1:200 in PBS). Slides were then washed in PBS and stained with Alexa-Fluor 555 goat anti-rabbit and Alexa-Fluor 488-donkey anti-goat (1:300, Invitrogen Canada Inc., Burlington, Canada) secondary antibodies. Incubation with secondary antibodies was for 1 hr at room temperature in a dark humidity chamber.

Semi-quantification of XB130 staining intensity for airway and alveolar space was performed on lung tissue stained for XB130 on a scale of 0-3 where 0 represents completely absent staining and 3 represents a very strong positive signal.

3. RT-PCR

Sample collection

Total RNA from snap-frozen organs or from right lung only for the LPS induced lung injury study was isolated using TRIZol Reagent (Invitrogen, Burlington, Canada), and purified with an RNeasy Mini Kit (Qiagen, Mississauga, Canada). Synthesis of cDNA from total RNA was carried out using a TaqMan Reverse Transcription Reagent kit (Applied Biosystems, Foster
cDNA synthesis was performed using a Thermocycler PTC-100 PCR machine (MJ Research Ramsey, Minnesota) with the protocol shown in Table 3.

### Table 3 Reverse transcriptase protocol

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25 °C</td>
<td>10 minutes</td>
</tr>
<tr>
<td>2</td>
<td>37 °C</td>
<td>90 minutes</td>
</tr>
<tr>
<td>3</td>
<td>85 °C</td>
<td>5 minute</td>
</tr>
<tr>
<td>4</td>
<td>4 °C</td>
<td>Indefinitely</td>
</tr>
</tbody>
</table>

Reaction protocol for generation of cDNA from total RNA.

**PCR Protocol**

Primers for XB130 were designed using the Primer Express 1.5 software (Applied Biosystems) and synthesized by ACGT Corp (Toronto, Canada). The amplification was performed in a real-time PCR machine (ABI PRISM 9700HT, Applied Biosystems, Foster City, CA). The relative expression levels of genes interested were normalized with the housekeeping gene, GAPDH.

### Table 4 RT-PCR primers

<table>
<thead>
<tr>
<th>Protein</th>
<th>Forward (5’→3’)</th>
<th>Reverse (5’→3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XB130</td>
<td>TCAGCATCTCCAGAC</td>
<td>GGCTGTTTCCTCTCT</td>
</tr>
<tr>
<td>GAPDH</td>
<td>GCACAGTCAAGGCCGAGAAT</td>
<td>GCCTTCTCCATGGTGTTGAA</td>
</tr>
</tbody>
</table>
D. Results

1. XB130 in mice

a) Tissue distribution in normal mice

Using both RT-PCR and western blot with our specific polyclonal anti-mouse XB130 antibody, we determined the distribution of XB130 in C57/B6 mice. At the mRNA level, XB130 appeared to have the highest expression in lung followed by spleen, with strong expression also found in kidney. XB130 mRNA was also detected in brain and thymus with other organs tested showing almost negligible expression. At the protein level strongest expression was found in the thymus and spleen with significant expression also in brain and lung. XB130 appeared to be less detectable in other organs tested.
Figure 5 XB130 expression in C57/B6 mice

Expression pattern of XB130 in C57/B6 mouse.

**Top:** Expression of XB130 mRNA in various organs expressed as a ratio to GAPDH mRNA

**Middle:** Western blot showing XB130 and GAPDH bands

**Bottom:** Quantification of western blot. Ratio of XB130 to the housekeeping protein GAPDH is shown
b) Lung Injury

(1) Indirect lung injury: Intestinal ischemia/reperfusion

Staining for XB130 in lung tissue from animals treated with ischemia/reperfusion appeared to show an increase in XB130 relative to control animals. It appeared that XB130 was primarily localized in the cytosol and apical membrane of the airway epithelial cells under control conditions. In response to the treatment, the expression levels of XB130 increased, largely in the peri-nuclear cytoplasmic area of airway epithelial cells as seen in Figure 6.
Figure 6 Expression of XB130 in the airway after intestinal ischemia-reperfusion

Representative comparison of XB130 in the airway after intestinal ischemia/reperfusion or control animals double stained with Pan-cytokeratin to highlight epithelial cells.
(2) LPS induced lung injury

Comparison of C57/B6 mice with or without intratracheal instillation of LPS (6 mg/Kg) was performed both at the protein and mRNA level. It was found that neither XB130 staining nor mRNA expression was changed significantly in response to the LPS treatment relative to saline treated control (Figure 7, Figure 8).
Figure 7 Quantification of XB130 staining after direct lung injury

Semi-quantitative analysis of XB130 staining intensity in lung tissue of LPS treated or saline treated control animals. Bars show mean +/- SD.
Figure 8 Quantification of XB130 mRNA expression after direct lung injury

Quantification of XB130 mRNA expression in lung tissue of LPS treated or saline treated control animals. Bars show mean +/- SD.
(3) Human lung transplantation

After analysis of lung biopsies at two critical time points in the transplantation procedure, it was found that regardless of the type of donors (cardiac or brain death) there was a significant decrease in XB130 2 hrs after reperfusion relative to the end CIT.

![Figure 9 Quantification of XB130 mRNA expression after human lung transplantation](image)

Quantification of XB130 expression in human lung transplant samples. CD. Cardiac death. BD. Brain death. CIT. Cold ischemic time. 2Hrs. 2 hrs post-reperfusion. Bars represent mean +/- SD. *=P<0.05
E. Discussion

1. Mouse and human XB130 distribution

Similar to the human mRNA expression data (55), highest expression of XB130 was found in the thymus and spleen by western blotting. Protein expression was clearly higher than all other organs tested (Figure 5). There has been no study in either mouse or humans to determine the role of XB130 in these organs. The strong physiological impact of XB130 that occurs in these organs was anticipated.

Contrary to protein levels, at the mRNA level, XB130 expression appears to be highest in lung. This could imply a higher degree of protein turnover in lung tissue compared to other organs. Spleen and thymus had significant expression of XB130 at the mRNA level, which is consistent with the measured protein expression levels in these organs.

XB130 protein expression in mouse brain and lung, like in humans, was found to have the next highest expression levels as seen when comparing the human expression data found in the Genecard database to the mouse multi-organ western blot (Figure 5). In vitro work on XB130 using lung epithelial cells has already indicated that XB130 is important for various lung cellular functions (1); however, it is unknown if the expression level found in vivo is sufficient to affect physiological function or play a role in injury/repair. The function of XB130 in neural tissue is even less clear as there are no studies directed at determining its function in the brain.

The most significant exception to the similar expression of XB130 between human and mouse is in liver tissue. Human liver appears to have a similar level of XB130 expression as that found in lung and brain, however the murine liver appears to lack XB130 altogether at both the protein and mRNA levels. One of the main functions of PI3K signaling, the main signaling
pathway of XB130, in the liver is to regulate insulin sensitivity (92). The role of XB130 in human liver should be studied.

It is interesting to note that in spite of high XB130 mRNA expression in kidney tissue, there is a curious lack of XB130 at the protein level. The significance of this observation is unknown at this time but could be indicative of very rate of degradation which could have physiological significance. XB130 mRNA in human kidney tissue appears to be consistent with the mRNA level of XB130 in mouse kidney as both Lodyga et al. and the Genecard database have indicated strong expression in this organ, but less than thymus and spleen (55).

2. Lung Injury

a) Intestinal Ischemia/Reperfusion

Intestinal ischemia/reperfusion is a model used to simulate extrapulmonary insults that induce acute lung injury. Such a model is useful to examine the mechanism of lung injury after trauma, organ transplantation, or sepsis (80, 93). In these clinical conditions, the mechanism of lung injury is still unclear. In response to intestinal ischemia/reperfusion, expression of XB130 appears increased, mainly in the peri-nuclear cytoplasmic area of airway epithelial cells. It is known that XB130 plays a role in modulating cell cycle progression, death, and migration ((1) and unpublished observations), all of which are occurring in cells which show strong XB130 signal. Unfortunately, without further study, the exact role of XB130 in modulating indirect lung injury, if any, is unclear. Because of the strong change in staining pattern, further study is warranted.
b) LPS induced lung injury

LPS has been used as a model of intrapulmonary acute lung injury (94). Based on our previous knowledge of the role of XB130 in cell death and cell cycle regulation ((1) and unpublished observations), it was thought that XB130 may play a role in mediating cell death in this model of lung injury. Quantification of XB130 staining intensity and mRNA expression from LPS treated or saline treated animals demonstrated that XB130 expression had a slight decreasing trend; however this trend did not reach statistical significance. This observation however does not preclude XB130 from playing a role in this type of lung injury. It is possible that normal expression of XB130 is sufficient to be beneficial or detrimental to the lung. Further study of this model involving tissue specific XB130 knockout mice will add much information to our understanding of the role of XB130 in vivo, ultimately improving our understanding of the mechanisms associated with acute lung injury.

c) Human Lung Transplantation

Lung transplantation is currently the only life saving treatment for patients with a variety of diseases. Unfortunately, the process of transplantation and specifically ischemia/reperfusion injury can cause significant damage to the graft, leading to post-operative complications and at times, even death (89). During the ischemia/reperfusion phase of lung transplantation, a significant amount of cell death, proliferation, and migration is occurring resulting from the injury induced by ischemia/reperfusion and the natural repair processes required to restore normal graft function (80, 89). In an effort to determine the role of XB130 in this direct form of ischemia/reperfusion induced lung injury, biopsies were collected from cardiac or brain death donors at the critical times during the ischemia and reperfusion phases on transplantation. It was found that XB130 is significantly down-regulated in response to reperfusion. Again it is not
known if this down-regulation is harmful or protective to the graft. The fact that XB130 expression is significantly changed after transplantation argues that this novel signal adaptor protein could be playing an important role in the post-transplant injury/recovery process and warrants further study.
Chapter 5.  XB130 in Embryological Development

A. Section Abstract

XB130 is a novel signal adaptor protein that has been shown to be differentially regulated in models of acute lung injury. Changes in XB130 expression are dependent on the type of model used to induce acute lung injury indicating a potential role for XB130 in the pathogenesis of acute lung injury. In order to further study the role of this protein in vivo, we attempted to knockout the XB130 gene in a mouse model. Heterozygous transgenic mice were crossed with deleter-cre mice and backcrossed onto a C57/B6 background for 10 generations. During the breeding process, there was a curious lack of XB130 transgenic mice homozygous for the mutant allele. The observation of an approximate 1:2 ratio of wildtype to heterozygous mice followed an expected Mendelian distribution of the mutant allele with the notable exception of complete knockout animals. It was hypothesized that XB130 could be required for embryological development and that mice lacking XB130 were not surviving to be genotyped at the time of weaning (3 weeks of age). Dissection of embryos from XB130 heterozygous breeding pairs was performed at various time points at or before E13.5. These dissections have not yet revealed a homozygous knockout animal. We believe that this does not preclude the possibility of XB130 being required for embryological development as this observation can be explained in part by the small size of the embryos prior to E13.5, and the associated difficulty in separating all maternal tissue from the embryo. Genotyping results show a skew towards excess heterozygous animals which is consistent with this hypothesis. We hope that in the future we can work in collaboration with groups which have more experience and technical expertise in this area to confirm our hypothesis that XB130 is required for embryonic development.
B. **Introduction**

1. **Transgenic mice**

   a) **Introduction to transgenic mouse generation**

   Prior to the development of technology to targeted manipulations in the genome, the information acquired from functional genomic research was limited. Most of the data came from the study of spontaneous mutations in animals or humans which allowed for the association or correlation of gene function to disease. Cell culture studies yielded some important functional genomic information, however it is well known that the function of a gene in vitro could be very different from the function of the same gene in vivo (95). In 2007, the Nobel Prize in physiology or medicine was awarded to Drs. Mario Capecchi, Martin Evans, and Oliver Smithies for their work in modifying the genome of mouse embryonic stem cells. They first developed the technology which allowed for a gene to be partly deleted (knockout) or inserted (knockin) at the embryonic stem cell stage, a change could be carried on to adult mice (96). Molecular genetics has the potential to contribute much information to the field of human medicine and improve our understanding of human biology.

   In spite of approximately 75 million years of separate evolution, humans and mice retain a high degree of genetic similarity. This is especially true when compared to other model systems currently being used by many researchers to mimic human diseases such as the Zebrafish (*Danio Rerio*), fruit fly (*Drosophila*), or bacterial cell culture. On top of their similarity to humans, mice are fecund, easy to feed, require little space for housing, and many assisted reproductive techniques are easily applicable. Modern techniques for genetic manipulation of the mouse genome in combination with the genetic identity of inbred strains mean that experiments performed by different researchers or even in different laboratories are
directly comparable (95). Finally, the mouse genome has been completely sequenced and for several different strains and is publicly available (63). In brief, the ease of care and housing plus the range of information and technology already available in combination with their similarity to humans, make mice an excellent means to model human biology (95).

**b) Basic strategies for targeted genetic manipulation**

Manipulation the mouse genome relies on two important principals. The first is the totipotent nature of embryonic stem cells. This property is fundamental because when manipulated and injected into a blastocyst, these cells will differentiate normally and integrate with the host cells to form a chimeric adult mouse. The ultimate goal is to have the chimeric distribution of host and manipulated cells extend to the reproductive cells of the chimeric mouse so that the mutation can be passed on to the next generation, ultimately producing a transgenic animal (97).

The second fundamental principal is the homologous recombination of similar sequences of DNA within the nucleus. Targeted manipulation techniques take advantage of this principal by using a short section of a gene (which includes selection and excision regions) to replace a region of the endogenous gene (knockin). When a short section of the gene of interest is microinjected into the nucleus, the endogenous sequence is replaces by the novel targeting sequence in approximately 1 in every 1000 cells (97).

With this efficiency of target manipulation, it is important to have a method to select the cells that have taken the novel sequence into their genomic DNA. To accomplish this, most novel targeting sequences contain a neomycin phosphotransferase gene. This technique conveys resistance to the aminoglycoside derivative G418, allowing the cells to grow in the presence of G418. The reverse caveat is that there must also be a negative selection technique which allows
for the exclusion of cells that have incorporated the novel targeting sequence into a region unrelated to the gene of interest through random integration. To accomplish this selection, it is common to use herpes simplex virus derived thymidine kinase (HSV-tk) expression in instances of random integration but not when the gene of interest has been targeted properly. This strategy utilizes the difference in phosphorylation specificity between the viral and mammalian form of this nucleoside kinase. When HSV-tk is expressed, it will act on nucleoside analogues such as gancyclovir. By this method, embryonic stem cells that have undergone targeted recombination will grow well in the presence of gancyclovir where as those that have undergone inappropriate, random integration and express HSV-tk will not. Taken together, growth in the presence of G418 as well as the absence of HSV-tk expression allows for effective selection of embryonic stem cells that have undergone targeted homologous recombination. The ability to culture embryonic stem cells, modify their genome, select for the modification of interest, and inject those cells into a blastocyst to create a chimeric mouse is the basis for the generation of transgenic mice (95).

(1) Complete KO

The Cre-loxP system for targeted gene removal is currently the most common in use for the generation of transgenic mice. Cre, a 38KD recombinase protein recognizes a 34 bp nucleotide sequence. This sequence is encoded by two 13 bp inverted repeats which are separated by an 8 bp spacer region. Typically, a targeting vector that is homologous in sequence to the portion of the gene of interest that is to be deleted is transfected into embryonic stem cells. This targeting vector contains loxP sites, which when removed, are able to induce the generation of premature stop codons in the gene of interest, preferably early in the coding sequence. Cells that have undergone successful recombination are selected for and transferred into pseudo-
pregnant females as described above. After the mouse line carrying the targeting vector is established it is crossed with a mouse which expresses Cre in reproductive cells, or ubiquitously. In this manner, Cre-loxP recombination occurs in all the cells of the mouse, resulting in the effective removal of protein expression from the gene of interest (98).

(2) Conditional and inducible KO

A conditional knockout animal differs from a complete knockout in that the gene of interest has only been deleted in a specific organ or tissue. This is accomplished by expressing Cre recombinase or Flippase in these mice under a tissue specific promoter, such as the Ick promoter for T cell specific expression (99). In this manner, Cre or Flp are only expressed in the tissue of interest, and floxed genes in other tissues are unaffected by the excision. Alternately, the inducible knockout technique can also allow for temporal control of gene deletion. In this situation, Cre of Flp expression requires the binding of a ligand, such as tetracycline or tamoxifen, to the promoter region. These mice will develop normally and the gene of interest will only be deleted when they are exposed to the ligand (100).

The most common application of this technique is to bypass the limitation of trying to study a gene of interest under experimental conditions that may be required for embryonic or early development or to investigate the role of a specific organ or tissue in a disease. An example of this is the study of neurofibromin (encoded by Nf1) a protein important for neurofibromatosis type 1 (NF1). NF1 is characterized by the abnormal proliferation and malignant transformation of neural-crest cell derivatives. Patients with NF1 also suffer from vascular abnormalities such as pulmonary valve and renal artery stenosis. Nf1 knockout mice die in mid-embryonic development (E11.5-E13.5) from cardiac defects (101). Because the cardiac defects occur in areas that are formed from both neural-crest and endothelial cells, the
mechanism of death is unknown. To further investigate, researchers developed conditional knockout mice lacking Nf1 in either neural-crest or endothelial cells. Through this technique they were able to demonstrate that cardiac defects only formed when Nf1 was knocked-out of endothelial cells and not neural-crest cells, however neural-crest cell knockouts still developed ganglioneuromas and gangliosarcomas. Because this tumor formation is also found in the human condition, these conditional animals are thought to be an excellent model for studying the human NF1 condition (101).

C. Methods

Targeting Strategy

To generate XB130 transgenic mice, we collaborated with Dr. Tak W Mak from the University of Toronto. Briefly, a targeting vector spanning the intron region between exons 3 and 4 to the intron region between exons 5 and 6 was used. This targeting vector included 4th and 5th exon homology regions as well as a positive selection vector (neomycin) sequence 3’ to the 5th exon. Loxp sites were 5’ to the 5th exon and 3’ to the neomycin sequence. The neomycin sequence was also flanked with flippase recombination sequences. The neomycin sequence was removed by flippase recombination while Cre-Loxp dependant recombination removed the 5th exon. Removal of this exon generated premature stop codons in the 6th exon, leading to the elimination of XB130 expression. The structure of the XB130 gene before and after recombination is shown in Figure 10.
1. WT

2. Targeting Vector

3. Targeted Locus

4. Locus after Flpe- fRT recombination

5. Locus after Cre-loxp recombination

6.8 kb

16 kb?

15 kb?

14 kb?

Exon 4
75 bp

1165 bp

95 bp

6119 bp

91 bp

Deleted region (855bp)

Exon 4

6524 bp

75 bp

91 bp

Premature Stop Codons
**Figure 10 Mutation of the murine XB130 gene**

Top panel: Diagrammatic representation of the targeting strategy and resulting mutation in XB130 transgenic mice.

E. EcorI restriction site. B. BamHI restriction site. a, b. Targeting sites for southern blot probes

Bottom panel: Simplified diagram of wild type and knockout alleles in XB130 mice. Normal XB130 gene is shown on the top line with the region targeted for deletion indicated. Bottom line shows the resulting recombination of the XB130 gene and the location of the premature stop codons induced by the frame shift mutation after deletion.
Genetic Background of mice

Heterozygous ES cells (C57/B6 background) were selected by southern blot and implanted into a pseudopregnant mouse of 129/Ola background. Chimeric progeny were crossed with a C57/B6 mouse and only the C57/B6 progeny were used to generate the XB130 transgenic line. F1 XB130 heterozygous mice were then crossed with deleter-Cre mice to remove XB130 expression throughout the mouse.

Backcross information

XB130 heterozygous females were backcrossed with normal C57/B6 males for 10 generations. All mice used for experimentation were from the 10th generation of backcrossing.

Genotyping

Genotyping is performed on 0.5 cm tail cuts taken from mice at the time of weaning (3 weeks of age) or on whole embryos at the developmental stage indicated in Table 7 and Table 8. A modified protocol from the REDExtract-N-Amp™ Tissue PCR Kit (Sigma, XNAT) was used. DNA was extracted according to kit protocol; however the reagent volumes used were modified slightly as shown in Table 5. Primers for the sequences 5’ AGACTTAGAGCCATGTCCTG 3’, 5’ AGAAGAGCCTTCAGACCTC 3’, and 5’ TTCTAAGCTTTCCATGACCT 3’ were used as forward primer 1, forward primer 2 and reverse primer respectively. Location of the primers within the XB130 gene are shown in Figure 11.
Table 5 Modified PCR reaction mixture for genotyping

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>REDExtract-N-Amp PCR Reaction Mix</td>
<td>10 µl</td>
</tr>
<tr>
<td>Forward Primer 1</td>
<td>2 µl (1:10 dilution in PCR grade H₂O)</td>
</tr>
<tr>
<td>Forward Primer 2</td>
<td>2 µl (1:10 dilution in PCR grade H₂O)</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>2 µl (1:10 dilution in PCR grade H₂O)</td>
</tr>
<tr>
<td>Tissue extract</td>
<td>6 µl</td>
</tr>
</tbody>
</table>

Table 6 PCR reaction protocol and primers

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
<th>Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturing</td>
<td>95 ºC</td>
<td>3 minutes</td>
<td>1</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95 ºC</td>
<td>1 minute</td>
<td></td>
</tr>
<tr>
<td>Annealing</td>
<td>55 ºC</td>
<td>1 minute</td>
<td>30</td>
</tr>
<tr>
<td>Extension</td>
<td>72 ºC</td>
<td>1 minute</td>
<td></td>
</tr>
<tr>
<td>Final Extension</td>
<td>72 ºC</td>
<td>10 minutes</td>
<td>1</td>
</tr>
<tr>
<td>Hold</td>
<td>4 ºC</td>
<td>Indefinitely</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The locations of the PCR primers in the XB130 gene are shown along with the expected lengths of the resulting amplified products in Figure 11.
Figure 11 Strategy for genotyping by PCR

Diagram of wild type and mutant XB130 genes shown with overlaid PCR primers used for genotyping. Forward primer 1 and reverse primer are marked with small arrows and result in the amplification of “Wild type” and “XB130KO” PCR products of size shown on the left. Forward primer 2 (exon primer) is placed directly in exon 5 and marked with a large arrow. When combined with the reverse primer as shown in the top line, this amplification results in a PCR product of length shown on the right (Exon Primer). Forward primer 2 has no binding site on the mutant XB130 gene and thus only the XB130KO PCR product can be found.

Agarose Gel Electrophoresis of PCR products

Agarose gel (1.5%) was prepared by placing 1.5 g Ultrapure Agarose (Invitrogen, Canada, 15510-027) into 100 ml TAE buffer (Invitrogen, Canada 15558042). The mixture was boiled in a microwave until the agarose was completely dissolved. The mixture was then
allowed to cool to 60ºC and 5 µL Ethidium Bromide was added. The mixture when then poured into 12 well gel molds and allowed to further cool and solidify.

Gels were then immersed in TAE buffer inside an electrophoresis tank (EmbiTec, San Diego CA) 3 µL DNA ladder (GeneRuler™ 100 bp DNA Ladder, ready-to-use, Fermentas International Inc. Burlington, Canada) was added to the first lane for reference and 10 µL of PCR amplified sample was added to each subsequent lane as needed. DNA migrated through the gel at 100 V for 20 min. DNA was then visualized under UV light (305 nm) in a FluorChem Xplor gel reader (Alpha Innotech) with Fluorchem FC2 AIC software (Alpha Innotech) using a 595 nm filter.

**PCR product sequencing**

PCR products from genotyping were purified from agarose gel using QIAprep Spin Miniprep Kit (Qiagen Mississauga, Ontario). Sequencing of the purified PCR product was performed by the University Health Network DNA Sequencing facility (Toronto, Canada). Sequence data was then compared to known database sequence of mouse AFAP1L2 found in the Ensembl database.

**Embryo dissection**

The dissected uterus was placed into a cell culture dish containing cold PBS. Under a dissecting microscope, the uterus was grasped immediately next to the deciduum with two pairs of fine forceps. The uterus was then torn from the top of the deciduum exposing the embryo. The embryo is then squeezed out of the deciduum by placing one pair of forceps on the overlying uterine wall and gently applying pressure. The embryo was then snap frozen for further analysis.
D. Results

b) Genotyping

(1) PCR product sequence compared to database

In order to confirm the specificity of our genotyping, PCR products were sequenced and compared to known database sequences. All PCR products were found to be over 80% identical to the database sequence.
Figure 12 Confirmation of specificity of genotyping by PCR

Summary of PCR products used for confirming sequence of PCR amplification. Length primers lane uses only forward primer 1 and reverse primer while Length Primers + Exon Primer lane uses all three primers. Percentage identity to database sequence is shown in table on the right. *= the length of the band is longer than the maximum length that can be sequenced due to technical limitations. Value indicated does not take this into account and represents the raw comparison value.
(2) Population analysis

Through normal genotyping, a record of all the mice breed was kept and their genotype shown in Table 7. The expected Mendelian ratio of 1:2:1 is shown for comparison; however at the time of weaning (3 weeks of age) there is an approximate 1:2:0 ratio of wildtype to heterozygous to knockout mice.

Table 7 Genotype of XB130 transgenic mice at 3 weeks of age

<table>
<thead>
<tr>
<th>Class</th>
<th>Wildtype</th>
<th>Heterozygous</th>
<th>Knockout</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected (144)</td>
<td>36</td>
<td>72</td>
<td>36</td>
</tr>
<tr>
<td>All (144)</td>
<td>41</td>
<td>103</td>
<td>0</td>
</tr>
<tr>
<td>Males (77)</td>
<td>22</td>
<td>58</td>
<td>0</td>
</tr>
<tr>
<td>Females (62)</td>
<td>19</td>
<td>45</td>
<td>0</td>
</tr>
</tbody>
</table>

Genotyping summary of XB130 transgenic mice breed as found at 3 weeks of age.

(3) Embryo dissections

Embryos were dissected and genotyped by PCR. In a few cases there were embryos that appeared to be in an advanced stage of resorption and could not be effectively dissected away from the maternal tissue for proper genotyping. These embryos are listed as “unknown” in Table 8.

Table 8 Genotype of XB130 transgenic embryos

<table>
<thead>
<tr>
<th>Stage</th>
<th>Wildtype</th>
<th>Heterozygous</th>
<th>Knockout</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>E7.5</td>
<td>6</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E9.5</td>
<td>2</td>
<td>21</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>E10.5</td>
<td>5</td>
<td>20</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>E13.5</td>
<td>4</td>
<td>26</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
Genotype summary of XB130 transgenic embryos at various stages of embryonic development.

**Figure 13 Representation of normal and abnormal E13.5 embryos**

Comparison of normal E13.5 embryos prior to dissection from uterus (left) and E13.5 embryos with two reabsorbed abnormalities (right)

**Figure 14 Structure of normal E13.5 embryo**
Figure 15 Size comparison of E7.5 and E9.5 embryos

Pictures of dissected E7.5 (left) and E9.5 (right) embryos in Eppendorf tubes demonstrating the small size of early stage mouse embryos.
(a) **Embryo Genotyping**

Genotyping was performed on whole embryos in an identical manner to tail cuts from 3 week old mice. Representative results from E9.5 embryos are shown in Figure 16.

![Figure 16 Representative genotyping results from E9.5 embryos](image)

Figure 16 **Representative genotyping results from E9.5 embryos**

Representative DNA gel showing the genotype of a litter of E9.5 XB130 embryos. Genotype of embryos is indicated below gel.
Figure 17 Western blot from E10.5 Embryos

Western blot was performed on whole embryo (E10.5) lysates. Arrow indicates the band corresponding to XB130 (top). All embryos express XB130 protein.

Western blot analysis of embryos dissected at E10.5 demonstrated that all embryos express XB130 protein (Figure 17). This confirms the genotyping results which showed all embryos have at least one copy of the normal XB130 gene.
E. Discussion

1. XB130 is essential for embryonic development

a) Population distribution

Genotyping of pups is performed at the time of weaning, when they are 3 weeks of age. After the establishment of a breeding colony, it was surprising to see that all mice genotyped at the time of weaning were either heterozygous or wild type. The distribution of wild type and heterozygous is approximately 1:2, which is consistent with an expected Mendelian ratio. The notable exception to this ratio is the lack of complete knockout animals (Table 7). This observation could be explained by two potential hypotheses. Firstly, technical problems associated with the genotyping process, mainly non-specific amplification, could explain why no knockouts were observed. The second and more exciting explanation was that mice lacking XB130 were not surviving to 3 weeks of age, and thus were not being identified by the current method of pup genotyping.

To answer the first question, results from PCR genotyping were sent for sequence analysis. It was found that there was a very high sequence identity between the amplified PCR products and the expected PCR sequence (Figure 12). This data confirms the high quality of genotyping, allowing the second hypothesis to be investigated.

In an attempt to identify a complete knockout animal, embryos were dissected at various stages of embryonic development. Unfortunately, as early as E7.5, knockout animals could not be detected, in spite multiple instances of empty decidua being identified at all stages later then E9.5 (Table 8). It is critical to observe an obvious skew in the distribution of genotypes at all stages of embryo development. The ratio of wild type to heterozygous is much different than the
expected Mendelian distribution of 1:2. Based on the size and difficulty of dissection of the early embryos that were analyzed (Figure 15), it is possible that the PCR reaction was, at times, contaminated by maternal tissue. Since all pregnant female mice in this experiment were heterozygous, even a small amount of maternal tissue remaining attached to the embryo would result in inaccurate genotyping. In this instance, the identification of a complete knockout embryo would be hidden by technical difficulties, and the lack of observation of a knockout embryo does not preclude its existence. It should be noted however, that when embryos were analyzed by western blot (Figure 17) it was found that all embryos tested expressed XB130, implying that wild type embryos may have been mistaken as heterozygous, and that it is less likely a knockout embryo was mistaken as heterozygous.

b) Potential role of XB130 in embryo development

Early in embryo development, there is a large amount of inter- and intra-cellular signaling occurring. There have been many studies indicating that PI3K signaling is present and functional in the pre-implantation embryo, as early as the one cell stage (102-105). It has also been shown that PI3K signaling is involved in haemopoiesis and embryonic stem cell fate determination in vitro (106, 107). Recently it was found that this signaling pathway is required for glucose metabolism in early embryo development and that disregulation of glucose metabolism was the mechanism of embryonic lethality in PI3K knockout embryos (104).

Since PI3K signaling is the canonical pathway for XB130 function (55), it is possible that the lack of XB130 affects downstream signaling in a way that is incompatible with life by disrupting stem cell differentiation, polarity orientation, or haemopoiesis. It should be noted however, that XB130 is thought to be only one of many proteins capable of influencing PI3K signaling as siRNA directed against XB130 does not completely inhibit the PI3K pathway.
(unpublished observations). With this in mind, it is not clear that influence over PI3K signaling is the method by which XB130 affects early embryological development. Further investigation into the role of XB130 in early embryos is required to definitively determine if PI3K signaling, or another pathway that XB130 is not yet known to be involved with, is the main signaling pathway for which XB130 is required for early embryological development.
Chapter 6. Future directions

A. How are XB130 knockout embryos dying?

The observation that XB130 is essential for embryological development implies an important, non-redundant role for this protein. The research into the role of XB130 in early embryo development has been limited; however the information that is known implies that XB130 functions in very early embryo development. The exact role of XB130 in early embryo development is still unknown. This represents potential for a novel pathway or novel mechanism modulation of a previously known pathway that has never before been studied. These results present an exciting opportunity for future study into the details of the role of XB130 in early development.

B. Conditional XB130 knockout mice

In spite of the focus of XB130 in vivo research being directed at embryo development, the potential role of this novel signal adaptor protein in lung injury and regeneration remains to be discovered. Literature review has indicated that the PI3K signaling pathway plays a role in modulating lung injury in certain animal models, and XB130 has the potential to modulate this pathway under these conditions. Lung injury studies aimed at investigating the role of XB130 in vivo demonstrated that the expression level changes depending on the type of injury studied. In the future, a conditional knockout lacking XB130 expression only in lung tissue could be developed to further elucidate the role of XB130 in acute lung injury. It is also possible that in spite of the lack of observable expression changes in the LPS-induced lung injury model, baseline XB130 may be sufficient to play a protective or harmful role. Again, a conditional XB130 knockout animal will be greatly beneficial in answering this question.
Chapter 7. References


