IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF *d*PALS2, THE *DROSOPHILA* ORTHOLOG OF MAMMALIAN PALS2

IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF *dPALS2*, THE *DROSOPHILA* ORTHOLOG OF MAMMALIAN PALS2

By

Kelly Teal, B.Sc. (Honours)

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Master of Science

McMaster University

MASTER OF SCIENCE (2005)   McMaster University
(Biology)   Hamilton, Ontario

TITLE: Identification and molecular characterization of *dPALS2*, the *Drosophila* ortholog of Mammalian PALS2

AUTHOR: Kelly Teal, B.Sc. (Honours), University of Guelph

SUPERVISOR: J. Roger Jacobs, Ph.D., Professor, McMaster University

NUMBER OF PAGES: xiv, 125

ABSTRACT

The proper organization of receptors and signal transduction protein complexes of epithelial and neuronal cells is crucial in tissue formation, cellular differentiation and proper overall development and function. Scaffolding proteins are major components involved in protein targeting and protein complex assembly. MAGUKs, a family of scaffolding proteins with multiple binding domains such as PDZ, SH3 and GUK, are important regulators of cellular polarity by recruiting and assembling signal and cytoskeletal components into large complexes. Cell polarity is established and maintained by the proper formation and placement of cellular junctions, which separate the plasma membrane into two distinct domains: apical and basolateral. Epithelial polarity determinants from the Bazooka, Crumbs and Scribble complexes establish the boundaries between the apical and basolateral membrane domains and situate the adherens junctions (AJ) at the interface between the two domains. In neuronal cells, the organization and polarization of the presynaptic and the postsynaptic membranes is organized by the CASK/VELI/MINT1/X11alpha complex. Both CASK and VELI also play a role in epithelial cells.

Two novel proteins, originally discovered by Far Western overlay assay in *Mus musculus*, have been identified as additional binding partners of VELI: PALS1 and PALS2. Both proteins are MAGUKs and are thought to compete with CASK for binding VELI via L27 domain dimerization. PALS1, a major component of the Crumbs complex, is essential for the formation of AJ and the establishment of cellular polarity. PALS2 has been shown to colocalize with E-cadherin below tight junctions and directly associate

with nectin-like molecule-2 (Necl-2) at extra junctional regions, however its function remains unknown.

Using *Drosophila melanogaster* as a model organism, we have identified the potential *Drosophila* ortholog of PALS2, termed dPALS2, and found that it is conserved across other species. We have done extensive sequence analysis of *dPALS2* at the nucleotide and amino acid level and determined the RNA transcript distribution and protein localization.

dPALS2 expression begins around stage 13 in embryonic tissues in a transverse striped pattern in the epithelia and continues in this striped pattern until the end of stage 17. dPALS2 is expressed in adult tissues but undetectable in larval tissues. Based on homology and the expression pattern, dPALS2 may play a role in cell adhesion or cell polarity, similar to the mammalian orthologs. However the striped expression pattern of dPALS2 is similar to segment polarity proteins thus implying dPALS2 may play a role in segment polarity.

ACKNOWLEDGEMENTS:

To my supervisor, Dr. Roger Jacobs, thank you for all the guidance and support you offered during this project. It has given me the opportunity to expand my research knowledge and skills, allowing me to become more confident and independent in a laboratory setting, for this I will be forever grateful. To my committee members, thank you for reviewing this thesis and providing assistance when it was most needed.

To the members of the Jacobs lab, past and present, thank you for your support and friendship. To Allison MacMullin, Katie Moyer, Leena Patel, Dr. Mihaela Geargescu and Ying Hsu Huang, words can not describe how much your continued support and friendship has meant. I have often wondered why I'm so incredibly lucky to have found such amazing friends and I now know that people enter our lives for a reason. Without your encouragement, emotional support, laughter, and friendship, I would never have been able to complete this degree. Each and every one of you has impacted my life in such a way that thank you just doesn't seem enough, but just know I am tremendously grateful for all you have done for me.

Finally, to my family and friends, thank you for the encouragement and support over the last few years. To my parents, who have stood behind me my entire life and encouraged me to be the person I am today, thank you. I dedicate this to my mom because she was my mentor, my inspiration, my pillar of strength and my best friend. She would have wanted me to finish this, which was the drive and determination that kept me going. To Chris, words can not express what you mean to me. Your love and support each and every day provides me with the strength, confidence, and determination to

accomplish my goals. You encouraged me and believed in me when I needed it the most and for that I am greatly appreciative. When the ground caved in and the world came crashing down, a foundation of love and strength remained – you.

TABLE OF CONTENTS:

CHAPTER 3: RESULTS

CHAPTER 4: DICUSSION

LIST OF FIGURES:

LIST OF ABBREVIATIONS:

aPKC	Atypical protein kinase C
AJ	Adheren junction
Baz	Bazooka
BCIP	5-Bromo-4-chloro-3-indolyl-phosphate,(X-phosphate, 4-toluidine salt)
CASK	Calcium/calmodulin dependent serine kinase
CNS	Central nervous system
Crb	Crumbs
DAB	3,3'-Diaminobenzidine tetrahydrochloride
DEPC	Diethyl pyrocarbonate
DIG	Digoxigenin
Dlg	Discs Large
Dlt	Discs Lost
ECR	Evolutionarily conserved region
GuK	Guanylate kinase
HRP	Horseradish peroxidase
IPTG	Isopropylthio-beta-D-galactoside
JAM-1	Junctional adhesion molecule 1
LAP	LRR and PDZ domain binding family
L27	Lin-2/Lin-7 binding domain
Lgl	Lethal giant larvae
LRR	Leucine- rich repeats
MAGUK	Membrane-associated Guanylate kinases
MRE	MAGUK recruitment domain
NBT	Nitroblue tetrazolium chloride
NMJ	Neuromuscular junction
PAGE	Polyacrylamide gel electrophoresis
PALS	Proteins associated with Lin-7
PATJ	Pals-associated tight junction protein
PBS	Phosphate buffered saline
PBT	PBS with Triton
PCR	Polymerase chain reaction
PDZ	Post synaptic density-95/Disc-large/Zona occludens-1
PSD-95	Post synaptic density-95
RT-PCR	Reverse transcription Polymerase chain reaction
SAR	Subapical region
Scrib	Scribble
SDS	Sodium dodecyl sulphate
Sdt	Stardust
SH3	Src Homology 3
SJ	Septate junctions
SSC	Sodium chloride-sodium citrate

TJ	Tight junctions
ZO	Zona occludens
ZA	Zonula adherens

CONTRIBUTIONS:

Dr. Mihaela Georgescu designed and prepared the dPALS2 His fusion construct used in the generation of a polyclonal antibody. Confocal images were processed and taken by Allison MacMullin. The 5'RACE data was provided by Katie Moyer.

Chapter 1

Introduction

The polarization of epithelial and neuronal cells is crucial in tissue formation, cellular differentiation and proper overall development and function in both vertebrates and invertebrates. Epithelial cells rely on the proper organization of polarity determinants along an apical-basolateral axis, separating the plasma membrane into two distinct domains: apical and basolateral (Woods and Bryant, 1993). The apical surface faces the extracellular environment or lumen while the basolateral domain is in contact with neighboring cells and the extracellular matrix (Muller, 2000). The maintenance and establishment of epithelial polarity is accomplished through cellular junctions. These junctions provide a strong connection between cells and a tight barrier to paracellular transportation and protein movement. Invertebrates have two types of cellular junctions - zonula adherens (ZA) and septate junctions (SJ). ZA are involved in cellular adhesion and signaling while SJ are mostly involved in controlling paracellular diffusion (Tepass, 2002). In vertebrates there are ZA and tight junctions (TJ), which function similarly to their invertebrate counterparts. The ZA in invertebrates are located apically to the SJ however in vertebrates the TJ are located apically to the ZA (Shingai *et al.*, 2003). Early in *Drosophila* development during cellularization, the ZA appear as spots along the lateral surface and become a continuous adhesive belt like structure at the apical region of the epithelial cell by gastrulation. The ZA defines the distinct regions of the apical and basolateral domains (Tepass, 2002; Muller, 2003).

Cell polarization is assembled and maintained by scaffolding protein complexes

containing PDZ binding domains. In epithelial cells, the correct localization of these

scaffolding proteins organizes and perpetuates polarity by assembling cellular junctions.

In neuronal cells, the definition of polarity is accomplished by forming the pre-synaptic

domain (axon) and the post-synatic domain (dendrite) through the localization of these

multi-protein complexes rather than through cellular junctions (Muller, 2000).

1.1 PDZ Domains:

PDZ domains are modular interaction domains involved in protein targeting and

protein complex assembly. The PDZ domain was initially identified in the synaptic

protein PSD-95/SAP90, the *Drosophila* septate junction protein Discs-large and the

epithelial tight junction protein ZO-1 as a triple repeat region of about 90 amino acids

(Sheng and Sala, 2001). PDZ domains are the most common protein domain and are

present in numerous proteins across a variety of organisms (Ponting, 1997). There have

been 138 PDZ domains identified in 86 proteins in the *Drosophila* genome (Scultz *et al.*,

2000). These proteins often contain multiple PDZ proteins in tandem repeats to enhance

target ligand binding and often contain other known interacting or signaling domains to

assist in the assembly of multi-protein complexes (Hung and Sheng, 2002).

PDZ domains are highly conserved, however some sequence variation reflects

diversity in functionality and binding partners. Even with sequence variation, the

secondary structure of all PDZ domains is similar with six anti-parallel β-strands (βA-βF)

and two α-helices (αA and αB) folding into a six stranded β sandwich (Doyle *et al.*,

1996; Morais *et al.*, 1996). The formation of a groove between βB strand and αA helix

functions as the ligand binding site, while the binding specificities is within the C-

terminal tail of the ligand (Sheng and Sala, 2001). The last four amino acids in the carboxyl tail of the ligand interacts directly with the PDZ binding groove. The side chain of the last amino acid (position 0) of the carboxyl tail projects into the hydrophobic pocket of the PDZ domain. Thus, PDZ domains generally interact with ligands containing a C-terminal valine but will also interact with other proteins containing hydrophobic residues such as leucine and isoleucine (Songyang, et al., 1997). The amino acid located at position -1 appears to have no PDZ binding specificity. However the - 2 residue of the ligand is extremely important for binding specificity and provides the basis for how PDZ domains are classified (Songyang et al., 1997).

The classification of PDZ domains is defined by the amino acid residue located at the - 2 position of a ligand. Class I PDZ domains select for ligands with a threonine or serine at the - 2 position. The side chain of this residue forms a hydrogen bond with a histidine at $\alpha B1$. Class II PDZ domains select for ligands with hydrophobic amino acids at the - 2 position to interact with a hydrophobic residue at $\alpha B1$. Class III PDZ domains select for ligands with negatively charged residues at the – 2 position to interact with a tyrosine at $\alpha B1$ (Songyang et al., 1997; Stricker et al., 1997). Even though the – 2 position of the C-terminal end of a ligand is considered the most important, the – 3 position also makes contact with the PDZ domain contributing to binding specificity (Songyang et al., 1997).

Although C- terminal ligands are the typical binding partners, PDZ domains are able to participate in other modes of protein interaction. PDZ domains can bind internal peptide sequences that bend into β-hairpin structures to mimic C-terminal ligands or form homo-dimers with other PDZ domains (Hsueh et al., 1997; Dong et al., 1999).

1.2 Membrane-Associated Guanylate Kinases (MAGUKs) are important types of PDZ Scaffolding Proteins:

Typically PDZ scaffolding proteins contain multiple interacting or signaling domains to assist in protein-protein interactions. The family of scaffolding proteins known as the membrane-associated guanylate kinases (MAGUKs) is an example of PDZ proteins with multiple binding domains. The MAGUKs are important regulators of polarity, recruiting and assembling signaling and cytoskeletal components into large complexes. The common domain structure of MAGUKs contains one to three PDZ domains, a Src (SH3) domain and a guanylate kinase domain (GUK) (Anderson, 1996; Dimitratos *et al.*, 1999). The PDZ domains of MAGUKs are essential for receptor and channel protein clustering (Kim *et al.*, 1995) while the SH3 and GUK domains are required for protein binding (Pawson, 1994). Some MAGUKs contain a fourth protein binding domain known as the HOOK domain, which binds actin-associated proteins, connecting MAGUKs to the cytoskeleton (Caruana, 2002).

MAGUKs are distinctive enough even with a common domain structure to be divided into four subfamilies based on the number of PDZ domains, the presence of other protein binding domains and sequence similarity (Dimitratos *et al.*, 1999). Figure 1 illustrates the four MAGUK subfamilies and their members. The Dlg (Discs Large) subfamily is defined by the MAGUK common domain structure of PDZ domain/SH3 domain/GUK domain, along with two additional N-terminal PDZ domains. The ZO-1 subfamily (ZO-1,2,3) is characterized by the presence of three PDZ domains, an SH3 domain, a GUK domain and proline rich regions at the C-terminus. Members of this subfamily associate at epithelial cell junctions, specifically tight junctions in mammals

Figure 1: Members of the MAGUK family. MAGUK proteins contain a core domain structure of a PDZ domain (red), an SH3 domain (yellow), and a GUK domain (green). The subfamilies of MAGUKs (Dlg, p55, Lin-2 and ZO-1) are defined by additional domains such as the HOOK or protein 4.1 binding domain (purple), the L27 domain (blue) and a calcium/calmodulin dependent kinase (Camk) domain (orange).
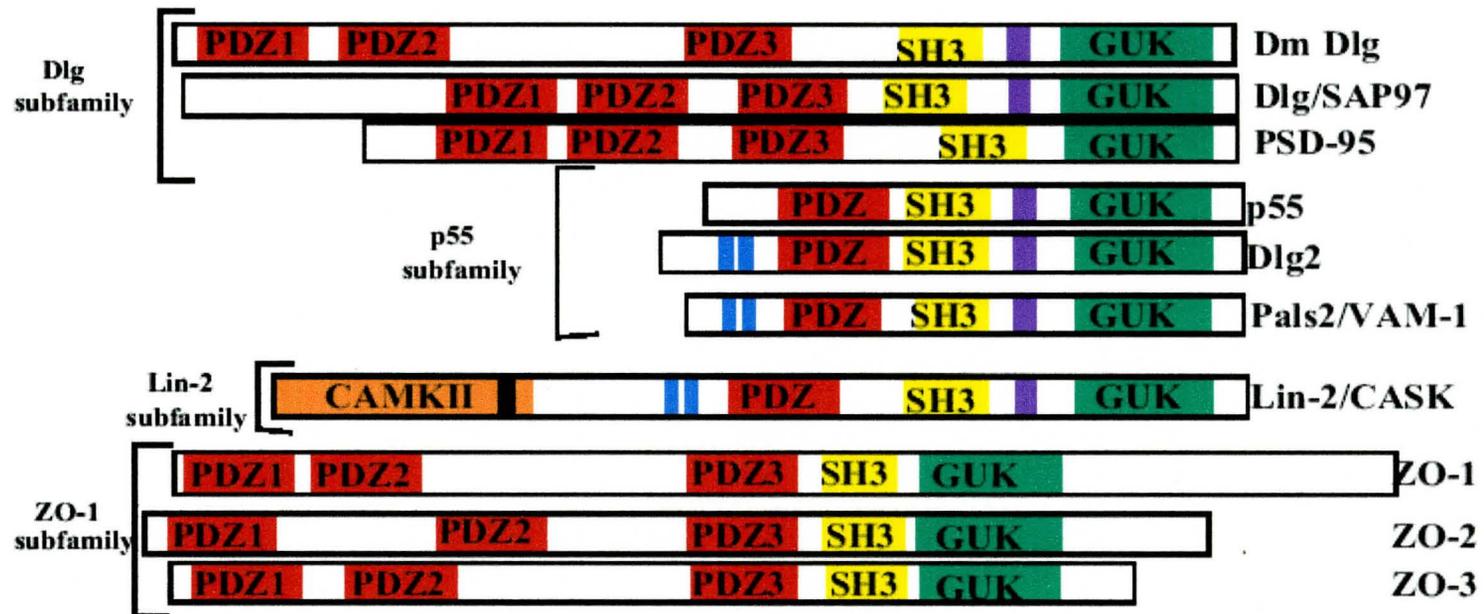
Caruana, 2002

and septate junctions in *Drosophila*. The p55-like subfamily, consisting of p55, Dlg2,

Dlg3, Pals1 and Pals2/VAM, contains the MAGUK common domain structure, plus a

repeated L27 (Lin-2/Lin-7 binding region) domain (Caruana, 2002). L27 domains,

termed for the dimerization of two *C. elegans* proteins, LIN-2 and Lin-7, are considered

essential binding domains in the assembly of scaffolding structures critical in polarity

(Doerks *et* al., 2000; Li *et* al., 2004; Harris *et al.*, 2002). Finally, the Lin-2/CASK

subfamily consists of an N-terminal calcium/calmodulin-dependent protein kinase, a

repeated L27 domain followed by a PDZ domain, an SH3 domain and a GUK domain

(Caruana, 2002).

1.3 MAGUK proteins are major components of Epithelial Polarity:

MAGUK scaffolding proteins assemble and localize to different regions along the

apicobasal axis of the plasma membrane to establish and maintain cellular structure and

polarity. Recent research has focused on three major protein complexes, the

Par3/Bazooka complex, the Crumbs complex and the Scribble complex, which are

considered pertinent in organizing and maintaining epithelial polarity.

1.3a The Par3/Bazooka Complex:

The first of these complexes is the Par3/Par6/atypical protein kinase C (aPKC)

complex and is essential for determining polarity in many cell types (Wodarz, 2002; Gao

et al., 2002; Suzuki *et al.*, 2001). The *par* genes (1-6) were originally identified in a

genetic screen for maternal–effect mutations affecting unequal partitioning during

anterior-posterior determination of the one cell-embryo in *C. elegans* (Kemphues *et al.*,

1988). Recent studies in mammals have highlighted the fundamental importance of this

protein complex.

Mammalian orthologs of the Par complex have been identified (Par3, atypical PKC isotype-specific interaction protein (ASIP), Par6 and aPKC) and are localized to TJ (Bilder *et al.*, 2001). Figure 2 illustrates these three proteins which function together in a tripartite complex to establish and maintain epithelial polarity by promoting proper TJ formation (Gao *et al.*,2002; Suzuki *et al.*, 2001; Yamanaka *et al.*,2001). Par3/ASIP is recruited first to the apical cortex of the epithelia and acts as a scaffold by targeting Par-6 and aPKC to the apical domain of epithelial cells.

Par-3/ASIP and Par-6 are both classified as PDZ domains containing three and one domains, respectively and Par-6 also contains an incomplete Cdc42/Rac1 interactive-binding domain (CRIB) (Burbelo, *et al.*, 1995; Joberty, *et al.*, 2000). The first PDZ domain of Par-3/ASIP protein interacts with the single PDZ domain of Par-6 (Joberty *et al.*, 2000; Lin *et al.*, 2000), while the N-terminus of Par-6 protein directly associates with the regulatory domain of aPKC. Cdc42-GTP binds the incomplete CRIB and a part of the neighboring PDZ domain of Par-6 and induces a conformational change in Par-6 leading to aPKC activation (Joberty, *et al.*, 2000; Yamanaka *et al.*, 2001; Garrand *et al.*, 2003). Par-3/ASIP has also been identified as a substrate of aPKC however the purpose of this phosphorylation is unclear (Hirose *et al.*, 2003; Lin *et al.*, 2000).

The *Drosophila* orthologs of the Par complex are known as Bazooka (Baz)/DmPar6/DaPKC respectively and co-localize at the subapical region (SAR), located above the ZA in epithelial cells. Figure 2 illustrates the Baz complex associating with each other in the same manner as the mammalian orthologs, however it is unknown if Cdc42-GTP regulation is present in *Drosophila*. DmPAR6 and DaPKC depend on Baz

Figure 2: Comparison of cell junctions in vertebrates and invertebrates. The zonula adherens (ZA) are common cell junctions in both vertebrates and invertebrates that function in cellular adhesion. In *Drosophila*, the sub-apical region (SAR) is located apically to the ZA, whereas in mammals, the tight junctions (TJ) are located at this position. In both instances, this is where the Bazooka/Par-3 and Crumbs complexes localize and function in ZA morphogenesis. In *Drosophila*, the septate junctions (SJ) are located basally to the ZA and function similar to TJ in paracellular diffusion. The Scribble complex localizes at the SJ.

NB: This figure indicates the Crumbs complex contains Discs lost (Dlt) however recent studies have demonstrated that Dlt is not the PDZ protein required for cellular polarity. Another protein thought to be encoded by Dlt has been identified and labeled after its vertebrate ortholog, Pals-associated tight junction protein (PATJ) (Peilage *et al.*, 2003).

Medina *et al.*, 2002
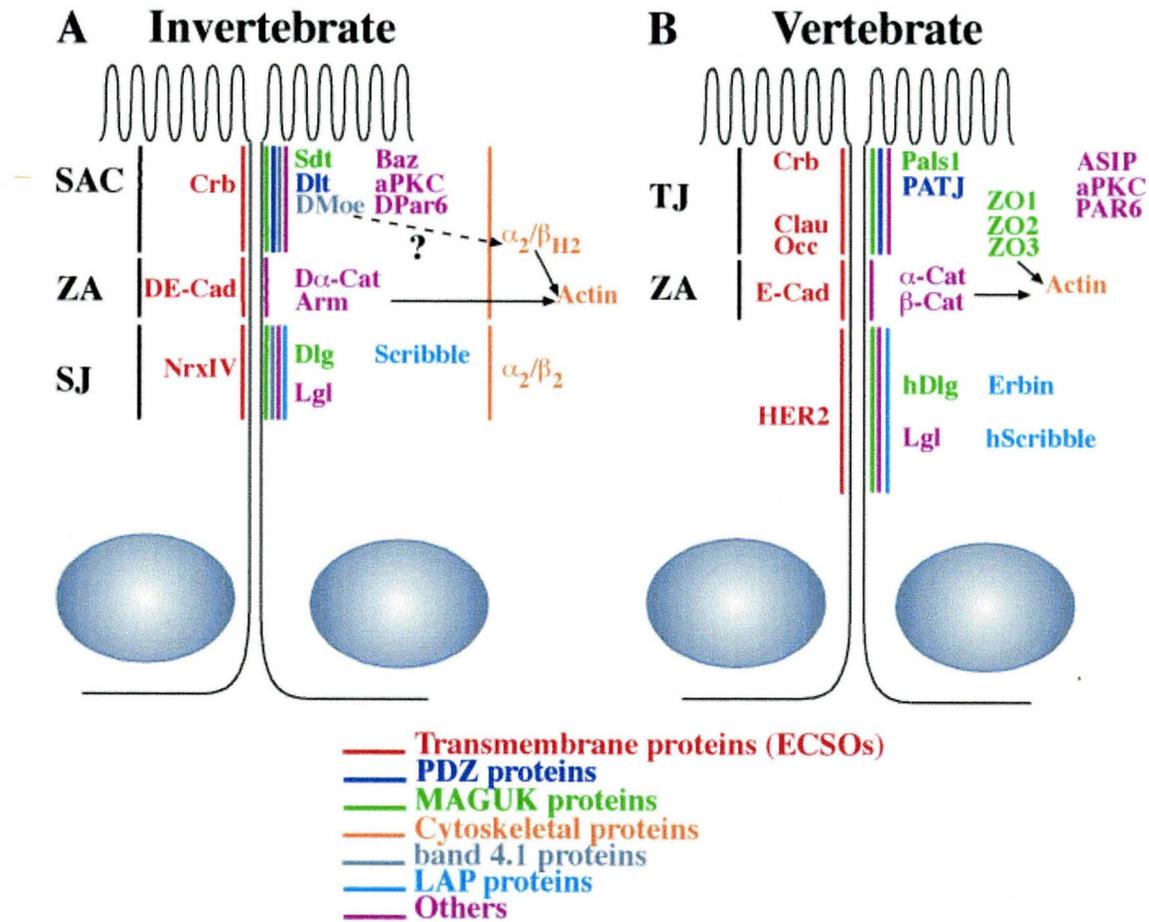
for proper targeting to the SAR and vise versa (Wodarz *et al.*, 2000; Petronczki *et al.*, 2001). In Baz loss of function mutants, cell polarity is interrupted exhibiting structural defects, irregular arrangement of the epithelial monolayer and incomplete ZA formation (Muller *et al.*, 1996). Furthermore, other apical markers such as Crumbs (Crb) are mislocalized. This suggests that the Baz complex is essential for the formation of ZA and the establishment of cellular polarity (Roh and Margolis, 2003).

1.3b The Crumbs Complex:

The next protein complex involved in epithelial polarity was originally identified in *Drosophila* as the Crb/Stardust (Std)/Discs Lost (Dlt) complex (Bhat *et al.*, 1999). However recent studies have demonstrated Dlt does not encode the PDZ protein required for cell polarity regulation but does encode a protein believed to be required for cell survival and cell cycle progression. Therefore the PDZ protein thought to be encoded by Dlt has been renamed after its vertebrate ortholog, Pals-associated tight junction protein (PATJ) (Pielage *et al.*, 2003).

The Crb/Std/PATJ complex also localizes to the sub-apical region of the plasma membrane (Figure 2). Crb is a transmembrane protein with a small cytoplasmic domain and a large extracellular domain (Tepass *et al.*, 1990). The cytoplasmic tail contains a 4.1 binding domain and a PDZ binding motif. The binding of Protein 4.1 to the 4.1 binding domain links Crb with the underlying spectrin/actin cytoskeleton, while the PDZ binding motif interacts with the PDZ domain of Std (Bachmann *et al.*, 2001; Hong *et al.*, 2001). The MAGUK protein Std contains the core domain structure of PDZ domain/SH3 domain/GUK domain, plus an L27 domain. This L27 domain interacts with PATJ

(Pals1-associated tight junction protein) via its L27 domain. The mammalian orthologs of the Crb complex localize to TJ, been identified as Crb3/Pals1/PATJ respectively and appear to associate and function in a similar manner as the *Drosophila* orthologs (Rho *et al.*, 2002; Rho *et al.* 2003).

As in the Baz complex, mislocalization of any one component of the tripartite Crb complex results in mislocalization of the other two proteins (Bachmann *et al.*, 2001; Hong *et al.*, 2001). Crb is an important apical surface determinant and is necessary for the correct placement and formation of ZA. Recent studies have demonstrated loss of epithelial organization and improper ZA formation in the absence of Crb or Std (Tepass, 1996; Grawe *et al.*, 1996; Muller *et al.*, 1996). The mislocalization of Baz also occurs in the absence of Crb implying the proper positioning of Baz at the subapical region is mediated by Crb. Thus the Crb and the Baz complexes colocalize and function together at the apical membrane to form the ZA (Bilder *et al.*, 2003).

1.3c The Scribble Complex:

The third protein complex involved in cellular polarity consists of Scribble (Scrib)/Discs Large (Dlg)/Lethal giant larvae (Lgl). Originally identified in the larval stages of *Drosophila*, Scrib, Dlg and Lgl were classified as neoplastic tumor suppressor genes (nTSGs) given that loss of function mutants demonstrated overproliferating cells with epithelia disorganization (Gateff and Schneiderman, 1967; Stewart *et al.*, 1972; Bilder and Perrimon, 2000). Scrib encodes a cytoplasmic LAP scaffold protein with 16 leucine rich repeats (LRR) near the N-terminus and 4 PDZ domains (Bilder and Perrimon, 2000). Proteins that contain both LRRs and PDZs are members of the LRR and PDZ domain-containing (LAP) family and are considered to function in structural

scaffolding (Santoni *et al.*, 2002).   Dlg is classified as a scaffolding protein of the MAGUK family, containing 3 PDZ domains, an SH3 domain and a GUK domain (Woods and Bryant, 1991).   Both Scrib and Dlg are restricted to the basolateral epithelial surfaces where they localize and associate together at the SJ (Mathew *et al.*, 2002) (Figure 2).   While Lgl, a myosin II binding protein that contains four WD-40 repeats (Strand *et al.*, 1994; Gateff, 1978; Mechler, *et al.*, 1985), localizes to the basolateral membranes (Manfruelli *et al.*, 1996; Bilder *et al.*, 2000) (Figure 2).   It is unclear which interactions specify Lgl localization, however in the absence of Scrib and Dlg, Lgl is lost from the plama membrane (Bilder *et al.*, 2000).   Thus a critical function of Scrib and Dlg is to recruit Lgl to the membrane, where Lgl can mediate its currently unknown function. Scrib, Dlg and Lgl are not limited to epithelial tissues, they are also found at synaptic junctions and in embryonic neuroblasts (Mathew *et al.*, 2002; Albertson and Doe, 2003).

Loss of Lgl or Dlg leads to the mislocalization of Scrib, while the loss of Dlg or Scrib leads to the mislocalization of Lgl.   These results, along with similar phenotypes seen in mutants, suggest that Scrib/Dlg/Lgl function together in a common pathway.   In the absence of Dlg, Scrib or Lgl, the apical surface expands causing the ZA and apical determining proteins such as Crb to mislocalize more laterally along the basolateral membrane (Woods *et al.*, 1997; Bilder *et al.*, 2000; Bilder and Perrimon; 2000).   There is a precise balance between the Scrib complex and the Crb complex for the accurate localization of ZA and establishment of apical and basalateral boundaries of the plasma membrane (Bilder *et al.*, 2003; Tanentzapf and Tepass, 2003).   However in Scrib complex mutants, the basolateral membrane remains organized and intact, suggesting the Scrib complex antagonizes apical proteins from expanding into the basolateral region.

Thus it would seem the Scrib complex maintains rather than establishes the basolateral membrane (Bilder *et al.*, 2003; Tanentzapf and Tepass, 2003).

The mammalian orthologs of the Scrib complex localize basally to the AJ and have been identified as Scrib, Vartul/mDlg,SAP97/mLgl, respectively (Lee *et al.*, 2002; Musch *et al.*, 2002; Nakagawa and Huibregtse, 2000). They associate and function in a similar manner as *Drosophila* orthologs in epithelial polarity.

1.4 The Baz, Crb and Scrib Complexes Function Together to Establish and Maintain Epithelial Polarity:

Recent research in the establishment and maintenance of epithelial polarity has focused on how the Baz, Crb, and Scrib complexes function together, rather than as separate components. The assembly of the spot adheren junctions (sAJ) and the localization of the Baz complex is the initial step in propagating apical epithelial polarity, while the Scrib complex distinguishes a basolateral domain by repressing the activity of Baz (Bilder *et al.*, 2003; Tanentzapf and Tepass, 2003). To counteract the repression of the Scrib complex, the Baz complex recruits the Crb complex to the SAR, and Crb acts to represses Scrib activity. Thus Baz/Dm-Par6/DaPKC and Crb/Std/PATJ function together to restrict Scrib basally while Scrib/Dlg/Lgl restricts the Crb complex apically. The delicate balance between these complexes establishes the boundaries of the apical and basalateral membrane domains and situates the ZA at the interface between the two domains (Bilder, 2004).

1.4a Par/Bazooka complex interacts with Crumbs complex:

Par-6 directly interacts with PALS1:

The molecular mechanisms underlying the regulatory hierarchy between these three complexes remains unclear, however, recent studies in vertebrates have revealed a direct interaction between the Par3/Par6/aPKC, Crb3/Pals1/PATJ and Scrib, Vartul/mDlg, SAP97/mLgl complexes. A newly discovered evolutionarily conserved region (ECR1) of the amino terminus of Pals1 functions as the Par6 binding site, as it mimics a protein carboxyl terminus and binds the PDZ domain of Par6 (Wang *et al.*, 2004; Penkert *et al.*, 2004). Crb has also been recognized in this interaction to promote recruitment of Par6 through Pals1 to the membrane suggesting a co-dependent process. Thus the disruption of one complex subsequently affects the localization of proteins in the other complex (Hurd *et al.*, 2003). The following reaction also exists in *Drosophila* orthologs, Stardust and DmPar6. However Par6/Pals1 binding may interfere with PATJ/Pals1 binding, suggesting these two interactions work independently rather than synergistically (Wang *et al.*,2003). Thus, Pals1 may have two conformational states at its amino terminus. In the first state, Pals1 interacts with PATJ with a properly folded L27N and heterodimerizes with the L27 domain of PATJ. In the second state, the L27N domain of Pals1 is largely unfolded, allowing for internal recognition of the ECR1 by the PDZ domain of Par-6 (Wang *et al.*, 2004). Some researchers believe that the interaction between Par-6 and Pals-1 is regulated by Cdc42-GTP (Wang *et al.*, 2004).

DaPKC interacts with Crumbs and PATJ:

Another recent study has shown the requirement for the kinase activity of DaPKC to maintain the position of apical determinants from the Crb complex. DaPKC directly

binds to both Crb and PATJ, and Crb is a phosphorylation target of DaPKC.   Early in development DaPKC kinase activity is necessary for the apical localization of PATJ. When the Crb complex accumulates at the apical plasma membrane, its hypothesized that PATJ has a greater affinity for Sdt bound to Crb than for DaPKC thus releasing PATJ from its interaction with DaPKC and permitting Crb to be phophorylated. The phosphorylation of Crb by DaPKC may be necessary for Crb to assume the proper conformational and/or to interact with proteins essential for Crb stabilization at the apical domain (Sotillos *et al.*, 2004).

1.4b The Par/Bazooka complex interacts with the Scribble complex:

Par-6 directly interacts with mLgl:

Par6 also directly binds mLgl in a similar manner as the Pals1/Par6 interaction. The binding is mediated through the amino-terminal region of mLgl, where the WD-40 repeats are located, to the Par-6 PDZ domain (Yamanaka *et al.*, 2003; Plant *et al.*, 2003; Betschinger *et al.*, 2003). This allows Lgl to become a substrate of aPKC. The phosphorylation of Lgl at conserved Protein Kinase C consensus sites by aPKC causes inactivation and dissociation of Lgl from the cytoskeleton selectively at the apical plasma membrane (Kalmes *et al.*, 1996; Betschinger *et al.*, 2003). Lgl localized basally remains unphosphorylated and maintains an association with the basalateral plasma membrane. Thus the apical localization of aPKC by the Baz complex prevents Lgl associations with the apical membrane, thereby restricting Lgl activity to the basolateral surface. Figure 3 illustrates these interactive relationships between the Baz, Crb and Scrib complexes.

These results emphasize the sophisticated relationships between the evolutionarily conserved proteins involved in establishing and maintaining epithelial polarity. The

Figure 3: A proposed Model for the Function of the Bazooka, Crumbs and Scribble complexes as epithelial polarity regulators. The three complexes appear to function together to establish and maintain polarity in epithelial cells. The Bazooka complex determines the apical domain by localizing to the sub-apical region (SAR), to recruit the Crumbs complex. The Scribble complex distinguishes the basolateral domain by repressing the activity of Bazooka. Par 1 and 14-3-3 proteins may also play a role in excluding the Bazooka complex from the basolateral membrane. The recruitment of the Crumbs complex is to counteract the repression of the Scribble complex. Thus the delicate balance between the Bazooka, Crumbs and Scribble complexes establishes the apical/basolateral borders of epithelia and situates the ZA at the interface between the two domains.
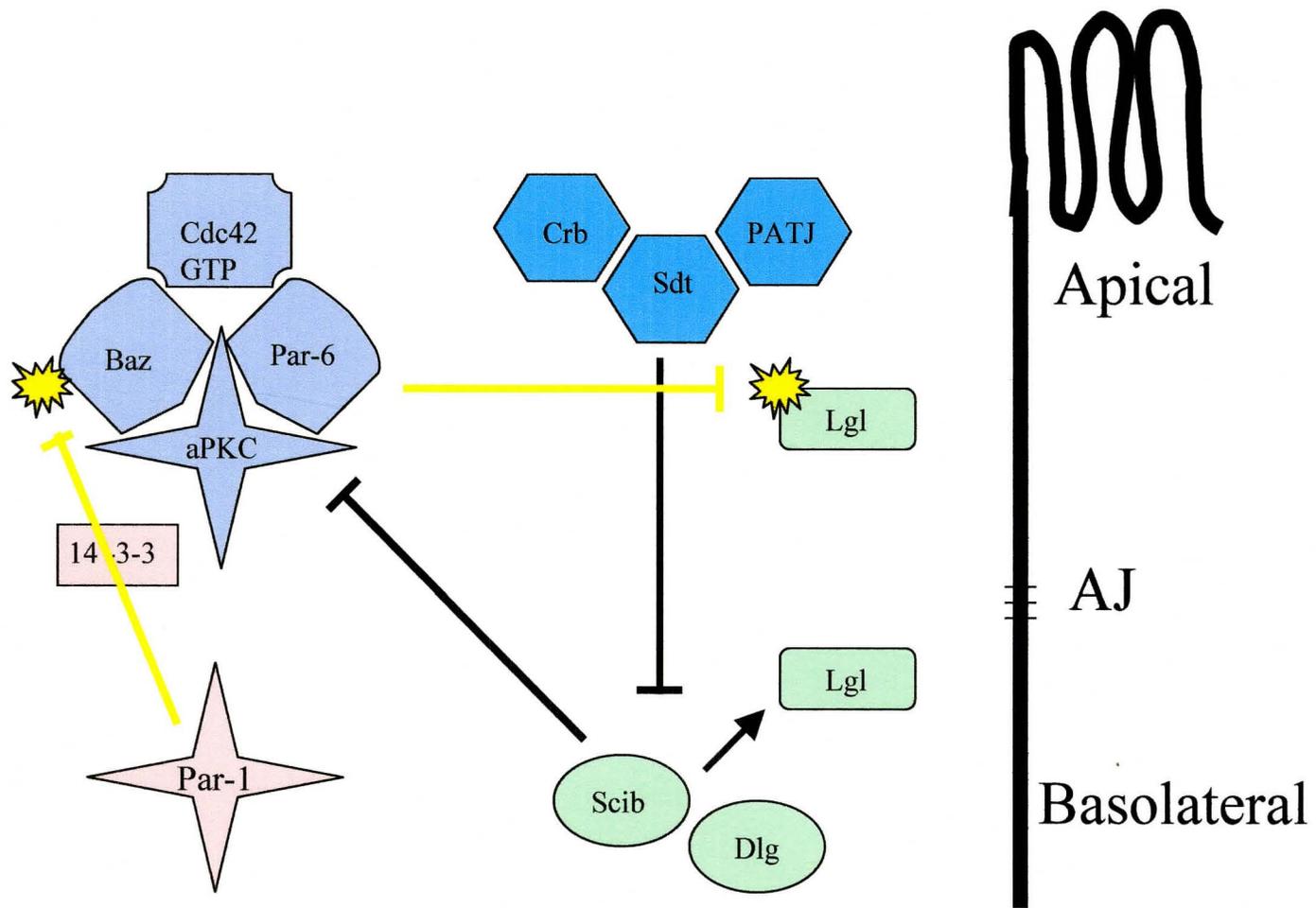
flexibility of the Par-6 PDZ domain able to bind to a large number of proteins (Cdc42, Par-3, Pals1/Sdt, Crb, and Lgl), along with the direct interactions between various members of the three complexes gives insight into the complexity and the yet unknown interactions necessary for proper epithelial formation (Penkert *et al.*, 2004). Further characterization and identification of proteins involved in the establishment and maintenance of apical/basolateral domains will provide valuable information regarding the structure/function relationships in epithelial polarity.

1.5 Proteins – Associated with Lin-7 (PALS)

Another PDZ protein complex consisting of LIN-2/LIN-7/LIN-10, originallyidentified in *C. elegans,* is responsible for the proper organization of epithelia cells. LIN-7 establishes a complex with LIN-2 and LIN-10 to localize LET-23, a homologue of the epidermal growth factor receptor. The proper basolateral positioning of LET-23 by this complex is necessary for the vulva induction in the developing nematode. The *Drosophila* orthologs of LIN-2, LIN-7 and LIN-10 have been identified and are known as dVELI, CAMGUK and dMINT, respectively. It is hypothesized these three proteins function together in a complex to localize proteins to the post-synaptic density (MacMullin, 2001). The mammalian orthologs of LIN-2, LIN-7 and LIN-10 have been identified and are known as Veli, CASK and Mint1, respectively. These proteins function together as a scaffolding complex in both neuronal and epithelial cell. Veli can bind CASK and Mint1 independently of the PDZ domains, which are free to recruit cell adhesion molecules such as neurexin and syndecan (Cohen *et al.*, 1998), receptors such as the NMDA receptor (Setou *et* al., 2000; Jo *et al.*, 1999) and signaling

proteins such as the GABA transporter (Perego *et al.*, 1999; Straight *et al.*, 2000). Thus this complex plays a fundamental role in sorting the interacting proteins to specialized membrane domains in neuronal and epithelial cells. The multiple binding partners of the Veli/CASK/Mint1 complex implies that these interactions may function together within the same cell, in a distinct sequential order or in a competitive manner. However the precise mechanism used by this complex to properly target their binding partners remains to be elucidated. Thus there must be additional binding partners and/or modes of regulation for cell type specificity to favor one interaction over another.

Two novel proteins, originally discovered by Far Western overlay assay in *Mus musculus*, have been identified as additional binding partners of Veli: PALS1 and PALS2 (Proteins Associated with LIN-7). Both proteins are members of the MAGUK protein family, with a domain structure of two L27 domains (L27N and L27C), a PDZ domain, a SH3 domain, a HOOK domain and a GUK domain. PALS1 and PALS2 are thought to compete with Lin-2/CASK for binding VELI and its multiple isoforms via L27 domain dimerization (Kamberov *et al.*, 2000).

PALS1, in addition to the core MAGUK domain structure, also contains an Unknown 1(U1) domain at the extreme N-terminus that bears no significant similarity to other proteins (Roh *et al.*, 2002). This U1 domain contains two conserved domains, identified as evolutionarily conserved region 1 and 2 (ECR1 and ECR2, respectively). The ECR1 of PALS1 has been shown to be necessary for binding Par-6 (Wang *et al.*, 2004). As mentioned above, cell polarity determinant, PALS1 and its *Drosophila* ortholog Std, localize apically to TJ and the SAR, respectively, to serves as an adapter proteins linking Crumb and PATJ (Roh *et al.*, 2002). Most recently, PALS1 has

demonstrated a direct link to the cytoskeleton by directly binding ezrin, a band 4.1

protein and localizing it to the apical membrane of gastric pariental cells (Cao *et al.*,

2005). Recently, Std has been shown to directly interact with dVELI, as in mammals, at

the SAR of epithelial cells via their L27 domains (Bachmann *et* al., 2004). This is in

striking contrast to the basolateral plasma localization seen in both mammals and

nematodes (Perego *et al.*, 1999; Straight *et al.*, 2000; Straight *et al.*, 2001; Simske, *et al.*,

1996; Kaech, *et al.*, 1998). Thus difference in expression, subcellular localization and

binding capacity of potential interaction partners may account for this discrepancy

(Bachmann, *et al.*, 2004).

PALS2, like PALS1 was originally identified in *Mus musculus*, as an alternative

binding partner for Veli. There are two isoforms of PALS2, PALS2α and PALS2β,

differing from each other by a fourteen amino acid insertion located between the PDZ

and the SH3 domain (Kamberov *et al.*, 2000). Studies have shown that PALS2

colocalizes with E-cadherin below the TJ, however its functionality remains unclear (Roh

et al., 2002). Recent studies have shown PALS2 directly associates with Nectin-like

Molecule-2 (Necl-2) at extra-junctional regions of the basolateral surface of epithelial

cells (Shingai *et al.*, 2003). Necin-like molecules (Necls) were identified and designated

based on their similar domain structure to nectins (Ikeda *et al.*, 2003). Nectins are Ca^{2+}

independent cell-cell adhesion molecules that function in the organization of cell

junctions, such as AJ, TJ, synaptic junctions and heterotypic junctions. The nectin family

has four members (nectin-1, 2, 3, 4), where all, except nectin-4, have a C-terminal PDZ

binding motif that interacts with afadin. Afadin is an actin binding protein that connects

nectin proteins to the cytoskeleton (Takai and Nakanishi, 2003). The C-terminus of

Necl-2 does not bind Afadin but directly binds the PDZ domain of PALS2. This association, along with Necl-2 directly binding band 4.1 protein DAL-1, suggests Necl-2 functions in cell adhesion and transmembrane proteins localization (Shingai *et al.*, 2003).

The human ortholog of PALS2 has recently been identified as VAM-1 (VELI associated MAGUK)/MPP6 (Multiple PDZ protein 6) (Tseng *et al.*, 2001; Katoh and Katoh, 2004). VAM-1/MPP6 may be a candidate tumor suppressor since its mapping to human chromosome 7p15-21 is a region reported to demonstrate changes in Wilm's tumors (Tseng *et al.*, 2001). However its function is still unknown.

1.6 *Drosophila* – an excellent model organism to study the structure and function of PALS2 with respect to epithelial polarity:

Drosophila melanogaster is an excellent model organism for genetic and biochemical research. The simple morphology, well annotated genome and ease of genetic manipulation make *Drosophila* an organism of choice for studying protein interactions. With high degree of conservation at the genomic level with mammals, *Drosophila* share many essential biological properties but in a much more genetically amenable organism.

Epithelial polarity is essential for the proper development and function of all organisms. The proper organization of polarity determinants divides the plasma membrane into apical and basolateral compartments, which is important for cell morphology, tissue physiology and cell signaling. Defects in epithelial polarity determinant CRB1, an isoform of CRB3, leads to retinal degeneration as a result of photoreceptor disorganization (Williams *et al.*, 1990; den Hollander *et al.*, 2001 & 2002).

Loss of any basolateral polarity proteins from the Scrib complex result in tumor growth from unregulated proliferation (Bilder 2004). Thus it is necessary to further identify and understand proteins involved in epithelial polarity. Studies of PALS2/VAM-1 have indicated a potential involvement in cellular polarity, however very little is currently known about its function. Since many of the mammalian epithelial polarity proteins have a *Drosophila* ortholog with similar structure and function, a PALS2/VAM-1 ortholog was sought to help understand its purpose and necessity in development.

The *Drosophila* genome was screened for the presence of an ortholog to PALS2 and revealed two proteins, CG9326 and CG13219, now known as Skiff. Both CG9326 and Skiff, containing an L27 domain, a PDZ domain, a SH3 domain and a Guk domain, showed 39% amino acid identity to PALS2. The sequence homology discovered by a blast search selected CG9326 as a possible PALS2 ortholog and a possible cell polarity protein.

The focus of this thesis will be to molecularly characterize the possible PALS2 ortholog, CG9326. Genomic analysis at the nucleotide and amino acid level along with domain homology will be assessed for similarities to other known PALS2 proteins in other species. Embryonic and larval expression patterns of CG9326 transcripts will also be studied. Generation of an antibody will allow for protein expression characterization. Therefore the identification and characterization of CG9326 will reveal initial insights into its function and possible role in epithelial polarity.

Chapter 2

Materials and Methods

2.1 Embryo, Larvae and Adult *Drosophila* Collections:

To collect embryos, adult flies were placed in a "house" which consisted of a 100 ml plastic beaker with air holes capped off by a 60 x 15 mm petri dish containing apple juice agar and a small amount of yeast paste. Plates placed on the house in the morning would be kept on for 8 hours at 25°C and then placed in 18°C overnight to slow the development of the embryos. The plate placed on the house in the evening would be kept on overnight at 25°C and then placed in 18°C for 5 hours the following morning. After incubation in 18°C, plates are then kept at 4°C. This method would give a sufficient mixture of mid-late stage embryos.

To collect 3rd instar larvae, ddH2O was added to a food vial where larvae were at wandering age (80-90 hours). The larvae were then collected on a larval sieve.

To collect adult flies, males and females approximately a week old were moved from a yeast food vial to an empty food vial for 2-3 hours to rid the flies of ingested yeast. They were then ready for use.

2.2 *Drosophila* embryonic, larval and adult total RNA preparations:

All solutions were treated with Diethyl Pyrocarbonate (DEPC) (Sigma Alderich, cat. # D5758) to a final concentration of 0.1%, incubated at 37°C and then autoclaved.

The total RNA preparations were achieved using TRIzol reagent (Life Technologies, cat. # 15596). *Drosophila* tissue (50-100mg) from each embryo, larvae and adult were frozen in liquid nitrogen and ground up with a mortar and pestle. The tissue was then transferred to a dounce homogenizer containing 2ml of TRIzol reagent and homogenized a few times. The extraction was transferred to 1.5ml eppendorf tubes and centrifuged at 4000 rpm for 4 minutes at 4°C to remove any tissue debris. Chloroform was added to a volume of 0.2ml per 1ml of TRIzol reagent, the tube was vortexed and centrifuged at 10000 rpm for 15 minutes at 4°C. The aqueous phase was transferred to a new tube and the RNA was precipitated with isopropyl alcohol at −80°C for 15 minutes followed by centrifugation at 15000 rpm for 30 minutes at 4°C. The supernatant was removed and the RNA pellet was washed with 70% ethanol and centrifuged again at 15000 rpm for 15 minutes at 4°C. The pellet was briefly air dried for 5 minutes and then resuspended in an appropriate amount of DEPC ddH_2O. The concentration and purity of the total RNA was determined by absorbance at 260 and 260/280.

2.3 Preparation of mRNA from total RNA preprations:

(all solutions are RNase free)

Purification of mRNA from the total RNA preparations were achieved by the QIAGEN Oligotex mRNA kit (QIAGEN, cat. # 70042).

2.4 Reverse Transcriptase Polymerase Chain Reaction (RT-PCR):

RT-PCR was carried out using Ready-to-go RT-PCR Beads (Amersham Pharmacia Biotech, cat. # 27-9259-01) that entailed a one step protocol to generate a PCR product from an RNA template. Each RT-PCR reaction when brought to a final volume of 50μl contained approximately 2 units of *Taq* DNA polymerase, 10mM Tris-HCL, 60mM KCl, 1.5mM MgCl2, 200μMof each dNTP, Moloney Murine Leukemia Virus (M-MulV) reverse transcriptase and an RNA inhibitor. For the reaction a total of 1μg of total RNA was added along with 0.5μg of first strand primer pd(T) 12-18, 10pmol of each forward and reverse primers (Appendix 1) and DEPC H2O up to 50μl final volume.

The reverse transcriptase reaction was initiated by incubation at 42°C for 25 minutes and then inactivated at 95°C for 5 minutes. The reaction was continued through a standard 3 step polymerase chain reaction (PCR) in a thermal cycler with denaturation at 95°C for 1 minute, annealing for 1 minute and polymerization at 70°C for 2 minutes. This reaction was cycled 35 times.

2.5 Preparation of *dPALS2 pBS SKII* contructs:

For the *dPALS2* construct, RT-PCR amplified an ~259 base pair region of cDNA of *dPALS2* that lies between the SH3 domain and the Guanylate kinase domain (Appendix 2). For the L27 construct, RT-PCR amplified an ~ 180 base pair region of the 5' UTR of *dPALS2* that coded for the L27 domain. Primers were designed for each reaction with specific restriction sites for subcloning into *pBS SKII*. The *dPALS2* construct used *XhoI* and *XbaII* while the L27 construct used *KpnI* and *XbaI* (Appendix 3).

The PCR products were cloned directed into *pBS SKII* and screened by restriction digest and sequence analysis for positive transformants.   These constructs were stored at –80C until required for DIG labelling.

2.6 Preparation of DIG labeled RNA probes:

The production of sense and anti-sense RNA probes labeled with digoxigenin were achieved by *in vitro* transcription using T7 and T3 polymerases and the DIG RNA labeling kit (Roche Diagnostics, cat. # 1277073).

DNA was purified from an overnight bacterial culture using Qiagen mini prep kit (cat. # 27106) and linearized with the appropriate restriction enzyme. Phenol/Chloroform extraction and precipitation with 3M sodium acetate and chilled 100% ethanol purified the linearized template DNA. For the labeling reaction, 1µg of purified linearized DNA was mixed with 2µl of transcription buffer, 1µl of RNase inhibitor, 2µl of NTP labeling mix, 2µl of T_7 or T_3 polymerase and DEPC treated ddH_2O to a final volume of 20µl. The reaction was incubated at 37°C for 2 hours and then stopped with 2µl of 200mM EDTA, pH 8.0.

2.7 Embryonic and Larval Fixation for RNA *in situ* hybridization:

CS-P embryos were collected as described in 2.1. The embryos were dechorionated by rinsing with 50% bleach for 5 minutes, collected on a nitex sieve and rinsed with embryo wash (7% NaCl, 0.05% Triton X-100). The embryos were then placed in a scintillation vial that contained 2ml of embryo wash, 2ml of 10%

formaldehyde, 1ml of 5X fixation buffer (800mM KCl, 200mM NaCl, 20mM EGTA,

5mM spermadine HCl, 2mM spermine, 150mM pipes, pH 7.4) and 5ml of heptane. The

vial was shaken well to ensure proper mixture and then place on a rotator for 20 minutes.

The bottom aqueous layer was removed, methanol was added and the vial was shaken

vigorously to crack the vitelline membrane of the embryos. The embryos that fell and

gathered at the bottom of the vial were collected, washed 3X with methanol, washed 4X

with 100% ethanol and stored at -20°C in ethanol until *in situ* hybridization

The brain, nerve cords and imaginal discs of CS-P larvae were dissected in 1X ice

cold PBS and fixed in 4% formaldehyde at 4°C overnight on a rotator. The next morning

the larval tissue was washed 4X with methanol, 5X with 100% ethanol and then stored at

–20°C in ethanol until *in situ* hybridization.

2.8 RNA *in situ* Hybridization:

(all solutions were DEPC treated)

The protocol for RNA *in situ* labeling was similar for both embryos and larvae

with the exception of some minor modifications. After fixation, the embryos and larval

parts were rinsed in 50% ethanol/50% xylene and then incubated in 100% xylene for 3

hours with shaking. Upon completion, the xylene was removed and the embryos and

larval parts were rinsed again with 50% ethanol/50% xylene. After several rinses with

100% ethanol and a single rinse in 50% methanol/50% PBT, the embryos were post fixed

for 10 minutes in PBT containing 2% formaldehyde while the larval parts were post fixed

in 50% methanol/50% PBT containing 5% formaldehyde. An additional 30 minutes

fixative step in PBT containing 5% formaldehyde was performed on the larval tissues

before both embryos and larval tissues were washed 3X for 2 minutes in PBT. This was

followed by a reaction of 50µg/ml of Proteinase K in PBT for 2 minutes for embryos and

7 minutes for larva. These reactions were stopped by adding PBT containing 2mg/ml

glycine, followed by a second post fixative reaction in PBT with 5% formaldehyde for 10

minutes. Both embryos and larvae were then washed 4X in PBT and 3X with 50%

PBT/50% prehybridization solution (50% formamide, 300mM NaCl, 10mM Tris-HCl

(pH 6.8), 10mM NaH2PO4 (pH 6.8), 1X Denharts solution, 5mM EDTA, 1mg/ml yeast

tRNA) in 2 minute intervals.

A single room temperature wash with prehybridization solution was performed

before embryos and larval parts were incubated at 53°C for 1 hour in prewarmed

prehybridization solution with rocking. After prehybridization, the embryos and larvae

were placed in hybridization solution (same as prehybridization solution except for the

addition of 10% w/v dextran sulfate and absence of tRNA) containing DIG labeled RNA

probes (range of dilutions from 1:500 to 1:1000) to incubate at 53°C overnight without

rocking.

The following day, embryos and larval tissue were placed in a series of post

hybridization solution (same as prehybridization solution except for the absence of

tRNA) and PBT washes at 53°C for 20 minutes where the gradient of post hybridization

solution decreased and PBT increased (80%, 60%, 40%, 20%). Two 20 minute washes

with PBT, 1 at 53°C and 1 at room temperature followed. Incubation for 20 minutes at

37°C with 20µg/ml of RNase A was needed to remove excess unbound probe. After

RNase A treatment, embryos and larval parts were washed 3X with room temperature

PBT and then incubated on a shaker with 1ml of 1:2000 dilution of anti-DIG antibody

(anti-Digoxigenin-AP, Fab fragments, Roche Diagnostics) for 1 hour.

The embryos and larvae tissue were then washed 4X with room temperature PBT

for 20 minutes each and washed 3X with detection solution (100mM NaCl, 50mM

MgCl2, 100mM Tris (pH 9.5), 0.1% Tween 20) for 5 minutes each. The reaction

solution (3.5ul 5-Bromo-4-chloro-3-indolyl-phosphate, 4-toluidine salt (X-phosphate, 4-

toluidine salt) (BCIP) and 4.5ul nitroblue tetrazolium chloride (NBT) per 1mL of

detection solution) replaced the detection solution and the embryos and larvae were

placed in the dark.

Once the preferred reaction had occurred, the reaction solution was removed and

then washed 5X with PBT and 2X with PBS. Final washes were through a glycerol/PBS

gradient where embryos and larval parts were stored in 70% glycerol/30% PBS.

2.9 Northern Blot Analysis:

(all solutions were DEPC treated)

A 1.2% agarose gel containing MOPS buffer (20mM Morpholinopropanesulfonic

acid, 5mM sodium acetate, 1mM EDTA, pH 7.0) and 6% formaldehyde was prepared. A

total of 5μg of adult, larval and embryonic mRNA were loaded into the agarose gel.

Along side the mRNA samples, 40ng of RNA molecular weight marker III (Roche

Diagnostics, cat. # 1062638) was also loaded. Both the mRNA samples and weight

marker were denatured at 65°C for 10 minutes and then cooled on ice before loading. The agarose gel was electrophoresed in MOPS buffer at 80V for 2 hours.

The gel was washed twice for 15 minutes with 20X SSC (3M NaCl, 0.3M Na-citrate) to remove the formaldehyde. A blot transfer was set up where Whatmann paper soaked with 20X SSC laid on top of an inverted glass container resting in a shallow reservoir of 20X SSC. The gel was placed face down on top of the wet sheet of Whatmann paper and a piece of positively charged nylon membrane cut to the size of the gel was placed on top. Two dry sheets of Whatmann paper cut to the size of the gel were added next. A 6 inch stack of paper towels, a glass plate and a 500g weight were the last added. The RNA transferred for 16 hours with 20X SSC as transfer buffer.

The membrane was washed briefly with 2X SSC to remove excess salt and then baked at 80°C for 2 hours to fix the RNA to the blot. Prehybridization of the membrane took place at 68°C with agitation in 25ml of DIG Easy Hyb Solution (Roche Diagnostics, cat. # 1603558) for 1 hour. The blot was incubated with the DIG labeled RNA probe at a concentration of 100ng/ml in Easy Hyb Solution overnight at 68°C, with agitation. After hybridization the membrane was washed twice for 5 minutes with room temperature low stringency buffer (2X SSC containing 0.1% SDS) and twice for 15 minutes with 68°C high stringency buffer (0.1X SSC containing 0.1% SDS).

After stringency washes, the membrane was rinsed briefly in washing buffer (Roche Diagnostics, cat. # 1585762) and incubated for 30 minutes in 50ml of blocking solution (Roche Diagnostics, cat. # 1585762) on a shaker. The membrane was then incubated with 30ml of a 1:5000 dilution of anti-DIG antibody (anti-Digoxigenin-AP,

Fab fragments, Roche Diagnostics) for 1 hour with shaking. After antibody incubation,

the membrane was washed twice for 15 minutes with washing buffer (Roche Diagnostics)

and equilibrated with detection buffer (Roche Diagnostics, cat. # 1585762) for 5 minutes.

The membrane was then placed in NBT/BCIP reaction solution in the dark for 12-16

hours.

2.10 Preparation of *dPALS2* His Fusion construct:

A 909 base pair region of *dPALS2* located after the PDZ domain was amplified by

RT-PCR. Restriction sites *EcoRI* and *BglII* were incorporated into the primers for

subcloning purposes (Appendix 4). The PCR product was directly cloned into *pCR2.1*

vector by utilizing the TA cloning Kit (Invitrogen, cat. # 45-0046) and screened by

restriction digestion and sequence analysis for potential positive transformants. Once a

dPALS2 pCR2.1 transformant was confirmed, *dPALS2* was shuttled into *pET29b* vector

via the *EcoRI* and *BglI* restriction sites. The *dPALS2 pET29b* plasmid was then

transformed into bacterial host strain BL21λDE3 (Invitrogen, cat. # C6000-03) for

protein induction.

2.11 Induction of dPALS2 His Fusion Protein:

An inoculated culture of *dPALS2 pET29b* was grown for 16 hours at 37°C on a

shaker. The following morning, the overnight culture was used to inoculate 1L of

LB containing Kanamycin (final concentration of 30ug/ml). After inoculation the culture

was incubated at 37°C with shaking until the $OD_{600} \sim 0.4$ was reached. The culture was

then induced with 0.5mM of IPTG (Isopropylthio-beta-D-galactoside) and allowed to

grow for an additional 4 hours at 37°C with shaking. The cells were then harvested by

centrifugation at 10000 rpm for 10 minutes and placed at –20°C until purification of

dPALS2 protein.

2.12 Purification of dPALS2 His Fusion Protein:

After the cells were harvested, the dPALS2 His fusion protein was purified by

two methods, a denaturation method with purification by affinity chromatography and a

refolding method with N-Lauroylsarcosine and DTT.

2.12a Purification by Affinity Chromatography:

For the denaturing method, the cells were resuspended in 40ml of ice cold binding

buffer (0.5M NaCl, 20mM Tris-HCl, 5mM imidazole, pH 7.9) per 100ml culture volume.

Complete EDTA-free protease inhibitor (Roche Diagnostics, cat. #1873580) and 0.1%

NP-40 (Roche Diagnostic, cat. #1754599) were present in the binding buffer to inhibit

serine and cysteine proteases and to reduce non-specific binding respectively. The cells

were sonicated on ice for 5 minutes with 30 second breaks between each minute.

Centrifugation at 5000 x g for 15 minutes collected the inclusion bodies and cellular

debris while leaving other proteins in solution. The supernatant was removed and the

pellet was resuspended in 20ml binding buffer per 100ml culture volume. The inclusion

bodies were collected again by centrifugation at 5000 x g for 15 minutes. The pellet was

resuspended in 5ml of 1X binding buffer containing 6M urea per 100ml culture volume

and incubated on ice for 1 hr to completely dissolve the protein. Insoluble material was

removed by centrifugation at 16000 x g for 30 minutes and the supernatant was stored at 4°C for purification by affinity chromatography.

To prepare for soluble target protein purification using affinity chromatography, 2ml of His-Bind Resin was added to a column. The resin was allowed to settle under gravity flow and then washed with 3 volumes of sterile deionized water, 5 volumes of charge buffer (50mM NiSO$_4$) and 3 volumes of binding buffer. Once the binding buffer had drained to the top of the column, the prepared soluble protein was loaded into the column. The column was washed with 10 volumes of binding buffer and 6 volumes of wash buffer (0.5M NaCl, 60mM imidazole, 20mM Tris-HCl, pH 7.9). The His tagged protein was then eluted from the column with 12 volumes of elution buffer (1M imidazole, 0.5M NaCl, 20mM Tris-HCl, pH 7.9).

The affinity purified His tagged dPALS2 protein was dialyzed in PBS multiple times and the Bradford Assay (Bradford, 1976) was used to determine the final concentration of purified protein.

2.12b Purification by Refolding method:

For the refolding method of protein purification, the harvested cells were resuspended in 0.1culture volume of IB wash buffer (20mM Tris-HCl pH 7.5, 10mM EDTA, 1% Triton X-100) that contained Complete EDTA-free protease inhibitor and 0.1% NP-40. The cells were sonicated on ice for 5 minutes with 30 second breaks between each minute. The inclusion bodies and cellular debris were collected by centrifugation at 10000 x g for 10 minutes. After removing the supernatant, the pellet was resuspended in 0.1 culture volume of IB wash buffer and centrifuged again at 10000

x *g* for 10 minutes. For every 15mg of pelleted inclusion body, 1ml of IB solubilization

buffer (50mM CAPS, pH 11.0) containing 0.3% N Lauroylsarcosine was used for

resuspension. This was incubated at room temperature for 15 minutes and then clarified

by centrifugation at 10000 x *g* for 10 minutes.

The His tagged purified dPALS2 protein was dialyzed for at least 3 hours at 4°C

in PBS containing 0.1mM of DTT and the dialysis continued through three additional

changes with PBS lacking DTT. The concentration of purified protein was determined as

described in 2.11a

2.13 Preparation of Polyclonal antiserum:

Using an 18 gauge needle, purified dPALS2 His fusion protein (final

concentration of 100ug/ml) in PBS was mixed with Freund's complete adjuvant (CFA)

until a homogeneous mixture was obtained. The mixture was then injected into 4 rats

(*Rattus norvegicus*) subcutaneously at 3 different sites where each site received 0.1ml.

To allow prime immunization, follow up booster shots were given to each rat every 14

days using incomplete Freund's adjuvant instead of CFA. Before each injection a tail

bleed of 0.5ml from each rat was taken. To prepare the sera, the bleeds were incubated at

37°C for 1 hour to allow maximum clotting and then placed on ice at 4°C overnight. The

following morning, the serum was centrifuged at 10000rpm at 4°C and the supernatant

was collected. Sodium azide was added to prevent bacterial contamination to a final

concentration of 0.02% and the sera was stored at –20°C.

2.14 Affinity Purification of dPals2 antibody:

dPALS2 polyclonal anti-serum was affinity purified with CNBr-activated sepharose 4B (Amersham Biosciences, cat. # 17-0430-01). The sepharose beads were first swollen in 1mM HCl for 15 minutes. The amount of beads was dependent on the amount of conjugated protein, approximately 5-10mg of protein per 1ml of beads (1g of sepharose beads yielded 3.5ml of gel). Then the swollen beads were washed on a sintered glass filter with 1mM HCl (200ml HCl/g of dry beads) and subsequently washed 5X in coupling buffer (0.1M NaHCO3, 0.5M NaCl, pH 8.3).

The beads were then incubated with the antigen solution in a 2:1 ratio overnight at 4°C. The following day, the beads were allowed to settle and then washed 5X with coupling buffer to remove excessive unbound antigen. Then 0.1M Tris-HCl (pH 8.0) was added to the beads and incubated for 2 hours at room temperature to block remaining active groups on the gel. Next, the beads were poured into a Biorad poly-prep chromatography column (BioRad, cat. # 731-1550) and allowed to settle. The column was washed with 5-10 column volumes of PBS and then the serum was added to the column. The serum was incubated with the beads overnight at 4°C on a rotator. The following morning, the serum was allowed to flow through the column by gravity flow. The column was washed of unbound antibody with several volumes of PBS.

The bound antibody was eluted from the column with 0.2M glycine (pH 2.3). The eluted antibody was neutralized with 1M Tris (150ul/mL of eluted solution) and then dialyzed overnight at 4°C in PBS. The following morning, add 0.02% sodium azide to the purified antibody and store at -80°C.

2.15 Preparation of Crude Protein from Drosophila Tissue:

Tissue from embryos, larvae and adults were collected and homogenized in cold RIPA lysis buffer (150mM NaCl, 1% NP-40, 0.5% deoxycholate, 0.1% SDS, 50mMTris-HCl pH 8.0, protease inhibitor cocktail tablets 1/10mL (Roche Diagnostics, cat. # 1836170)). A volume of 2mL of RIPA buffer was used for every 1g of tissue. Once homogenized, the tissues were incubated on ice for 10 minutes and then centrifuged for 5 minutes at 3500 rpm. The supernatant was collected and stored at –80°C.

2.16 Western Blot Analysis:

The expression levels of dPALS2 were analyzed in all stages of *Drosophila* development by Western Blot. Crude protein samples containing 30-50ug of protein were mixed with 5X SDS loading buffer (2% SDS, 10% glycerol, 100mM DTT, 60mM Tris (pH 8.0), 0.01% Bromophenol blue, 5% 2-mereaptoethanol) and boiled for 5 minutes. The protein samples were then loaded onto a 14% SDS PAGE gel and run at 120V for 90 minutes. After eletrophoresis, the protein was transferred to a PVDF membrane (Pall Life Sciences, P/N 66543). Prior to the assembly of the transfer stack, the sponges, 3MM whatman paper, PVDF membrane and the gel were soaked in transfer buffer (15.6mM Tris, 120mM glycine, 20% methanol, 0.02% SDS) for 30 minutes. The transfer stack was assembled and the protein was transferred for 90 minutes at 100V at 4°C.

Following the transfer, the membrane was washed briefly in post blot buffer (20mM Tris-HCl, pH 7.4, 150mM NaCl) and then blocked in 5% skim milk for 1 hour at

room temperature. The membrane was incubated overnight at 4°C in the appropriate dilution of primary antibody.

The following morning, the membrane was washed 3X with wash buffer to remove unbound primary antibody. Then the membrane was incubated in a 1:15000 dilution of horseradish peroxidase conjugated secondary antibody for 30 minutes at room temperature. The membrane was washed with wash buffer to removed unbound secondary antibody and then prepared for autoradiographic developing using Kodak X-ray film.

The membrane was incubated in equal volumes of reagent 1 mixed with reagent 2 from the ECL Western Blotting Kit (Amersham BioSciences, cat. # RPN2106) for 1 minute. The membrane was then exposed to the Kodak X-OMAT AR film in the dark at various time intervals and then developed to visualize the optimal signal.

2.17 Antibody staining of *Drosophila* embryos:

CS-P embryos were collected as described previously in 2.1. The embryos were dechorionated by rinsing with 50% bleach for 5 minutes, collected on a nitex sieve and rinsed in ddH$_2$O. The embryos were then placed in a scintillation vial that contained 5ml of heptane, 4.5ml of PBS and 0.5ml of 37% formaldehyde. The vial was shaken well to ensure proper mixture and then place on a rotator for 20 minutes. The bottom aqueous layer was removed, methanol was added and the vial was shaken vigorously to crack the vitelline membrane of the embryos. The embryos that fell and gathered at the bottom of the vial were collected, transferred to a glass tube and washed 3X with methanol. The

methanol was removed and the embryos were washed 3X with PBS with 0.2 Triton-X (PBT), then placed on a rotator for 20 minutes in PBT.

The embryos were removed from the rotator and washed once more with PBT. The embryos were then blocked with a 1:20 dilution of normal goat serum (Sigma Alderich) in PBT at room temperature for 30 minutes on an orbital shaker to reduce non-specific binding. Following the blocking step, primary antibody was added to the embryos at an appropriate dilution and incubated overnight at 4°C on an orbital shaker. The next morning the embryos were washed 5X in PBT and then placed on a rotator for 4 hours at room temperature with PBT washes every 30 minutes to eliminate excessive primary antibody. The embryos were then re-blocked and placed in an appropriate dilution of horseradish peroxidase conjugated secondary antibody for 2 hours at room temperature on an orbital rotator. The embryos were then washed 5X with PBT and placed on a rotator overnight at 4°C to eliminate excessive secondary antibody.

The following morning, the embryos were incubated in 0.33mg/ml of diaminobenzidine (DAB) in PBT for 2 minutes and then the reaction of the horseradish peroxidase was initiated by the addition of $3\mu l$ of 0.03% hydrogen peroxide. Once the desired signal was obtained, the reaction was stopped with PBT and the embryos were dehydrated down an ethanol gradient and stored in methyl salicylate.

2.18 Antibody Staining of Larval CNS and Imaginal Discs:

The nervous system and imaginal discs of 3rd instar larvae were dissected in cold PBS and fixed in 4% paraformaldehyde in PBS overnight at 4°C on a rotator. The

following morning the fixative was removed and the larval tissues were washed with PBT multiple times. The larval tissues were placed on a rotator in PBT for 2 hours at room temperature, changing the PBT every half an hour. The larval tissues were blocked with a 1:20 dilution of normal goat serum for 40 minutes to remove non-specific binding. The primary antibody was then added at the appropriate dilution and the larval tissues were placed overnight at 4°C on a shaker.

The following morning the primary antibody was removed and the tissues were washed 5X with PBT. The larval tissues were further washed in PBT on a rotator for 2 hours and then blocked again. The horseradish peroxidase secondary antibody was added at the appropriate dilution and incubated for 2 hours at room temperature. The larval tissues were washed 5X with PBT and placed on a rotator overnight at 4°C.

The next day, the larval tissues were incubated in 0.33mg/ml of DAB in PBT as described in 2.15. After all traces of DAB were removed, the larval tissues were placed and stored in 70% glycerol/PBS.

Chapter 3

Results

A bioinformatics screen for a *Drosophila* ortholog of mouse PALS2 revealed two candidate genes, CG9326 and CG13219, now known as Skiff. Both CG9326 and Skiff, containing an L27 domain, a PDZ domain, an SH3 domain, and a GUK domain, showed 39% amino acid identity to PALS2 (MacMullin, 2001). Based on sequence homology, we chose CG9326, which will now be referred to as dPALS2, as the potential *Drosophila* ortholog of PALS2.

3.1 Identification and sequence analysis of *Drosophila* PALS2:

Genomic Structure of *dPALS2*:

The genomic sequence of *dPALS2* is 2224 base pairs in length and contains 6 introns and 7 exons. The introns are scattered throughout the gene and range in size from 55 base pairs to 449 base pairs. *dPALS2* is located on the left arm of the 2nd chromosome at 38E10. Appendix 5 illustrates the genomic structure and the resultant protein product of dPALS2.

dPALS2 transcripts and protein products:

There are three predicted transcripts of *Drosophila PALS2*, designated by Flybase (Drysdale *et al.*, 2005) as b, c, and d (Appendix 6), that translate into different protein products (Appendix 7). Transcripts b and d differ from transcript c in the 5' UTR due to the presence of variable splice sites. The 5' UTR of transcript c is shorter than transcript

b and d. However the cDNA and amino acid sequence of transcript b and c are identical

and differ from transcript d by an absence of twenty-one amino acids. Therefore the

cDNA for transcripts b and c is 1407 base pairs in length, corresponding to 469 amino

acids, where the cDNA for transcript d is 1344 base pairs coding for 448 amino acids.

Protein Domains of dPALS2:

The three predicted transcripts of *dPALS2*, all contain a 159 base pair region

within the 5'UTR that includes an untranslated ORF coding for a conserved protein

domain known as an L27 domain. This domain was identified as a conserved binding

site in LIN-2 and LIN-7 for the interaction of these two proteins (Doerks et al., 2000).

This domain appears to not be translated because of stop codons located at –606, -423

and -303 in transcripts b and d and –465, -417 and -330 in transcript c (Appendix 6).

All protein products of dPALS2 contain a PDZ domain, an SH3 domain and a

GUK domain (Appendix 7). The PDZ domain is a protein binding domain that targets

the C-terminal ends of other proteins (Sheng and Sala, 2001). dPALS2 PDZ domain

contains 78 amino acids. The SH3 and GUK domains are enzymatically inactive but may

function as protein binding domains (Pawson, 1994). The SH3 and GUK domains of

Drosophila dPALS2 are 66 and 107 amino acids respectively.

Comparison of Amino Acid Sequences between PALS2 Orthologs

The entire dPALS2 protein and its individual protein binding domains were

compared with other PALS2 proteins from species: *M. musculus, H. sapiens, X. laevis, A.*

mellifera, and *A. gambiae* (Figure 4). The *Drosophila* PALS2 shows the highest identity

with *A.* gambiae (mosquito) where the overall amino acid identity is 54 % and individual

domain identities are 48% for the L27 domain, 74% for the PDZ domain, 73% for the

SH3 domain and 72% for the GUK domain. The mosquito protein only contains a single

L27 domain where all other PALS2 proteins contain two L27 domains: L27N and L27C

(Kamberov *et al.*, 2000). It appears the dPALS2 L27 domain, which is located in the 5'

UTR, has a higher degree of identity with the L27C domain than the L27N domain in the

other PALS2 orthologs. The other three protein binding domains are highly conserved in

all proteins. Compared to the *Drosophila* protein, the PDZ domain shows an identity

ranging from 46%-74%, the SH3 domain shows an identity ranging from 46%-73% and

the GUK domain shows an identity ranging from 42%-72%. dPALS2 appears to show a

higher degree of homology to *A. mellifera* (honey bee) and *A. gambiae* (mosquito) than to

the mammalian orthologs and shows the least homology to *X. laevis* (Xenopus).

Appendix 8 illustrates the amino acid alignment of PALS2 proteins.

3.2 5'RACE and Northern Blot analysis were unable to confirm the number of *dPALS2* transcripts:

5'RACE identified only one *dPALS2* transcript – transcript c:

5' RACE (Rapid amplification of cDNA ends) was performed by Katie Moyer to clarify

the number of transcripts, the length of the 5'UTR and the location of the coding

sequence start site. She identified only one *dPALS2* transcript (Figure 5) and it matched

transcript c. The length and sequence of the 5'UTR that contains the L27 domain and the

coding sequence start site of transcript c were also confirmed (Figure 6). However

5'RACE was unable to detect *dPALS2* transcripts b and d.

Figure 4: *Drosophila* **PALS2 and its orthologs illustrating identity at that amino acid level.** The precent identity for the entire protein as well as indicated protein domains is shown. Each domain illustrated indicates the identity for species *Homo sapien*, *Mus musculus* (α and β isoforms), *Apis mellifera*, *Anopheles gambiae*, and *Xenopus laevis* as compared to domains in *Drosophila*. The L27 domain of *Drosophila* which is located in the 5' UTR was compared to all L27 domains present in other species. L27 = LIN-2/LIN-7 binding domain, PDZ = PDZ domain, SH3 = SH3 domain, GuK = guanylate kinase domain.

| L27 | PDZ | SH3 | GUK | *D. mellanogaster*
469aa |

36% | 9% | 26% | 46% | 50% | 47% | *M. musculus*
Pals2α 539aa

36% | 9% | 26% | 46% | 50% | 47% | *M. musculus*
Pals2β 553aa

36% | 7% | 26% | 46% | 50% | 47% | *H. sapiens*
VAM1 540aa

36% | 11% | 21% | 47% | 46% | 42% | *X. laevis*
538aa

47% | 10% | 28% | 52% | 68% | 53% | *A. mellifera*
570aa

54% | 48% | 74% | 73% | 72% | *A. gambiae*
570aa
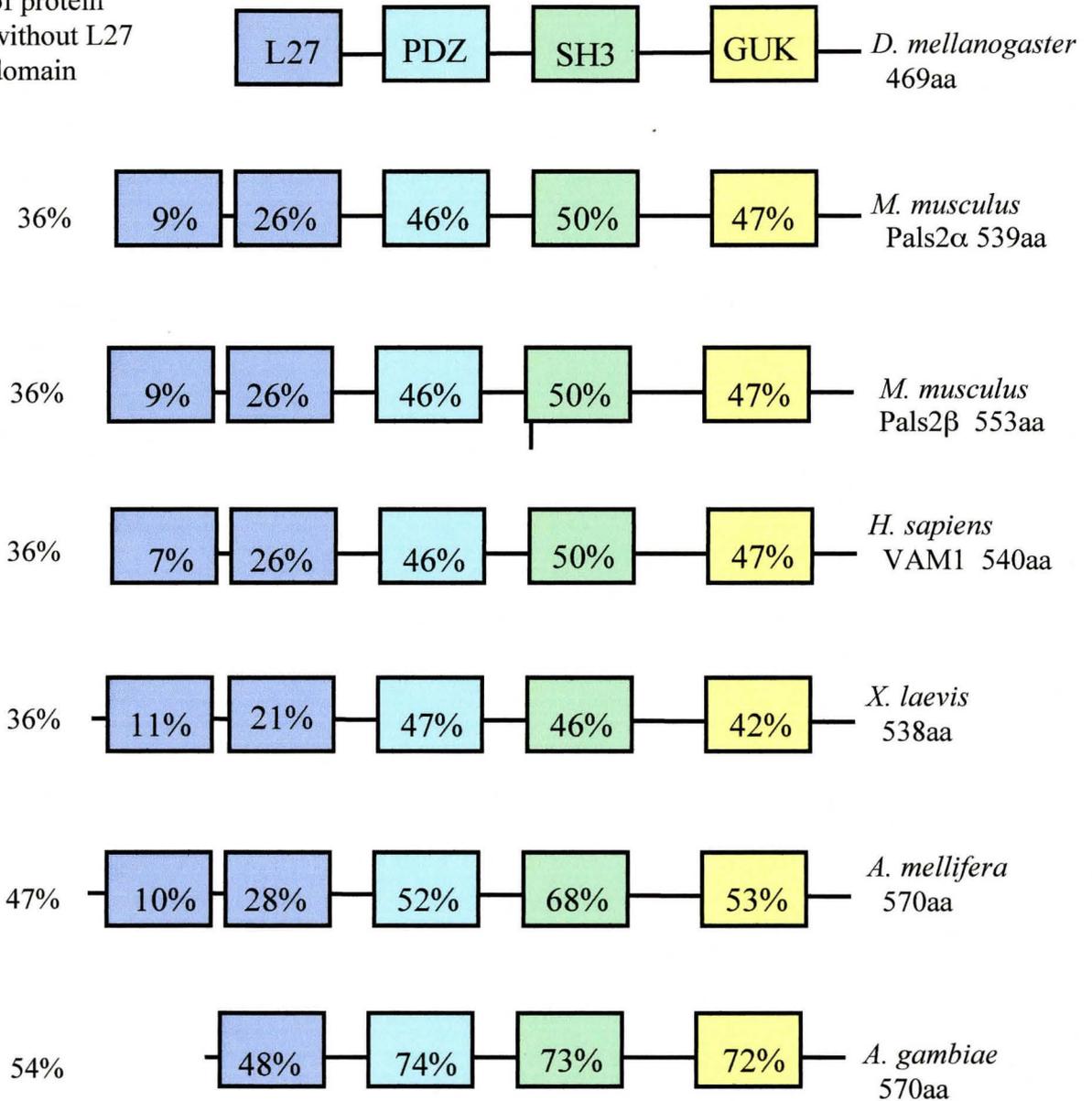
Northern Blot analysis failed to confirm the number of dPALS2 transcripts:

Since three dPALS2 transcripts, b, c, and d, have been predicted, RNA was

extracted, purified and subsequently utilized in northern blot analysis to verify the 5'

RACE results. RACE tends to select the most efficient transcript (shortest or most

abundant transcript), so northern blot analysis will confirm the total number of transcripts

and their approximate size. As described in the Materials and Methods section, a DIG

labeled probe designed to target *dPALS2* transcripts was generated from the *pBS SKII*

dPALS2 construct (Appendix 1). The northern blot was attempted and repeated

numerous times with inconclusive results.

3.3 *In* situ Hybridization illustrates the expression patterns of *dPALS2* transcripts:

dPALS2 RNA expression is first detected in embryonic stage 10:

Once *dPALS2* was identified, the expression patterns of mRNA transcripts were

pursued in various stages of *Drosophila* development. During *in* situ

hybridization, embryos were collected, fixed and probed as described in the Materials and

Methods section to detect RNA expression patterns. *Robo* sense and anti-sense DIG

labeled probes were used as positive controls to control for errors in procedure.

Embryonic *in situ* hybridization was performed on all developmental stages and

detected early *dPALS2* RNA expression in stage 10-11 embryos. RNA expression

appeared in the epithelia as a transverse striped pattern with strong RNA distribution in

the hindgut of the embryo. Figure 7 displays the sense control probe in panels A-C,

illustrating stage 10-11 embryos sagittal, dorsal and ventral respectively. Comparing

Figure 5: 5'RACE identified only one *dPALS2* transcript. Total RNA isolated from embryo and adult stages of *Drosophila* development was used in RACE kit from Ambion, where 5' and 3' RACE was performed to determine the number of *dPALS2* transcripts. (a) Reverse transcription and PCR amplification of the 5' end of *dPALS2* from adults and embryos produced a band of ~ 1.2 kb. Sequencing confirmed the 1.2 kb band to be transcript c, where the L27 domain is encoded in the 5' UTR. The other RT-PCR product located above the 1.2 kb band on the gel at ~1.6 kb was also sequenced but produced inconclusive results. (b) Reverse transcription and PCR amplification of the 3' end of *dPALS2* from adult and embryos produced a single band, which was expected since all three predicted transcripts have similar 3' ends. Asterisks indicate background due to non-specific binding or primers during PCR reaction. (Data from K. Moyer)
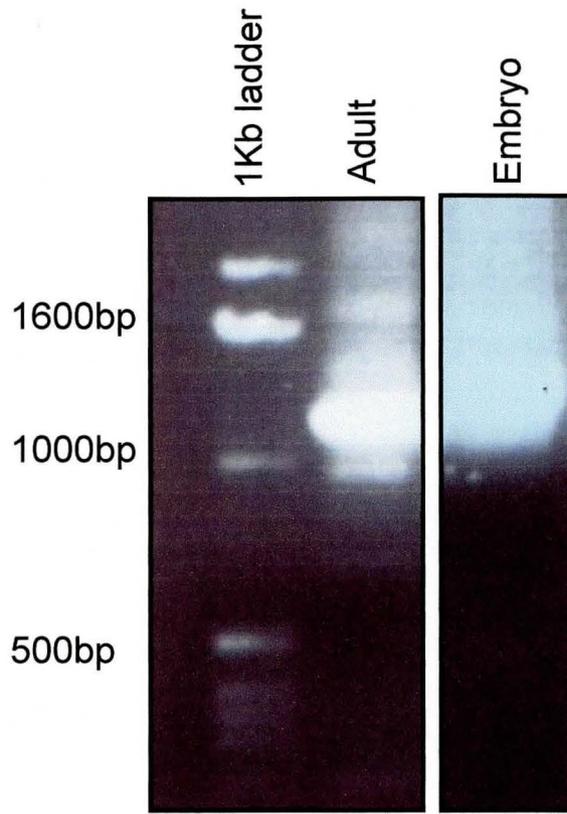
Figure 6: The nucleotide and amino acid sequence of *dPALS2*. *dPALS2* (transcript c)
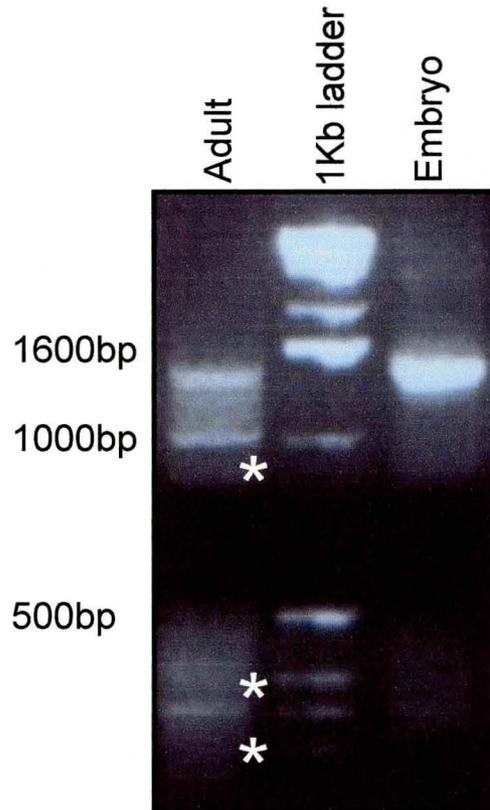
is 1407 base pairs in length producing 469 amino acids. Protein domains are color coded.

Purple = PDZ domain, Blue= SH3 and Red = GUK domain.

```
atgccagtggagaccatcaagatggtgggtctgcgccgagatcccagcaagccgctgggc
 M   P   V   E   T   I   K   M   V   G   L   R   R   D   P   S   K   P   L   G
ctgaccgtcgaactggacgaattcaagcagctggtcgtggccaggattctggcggccggg
 L   T   V   E   L   D   E   F   K   Q   L   V   V   A   R   I   L   A   G   G
gtgatcgacaaacagagcatgctgcacgttggcgatgtcatcctagaggtgaacggtacg
 V   I   D   K   Q   S   M   L   H   V   G   D   V   I   L   E   V   N   G   T
cccgttcgcactcccgatgagctgcaggtggaggtgtcgcgggccaaggagaatctcacc
 P   V   R   T   P   D   E   L   Q   V   E   V   S   R   A   K   E   N   L   T
ctcaagatcgggccgaacgtggacgaggagatcaagagcggtcgctatactgtgagtggg
 L   K   I   G   P   N   V   D   E   E   I   K   S   G   R  ·Y   T   V   S   G
ggtcaggtaaaacagaatggcatcgcgagtctcgagacgggcaagaaactgacgtgctac
 G   Q   V   K   Q   N   G   I   A   S   L   E   T   G   K   K   L   T   C   Y
atgcgtgccctttttcacatacaatccatccgaagattccttgctgccatgcaggatatt
 M   R   A   L   F   T   Y   N   P   S   E   D   S   L   L   P   C   R   D   I
ggattgcccttcaagtcgggcgacattttgcagatcatcaacgtaaaggatcccaactgg
 G   L   P   F   K   S   G   D   I   L   Q   I   I   N   V   K   D   P   N   W
tggcaggccaagaacattactgccgaatctgacaaaattggactcataccatcccaagag
 W   Q   A   K   N   I   T   A   E   S   D   K   I   G   L   I   P   S   Q   E
ctggaggagcggcgcaaagccttcgtggcccccgaggcggactacgttcacaagattggc
 L   E   E   R   R   K   A   F   V   A   P   E   A   D   Y   V   H   K   I   G
atttgcggaacaagaatctcgaagcgaaagcgcaagaccatgtaccgatctgtggccaat
 I   C   G   T   R   I   S   K   R   K   R   K   T   M   Y   R   S   V   A   N
tgcgagttcgacaaggcggagctgctgctctatgaggaggtcacgcggatgccgccgttc
 C   E   F   D   K   A   E   L   L   L   Y   E   E   V   T   R   M   P   P   F
cgcaggaaaaccctggttctcatcggtgtttccggggtgggaagacgtacgctcaagaat
 R   R   K   T   L   V   L   I   G   V   S   G   V   G   R   R   T   L   K   N
cggctgatcaacagcgatgtggacaagttcggagcggtcattccacataccagtcgcccc
 R   L   I   N   S   D   V   D   K   F   G   A   V   I   P   H   T   S   R   P
aagcgcgccttggaggagaacggcagtagctactggttcatggaccgcgaggaaatggaa
 K   R   A   L   E   E   N   G   S   S   Y   W   F   M   D   R   E   E   M   E
gaggccgtgcggaacaacgagttcctggagtacggcgagcacaatggcaatctgtacggc
 E   A   V   R   N   N   E   F   L   E   Y   G   E   H   N   G   N   L   Y   G
acgcatttgcagtccatcaaggatgtgatcaacagtgggcgcatgtgcatcctggattgt
 T   H   L   Q   S   I   K   D   V   I   N   S   G   R   M   C   I   L   D   C
gcaccgaatgccctgaagatcctgcacaacagccaggaactgatgccttttgtcatcttt
 A   P   N   A   L   K   I   L   H   N   S   Q   E   L   M   P   F   V   I   F
gtggcagcgcctggcatggagcagctcaagaccatctatgcggatcgcagggccacgggc
 V   A   A   P   G   M   E   Q   L   K   T   I   Y   A   D   R   R   A   T   G
tccaaccggaatttatcttttgatcgccagagttccataagattcagctccagacgcgcc
 S   N   R   N   L   S   F   D   R   Q   S   S   I   R   F   S   S   R   R   A
cgtacgctcgagtccctagcgtcgttgtatgaggacgacgaccttgtcgccaccgtcgag
 R   T   L   E   S   L   A   S   L   Y   E   D   D   D   L   V   A   T   V   E
gagagcagcttcgtccagcgcaagtatgagaaatacttcgacatggtcatcgtgaacgag
 E   S   S   F   V   Q   R   K   Y   E   K   Y   F   D   M   V   I   V   N   E
gacttcgatgagacgttccgccaggtggtggaaacgctggatcagatgagccacgaggag
 D   F   D   E   T   F   R   Q   V   V   E   T   L   D   Q   M   S   H   E   E
cagtgggtgcctgtcaactggatctac
 Q   W   V   P   V   N   W   I   Y
```

panels A-C to panels D-F, illustrating stage 10-11 embryos and panels G-I, illustrating

stage 12-13 embryos, *dPALS2* RNA is expressed in a distinct striped pattern. This

expression pattern is similar in stage 14 embryos. Mammalian PALS2 has been shown to

colocalize with E-cadherin below tight junctions (TJ) and associate directly with nectin

like molecule 2 (necl-2) at basolateral extra junctional regions in epithelial cells. Both E-

cadherin and necl-2 are cell adhesion molecules involved in the structural and functional

organization of epithelial cells (Roh *et al.*, 2002; Shingai *et al.*, 2003). This expression

pattern of *dPALS2* within the epithelia suggests a similar function to that of the

mammalian counterpart.

The RNA expression pattern of *dPALS2* in later staged embryos is similar to the

expression pattern seen in stages 10-14. Figure 8 displays *in situ* hybridization of stage

15-16 embryos, where panels A-C illustrate the sense probe controls, sagittal, dorsal and

ventral respectively, and panels D-G display the *dPALS2* RNA expression pattern, in the

epithelia as a transverse striped pattern. Background staining is visible in the trachea and

the midline (indicated by arrows) which is standard in embryonic *in situ* hybridizations,

when the probe concentration is too high.

dPALS2 RNA expression not detected in Larval Central Nervous System:

Since *dPALS2* RNA was detected in late staged embryos and since there are some

polarity proteins such as Scribble and Discs large, that play an organizing role in both

embryonic and larval stages, an *in situ* hybridization of larval tissues was completed to

show *dPALS2* RNA expression. 3rd instar larvae were dissected, fixed and probed as

described in the Materials and Methods sections. As seen in Figure 9, *dPALS2* RNA was

Figure 7: *In situ* Hybridization illustrates RNA transcript distribution in Embryonic tissues, stages 10-14. Various stages of embryos were collected, fixed and probed with DIG labeled sense and anti-sense probes to detect the distribution of *dPALS2* RNA in embryonic development. Panels A-C illustrate the sense control probe in embryonic stage 10-11. D-F exhibit embryonic stage 10-11, whereas G-I exhibit embryonic stage 12-13, probed with an anti-sense DIG labeled probe. *dPALS2* transcripts are first detectable at embryonic stage 10 in the epithelia as a transverse striped pattern with strong RNA distribution in the hindgut of the embryo. This RNA distribution is consistent in stage 14.

Sense

Anti-sense
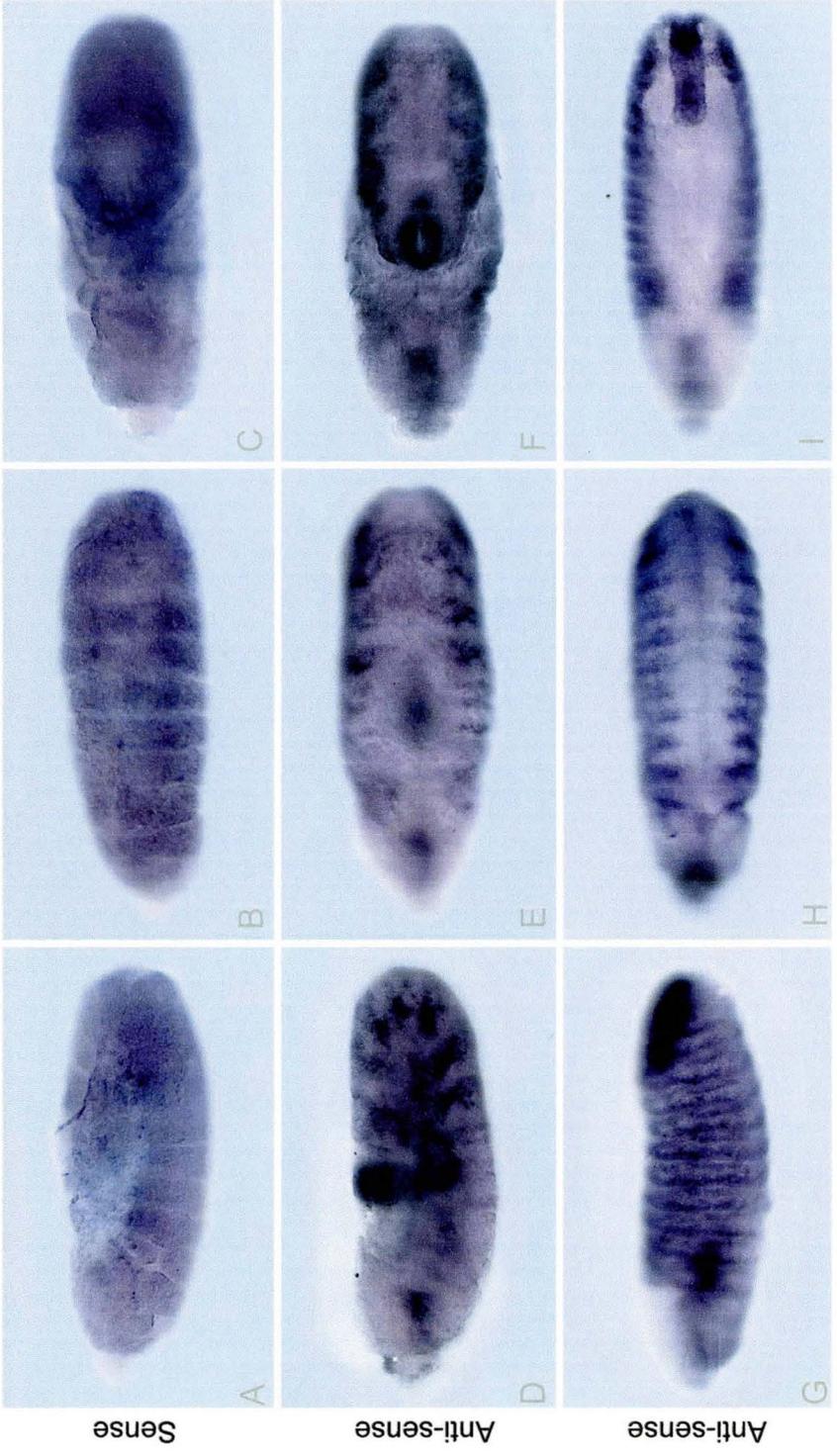
Anti-sense

Figure 8: *In situ* **Hybridization illustrates** *dPALS2* **RNA transcript distribution in Embryonic tissues, stages 15-17.** Various stages of embryos were collected, fixed and probed with DIG labeled sense and anti-sense probes to detect the distribution of *dPALS2* RNA in embryonic development. Panels A-C illustrate the sense control probe in embryonic stage 15-16. D-G depicts embryonic stage 15-16 probed with an anti-sense DIG labeled probe. *dPALS2* RNA expression in late staged embryos is visible in the epithelia as a transverse striped pattern, which is similar to stages 10-14. Background staining is indicated by arrows and is visible in the trachea and the midline.

Sense

Anti-sense
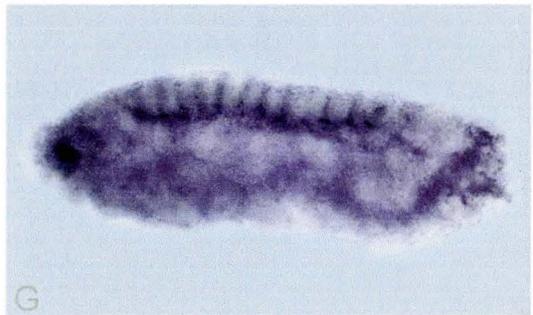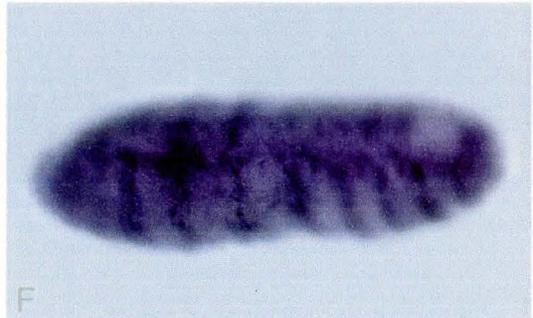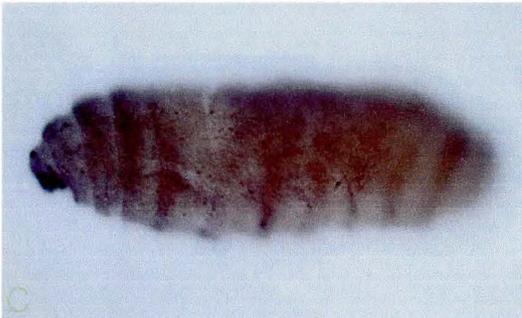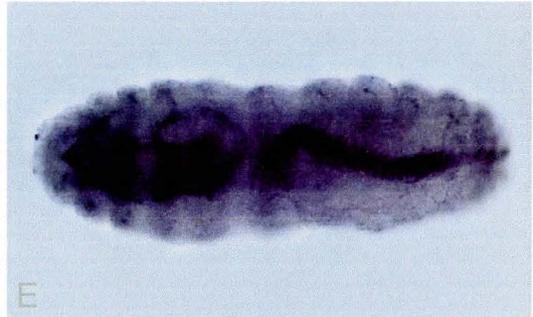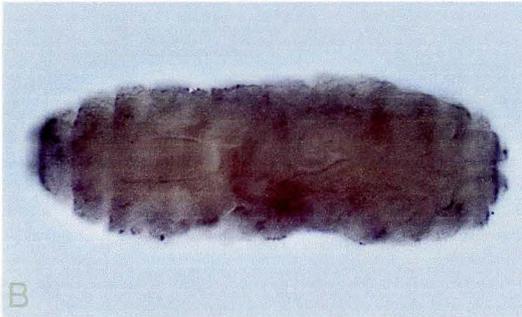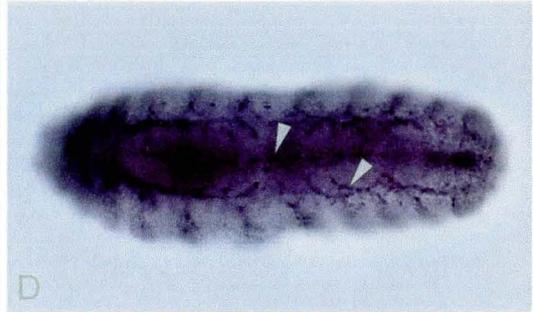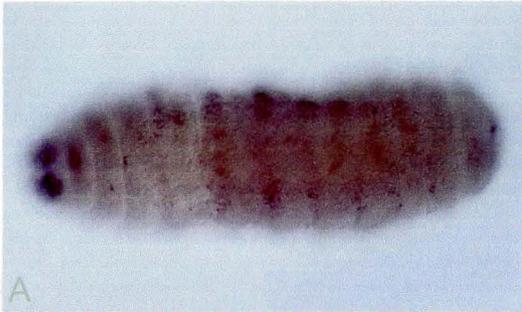
not detectable in the central nervous system (CNS) of larvae. Thus *dPALS2* was either

expressed at such low levels, that *in situ* hybridization was unable to detect any RNA

expression or *dPALS2* is not expressed in the larval CNS.

*d*PALS2 RNA expression not detected in Imaginal Discs:

When the central nervous sytem was dissected from 3rd instar larvae, some

imaginal discs were also dissected and screened for *dPALS2* RNA expression by *in situ*

hybridization since *dPALS2* was visualized in the epithelia during embryonic

development. Imaginal discs are primordial epithelial tissues set aside during

embryogenesis to proliferate during larval development to develop into adult structures

such as the legs, wings, eyes and antennae (Bilder, 2004). As seen in Figure 9, *dPALS2*

RNA expression was not detectable in the leg, wing or the eye-antennal imaginal discs.

It is possible *dPALS2* is expressed at such low levels in the imaginal discs, that *in situ*

hybridization failed to detect any RNA expression or *dPALS2* is not expressed in the

imaginal discs.

3.4 *In situ* Hybridization shows no expression pattern for a *dPALS2* L27 transcript:

Since the presence of an L27 domain within the *dPALS2* transcript has remained

enigmatic, an L27 *pBS SKII* construct (Appendix 2) was created to generate DIG labeled

probes for the L27 domain in *dPALS2* transcripts. As described in the Materials and

Methods, sense and antisense probes were generated from the L27 construct and used in

embryonic and larval *in situ* hybridization. Both developmental stages were examined

Figure 9: *In situ* **Hybridization demonstrates no distinct pattern of** *dPALS2* **RNA in larval tissues.** 3rd instar larvae were dissected, fixed and probed with DIG labeled sense and anti-sense probes to detect the distribution of *dPALS2* RNA in larval development. Panels A-D show the sense control on larval tissues. Panels E-H show the anti-sense DIG labeled tissues. CNS = Central Nervous System, L = Leg imaginal discs, W = Wing imaginal discs, E-A = Eye-antennal imaginal discs. There was no distinct pattern of *dPALS2* RNA in larval tissues.

CNS   L   W   E-A
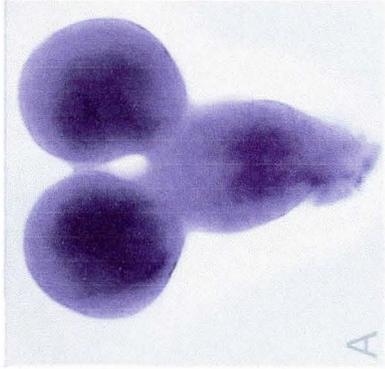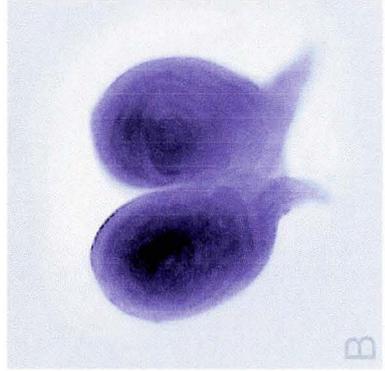
Sense

Anti-sense
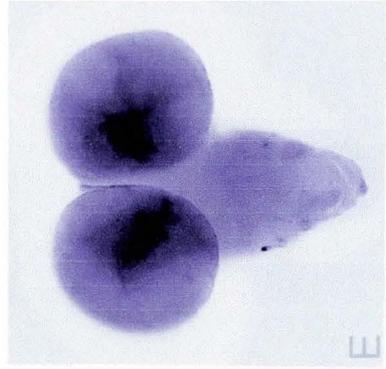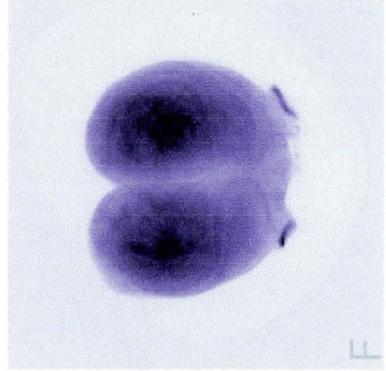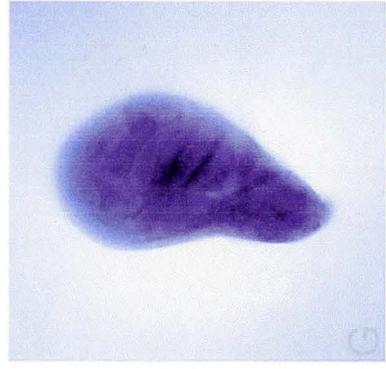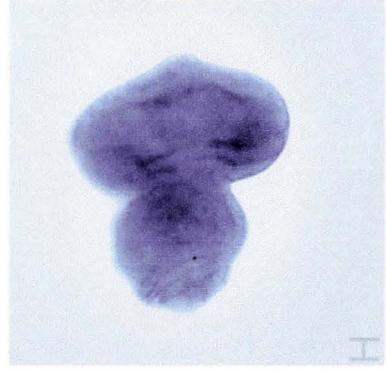
since the number of *dPALS2* transcripts and the possibility of L27 domain expression had not been fully determined.   Both embryonic and larval *in situ* hybridization were repeated many times with no detectable expression pattern. Robo sense and anti-sense probes were tested simultaneously as positive and negative controls. Figure 10 displays stage 13 embryos with no distinct expression pattern compared with the sense control. Various stages of embryos were analyzed, however it appeared a *dPALS2* L27 transcript was undetectable in embryonic development. Figure 11 displays the CNS, the leg, wing and eye-antennal imaginal discs of dissected 3rd instar larvae with no distinct expression pattern visible when compared with the sense control. It appears a *dPALS2* L27 transcript is undetectable in larval development. Thus *dPALS2* L27 domain is expressed at extremely low levels, it is not expressed at all or it is only expressed in stages or tissues that were not examined.

The lack of *dPALS2* L27 transcript expression seen in both the embryonic and larval *in situ* hybridization conflicts with the results of the 5'RACE, which showed the L27 domain was present in the 5' UTR of the mRNA.

3.5 The Protein Distribution and Expression Pattern of dPALS2:

Analysis by Western Blot supports dPALS2 protein expression in embryos and adults but not in larvae:

dPALS2 antibody is required to determine the size and number of dPALS2 protein products and where they are expressed throughout *Drosophila* development.

Figure 10: *In situ* **Hybridization does not detect** *dPALS2* **L27 domain in embryonic**

tissues. Various stages of embryos were collected, fixed and probed with DIG labeled

sense and anti-sense probes to detect the distribution of *dPALS2* L27 domain in

embryonic development. Panel A shows the sense control in stage 13 embryos while

panel B shows stage 13 embryos probed with anti-sense DIG labeled probe. There was

no distinct pattern for *dPALS2* L27 domain in any stage of embryonic development.

Sense
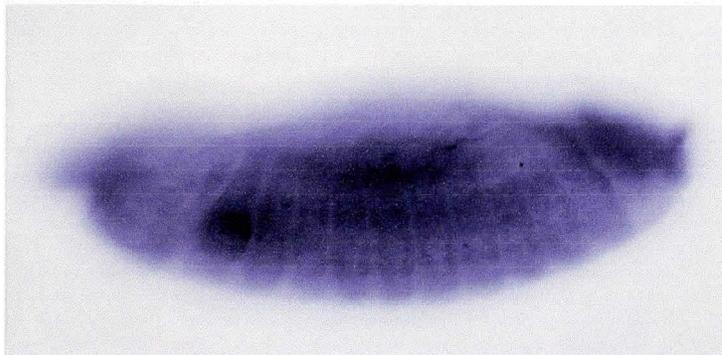
Anti-sense
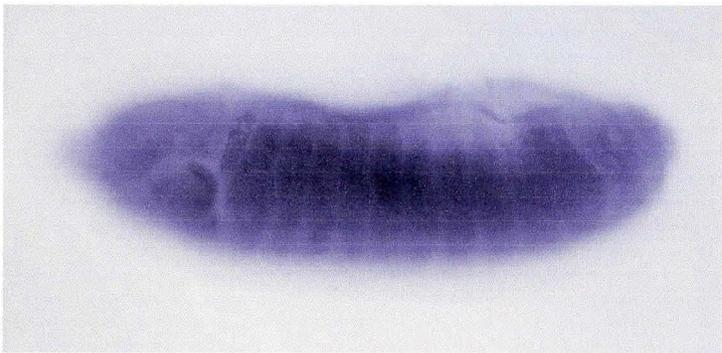
Figure 11: *In situ* **Hybridization does not detect** *dPALS2* **L27 domain in larval**

tissues. 3rd instar larvae were dissected, fixed and probed with DIG labeled sense and

anti-sense probes to detect the distribution of *dPALS2* L27 domain in larval development.

Panels A-D show the sense control on larval tissues. Panels E-H show the anti-sense

DIG labeled tissues. CNS = Central Nervous System, L = Leg imaginal discs, W = Wing

imaginal discs, E-A = Eye-antennal imaginal discs. There was no distinct pattern in

larval tissues for *dPALS2* L27 domain.

|	CNS	L	W	E-A
Sense	A	B	C	D
Anti-sense	E	F	G	H
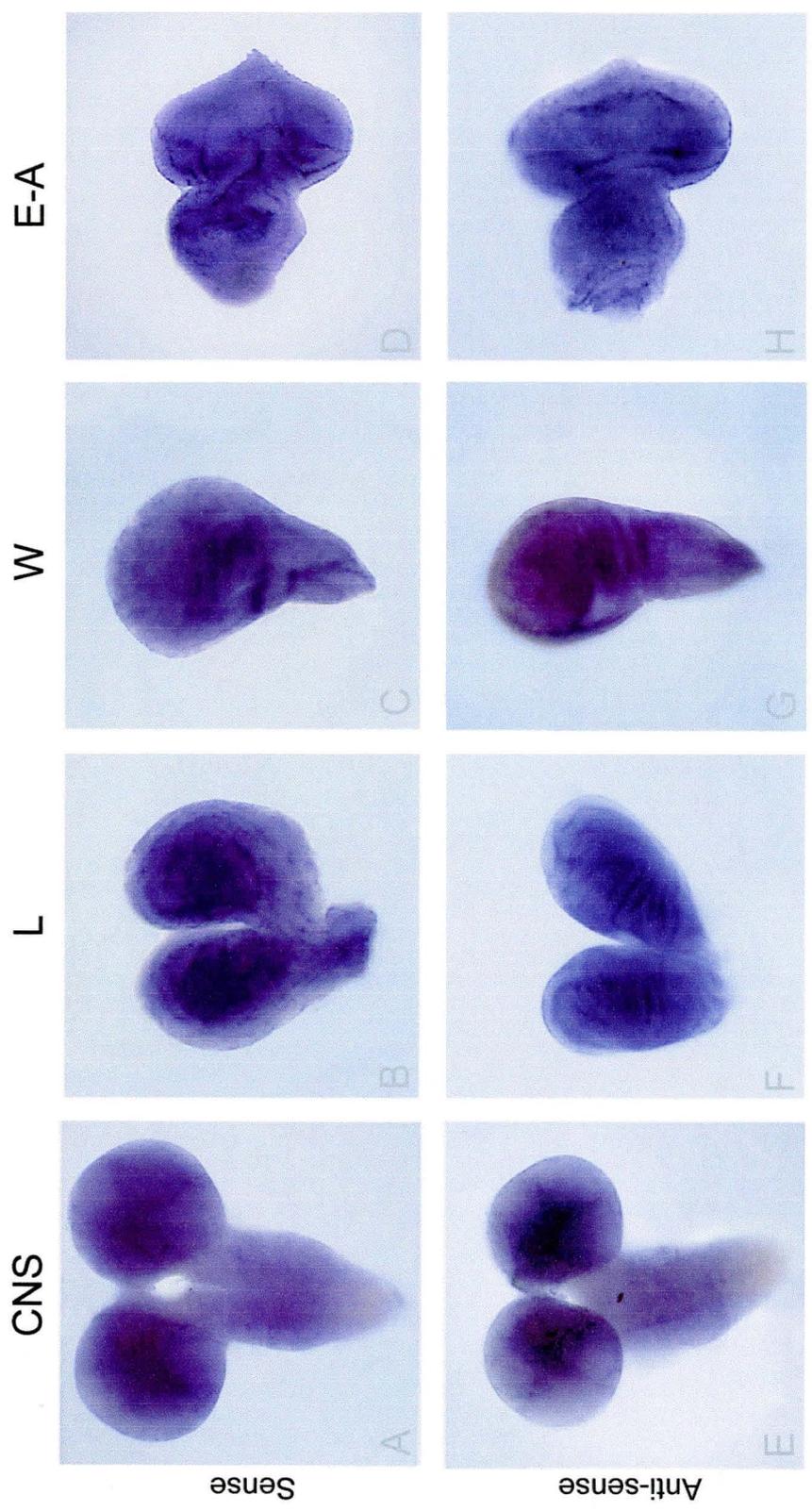
dPALS2 antibody was generated in rats with a His fusion protein (Appendix 4) and purified with CN-Br affinity purification (see Materials and Methods).

To determine the size and number of dPALS2 protein products, whole protein lysates were extracted from embryonic, larval and adult tissues of *Drosophila* for analysis by Western Blot. Purified dPALS2 antibody of varying concentrations was utilized to determine the optimum concentration required for detection. The autoradiograph was developed using the ECL detection kit. As seen in Figure 12, dPALS2 protein was detected in embryonic and adult tissues. dPALS2 appears on the blot as a single band indicating an approximate size of 54kDa which was predicted based on amino acid content (52kDa). There appears to be proportionately more dPALS2 protein in embryonic tissues as compared to adult tissues. A detectable level of dPALS2 protein was unable to be seen in larval tissues, even after prolonged exposure of the autoradiograph to the blot. The antigen used in antibody generation was used as a positive control to show dPALS2 antibody specificity. A second western blot using actin antibody was performed to show protein extracts were not degraded. Actin, a cytoskeletal protein ubiquitously expressed throughout all stages of development, was detected at 42kDa in adults, larvae and embryos. The lack of dPALS2 detected in larval tissues is not due to protein degradation but possibly because dPALS2 protein is not expressed in larval tissues or it is expressed at such low levels, that western blot analysis failed to detect protein expression.

Figure 12: Western Blot Analysis of dPALS2 protein distribution in various stages of *Drosophila* development. Protein extracted from *Drosophila* embryos, larvae and adults were loaded onto a SDS-PAGE gel, transferred to a membrane, incubated with purified dPALS2 antibody (dilution factor of 1:1000) and developed the autoradiograph using the ECL detection kit. (**A**) dPALS2 protein expression was detected in adult and embryo showing a size of approximately 54kDa which was predicted based on amino acid content (approximately 52kDa). dPALS2 was not detected in larval protein extracts. The antigen used in antibody generation was used as a positive control to show dPALS2 antibody specificity. (**B**) A second positive control was performed using Actin antibody to show protein extracts were not degraded. Actin, a cytoskeletal protein ubiquitously expressed throughout all developmental stages, was detected at 42kDa in adults, larvae and embryos. The undetectable dPALS2 protein in larval tissues is not due to protein degradation.

A

dPALS2

134 kDa

83 kDa

41 kDa

17 kDa

+ve Control   Adult   Larvae   Embryo

B

actin

134 kDa
83 kDa

41 kDa

17 kDa

Adult   Embryo   Larvae
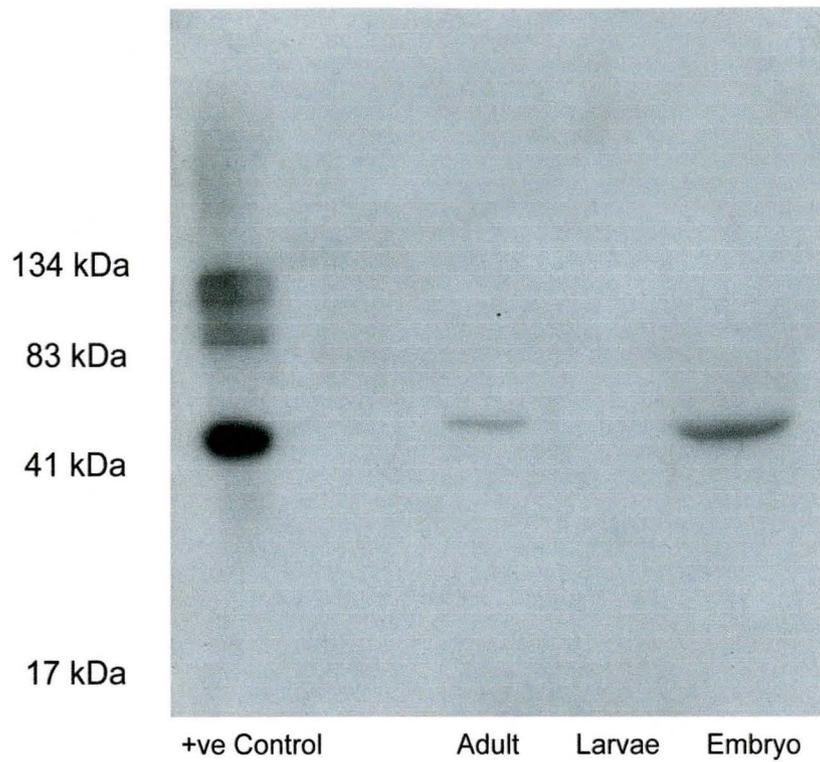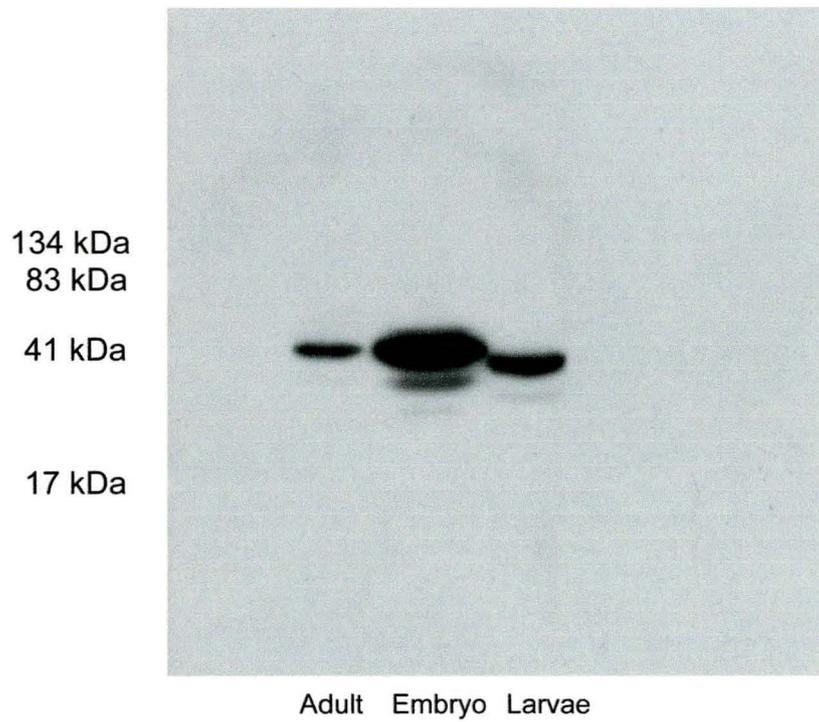
The absence of dPALS2 protein in larval tissue lysates supports the *in situ* data, which also failed to detect *dPALS2* transcripts in larvae. However, western blot analysis does not illustrate the spatial location of dPALS2 protein throughout development and must be determined by immunohistochemistry.

Immunohistochemistry illustrates the expression patterns of dPALS2 protein in embryonic development:

In previously mentioned data, *dPALS2* RNA expression initiates in embryonic stage 10 in the epithelia as a transverse striped pattern with strong RNA distribution in the hindgut of the embryo and continued the striped pattern until the end of stage 17. Whole embryos were collected, fixed and stained with purified dPALS2 antibody (see Material and Methods). The stained embryos were mounted on slides and visualized by microscopy to identify dPALS2 protein expression.

Immunohistochemistry was performed on all stages of embryonic development and identified dPALS2 protein expression in the epithelia as a transverse striped pattern. Figure 13 depicts the pre-immunization in panels A-C and the post-immunization in panels D-F of stage 13 embryos, sagittal, dorsal, and ventral, respectively. Comparing panels A-C to panels D-F, there is a distinct striped pattern of dPALS2 in the epithelia. Expression of dPALS2 protein begins at stage 13 and continues through to the end of stage 17, with the pattern remaining unchanged. Figure 14 depicts the pre-immunization in panels A-C and the post-immunization in panels D-F of stage 15-16 embryos, sagittal, dorsal, and ventral, respectively. There is no dPALS2 protein expressed in the hindgut as

Figure 13: *Drosophila* embryos, stages 13-14, express dPALS2 in the epithelia as a striped pattern. Various stages of embryos were collected, fixed and incubated with purified dPALS2 antibody. Both pre-immunization (panels A-C) and post-immunization (panels D-F) are shown for stage 13 embryos. dPALS2 protein is first detected in stage 13 embryos in the epithelial as a distinct striped pattern. Stage 14 embryos show a similar expression pattern.

Figure 14: ***Drosophila*** **embryos, stages 15-16, also express dPALS2 in the epithelia as a striped pattern.** Various stages of embryos were collected, fixed and incubated with purified dPALS2 antibody. Both pre-immunization (panels A-C) and post-immunization (panels D-F) are shown for stage 15-16 embryos. dPALS2 protein also localizes to the epithelial in a distinct striped pattern in late staged embryos. Stage 17 embryos show a similar expression pattern.

Pre-immune

Post-immune
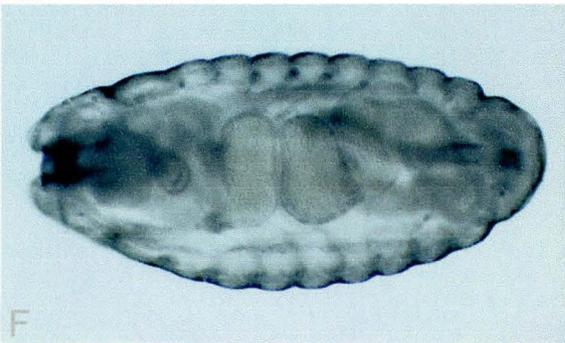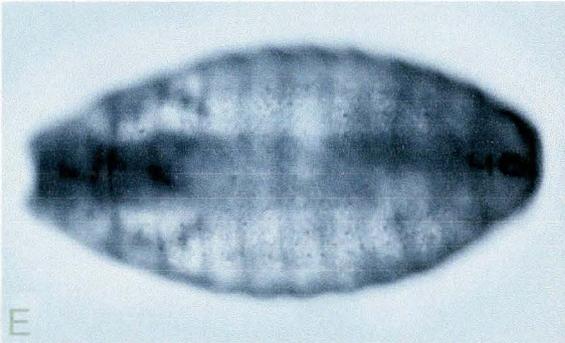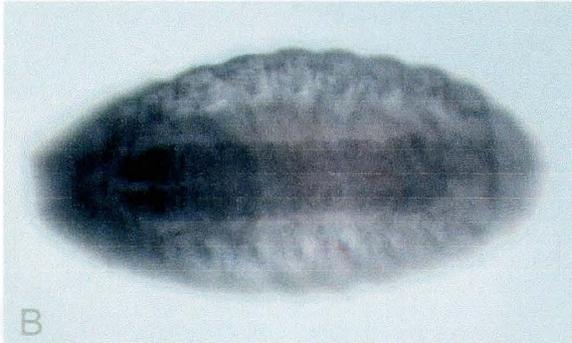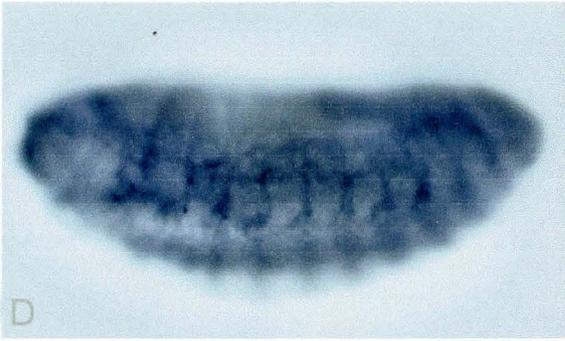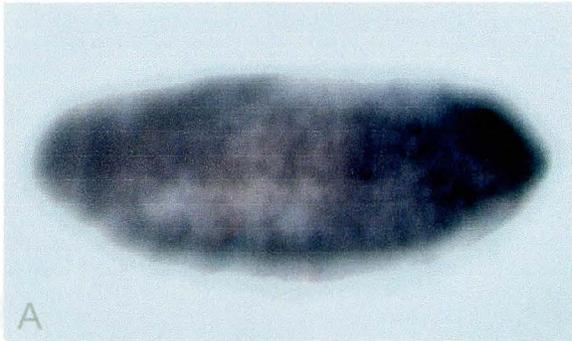
seen in stage 10-14 embryos for *dPALS2* RNA. Thus the temporal location of dPALS2

protein in epithelia supports the notion of similar function to mammalian counterparts.

Larval Tissues shows no Distinct Pattern of dPALS2 Protein Expression:

Since dPALS2 could not be detected in larval tissues by Western Blot analysis,

immunohistochemistry was utilized to confirm larval tissues lacked any detectabled

PALS2 protein expression. 3rd instar larvae were dissected, fixed and stained with

purified dPALS2 antibody. Once stained, the central nervous system and some of the

imaginal discs were mounted onto slides and visualized by microscopy to detect dPALS2

distribution. Figure 15 shows pre-immunization in panels A-D and post-immunization in

panels E-H of the central nervous system and the imaginal discs (leg, wing disc, and eye-

antennal). When compared, panels A-D and panels E-H, dPALS2 protein was

undetectable in the larval tissues tested. It is possible dPALS2 protein is expressed at

such levels in larval tissues, that immunohistochemistry failed to detect any protein

expression.

3.6 dPALS2 is a protein possibly involved in segment polarity:

The location of dPALS2 in a striped pattern in the epithelia is similar to the

expression pattern observed in segment polarity genes *wingless* and *engrailed* suggesting

a possible role in establishing segmental patterning in *Drosophila* embryos. Segment

polarity genes mediate the required cell interactions to establish segment borders and

control the anterior-posterior patterning within each segment during *Drosophila*

development (Bejsovec and Wieschaus, 1993; Bhat, 1999). Thus to investigate dPALS2

Figure 15: dPALS2 protein is not detectable in larval tissues. 3rd instar larvae were dissected, incubated with purified dPALS2 antibody. Panels A-D show pre-immunization controls while Panels E-H show post-immunization results. CNS = Central Nervous System, L = Leg imaginal discs, W = Wing imaginal discs, E-A = Eye-antennal imaginal discs. There was no indication of dPALS2 protein expression in any larval tissue.

CNS L W E-A

Pre-immune

A B C D

Post-immune

E F G H

as a potential segment polarity protein, the expression patterns of both dPALS2 and Engrailed were analyzed together.   Figure 16 illustrates confocal images (magnification 63X) of stage 13 embryos labeled with dPALS2 (green) and Engrailed (red) antibodies. Engrailed is a transcription factor expressed in a striped pattern at the posterior end of each segment, making it an excellent marker for segmental patterning (Bhat, 1999).

dPALS2, a cytoplasmic protein, appeared to be expressed in a gradient striped pattern throughout the entire epithelia, while Engrailed, a nuclear protein, demonstrated a defined striped pattern.   Double labeling was not to illustrate co-localization, which would yield a yellow colour, but to demonstrate dPALS2 expression with respect to a segment polarity protein.   In the merge image, there are epithelia cells that show co-expression of both proteins, where Engrailed (red) is localized in the center (nucleus) of the cell and dPALS2 (green) is localized in the cytoplasm of the cell, both having similar striped expression patterns.   Further studies in immunohistochemistry with other segment polarity proteins such as Shaggy or Fused that are localized in the cytoplasm, will hopefully provide a better understanding of dPALS2 expression with respect to segment polarity proteins.   It would also be beneficial to look at the expression pattern of dPALS2 at an even higher magnification, to visualize a single epithelial cell to predict its exact location within the cell.   Co-localization with apical markers such as members from the Crumbs and Bazooka complexes, basolateral markers such as Discs large or Scribble or cell adhesion markers such as DE-cadherin, will establish dPALS2 expression with relation to other epithelial markers to predict a possible function in the development of *Drosophila.*

Figure 16: dPALS2 is a protein expressed in the epithelia and may be involved in segment polarity. The location of dPALS2 in a striped pattern in the epithelia is similar to the expression pattern observed in segment polarity genes *wingless* and *engrailed* suggesting a possible role in establishing segmental patterning in *Drosophila* embryos. To investigate dPALS2 as a potential segment polarity protein, the expression patterns of both dPALS2 and Engrailed were compared at the embryonic level. Confocal images of stage 13 embryos were taken by Allison MacMullin (magnification 63X) to show dPALS2 (green) and Engrailed (red) alone as well as a merged together to show how their expression patterns relate to each other. dPALS2, a subcellular protein, appeared as a gradient striped pattern throughout the entire epithelia, while Engrailed, a nuclear protein, demonstrated a defined striped pattern. In the merge image, there are epithelia cells that show co-expression of both proteins, where Engrailed is localized in the center (nucleus) of the cell and dPALS2 is localized in the cytoplasm of the cell, both having similar striped expression patterns.

PALS 2
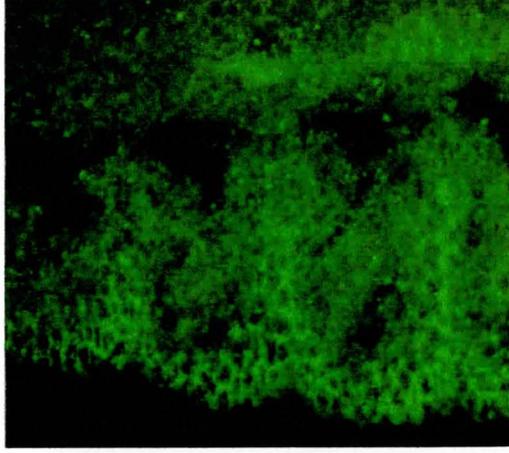
ENG
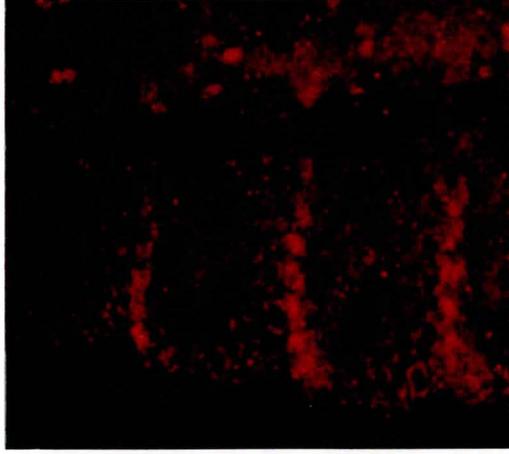
Merge
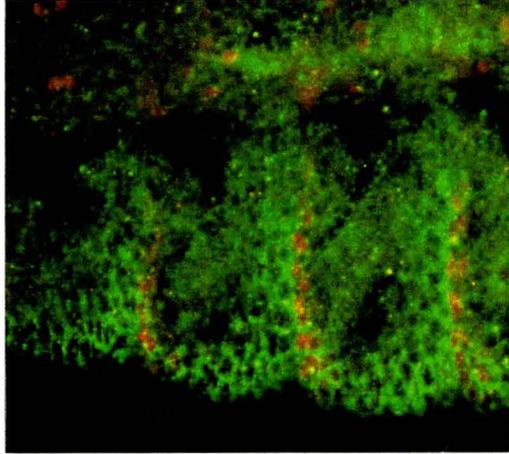
Chapter 4

Discussion

PALS (Proteins Associated with Lin-7) were originally identified in mouse as additional binding partners to VELI, a protein that functions with CASK and MINT1 as a scaffolding complex in neuronal and epithelia cells. Since the VELI/CASK/MINT1 complex is conserved across *C. elegans*, mammals and *Drosophila*, the identification of PALS orthologs could provide insight of their function and determine if the multiple binding partners function together within the same cell, in a distinct sequential order or in a competitive manner. The PALS1 *Drosophila* ortholog has been identified as Stardust, a protein involved in establishing and maintaining epithelial polarity. Since a PALS2 ortholog had not been discovered, the main focus of this thesis was to identify and characterize a *Drosophila* ortholog.

Two genes, CG9326 and CG13219 (MacMullin, 2001), now known as Skiff (Drysdale *et al.*, 2005) were identified as potential *Drosophila* orthologs to mammalian PALS2. Both CG9326 and Skiff showed 39% identity and contained the same domain structure of an L27 domain, a PDZ domain, a SH3 domain and a GUK domain. We selected CG9326, which we have termed dPALS2, as the potential PALS2 ortholog (MacMullin, 2001). The chromosomal location of dPALS2 was determined to be 2L38E10 (Drysdale *et al.*, 2005).

Since very little is known about mammalian PALS2, the objective was to characterize *Drosophila* PALS2 by comparing homology of the protein to known orthologs and by analyzing RNA and protein expression patterns.

4.1 Identification of one *Drosophila PALS2* transcript – transcript c confirmed the sequence of cDNA, where the 5' UTR encodes an L27 domain:

Originally there was only one *dPALS2* transcript predicted by Flybase when the *Drosophila* genome was fully sequenced (Drysdale *et al.*, 2005). This dPALS2 protein encoded an L27 domain, a PDZ domain, an SH3 domain and a GUK domain (Appendix 9). Later, Flybase predicted three *dPALS2* transcripts (denoted b, c, and d) where all three contained a PDZ domain, an SH3 domain and a GUK domain while the L27 domain was no longer present in the predicted protein product because it was encoded in the 5'UTR (Drysdale *et al.*, 2005). Transcript b and c produced the same protein product and transcript d produced a different protein product due to a 21 amino acid absence between the PDZ and SH3 domain.

5' RACE completed by Katie Moyer identified only one *dPALS2* transcript, which matched transcript c, and confirmed the L27 domain was not part of the protein product because it was encoded in the 5'UTR. An RNA *in situ* hybridization was performed, using a DIG labeled probe for the L27 domain of *dPALS2* transcripts, and showed no visible expression in embryonic or larval tissues for *dPALS2* L27 domain. This conflicts with the results seen in the 5' RACE, since the *in situ* hybridization should have detected the L27 domain even if it was located in the 5' UTR. It is possible the L27

probe, which was 180bp in length, was not long enough to detect *dPALS2* L27

transcripts. The recommended length for DIG labeled probes for *in situ* hybridization is

100-300bp, however it is possible a longer probe is required for detecting rare transcripts

(Roche Diagnostics).

DIG labeled Northern Blot analysis was performed to verify the results of 5'

RACE and to confirm the number of *dPALS2* transcripts and their approximate sizes.

Critical parameters of the northern blot were altered and tweaked to optimize protocol

conditions. Total RNA extracts from all stages of *Drosophila* development were utilized

in the first attempts, but eventually mRNA was used to increase the likelihood of

detecting *dPALS2* transcripts. The concentration of DIG labeled RNA probe was

increased to a maximum of 100ng/ml and the temperature for hybridization was reduced,

in order to increase the efficiency of probe binding. Each and every step of the procedure

was carefully analyzed and followed to ensure experimental error was not the reason for

failure. However after multiple attempts the northern blot produced inconclusive results.

It is possible *dPALS2* transcripts are rare and DIG labeling is not sensitive enough.

Perhaps a radioactive labeled northern blot should be attempted.

This poses a puzzling question of why a recognized protein binding domain, that

is functionally present in all orthologs of dPALS2 would be encoded in the DNA but not

transcribed into protein. The L27 domain was initially shown to be necessary for binding

Lin-7 to PALS (proteins associated with Lin-7) proteins, allowing their other binding

domains (PDZ, SH3, GUK) to associate with other proteins to build scaffolding

complexes. Many proteins with L27 domains exhibit high diversity in terms of location,

copy number and primary sequence. However the primary role of L27 domains are to

form and maintain multiple protein complexes (Li *et al.*, 2004; Feng *et al.*, 2004).

Mammalian PALS2, Mammalian PALS1 and its *Drosophila* ortholog Stardust all contain

L27 domains with conclusive data showing their association with LIN-7/dVELI via their

L27 domains (Kamberov *et al.*, 2000). Then why would the *Drosophila* ortholog of

mammalian PALS2, dPALS2, not contain an L27 domain? Especially considering a

recent mass Yeast two-hybrid screen identified a direct interaction between dPALS2 and

dVELI (CG7662). A group of researchers from CuraGen Corporation, Wayne State

University and Yale University used all known or predicted transcripts of *Drosophila*

melanogaster (10 623 in total) against DNA libraries and produced a draft map of 7048

proteins and 20 405 interactions (Giot *et al.*, 2003). dPALS2 was also shown to interact

with putative genes, CG9097 (bab1), CG8288 (mRpL3), CG5273 and CG12019 (Cdc37),

however the interactions identified in this map may not be entirely accurate (Appendix

10). For example, dVELI is known to interact with Stardust and Camguk, the *Drosophila*

ortholog of CASK, and yet the map does not identify them as potential interacting

proteins. dVELI and bab1 (bric-a-brac1), a transcription factor involved in proximal-

distal patterning (Godt, 1993), both show high to moderate confidence of protein

interaction with dPALS2, while the other putative genes show moderate to low

interaction confidence (Giot *et al.*, 2003). Unfortunately it can not be confirmed that

dPALS2 is interacting with dVELI via L27 dimerization because the data does not

identify the binding domains involved in these interactions.

The question that still remains then if dPALS2 does not contain an L27 domain how is it interacting with dVELI? It is possible that over time, evolution has eliminated the need for a functional L27 domain in *Drosophila* PALS2. Upon analysis of the predicted *dPALS2* transcripts, multiple start sites (AUG) for transcription were identified in the 5' UTR with stop codon sequences immediately following. The actual start site for *dPALS2* transcription is not until after the L27 domain and before the PDZ domain. Therefore it is possible over time, evolution has produced dPALS2 protein without a functional L27 domain, while proteins like Stardust and Skiff still require an L27 domain for proper function. Perhaps a dPALS2 protein with an L27 domain exists and we have not identified it yet, or dPALS2 and dVELI are interacting by way of another protein or protein binding domain (PDZ homodimerization), or dPALS2 does not interact with dVELI at all. The only conclusion we can confidently state is that dPALS2 definitely produces at least one transcript – transcript c, where the 5' UTR encodes an L27 domain. However further research must be continued to verify the number of dPALS2 transcripts and to confirm whether or not an L27 domain exists in the functional protein product.

4.2 PALS2 is conserved in *Drosophila melanogaster*:

When homology was compared across a variety of species, dPALS2 showed a 36% amino acid identity with mammalian PALS2 (*M. musculus* and *H. sapiens*). dPALS2 was also compared to the recently annotated genomes of *A. mellifera* (honeybee) and *A. gambiae* (mosquito) which showed a 47% and 54% amino acid identity, respectively. The overall domain structure of dPALS2 is the same in the other

known PALS2 orthologs with a single PDZ domain, SH3 domain and GUK domain.

These domains are fairly well conserved from *Drosophila* to the PALS2 proteins of other

species, with the amino acid identity ranging from 46-74% for the PDZ, 46-73% for the

SH3 and 42-72% for the GUK. The L27 domain encoded in the 5' UTR of *dPALS2* was

compared to the L27 domains (L27N and L27C) of other known PALS2 orthologs, and

found to be more similar to the L27C (26-48%) than to the L27N (7-11%). It is unusual

that the dPALS2 L27 domain is not part of the protein product when the percent identities

show it is fairly conserved. *Drosophila* PALS2 appears to be more similar to *A. gambiae*

since its overall amino acid identity is the highest, it contains only one L27 domain and

each domain structure is highly conserved. Thus the similarities observed between

dPALS and its orthologs suggest a conserved function.

4.3 *Drosophila* PALS2 is expressed in epithelia of embryonic tissues:

Once *Drosophila* PALS2 was identified and characterized, the goal was to

analyze RNA and protein expression patterns throughout *Drosophila* development.

Embryonic RNA *in situ* hybridization showed the spatial distribution of *dPALS2* as a

transverse striped pattern in the epithelia. *dPALS2* expression first appeared around stage

10-11 and remained consistent until stage 15-17. Along with the striped pattern there

was also strong *dPALS2* expression in the hindgut in stage 10-14 embryos. Larval RNA

in situ hybridization showed no specific *dPALS2* distribution in the central nervous

system (CNS) or imaginal discs. It would be expected that if *dPALS2* transcripts were

expressed in the epithelia then *dPALS2* distribution would be visible in imaginal discs

since they are primordial cells for adult epithelia. Though it is possible *dPALS2*

transcripts are expressed at very low levels in larval tissues and can not be detected.

dPALS2 protein distribution was similar to the striped pattern observed in RNA *in*

situ hybridization and thus supports the function of dPALS2 in epithelia cells. dPALS2

protein expression began at stage 13 and remained consistent until stage 17 as a

transverse striped pattern in the epithelia, however dPALS2 was undetected in the

hindgut of the embryo where *in situ* hybridization showed RNA expression. DPALS2

protein was also undetectable in the CNS or imaginal discs of larval tissues. Western blot

analysis detected a 54kDa band corresponding to dPALS2 in embryonic and adult stages

of development. There was no protein detected in the larval stages. It is possible as

mentioned before that dPALS2 protein is expressed at very low levels and can not be

detected. The expression pattern of dPALS2 implies a role in the epithelia, possibly

involved in cell adhesion or cell polarity, similar to that of the mammalian orthologs or

segment polarity (Roh *et al.*, 2002; Shingai *et al.*, 2003).

4.4 *Drosophila* PALS2 may play a role in Segment Polarity:

Since dPAL2 appeared as a transverse striped pattern in the epithelia, it was

suggested that dPALS2 might be involved in establishing the segments of the developing

fly, since its expression pattern resembled that of segment polarity genes, *wingless* and

engrailed. Segment polarity genes mediate the required cell interactions to establish

segment borders, control the anterior-posterior patterning within each segment, and

maintain parasegment boundaries by activating cell adhesion molecules during

Drosophila development (Bejsovec and Wieschaus, 1993; Bhat, 1999; Sanson, 2001). Segmentation involves the progressive subdivision of the anterior/posterior axis by the gradient expression of transcription factors from the maternal, gap, pair-rule and segment polarity genes (Ingham and Martinex Arias, 1992; St. Johnston and Nusslein-Volhard, 1992). The expression pattern of segment polarity proteins are induced and defined by the pair-rule gene products, which are responsible for the establishment of the parasegments. Each segment is comprised of the posterior compartment of one parasegment and the anterior compartment of the adjacent parasegment (DiNardo *et al.*, 1994; Sanson, 2001).

The major signaling proteins, Wingless and Hedgehog, are produced at the boundary between the anterior and the posterior compartments to maintain and organize the pattern within each segment. Wingless is secreted by cells in the posterior compartment of each parasegment to induce the secretion of Hedgehog from the adjacent anterior compartment. Hedgehog then acts in a positive feed back loop by inducing Wingless expression and secretion from the posterior compartment (Sanson, 2001). As well as acting on the immediately adjacent cells to maintain the signaling loop, both proteins also diffuse away from their source, establishing gradients over the remaining cells of the parasegments. These gradients have the potential to act as morphogens, to induce the expression and localization of other segment polarity proteins that play a role in patterning the parasegments (Martinez Arias, 1993; Sanson, 2001). Thus segment polarity proteins are expressed in a gradient as transverse stripes along the anteroposterior axis of the embryo defining the parasegmental boundaries (Sanson, 2001).

To investigate dPALS2 as a potential segment polarity protein, the expression patterns of dPALS2 and Engrailed were analyzed together by confocal imagery. dPALS2 was expressed in a gradient striped pattern throughout the entire epithelia while Engrailed, a transcription factor also appeared in a striped pattern. Double labeling with dPALS2 and Engrailed antibodies showed similar striping patterns however Engrailed was localized to the center (nucleus) of epithelial cells, while dPALS2 was localized to the cytoplasm of the cell. The similar expression pattern of dPALS2 with Engrailed does imply a potential role for dPALS2 in segment polarity. It would be beneficial to analyze the expression of dPALS2 in relation to other segment polarity proteins that are expressed subcelluarly such as Shaggy or Fused.

Some segment polarity proteins such as Frizzled and Disheveled also function in the planar cell polarity (PCP) pathway and the Wingless (Wnt/Wg) signaling pathway (Barth *et al.*, 1997; Mlodzik, 2002). The PCP pathway polarizes whole fields of cells along the plane of epithelia. Proteins of the PCP pathway become asymmetrically localized along the proximal – distal axis controlling cell fates and cell movements. However almost all the research on PCP demonstrates this asymmetric localization in the wing imaginal disc or the photoreceptors in the adult eye (Mlodzik, 2002). Since dPALS2 protein was undetectable in larval tissues, it is unlikely dPALS2 is a protein involved in the planar cell polarity pathway.

4.5 *Drosophila* PALS2 may play role in Cell Adhesion and/or Epithelia Polarity:

dPALS2 is a MAGUK scaffolding protein expressed throughout the embryonic

epithelium. Sites of cell adhesion and the plasma membrane are rich in MAGUK

scaffolding proteins, which establish and maintain the structural specialization of the cell

membrane (Dimitratos *et al.*, 1999). Below I explore possible functions of dPALS2,

based upon what we know of the mammalian protein.

Mammalian PALS2 has been shown to co-localize with E-cadherin at adherens

junctions (AJ) (Roh *et al.*, 2002). E-cadherin is a Ca^{2+} dependent cell-cell adhesion

molecule that contains three domains, the cytoplasmic domain, the hydrophobic

transmembrane domain and the extracellular domain (Gooding *et al.*, 2004). The

extracellular domain allows interactions with other cadherins, resulting in adhesive

activity. The cytoplasmic domain interacts with intracellular proteins β and α catenin,

which mediate the connections between cadherins and the actin cytoskeleton to regulate

cadherin adhesive activity (Gooding *et al.*, 2004; Vasioukhin *et al.*, 2000). Loss of

function mutants of E-cadherin and the catenins result in tissue disorder, cellular de-

differentiation and maliganancy (Humphries *et al.*, 1998; Takeichi, 1991).

Mammalian PALS2 has also been shown to directly associate with Nectin-like

Molecule-2 (Necl-2), a cell-cell adhesion molecule at extra-junctional regions of epithelia

cells (Shingai *et al.*, 2003). Nectin-like molecules (Necls) were identified and designated

based on their similar domain structure to nectins (Ikeda *et al.*, 2003). Nectins are Ca^{2+}

independent cell-cell adhesion molecules that are indirectly associated with the cadherin-

catenin system and function in the organization of cell junctions (Aoki, *et al.*, 1997;

Lopez *et al.*, 1998; Takahaski *et al.*, 1999). The nectin family interacts with Afadin, an

actin binding protein that connects nectin to the cytoskeleton (Takai and Nakanishi, 2003;

Ozawa, 2002). Necl-2 does not bind Afadin but does directly bind the PDZ domain of PALS2. Necl-2 directly binds actin associated band 4.1 protein DAL-1 suggesting Necl-2 functions in cell-cell adhesion and transmembrane protein localization (Shingai *et al.*, 2003).

If mammalian PALS2 functions in cell adhesion, it is possible that loss of function mutants would reveal overlapping phenotypes with cell adhesion mutants. The human ortholog of PALS2, identified as VAM-1 (VELI associated MAGUK)/MPP6 (Multiple PDZ protein 6), may be a candidate tumor suppressor since its mapping to human chromosome 7p15-21 is a region reported to demonstrate changes in Wilm's tumors (Tseng *et al.*, 2001; Katoh and Katoh, 2004).

The cadherin-catenin complex is present in *Drosophila* and also functions in cell-cell adhesion. The *Drosophila* orthologs of E-cadherin, β-catenin and α-catenin have been identified as DE-cadherin, armadillo and *Drosophila* α-catenin respectively and form a complex at AJ similar to that of their mammalian counterparts (Gooding *et al.*, 2004). Since dPALS2 appears to be expressed throughout the epithelia and is considered conserved when compared to its orthologs, it is possible that dPALS2 may play a role in cell adhesion. If this is the case, then dPALS2 would likely be associated with the DE-cadherin/armadillo/*Drosophila* α-catenin complex at AJ. However it is possible dPALS2 may not be directly involved in cell adhesion but rather the maintenance of cell adhesion through epithelia polarity.

Since PALS2 has been shown to co-localize with E-cadherin, which is a major constituent of the AJ, it is possible PALS2 functions with cellular polarity proteins that

establish and maintain the AJ. The formation of the AJ indicates the distinct apical and basolateral membrane domains and thus developing epithelia polarity (Tepass and Hartenstein, 1994; Tepass, 2002).

A number of conserved proteins (in both mammals and *Drosophila)* contribute to the formation of AJ but are not actual AJ components such as E-cadherin and catenins. The Crumbs/Stardust/PATJ, Bazooka/DmPar-6/DaPKC and Scribble/Dlg/Lgl complexes are prominent regulators of AJ formation and epithelia polarity (Bachmann *et al.*, 2001; Hong *et al.*, 2001; Roh *et al.*, 2002). Baz, Crb, Scirb loss of function mutants not only cause cell polarity defects but also incomplete AJ formation. If PALS2/dPALS2 are associated with cell polarity, it is possible that loss of function mutants would show overlapping phenotypes with polarity mutants. Since dPALS2 is expressed throughout the epithelia and is conserved when compared to its orthologs, it is possible dPALS2 may function in establishing and/or maintaining cell polarity with the Baz, Crb and/or Scrib complexes.

It would be beneficial to examine the expression pattern of dPALS2 at a higher resolution. Co-localization with apical markers such as members from the Crumbs and Bazooka complexes, basolateral markers such as Discs large or Scribble or cell adhesion markers such as DE-cadherin, will establish dPALS2 expression with relation to other epithelial markers to predict a possible function in the development of *Drosophila.*

4.6 Direction of Future Research :

Since very little is known about PALS2 in both mammals and *Drosophila*, the function of dPALS2 in *Drosophila* development must be speculated. Mammalian orthologs of dPALS2 have been shown to associate with extra-junctional proteins E-cadherin and Necl-2, implying a role in cell-adhesion and/or epithelial polarity, while the expression data of dPALS2 implies a definite role in the epithelia, possibly segment polarity. It is essential to generate a *dPALS2* mutant by P element mutagenesis, which is currently being conducted in Dr. Jacobs' lab, to determine its function and role in the epithelia. If dPALS2 is indeed involved in segment polarity, cell adhesion and/or cell polarity, loss of *dPALS2* function should result in segmentation defects, or epithelial cells lacking adhesion to neighboring cells and/or undefined regions of apical and/or basolateral domains. *dPALS2* mutants should be labeled with segment and epithelial polarity determinants and/or cell adhesion antibodies to observe any mislocalization or decreased levels of polarity determinants or cell adhesion components.

Determination of dPALS2 function can also achieved by mis-expression or over-expression of dPALS2 using a UAS-dPALS2 construct, which is also currently being carried out in Dr. Jacobs' lab. The UAS-GAL4 system is a tool used in *Drosophila* to control the ectopic expression of genes to a specific set of cells (Brand and Perrimon, 1993). It is expected that over-expression and/or mis-expression phenotypes of dPALS2 will give insight into its effects on cell-adhesion and/or cell polarity.

In addition, it is necessary to determine interacting partners of dPALS2 and how they associate to identify the possible functions of dPALS2. A yeast two hybrid assay is

currently in progress in Dr. Jacobs' lab where *dPALS2* is used as bait against an adult

Drosophila cDNA library to examine possible protein interactions. The discovery of

dPALS2 binding partners will hopefully reveal more information about its function.

Finally Northern Blot analysis should be attempted again to verify the number of

dPALS2 transcripts. Even though 5'RACE only identified one transcript, another

technique such as Northern Blot analysis could verify this finding.

Very little is known about PALS2 other than it may play a role in segmentation,

cell-adhesion and/or cell polarity. Since dPALS2 shows significant similarity to its

mammalian counterparts, it is hoped that further research will provide insight into its

function and purpose in establishing and maintaining the epithelia.

References:

Albertson, R, and Doe, CQ, (2003) Dlg, Scrib and Lgl regulate nuroblast cell size and mitotic spindle asymmetry. *Nat. Cell Biol.* **5**: 166-170.

Anderson, JM, (1996) Cell signaling: MAGUK magic. *Curr. Biol.* **6**: 382-384.

Aoki, J, Koike, S, Ise, I, Suwa, H, Tanaka, T, Miyasaka, M, and Nomoto, A, (1997) Mouse homology of poliovirus receptor-related gene 2 product, mPRR2, mediates homophilic cell aggregation. *Exp. Cell Res.* **235**: 374-384.

Bachmann, A, Schneider, M, Theilenberg, E, Grawe, F, and Knust, E, (2001) *Drosophila* Stardust is a partner of Crumbs in the control of epitheilial cell polarity. *Nature* **414**: 638-643.

Barth, A, Nathke, I, and Nelson, WJ, (1997) Cadherins, catenins and APC protein: interplay between cytoskeletal complexes and signaling pathways. *Curr. Opin. Cell Biol.* **9**: 683-690.

Bejsovec, A, and Wieschaus, E, (1993) Segment polarity gene interactions modulate epidermal patterning in *Drosophila* embryos. *Dev.* **119**: 501-517.

Benton, R, and St. Johnston, D, (2003) *Drosophila* PAR-1 and 14-3-3 inhibit Bazooka/PAR-3 to establish complementary cortical domains in polarized cells. *Cell* **115**: 691-704.

Betschinger, J, Mechtler, K, and Knoblich, JA, (2003) The Par complex directs asymmetric cell division by phosphorylating the cytoskeltal protein Lgl. *Nature* **422**: 326-330.

Bhat, KM, (1999) Segment polarity genes in neuroblast formation and identity specification during *Drosophila* neurogenesis. *Bioessays* **21**: 472-485.

Bhat, MA, Izaddoost, S, Lu, Y, Cho, KO, Choi, KW, and Bellen, HJ, (1999) Discs Lost, a novel multi-PDZ domain protein, establishes and maintains epithelial polarity. *Cell* **96**: 833-845.

Bilder, D, and Perrimon, N, (2000) Localization of apical epithelial determinants by the basolateral PDZ protein Scribble. *Nature* **403**: 676-680.

Bilder, D, Li, M, and Perrimon, N, (2000) Cooperative regularion of cell polarity and growth by *Drosophila* tumor suppressors. *Science* **289**:113-116.

Bilder, D, (2001) PDZ proteins and polarity: functions from the fly. *Trends Genet.* **17**: 511-519.

Bilder, D, Schober, M, and Perrimon, N, (2003) Integrated activity of PDZ protein complexes regulates epithelial polarity. *Nat. Cell Biol.* **5**:53-58.

Bilder, D, (2004) Epithelial polarity and proliferation control: links from the *Drosophila* neoplastic tumor supressors. *Gen. Dev.* **18**: 1909-1925.

Bradford, MM, (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **7**: 248-254.

Brand, AH and Perrimon, N, (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development.* **118**: 401-415.

Burbelo, PD, Drechsel, D, and Hall, A, (1995) A conserved binding motif defines numerous candidate target proteins for both Cdc42 and Rac GTPases. *J. Biol. Chem.* **270**: 29071-29074.

Caruana, G, (2002) Genetic studies define MAGUK proteins as regulators of epithelial cell polarity. *Int. J. Dev. Biol.* **46**: 511-518.

Cao, X, Ding, X, Guo, Z, Zhou, R, Wang, F, Fog, V, Long, F, Bi, F, Fan, D, Forte, JG, Margolis, B, Teng, M, and Yao, X, (2005) Pals1 specifies the localization of ezrin to the apical membrane of gastric pariental cells. *J. Biol. Chem.* **280**: 13584-13592.

Cohen, AR, Woods, DF, Marfatia, SM, Walther, Z, Chishti, AH, Anderson, JM, and Wood, DF, (1998) Human CASK/LIN-2 binds syndecan-2 and protein 4.1 and localizes to the basolateral membrane of epithelial cells. *J. Cell Biol.* **142**: 129-138.

den Hollander, AI, Heckenlively, JR, van den Born, LI, de Kok, YJ, van der Velde-Visser, SD, Kellner, U, Jurklies, B, van Schooeveld, MJ, Blankenagel, A, Rohrschneider, K, Wissinger, B, Cruysberg, JR, Deutman, AF, Brunner, HG, Apfelstedt-Sylla, E, Hoyng, CB, and Cremers, FP, (2001) Leber congenital amaurosis and retinitis pigmentosa with Coats-like exudative vasculopathy are associated with mutations in the crumbs homologue 1 (CRB1) gene. *Am. J. Hum. Genet.* **69**: 198-203.

den Hollander, AI, Ghiani, M, de Kok, YJ, Wijnholds, J, Ballabio, A, Cremers, FP, and Broccoli, V, (2002) Isolation of Crb1, a mouse homologue of *Drosophila* Crumbs, and analysis of its expression pattern in eye and brain. *Mech. Dev.* **110**: 203-207.

Dimitratos, SD, Woods, DF, Stathakis, DG, and Bryant, PJ, (1999) Signaling pathways are focused at specialized regions of the plasma membrane by scaffolding proteins of the MAGUK family. *Bioessays* **21**: 912-921.

DiNardo, S, Heemskerk, J, Dougan, S, and O'Farrell, PH, (1994) The making of a maggot: patterning the *Drosophila* embryogenesis. *Curr. Opin. Genet. Dev.* **4**: 529-534.

Doerks, T, Bork, P, Kamberov, E, Makarova, O, Muecke, S, and Margolis, B, (2000) L27, a novel heterodimerization domain in receptor targeting proteins lin-2 and Lin-7. *Trends Biochem. Sci.* **25**: 317-318.

Dong, H, Zhang, P, Song, I, Petralia, RS, Liao, D, and Huganir, RL, (1999) Characterization of the glutamate receptor-interacting proteins GRIP1 and GRIP2. *J. Neurosci.* **19**: 6930-6941.

Doyle, DA, Lee, A, Lewis, J, Kim, E, Sheng, M, and Mackinnon, R, (1996) Crystal structure of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell* **85**: 1067-1076.

Drysdale, RA, Crosby, MA, and The Flybase Consortium, (2005) Flybase: genes and gene models. *Nucleic Acids Res.* **33**: D390-D395. http://flybase.org

Feng, W, Long, J, Fan, J, Suetake, T, and Zhang, M, (2004) The tetrameric L27 domain complex as an organization platform for supramolecular assemblies. *Nat. Struct. Mol. Biol.* **11**: 475-480.

Garrand, SM, capaldo, CT, Gao, L, Rosen, MK, Macara, IG, and Tomchick, DR, (2003) Structure of Cdc42 in a complex with the GTPase–binding domain of cell polarity protein, Par-6. *EMBO. J.* **22**: 1125-1133.

Gao, L, Joberty, G, and Masara, LG, (2002) Assembly of epithelial tight junctions is negatively regulated by Par6. *Curr. Biol.* **12**:221-225.

Gateff, E, and Schneiderman, HA, (1967) Developmental studies of a new mutant of *Drosophila melanogaster*: Lethal malignant brain tumor (l(2)gl 4. *Am. Zool.* **7**: 760. **Gateff, E,** (1978) Malignant neoplasms of genetic origin in *Drosophila melanogaster*. *Science* **200**: 1448-1459.

Giot, L, Bader, JS, Brouwer, C, Chaudhuri, A, Kuang, B, Li, Y, Hao, YL, Ooi, EC, Godwin, B, Witols, E, Vijayadamodar, G, Pochart, P, Machineni, H, Welsh, M, Kong, Y, Zerhusen, B, Malcolm, R, Varrone, Z, Collis, A, Minto, M, Burgess, S, McDaniel, L, Stimpson, E, Spriggs, F, Williams, J, Neurath, K, Ioime, N, Agee, M, Voss, E, Furtak, K, Renzulli, R, Aanensen, N, Carrolla, S, Bickelhaupt, E, Lazovatsky, Y, DaSilva, A, Zhong, J, Stanyon, CA, Finley Jr., RL, White, KP, Braverman, M, Jarvie, T, Gold, S, Leach, M, Knight, J, Shimkets, RA, McKenna, MP, Chant, J, and Rothberg, JM, (2003) A protein interaction map of *Drosophila melanogaster*. *Science* **302**: 1727-1736.

Godt, D, Couderc, JL, Cramton, SE, and Laski, FA, (1993) Pattern formation in the limbs of *Drosophila*: *bric a brac* is expressed in both a gradient and a wave-like pattern

and is required for specification and proper segmentation of the tarsus. *Development* **119**: 799-812.

Gooding, JM, Yap, KL, and Ikura, M, (2004) The cadherin-catenin complex as a focal point of cell adhesion and signaling: new insights from three dimensional structures. *Bioessays* **26**: 497-511.

Grawe, F, Wodarz, A, Lee, B, Knust, E, and Skaer, H, (1996) The *Drosophila* genes crumbs and stardust are involved in the biogenesis of adherens junctions. *Development* **122**:951-959.

Harris, BZ, Venkatasubrahmanyam, S, and Lim, WA, (2002) Coordinated folding and association of the LIN-2,-7 (L27) domain. *J. Biol. Chem.* **277**: 34902-34908.

Hirose, T, Izumi, Y, Nagashima, Y, Tamai, Nagai, Y, Kurihara, H, Sakai, T, Suzuki, Y, Yamanaka, T, Suzuki, A, Mizuno, K, and Ohno, S, (2003) Involvement of ASIP/PAR-3 in the promotion of epithelial tight junction formation. *J. Cell Sci.* **115**: 2485-2495.

Hong, Y, Stronach, B, Perrimon, N, Jan, LY, Jan, YN, (2001) *Drosophila* Stardust interacts with Crumbs to control polarity of epithelia but not neuroblasts. *Nature* **414**: 634-638.

Hsueh, Y, Butz, S, and Sudhof, TC, (1997) Disulfide-linked head-to-head multimerization in the mechanism of ion channel clustering by PSD-95. *Neuron* **18**: 803-814.

Humphries, MJ, and Newham, P, (1998) The structure of cell-adhesion molecules. *Trends Cell Biol.* **8**: 78-83.

Hung, AY, and Sheng, M, (2002) PDZ domains: structural modules for protein complex assembly. *J. Biol. Chem.* **277**: 5699-5702.

Hurd, TW, Gao, L, Roh, MH, Macara, IG, and Margolis, B, (2003) Direct interaction of two polarity complexes implicated in epithelial tight junction assembly. *Nature Cell Biol.* **5**: 137-142.

Ikeda, W, Kakunaga, S, Itoh, S, Shingai, T, Fukami, T, Satoh, K, Inoue, Y, Hamagruchi, A, Morimoto, K, Takeuchi, M, Imai, T, and Takai, Y, (2003) Tage4/Nectin-like molecule-5 heterophilically trans-interacts with cell adhesion molecule Nectin-3 and enhances cell migration. *J. Biol. Chem.* **278**: 28167-28172.

Ingham, R, and Martinez Arias, A, (1992) Boundaries and fields in early embryos. *Cell* **68**: 221-235.

Jo, K, Derin, R, Li, M, and Bredt, DS, (1999) Characterization of MALS/Velis-1, -2,

and –3: a family of mammalian LIN-7 homologs enriched at brain synapses in association with the postsynaptic density-95/NMDA receptor postsynaptic complex. *J. Neurosci.* **19**: 4189-4199.

Joberty, G, Petersen, C, Gao, L, and Macara, IG, (2000) The cell-polarity protein Par6 links Par3 and atypical protein kinase C to Cdc42. *Nature Cell Biol.* **2**: 531-539.

Kaech, SM, Whitfield, CW, and Kim, SK, (1998) Basolateral membrane localization of the receptor tyrosine kinase LET-23 is mediated by the LIN-2/LIN-7/LIN-10 protein complex in the vulval precursor cells of *C. elegans*. *Cell* **94**: 761-771.

Kalmes, A, Merdes, G, Neumann, B, Strand, D, and Mechler, BM, (1996) A serine-kinase associated with the p127-l(2)gl tumor suppressor of *Drosophila* may regulate the binding of p127 to nonmuscle myosin II heavy chain and the attachment of p127 to the plasma membrane. *J. Cell Sci.* **109**: 1359-1368.

Kamberov, E, Makarova, O, Roh, M, Liu, A, Karnak, D, Straight, S, Margolis, B, (2000) Molecular Cloning and Characterization of Pals, Proteins Associated with mLIN-7. *J. Biol. Chem.* **275**: 11425-11431.

Katoh, M, and Katoh, M, (2004) Identification and characterization of human MPP7 gene and mouse Mpp7 gene in *silico. Int. J. Mol. Med.* **13**: 333-338.

Kemphues, KJ, Priess, JR, Morton, DG, Cheng, NB, (1988) Identification of genes required for cytoplasmic localization in early *C. elegans* embryos. *Cell* **52**: 311-320.

Kim, E, Niethammer, M, Rothschild, A, Jan, YN, and Sheng, M, (1995) Clustering of Shaker-type K^+ channels by interaction with a family of membrane-associated guanylate kinases. *Nature* **378**: 85-88.

Lee, S, Fan, S, Makarova, O, Straight, S, and Margolis, B, (2002) A novel and conserved protein-protein interaction domain of mammalian Lin-2/CASK binds and recruits SAP97 to the lateral surface of epithelia. *Mol. Cell Biol.* **22**: 1778-1791.

Li, Y, Karnak, D, Demeler, B, Margolis, B, and Lavie, A, (2004) Structural basis for L27 domain-mediated assembly of signaling and cell polarity complexes. *EMBO J.* **23**: 2723-2733.

Lin, D, Edwards, AS, Fawcett, JP, Mbamualu, G, Scott, JD, and Pawson, T, (2000) A mammalian PAR-3-PAR-6 complex implicated in Cdc42/Rac1 and aPKC signaling and cell polarity. *Nature Cell Biol.* **2**: 540-547.

Lopez, M, Aoubala, M, Jordier, F, Isnardon, D, Gomez, S, and Dubreuil, P, (1998) Complementary DNA characterization and chromosomal localization of a human gene related to the poliovirus receptor-encoding gene. *Blood.* **92**: 4602-4611.

MacMullin, AA, (2001) Identification and molecular characterization of *dveli*, the *Drosophila* ortholog of *C. elegans lin-7*. Master thesis. McMaster University.

Manfruelli, P, Arquier, N, Hanratty, WP, and Semeriva, M, (1996) The tumor suppressor gene, lethal(2)giant larvae (l(2)gl), is required for cell shape change of epithelial cells during *Drosophila* development. *Development* **122**: 2283-2294.

Martinez Arias, A, (1993) Development and patterning of the larval epidermis of *Drosophila*. In **Bate, M, and Martinez Arias, A,** (eds), The Development of *Drosophila melanogaster*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 517-608.

Mathew, D, Gramates, LS, Packard, M, Thomas, U, Bilder, D, Perrimon, N, Gorczyca, M, and Budnik, V, (2002) Recruitment of Scribble to the synaptic scaffolding complex requires GUK-holder, a novel DLG binding protein. *Curr. Biol.* **12**: 531-539.

Mechler, BM, McGinnis, W, and Gehring, WJ, (1985) Molecular cloning of lethal(2)giant larvae, a recessive oncogene of *Drosophila melanogaster*. *EMBO J.* **4**: 1551-1557.

Miller, JR, Hocking, AM, Brown, JD, and Moon, RT, (1999) Mechanism and function of signal transduction by the Wnt/beta-catenin and Wnt/Ca^{2+} pathways. *Oncogene*. **18**: 7860-7872.

Mlodzik, M, (2002) Planar cell polarization: do the same mechanisms regulate *Drosophila* tissue polarity and vertebrate gastrulation? *Trends Genet.* **18**: 564-571.

Morais Cabral, JH, Petosa, C, Sutcliffe, MJ, Raza, S, Byron, O, Poy, F, Marfatia, SM, Chishti, AH and Liddington, RC, (1996) Crystal structure of a PDZ domain. *Nature* **382**: 649-652.

Muller, HA, and Wieschaus, E, (1996) Armadillo, bazooka and stardust are critical for early stages in formation of the zonula adherens and maintenance of the polarized blastoderm epithelium in *Drosophila*. *J. Cell Biol.* **134**: 149-163.

Muller, HAJ, (2000) Genetic control of epithelial cell polarity: lessons in *Drosophila*. *Dev. Dyn.* **218**: 52-67.

Muller, HA, (2003) Epithelial polarity in flies. More than just crumbs. *Dev. Cell.* **4**: 1-3.

Musch, A, Cohen, D, Yeaman, C, Nelson, WJ, Rodriguez-Boulan, E, and Brennwald, PJ, (2002) Mammalian homolog of *Drosophila* tumor suppressor lethal (2) giant larvae interactsw with basolateral exocytic machinery in Madin-Darby canine kidney cells. *Mol. Biol. Cell* **13**: 158-168.

Nakagawa, S, and Huibregtse, JM, (2000) Human scribble (Vartul) is targeted for ubiquitin-mediated degradation by the high risk papillomavirus E6 proteins and E6AP ubiquitin-protein ligase. *Mol. Cell Biol.* **20**: 8244-8253.

Pawson, T, (1994) SH2 and SH3 domains in signal transduction. *Adv. Cancer Res.* **64**: 87-110.

Peifer, M, and Polakis, P, (2000) Wnt signaling in oncogenesis and embryogenesis – a look outside the nucleus. *Science* **287**: 1606-1609.

Penkert, RR, Divittorio, HM, and Prehoda, KE, (2004) Internal recognition through PDZ domain plasticity in the Par-6-Pals-1 complex. *Nat. Struc. Mol. Biol.* **11**: 1122-1127.

Perego, C, Vanoni, C, Villa, A, Longhi, R, Kaech, SM , Frohli, E, Hajnal, A, Kim, SK, and Pietrini, G (1999) PDZ-mediated interactions retain the epithelial GABA transporter on the basolateral surface of polarized epithelial cells. *EMBO J.* **18**: 2384-2393.

Petronczki, M, and Knoblich, JA, (2001) DmPAR-6 directs epithelial polarity and asymmetric cell division of neuroblasts in *Drosophila. Nat. Cell Biol.* **3**: 43-49.

Pielage, J, Tobias, S, Bunse, I, and Klambt, C, (2003) The *Drosophila* cell survival gene *discs lost* encodes a cytoplasmic codanin-1-like protein, not a homolog of tight junction PDZ protein Patj. *Dev. Cell* **5**: 841-851.

Plant, PJ, Fawcett, JP, Lin, DCC, Holdorf, AD, Binns, K, Kulkarni, S, and Pawson, T, (2003) Apolarity complex of mPar-6 and atypical PKC binds, phosphorylates and regulates mammalian Lgl. *Nat. Cell Biol.* **5**: 301-308.

Ponting, CP, (1997) Evidence for PDZ domains in bacteria, yeast and plants. *Protein Sci.* **6**: 464-468.

Rho, MH, Makarova, O, Liu, CJ, Shin, KY, Lee, S, Laurinec, S, Goyal, M, Wiggins, R, and Margolis, B, (2002) The Maguk protein, Pals1, functions as an adapter, linking mammalian homologues of Crumbs and Discs Lost. *J. Cell Biol.* **157**: 161-172.

Roh, MH, and Margolis, B, (2003) Composition and function of PDZ protein complexes during cell polarization. *Am. J. Physiol. Renal. Physiol.* **285**: F377-F387.

Roh, MH, Shuling, F, Lui, CJ, and Margolis, B, (2003) The Crumbs2-Pals1 complex participates in the establishment of polarity in mammalian epithelial cells. *J. Cell Sci.* **116**: 2895-2906.

Sanson, B, (2001) Generating patterns from fields of cells. *EMBO reports.* **2**: 1083-1088.

Santoni, MJ, Pontarotti, P, Birnbaum, D, and Borg, JP, (2002) The LAP family: A

phylogenetic point of view. *Trends Genet.* **18**: 494-497.

Schultz, J, Copley, RR, Doerks, T, Ponting, CP, and Bork, P, (2000) SMART: a Web-Based Tool for the Study of Genetically Mobile Domains. *Nucleic Acids Res.* **28**: 231-234.

Setou, M, Nakagawa, T, Seog, DH, and Hirokawa, N, (2000) Kinesin superfamily motor protein KIF17 and mLin-10 in NMDA receptor containing vesicle transport. *Science* **288**: 1796-1802.

Sheng, M and Sala, C, (2001) PDZ Domains and the Organization of Supramolecular Complexes. *Annu. Rev. Neurosci.* **24**: 1-29.

Shingai, T, Ikeda, W, Kakunaga, S, Morimoto, K, Takekuni, K, Itoh, S, Satoh, K, Takeuchi, M, Imai, T, Monden, M, and Takai, Y, (2003) Implications of Nectin-like molecule-2/IGSF4/RA175/SgIGSF/TSLC1/SynCAM1 in cell-cell adhesion and transmembrane protein localization in epithelial cells. *J. Biol.Chem.* **278**: 35421-35427.

Simske, JS, Kaech, SM, Harp, SA, and Kim, SK, (1996) LET-23 receptor localization by the cell junction protein LIN-7 during *C. elegans* vulval induction. *Cell* **85**: 195-204.

Sotillos, S, Diaz-Meco, MT, Caminero, E, Moscat, J, and Campuzano, S, (2004) DaPKC dependent phosphorylation of Crumbs in required for epithelial cell polarity in *Drosophila. J. Cell Biol.* **166**: 549-557.

Songyang, Z, Fanning, AS, Fu, S, Xu, J, Marfatia, SM, Chishti, AH, Crompton, A, Chan, AC, Anderson, JM, and Cantley, LC, (1997) Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* **275**: 73-77.

St. Johnston, D, and Nusslein-Volhard, C, (1992) The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**: 201-219.

Stewart, M, Murphy, C, and Fristrom, FW, (1972) The recovery and preliminary characterization of X chromosome mutants affecting imaginal discs of *Drosophila melanogaster*. *Dev. Biol.* **27**: 71-83.

Straight, SW, Karnak, D, Borg, JP, Kamberov, E, Dare, H, Margolis, B, and Wade, JB, (2000) mLin-7 is localized to the basolateral surface of renal epithelia via its NH(2) terminus. *Am. J. Physiol. Renal Physiol.* **278**: F464-F475.

Straight, SW, Chen, L, Karnak, D, And Margolis, B, (2001) Interaction with mLin-7 alters the targeting of endocytosed transmembrane proteins in mammalian epithelial cells. *Mol. Biol. Cell* **12**: 1329-1340.

Strand, D, Raska, I, and Mechler, BM, (1994) The *Drosophila* lethal(2)giant larvae

tumor suppressor protein is a component of the cytoskeleton. *J. Cell Biol.* **127**: 1345-1360.

Stricker, NL, Christopherson, KS, Yi, BA, Schatz, PJ, Raab, RW, Dawes, G, Bassett, DE Jr, Bredt, DS, and Li, M, (1997) PDZ domain of neuronal nitric oxide synthase recognizes novel C-terminal peptide sequences. *Nat. Biotechnol.* **15**: 336-342.

Suzuki, A, Yamanaka, T, Hirose, N, Manabe, K, Mizuno, M, Shimizu, K, Akimoti, Y, Izumi, T, Oshnishi, and Ohno, S, (2001) Atypical protein kinase C is involved in the evolutionarily conserved par protein complex and plays a critical role in establishing epithelial-specific junctional complexes. *J. Cell Biol.* **152**: 1183-1196.

Takahashi, K, Nakanishi, H, Miyahara, M, Mandai, K, Satoh, K, Satoh, S, Nishioka, J, Aoki, J, Nomoto,A, Mizoguchi, A, and Takai, Y, (2000) Nectin/PRR: an immunoglobulin-like cell adhesion molecule recruited to cadherin-based adherens junctions through interactions with Afadin, a PDZ domain-containing protein. *J. Biol. Chem.* **145**: 539-549.

Takai, Y, and Nakanishi, H, (2003) Nectin and afadin: novel organizers of intercellular junctions. *J. Cell Sci.* **116**: 17-27.

Takeichi, M, (1991) Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* **251**: 1451-1455.

Tanentzapf, G, and Tepass, U, (2003) Interactions between the crumbs, lethal giant larvae and bazooka pathways in epithelial polarization. *Nat. Cell Biol.* **5**: 46-53.

Tepass, U, Theres, C, and Knust, E, (1990) Crumbs encodes an EGF-like protein expressed on apical membranes of *Drosophila* epithelial cells and required for organization of epithelia. *Cell* **61**: 787-799.

Tepass, U, and Hartenstein, V, (1994) The development of cellular junctions in the *Drosophila* embryo. *Dev. Biol.* **161**: 563-596.

Tepass, U, (1996) Crumbs, a component of the apical membrane, is required for zonula adherens formation in primary epithelia of *Drosophila*. *Dev. Biol.* **177**: 217-225.

Tepass, U, (2002) Adherens junctions: new insight into assembly, modulation and function. *Bioessays* **24**: 690-695.

Tseng, T, Marfatia, SM, Bryant, PJ, Pack, S, Zhuang, Z, O'Brien, J, Lin, L, Hanada, T, and Chishti, A, (2001) VAM-1: a new member of the MAGUK family binds to human Veli-1 through a conserved domain. B*iochim. Biophys. Acta.* **1518**: 249-259.

Vasioukhin, V, Bauer, C, Yin, M, and Fuchs, E, (2000)Directed actin polymerization

is the driving force for epithelial cell-cell adhesion. *Cell* **100**: 209-219.

Wang, Q, Toby, HW, and Margolis, B, (2004) Tight junction protein Par6 interacts with an evolutionarily conserved region in the amino terminus of Pals1/stardust. *J. Biol. Chem.* **279**: 30715-30721.

Williams, DS, Arikawa, K, and Paallysaho, T, (1990) Cytoskeletal components of the adherens junctions between the photoreceptors and the supportive Muller cells. *J. Comp. Neurol.* **295**: 155-164.

Wodarz, A, Ramrath, A, Grimm, A, and Knust, E, (2000) *Drosophila* atypical protein kinase C associates with bazooka and controls polarity of epithelia and neuroblasts. *J.Cell Biol.* **150**:1361-1374.

Wodarz, A, (2002) Establishing cell polarity in development. *Nature Cell Biol.* **4**: E39-E44.

Woods, DF and Bryant, PJ, (1991)The discs-large tumor suppressor gene of *Drosophila* encodes a guanylate kinase homolog localized at the septate junctions. *Cell* **66**: 451-464.

-**Woods, DF, and Bryant, PJ,** (1993) Apical junctions and cell signaling in epithelia. *J. Cell Sci.* **17**: 171-181.

Woods, DF, Wu, JW, and Bryant, PJ, (1997) Localization of protein to the apico-lateral junctions of *Drosophila* epithelia. *J. Cell Biol.* **134**: 1469-1482.

Yamanaka, T, Horikoshi, Y, Suzuki, A, Sugiyama, Y, Kitamura, K, Maniwa, R, Nagai, Y, Yamashita, A, Hirose, T, Ishikawa, H, and Ohno, S, (2001) PAR-6 regulates a PKC activity in a novel way and mediates cell-cell contact- induced formation of the epithelial junctional complex. *Gene Cell* **6**: 721-731.

Yamanaka, T, Horikoshi, Y, Sugiyama, T, Ishiyama, C, Suzuki, A, Hirose, T, Iwamatsu, A, Shinohara, A and Ohno, S, (2003) Mammalian Lgl forms a protein complex with Par-6 and aPKC independently of Par-3 to regulate epithelial polarity. *Curr. Biol.* **13**: 734-743.

Yamamoto, Y, Mandai, K, Okabe, N, Hoshino, T, Nakanishi, H, and Takai, Y, (2002) Localization of mLin-7 at nectin-based cell-cell junctions. *Oncogene* **21**: 2545-2554.

Appendix 1: Oligonucleotide primers utilized in amplification regions of *dPALS2* through RT-PCR. A table categorizes all primers used, the oligonucleotide sequences, the region amplified and the length of the PCR product. The restriction enzyme cut sites are illustrated in bold. All primers were manufactured at MOBIX at McMaster University.

Primer ID# (Name)	Primer Sequence	Region of Amplification	Product Length
Fwd PALSII probe Rev PALSII probe	5' GAG GCG GAC TAC GTT **CTC GAC** 3' **(XhoI)** 5' CTG GTA **TCT AGA** ATG ACC GCT CC 3' **(XbaI)**	Forward primer begins amplification at bp 574 of *dpals2* cDNA. Reverse primer begins amplification at bp 833 of *dpals2* cDNA.	Primers were used to amplify a region of *dpals2* cDNA without conserved protein domains for RNA i*n situ*. Product length is 259bp.
Fwd L27 probe Rev L27 probe	5' CGC TCA **TCT GCA** GAA CAA GCC CAT C 3' **(PstI)** 5' GAT CCC **CCG GCC** GCT TTA AGT C 3' **(EagI)**	Forward primer begins 255bp upstream of start codon. Reverse primer begins 75 bp upstream of start codon.	Primers were used to amplify the L27 domain in the 5'UTR of *dpals2*. Product length is 180bp.
Fwd AB Rev AB	5' GCA **AGA TCT** AGT GGA CGA CGA AAT AAT CAA G 3' **(BglII)** 5' GAT **GAA TTC** CGG TTG GAG CCC GTG G 3' **(EcoRI)**	Forward primer begins amplification at bp 249 of dpals2 cDNA Reverse primer begins amplification at bp 1134 of dpals2 of cDNA	Primers were used to amplify dpals2 cDNA without the conserved PDZ domain for protein induction. Product length is 909bp.

Appendix 2: The location of *dPALS2* cDNA inserted into pBluescript KSII multiple cloning site. A 259 base pair region of *dPALS2* that lies between the SH3 and GuK domain was cloned into *pBS KSII* at XbaII and XhoI restriction sites. The T7 and T3 promoters within the *pBS KSII* were used for sequencing as well as the creation of DIG labeled sense and anti-sense RNA probes for *in situ* hybridization.

pBluescript KSII (+/-) Muliple Cloning Site
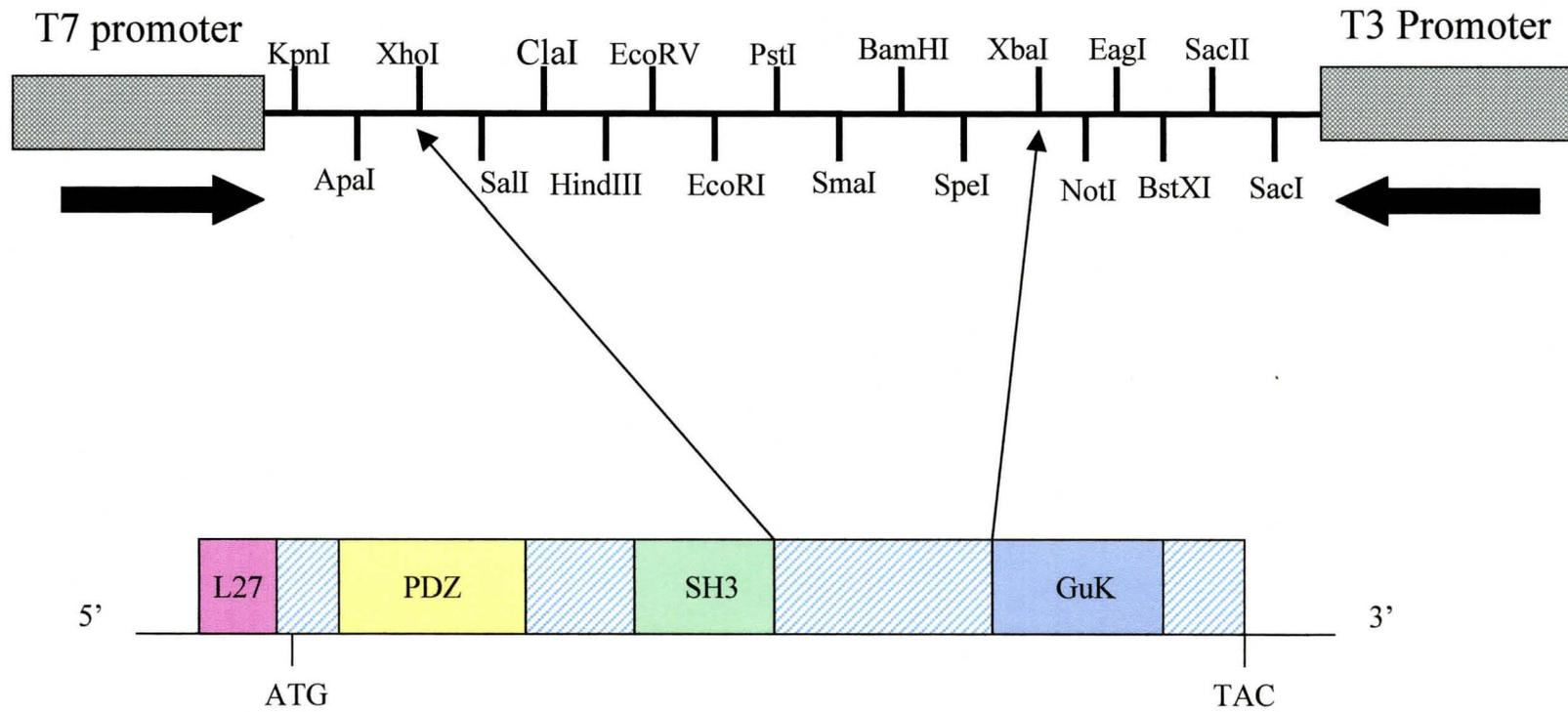
Appendix 3: The L27 domain (located in the 5' UTR) of *dPALS2* inserted into pBluescript KSII multiple cloning site. A180 base pair region in the 5' UTR of *dPALS2* that codes an L27 protein domain was cloned into *pBS KSII* at KpnI and XbalI restriction sites. The T7 and T3 promoters within in *pBS KSII* were used for sequencing as well as the creation of DIG labeled sense and anti-sense RNA probes for *in situ* hybridization.

pBluescript KSII (+/-) Muliple Cloning Site
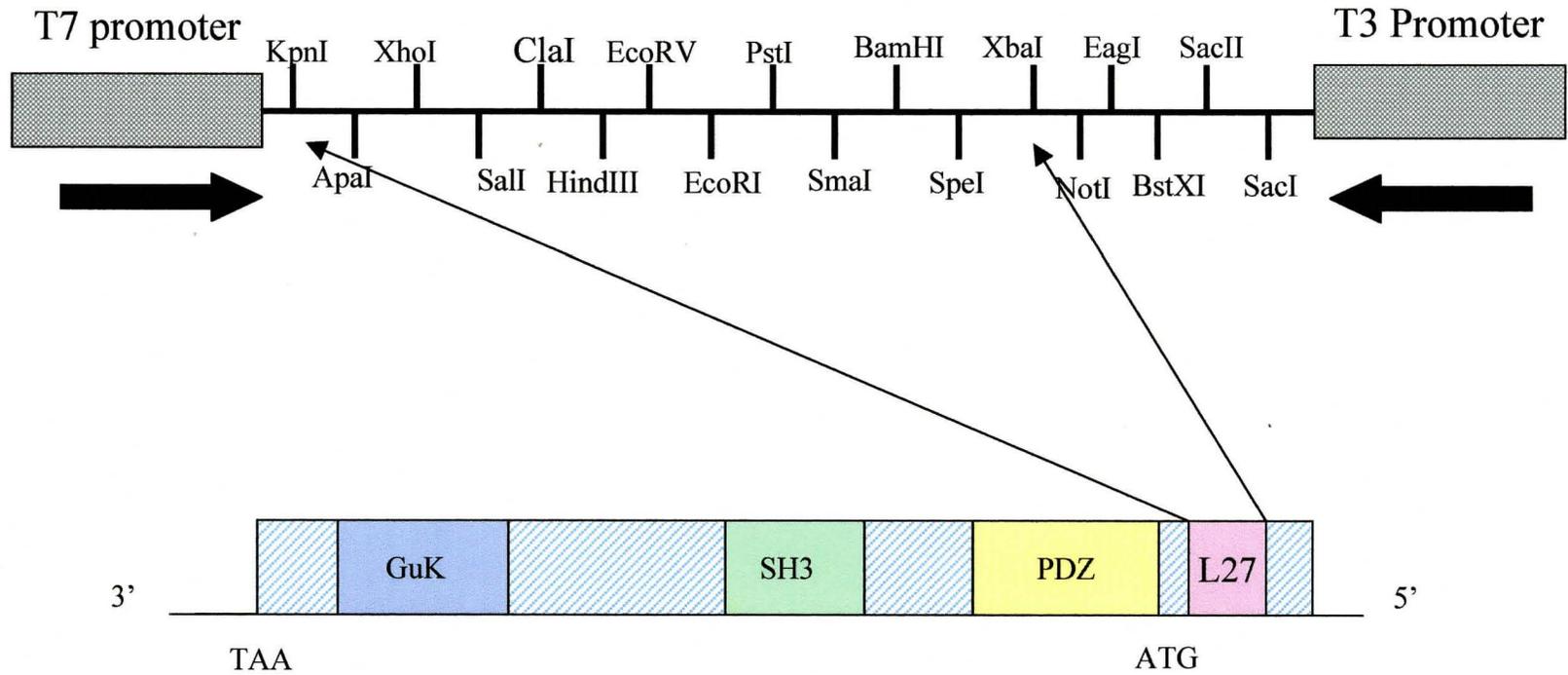
Appendix 4: The location of *dPALS2* cDNA inserted into *pET29b+* multiple cloning site for dPALS2-His fusion protein construct. A 909 base pair region of *dPALS2* that begins immediately after the PDZ domain and finishes at the end of the GUK domain was cloned into *pET29b+* at the EcoRI and BglII sites. This construct was used to induce protein expression for antibody production.

pET-29b(+) Multiple Cloning Site
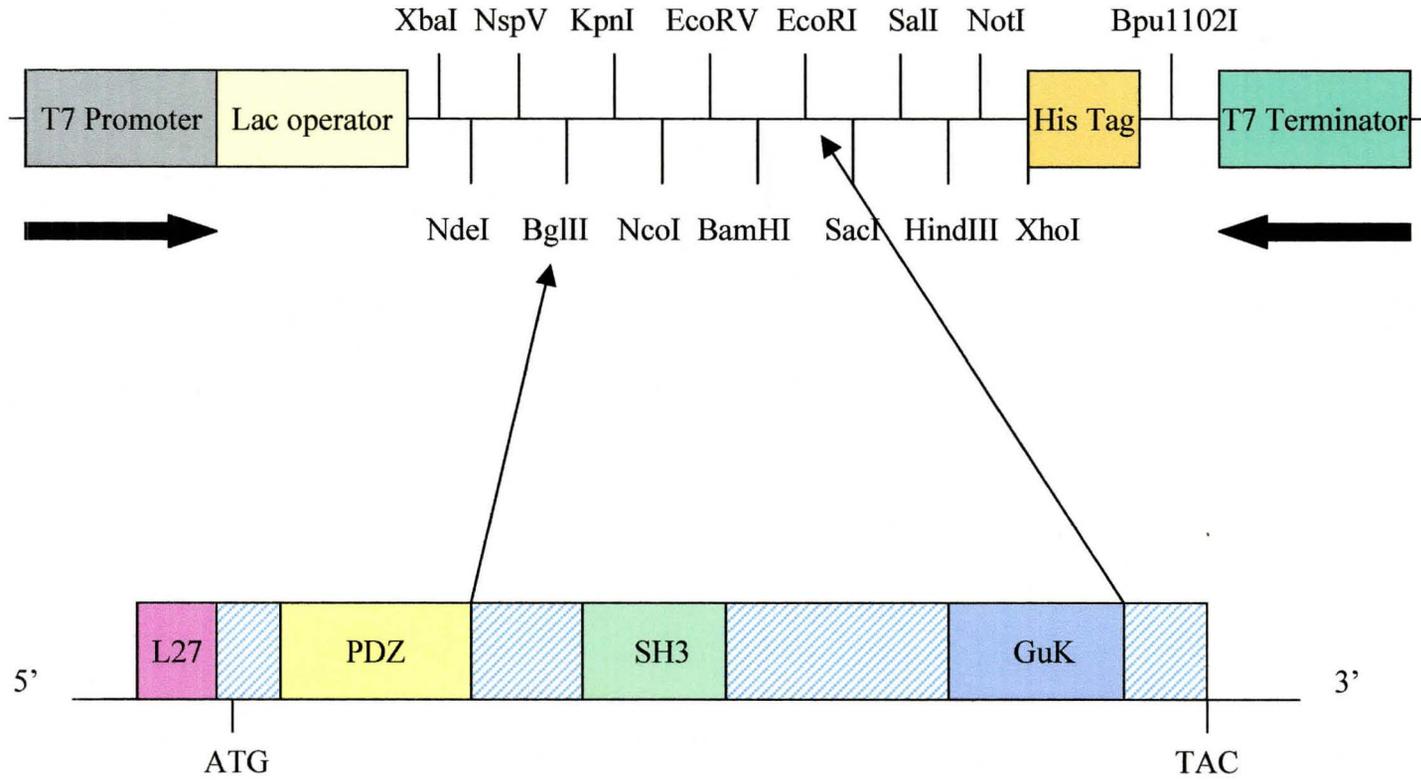
Appendix 5: *Drosophila PALS2* **genomic structure.** *dPALS2* is located on the left arm of the 2nd chromosome at 38E10. The genomic structure contains six introns (indicated by black lines) and seven exons, all varying in size. A region within the 5'UTR was included to show the location of the L27 domain. Protein domains within the exons are color coded. Pink = L27 domain, Yellow = PDZ domain, Green = SH3 domain, Purple = GuK domain. The dPALS2 protein is illustrated at the bottom.
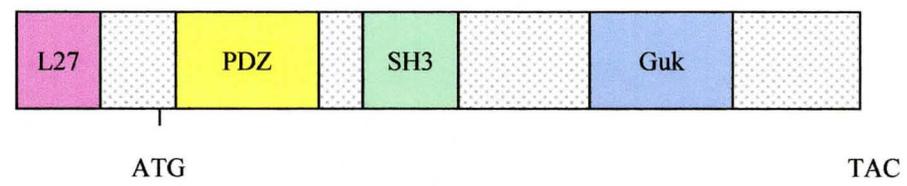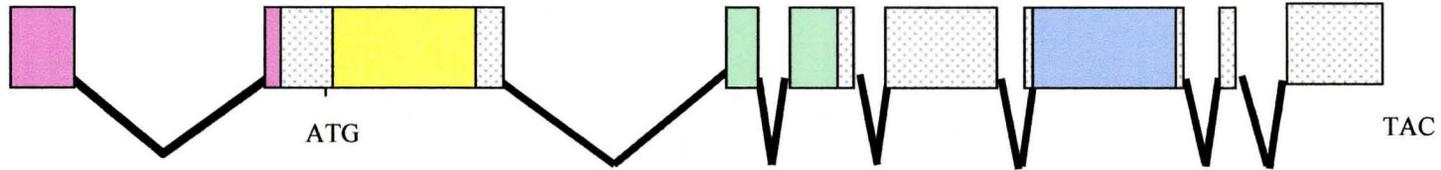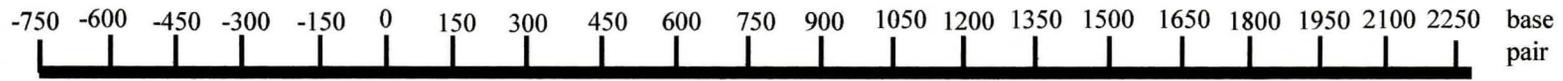
**Appendix 6: Nucleotide sequences of the three predicted transcripts of *Drosophila*
PALS2.** There are three predicted transcripts *dPALS2*, designated by flybase as b, c, and
d. All transcripts contain a PDZ (purple), SH3 (blue) and GuK (red) domain within the
cDNA and also contain an L27 domain (green) in the 5' UTR. The L27 domain appears
to not be translated due to stop codons (orange) located at –606, -423 and -303 in
transcripts b and d and –465, -417, and –330 in transcript c. The multiple start codons in
the 5' UTR are marked in pink. The actual start site for *dPALS2* transcription is located
after the L27 domain and before the PDZ domain (dark red). Transcripts b and d differ
from transcript c in the 5'UTR due to the presence of variable splice sites. The 5' UTR
of transcript c is shorter than transcript b and d. However the cDNA of transcript b and c
is identical and differ from transcript d by a 63 nucleotide absence between the PDZ and
SH3 domains. Therefore the cDNA of transcript b and c is 1407 base pairs in length and
the cDNA of transcript d is 1344 base pairs in length.

```
Transcriptb  AGTTAGCCTCTTTCTCACAAACAAGTCGGTTGTTTTTGCGCTGTTGGAGGAGT
Transcriptd  AGTTAGCCTCTTTCTCACAAACAAGTCGGTTGTTTTTGCGCTGTTGGAGGAGT
Transcriptc  -----------------------------------------------------

Transcriptb  GAAATTCTTAGCCGCGCTATTGTTTTTGTTGTTTGCGTGGCGTTGTGTTTATA
Transcriptd  GAAATTCTTAGCCGCGCTATTGTTTTTGTTGTTTGCGTGGCGTTGTGTTTATA
Transcriptc  ----------------------------------------------------

Transcriptb  ACAATAACAACAAGTGATTATCACTGGAAGACATGTTCTGTTCAAGTAGCTAA
Transcriptd  ACAATAACAACAAGTGATTATCACTGGAAGACATGTTCTGTTCAAGTAGCTAA
Transcriptc  ----------------------------------------------------

Transcriptb  GCATATGGCTGGCTAAGAAAAAAAACACAAGGGTTATAACCCTTTTTTGGCCG
Transcriptd  GCATATGGCTGGCTAAGAAAAAAAACACAAGGGTTATAACCCTTTTTTGGCCG
Transcriptc  ----------------------------------------------------

Transcriptb  AAATACGAGACTGTCCTCCGGGGTTACTTCATTCGGTCTGTTATTTTGGCCAC
Transcriptd  AAATACGAGACTGTCCTCCGGGGTTACTTCATTCGGTCTGTTATTTTGGCCAC
Transcriptc  ----------------------------------------------------

Transcriptb  CTCCCAGTTCCTTCAAAATGGTTCGGTGGAGCGCGCGGTCGCGGCGCCAAGCT
Transcriptd  CTCCCAGTTCCTTCAAAATGGTTCGGTGGAGCGCGCGGTCGCGGCGCCAAGCT
Transcriptc  ---------------------------------------------------GT

Transcriptb  CGCCAGGATAAAGTGGAGCTATTGGCCAGAAACAACAAAGTGAATGGCGAGGA
Transcriptd  CGCCAGGATAAAGTGGAGCTATTGGCCAGAAACAACAAAGTGAATGGCGAGGA
Transcriptc  AACTTGGCTCACGCGGAGCTTTTTTCACGG--CAAGCCAGCGGCAAGCGA--A

Transcriptb  CGATACCTCCGATAACGCCGCCTTTCGCAACTCGACTGACCTATCCGACCACG
Transcriptd  CGATACCTCCGATAACGCCGCCTTTCGCAACTCGACTGACCTATCCGACCACG
Transcriptc  CTGAGCAGCAAACAAAGCTAGC--------CCTAACTTGCTCACTAGCTCTCG

Transcriptb  ATGAGATCTTCCTCAAGGGATTGCTGCGTAGTAACTCCAACACGCCCCACAAG
Transcriptd  ATGAGATCTTCCTCAAGGGATTGCTGCGTAGTAACTCCAACACGCCCCACAAG
Transcriptc  CCG---CTTGCTTCCATAGCTCGTGAGTCAGTTA---------GTTGCGAACA

Transcriptb  GAACTGATGGCAACTGAATCCGACCGAGCCGCAGCCAGTGCCACTATTCCTGC
Transcriptd  GAACTGATGGCAACTGAATCCGACCGAGCCGCAGCCAGTGCCACTATTCCTGC
Transcriptc  GTGCTAA-----GCTGAATCCGACCGAGCCGCAGCCAGTGCCACTATTCCTGC

Transcriptb  CCGCTCATCTGAACAACAAGCCCATCTGCGACGACATCATCCGCAAGTTCTCA
Transcriptd  CCGCTCATCTGAACAACAAGCCCATCTGCGACGACATCATCCGCAAGTTCTCA
Transcriptc  CCGCTCATCTGAACAACAAGCCCATCTGCGACGACATCATCCGCAAGTTCTCA

Transcriptb  CCCTCGAGGCGCCTGGAGTCGCGAGAGCTCGCCAAGCTTCTCGCCCAGCCGCA
Transcriptd  CCCTCGAGGCGCCTGGAGTCGCGAGAGCTCGCCAAGCTTCTCGCCCAGCCGCA
Transcriptc  CCCTCGAGGCGCCTGGAGTCGCGAGAGCTCGCCAAGCTTCTCGCCCAGCCGCA

Transcriptb  TTTTCGTGCTCTTTTGCGCGCCCATGATGAGATAGGAGCACTCTACGAGCAA
Transcriptd  TTTTCGTGCTCTTTTGCGCGCCCATGATGAGATAGGAGCACTCTACGAGCAA
Transcriptc  TTTTCGTGCTCTTTTGCGCGCCCATGATGAGATAGGAGCACTCTACGAGCAA

Transcriptb  CGACTTAAAGCTGCCGGGGGATCCACCAGCCAACTAGAGATCGCCAGCCAAC
Transcriptd  CGACTTAAAGCTGCCGGGGGATCCACCAGCCAACTAGAGATCGCCAGCCAAC
Transcriptc  CGACTTAAAGCTGCCGGGGGATCCACCAGCCAACTAGAGATCGCCAGCCAAC
```

```
Transcriptb   GCCAAACGGGAGGCTACCTGTTCACCGAGGACGTCCTCAACACCAAGATGCC
Transcriptd   GCCAAACGGGAGGCTACCTGTTCACCGAGGACGTCCTCAACACCAAGATGCC
Transcriptc   GCCAAACGGGAGGCTACCTGTTCACCGAGGACGTCCTCAACACCAAGATGCC

Transcriptb   AGTGGAGACCATCAAGATGGTGGGTCTGCGCCGAGATCCCAGCAAGCCGCTG
Transcriptd   AGTGGAGACCATCAAGATGGTGGGTCTGCGCCGAGATCCCAGCAAGCCGCTG
Transcriptc   AGTGGAGACCATCAAGATGGTGGGTCTGCGCCGAGATCCCAGCAAGCCGCTG

Transcriptb   GGCCTGACCGTCGAACTGGACGAATTCAAGCAGCTGGTCGTGGCCAGGATTC
Transcriptd   GGCCTGACCGTCGAACTGGACGAATTCAAGCAGCTGGTCGTGGCCAGGATTC
Transcriptc   GGCCTGACCGTCGAACTGGACGAATTCAAGCAGCTGGTCGTGGCCAGGATTC

Transcriptb   TGGCGGGCGGGGTGATCGACAAACAGAGCATGCTGCACGTTGGCGATGTCAT
Transcriptd   TGGCGGGCGGGGTGATCGACAAACAGAGCATGCTGCACGTTGGCGATGTCAT
Transcriptc   TGGCGGGCGGGGTGATCGACAAACAGAGCATGCTGCACGTTGGCGATGTCAT

Transcriptb   CCTAGAGGTGAACGGTACGCCCGTTCGCACTCCCGATGAGCTGCAGGTGGAG
Transcriptd   CCTAGAGGTGAACGGTACGCCCGTTCGCACTCCCGATGAGCTGCAGGTGGAG
Transcriptc   CCTAGAGGTGAACGGTACGCCCGTTCGCACTCCCGATGAGCTGCAGGTGGAG

Transcriptb   GTGTCGCGGGCCAAGGAGAATCTCACCCTCAAGATCGGGCCGAACGTGGACG
Transcriptd   GTGTCGCGGGCCAAGGAGAATCTCACCCTCAAGATCGGGCCGAACGTGGACG
Transcriptc   GTGTCGCGGGCCAAGGAGAATCTCACCCTCAAGATCGGGCCGAACGTGGACG

Transcriptb   AGGAGATCAAGAGCGGTCGCTATACTGTGAGTGGGGGTCAGGTAAAACAGAA
Transcriptd   AGGAGATCAAGAGCGGTCGCTATACT-------------------------
Transcriptc   AGGAGATCAAGAGCGGTCGCTATACTGTGAGTGGGGGTCAGGTAAAACAGAA

Transcriptb   TGGCATCGCGAGTCTCGAGACGGGCAAGAAACTGACGTGCTACATGCGTGCC
Transcriptd   -------------------------------------TGCTACATGCGTGCC
Transcriptc   TGGCATCGCGAGTCTCGAGACGGGCAAGAAACTGACGTGCTACATGCGTGCC

Transcriptb   CTTTTCACATACAATCCATCCGAAGATTCCTTGCTGCCATGCAGGGATATTG
Transcriptd   CTTTTCACATACAATCCATCCGAAGATTCCTTGCTGCCATGCAGGGATATTG
Transcriptc   CTTTTCACATACAATCCATCCGAAGATTCCTTGCTGCCATGCAGGGATATTG

Transcriptb   GATTGCCCTTCAAGTCGGGCGACATTTTGCAGATCATCAACGTAAAGGATCC
Transcriptd   GATTGCCCTTCAAGTCGGGCGACATTTTGCAGATCATCAACGTAAAGGATCC
Transcriptc   GATTGCCCTTCAAGTCGGGCGACATTTTGCAGATCATCAACGTAAAGGATCC

Transcriptb   CAACTGGTGGCAGGCCAAGAACATTACTGCCGAATCTGACAAAATTGGACTC
Transcriptd   CAACTGGTGGCAGGCCAAGAACATTACTGCCGAATCTGACAAAATTGGACTC
Transcriptc   CAACTGGTGGCAGGCCAAGAACATTACTGCCGAATCTGACAAAATTGGACTC

Transcriptb   ATACCATCCCAAGAGCTGGAGGAGCGGCGCAAAGCCTTCGTGGCCCCCGAGG
Transcriptd   ATACCATCCCAAGAGCTGGAGGAGCGGCGCAAAGCCTTCGTGGCCCCCGAGG
Transcriptc   ATACCATCCCAAGAGCTGGAGGAGCGGCGCAAAGCCTTCGTGGCCCCCGAGG

Transcriptb   CGGACTACGTTCACAAGATTGGCATTTGCGGAACAAGAATCTCGAAGCGAAA
Transcriptd   CGGACTACGTTCACAAGATTGGCATTTGCGGAACAAGAATCTCGAAGCGAAA
Transcriptc   CGGACTACGTTCACAAGATTGGCATTTGCGGAACAAGAATCTCGAAGCGAAA
```

```
Transcriptb   GCGCAAGACCATGTACCGATCTGTGGCCAATTGCGAGTTCGACAAGGCGGAG
Transcriptd   GCGCAAGACCATGTACCGATCTGTGGCCAATTGCGAGTTCGACAAGGCGGAG
Transcriptc   GCGCAAGACCATGTACCGATCTGTGGCCAATTGCGAGTTCGACAAGGCGGAG

Transcriptb   CTGCTGCTCTATGAGGAGGTCACGCGGATGCCGCCGTTCCGCAGGAAAACCC
Transcriptd   CTGCTGCTCTATGAGGAGGTCACGCGGATGCCGCCGTTCCGCAGGAAAACCC
Transcriptc   CTGCTGCTCTATGAGGAGGTCACGCGGATGCCGCCGTTCCGCAGGAAAACCC

Transcriptb   TGGTTCTCATCGGTGTTTCCGGGGTGGGAAGACGTACGCTCAAGAATCGGCT
Transcriptd   TGGTTCTCATCGGTGTTTCCGGGGTGGGAAGACGTACGCTCAAGAATCGGCT
Transcriptc   TGGTTCTCATCGGTGTTTCCGGGGTGGGAAGACGTACGCTCAAGAATCGGCT

Transcriptb   GATCAACAGCGATGTGGACAAGTTCGGAGCGGTCATTCCACATACCAGTCGC
Transcriptd   GATCAACAGCGATGTGGACAAGTTCGGAGCGGTCATTCCACATACCAGTCGC
Transcriptc   GATCAACAGCGATGTGGACAAGTTCGGAGCGGTCATTCCACATACCAGTCGC

Transcriptb   CCCAAGCGCGCCTTGGAGGAGAACGGCAGTAGCTACTGGTTCATGGACCGCG
Transcriptd   CCCAAGCGCGCCTTGGAGGAGAACGGCAGTAGCTACTGGTTCATGGACCGCG
Transcriptc   CCCAAGCGCGCCTTGGAGGAGAACGGCAGTAGCTACTGGTTCATGGACCGCG

Transcriptb   AGGAAATGGAAGAGGCCGTGCGGAACAACGAGTTCCTGGAGTACGGCGAGCA
Transcriptd   AGGAAATGGAAGAGGCCGTGCGGAACAACGAGTTCCTGGAGTACGGCGAGCA
Transcriptc   AGGAAATGGAAGAGGCCGTGCGGAACAACGAGTTCCTGGAGTACGGCGAGCA

Transcriptb   CAATGGCAATCTGTACGGCACGCATTTGCAGTCCATCAAGGATGTGATCAAC
Transcriptd   CAATGGCAATCTGTACGGCACGCATTTGCAGTCCATCAAGGATGTGATCAAC
Transcriptc   CAATGGCAATCTGTACGGCACGCATTTGCAGTCCATCAAGGATGTGATCAAC

Transcriptb   AGTGGGCGCATGTGCATCCTGGATTGTGCACCGAATGCCCTGAAGATCCTGC
Transcriptd   AGTGGGCGCATGTGCATCCTGGATTGTGCACCGAATGCCCTGAAGATCCTGC
Transcriptc   AGTGGGCGCATGTGCATCCTGGATTGTGCACCGAATGCCCTGAAGATCCTGC

Transcriptb   ACAACAGCCAGGAACTGATGCCTTTTGTCATCTTTGTGGCAGCGCCTGGCAT
Transcriptd   ACAACAGCCAGGAACTGATGCCTTTTGTCATCTTTGTGGCAGCGCCTGGCAT
Transcriptc   ACAACAGCCAGGAACTGATGCCTTTTGTCATCTTTGTGGCAGCGCCTGGCAT

Transcriptb   GGAGCAGCTCAAGACCATCTATGCGGATCGCAGGGCCACGGGCTCCAACCGG
Transcriptd   GGAGCAGCTCAAGACCATCTATGCGGATCGCAGGGCCACGGGCTCCAACCGG
Transcriptc   GGAGCAGCTCAAGACCATCTATGCGGATCGCAGGGCCACGGGCTCCAACCGG

Transcriptb   AATTTATCTTTTGATCGCCAGAGTTCCATAAGATTCAGCTCCAGACGCGCCC
Transctiptd   AATTTATCTTTTGATCGCCAGAGTTCCATAAGATTCAGCTCCAGACGCGCCC
Transcriptc   AATTTATCTTTTGATCGCCAGAGTTCCATAAGATTCAGCTCCAGACGCGCCC

Transcriptb   GTACGCTCGAGTCCCTAGCGTCGTTGTATGAGGACGACGACCTTGTCGCCAC
Transcriptd   GTACGCTCGAGTCCCTAGCGTCGTTGTATGAGGACGACGACCTTGTCGCCAC
Transcriptc   GTACGCTCGAGTCCCTAGCGTCGTTGTATGAGGACGACGACCTTGTCGCCAC

Transcriptb   CGTCGAGGAGAGCAGCTTCGTCCAGCGCAAGTATGAGAAATACTTCGACATG
Transcriptd   CGTCGAGGAGAGCAGCTTCGTCCAGCGCAAGTATGAGAAATACTTCGACATG
Transcriptc   CGTCGAGGAGAGCAGCTTCGTCCAGCGCAAGTATGAGAAATACTTCGACATG

Transcriptb   GTCATCGTGAACGAGGACTTCGATGAGACGTTCCGCCAGGTGGTGGAAACGC
Transcriptd   GTCATCGTGAACGAGGACTTCGATGAGACGTTCCGCCAGGTGGTGGAAACGC
Transcriptc   GTCATCGTGAACGAGGACTTCGATGAGACGTTCCGCCAGGTGGTGGAAACGC
```

```
Transcriptb    TGGATCAGATGAGCCACGAGGAGCAGTGGGTGCCTGTCAACTGGATCTACTA
Transcriptd    TGGATCAGATGAGCCACGAGGAGCAGTGGGTGCCTGTCAACTGGATCTACTA
Transcriptc    TGGATCAGATGAGCCACGAGGAGCAGTGGGTGCCTGTCAACTGGATCTACTA


Transcriptb    GAATCCCATGACCCTTTGCAATTGCGCCTTTATGCATATTGAAGAATATATA
Transcriptd    GAATCCCATGACCCTTTGCAATTGCGCCTTTATGCATATTGAAGAATATATA
Transcriptc    GAATCCCATGACCCTTTGCAATTGCGCCTTTATGCATATTGAAGAATATATA

Transcriptb    CTATAAAAATCGAATGCCTATCATCACAATTAAAGGGATAATTCCAGTAATT
Transcriptd    CTATAAAAATCGAATGCCTATCATCACAATTAAAGGGATAATTCCAGTAATT
Transcriptc    CTATAAAAATCGAATGCCTATCATCACAATTAAAGGGATAATTCCAGTAATT

Transcriptb    GAAGCTTAAACTCGCAACACAATACAAATTTAGCGAAGAGCATTTAAAACGC
Transcriptd    GAAGCTTAAACTCGCAACACAATACAAATTTAGCGAAGAGCATTTAAAACGC
Transcriptc    GAAGCTTAAACTCGCAACACAATACAAATTTAGCGAAGAGCATTTAAAACGC

Transcriptb    TAGCGATTTCAATATTTATTGTTGATTAGACGTGTTTTGCTTGAGGAATGTA
Transcriptd    TAGCGATTTCAATATTTATTGTTGATTAGACGTGTTTTGCTTGAGGAATGTA
Transcriptc    TAGCGATTTCAATATTTATTGTTGATTAGACGTGTTTTGCTTGAGGAATGTA


Transcriptb    ATGCGTTTTATTAACTATGTTTATCTCAATTACTGTCATTTTCCTCTTTAAA
Transcriptd    ATGCGTTTTATTAACTATGTTTATCTCAATTACTGTCATTTTCCTCTTTAAA
Transcriptc    ATGCGTTTTATTAACTATGTTTATCTCAATTACTGTCATTTTCCTCTTTAAA

Transcriptb    TATAAATCATGTATATATGCAATGTTTGTATAAGTTGAAATGCGCCCTCCGT
Transcriptd    TATAAATCATGTATATATGCAATGTTTGTATAAGTTGAAATGCGCCCTCCGT
Transcriptc    TATAAATCATGTATATATGCAATGTTTGTATAAGTTGAAATGCGCCCTCCGT

Transcriptb    TTACAATTCAGTTTTCCTTAAGTATAAGCCTAATTTTACGAAACGAATTCGT
Transcriptd    TTACAATTCAGTTTTCCTTAAGTATAAGCCTAATTTTACGAAACGAATTCGT
Transcriptc    TTACAATTCAGTTTTCCTTAAGTATAAGCCTAATTTTACGAAACGAATTCGT

Transcriptb    CCATTTTACCATCTCTTGTTTTATTTTTACAAATAATGAATGATGTCAACGC
Transcriptd    CCATTTTACCATCTCTTGTTTTATTTTTACAAATAATGAATGATGTCAACGC
Transcriptc    CCATTTTACCATCTCTTGTTTTATTTTTACAAATAATGAATGATGTCAACGC

Transcriptb    ATATCTCTTCAGTTTATTATTTTAACGAATCTCCTCAGTACAACAACCTTCT
Transcriptd    ATATCTCTTCAGTTTATTATTTTAACGAATCTCCTCAGTACAACAACCTTCT
Transcriptc    ATATCTCTTCAGTTTATTATTTTAACGAATCTCCTCAGTACAACAACCTTCT

Transcriptb    GCCAATATTTTCGTGTAGTTTTTAGATTTGTTTGAATAAAATGTGTAAGTTT
Transcriptd    GCCAATATTTTCGTGTAGTTTTTAGATTTGTTTGAATAAAATGTGTAAGTTT
Transcriptc    GCCAATATTTTCGTGTAGTTTTTAGATTTGTTTGAATAAAATGTGTAAGTTT

Transcriptb    AACGAACTAGTTAAATGTATTGATGTGGAGTTAAACAAACGAAATACGAATA
Transcriptd    AACGAACTAGTTAAATGTATTGATGTGGAGTTAAACAAACGAAATACGAATA
Transcriptc    AACGAACTAGTTAAATGTATTGATGTGGAGTTAAACAAACGAAATACGAATA

Transcriptb    AATTTGGCTCAAAGAAA
Transcriptd    AATTTGGCTC-------
Transcriptc    AATTTGGCTCAAAGAAA
```

Appendix 7: Amino acid sequences of the three predicted protein products of

***Drosophila* dPALS2.** There are three resultant protein products of dPALS2 from the

predicted transcripts by flybase, designated b, c, and d. All three proteins contain a PDZ

(purple), SH3 (blue) and GuK (red) domain. Protein b and c are identical and differ from

protein d in a 21 amino acid absence between the PDZ and SH3 domains. Therefore

protein b and c are 469 amino acids in length and protein d is 448 amino acids in length.

```
proteinb  MPVETIKMVGLRRDPSKPLGLTVELDEFKQLVVARILAGGVIDKQSMLHVGDVILEVNGT
proteind  MPVETIKMVGLRRDPSKPLGLTVELDEFKQLVVARILAGGVIDKQSMLHVGDVILEVNGT
proteinc  MPVETIKMVGLRRDPSKPLGLTVELDEFKQLVVARILAGGVIDKQSMLHVGDVILEVNGT

proteinb  PVRTPDELQVEVSRAKENLTLKIGPNVDEEIKSGRYTVSGGQVKQNGIASLETGKKLTCY
proteind  PVRTPDELQVEVSRAKENLTLKIGPNVDEEIKSGRY--------------------TCY
proteinc  PVRTPDELQVEVSRAKENLTLKIGPNVDEEIKSGRYTVSGGQVKQNGIASLETGKKLTCY

proteinb  MRALFTYNPSEDSLLPCRDIGLPFKSGDILQIINVKDPNWWQAKNITAESDKIGLIPSQE
proteind  MRALFTYNPSEDSLLPCRDIGLPFKSGDILQIINVKDPNWWQAKNITAESDKIGLIPSQE
proteinc  MRALFTYNPSEDSLLPCRDIGLPFKSGDILQIINVKDPNWWQAKNITAESDKIGLIPSQE

proteinb  LEERRKAFVAPEADYVHKIGICGTRISKRKRKTMYRSVANCEFDKAELLLYEEVTRMPPF
proteind  LEERRKAFVAPEADYVHKIGICGTRISKRKRKTMYRSVANCEFDKAELLLYEEVTRMPPF
proteinc  LEERRKAFVAPEADYVHKIGICGTRISKRKRKTMYRSVANCEFDKAELLLYEEVTRMPPF

proteinb  RRKTLVLIGVSGVGRRTLKNRLINSDVDKFGAVIPHTSRPKRALEENGSSYWFMDREEME
proteind  RRKTLVLIGVSGVGRRTLKNRLINSDVDKFGAVIPHTSRPKRALEENGSSYWFMDREEME
proteinc  RRKTLVLIGVSGVGRRTLKNRLINSDVDKFGAVIPHTSRPKRALEENGSSYWFMDREEME

proteinb  EAVRNNEFLEYGEHNGNLYGTHLQSIKDVINSGRMCILDCAPNALKILHNSQELMPFVIF
proteind  EAVRNNEFLEYGEHNGNLYGTHLQSIKDVINSGRMCILDCAPNALKILHNSQELMPFVIF
proteinc  EAVRNNEFLEYGEHNGNLYGTHLQSIKDVINSGRMCILDCAPNALKILHNSQELMPFVIF

proteinb  VAAPGMEQLKTIYADRRATGSNRNLSFDRQSSIRFSSRRARTLESLASLYEDDDLVATVE
proteind  VAAPGMEQLKTIYADRRATGSNRNLSFDRQSSIRFSSRRARTLESLASLYEDDDLVATVE
proteinc  VAAPGMEQLKTIYADRRATGSNRNLSFDRQSSIRFSSRRARTLESLASLYEDDDLVATVE

proteinb  ESSFVQRKYEKYFDMVIVNEDFDETFRQVVETLDQMSHEEQWVPVNWIY
proteind  ESSFVQRKYEKYFDMVIVNEDFDETFRQVVETLDQMSHEEQWVPVNWIY
proteinc  ESSFVQRKYEKYFDMVIVNEDFDETFRQVVETLDQMSHEEQWVPVNWIY
```

Appendix 8: The alignment of PALS2 proteins across various species. The amino acid sequences of PALS2 proteins were aligned to demonstrate the similarities within the protein across different species. Each domain is designated by a colored line, L27 domain = green, PDZ domain = purple, SH3 domain = blue, GUK domain = red. The amino acids are color coded to indicate residues identical for at least four amino of the seven aligned sequences – blue if the aligned sequences include the *Drosophila* protein and purple if the aligned sequences do not.

```
mousealpha  ------------------------------------------------------------
mousebeta   ------------------------------------------------------------
humanvam1   ------------------------------------------------------------
xenopus     ------------------------------------------------------------
drosophila  ------------------------------------------------------------
mosquito    ------------------------------IFPDCANSVT------FLLLVCSP
honeybee    MKNSKSLMDIEDICGCSKLPSIIPRKELPHLRSTSVSLPLAHDEVTRCVMEKCLVTVPAA

mousealpha  MQQVLENLTELPSSTGAEEIDLIFLKGIMENPIVKSLAKAHERLEDSKLEAVSDNNLELV
mousebeta   MQQVLENLTELPSSTGAEEIDLIFLKGIMENPIVKSLAKAHERLEDSKLEAVSDNNLELV
humanvam1   MQQVLENLTELPSSTGAEEIDLIFLKGIMENPIVKSLAKARERLEDSKLEAVSDNNLELV
xenopus     MQQVLDNLTDLPTSTGAEEIDIIFLKGIMENPIVRSLAKAHERLEDTKLEAVSENNVELL
drosophila  ------------------------------------------------------------
mosquito    RKQIQDKLEEAP--APLHPVRVDNAPLLKEVIDRCSLSRNPHARELARIFRYPHFRALLE
honeybee    FMHVRDNLEELG--KVADDTDLLFLKGLLDSPVVTSLVKVQERLEDPPLHVEPVCSSVCD

mousealpha  NEILEDITPLISVDENVAELVGILKEPHFQSLLEAHDIVASKCYDSPPSSPEMNIPSLNN
mousebeta   NEILEDITPLISVDENVAELVGILKEPHFQSLLEAHDIVASKCYDSPPSSPEMNIPSLNN
humanvam1   NEILEDITPLINVDENVAELVGILKEPHFQSLLEAHDIVASKCYDSPPSSPEMNNSSINN
xenopus     NDIITDISSFEDEDKNVQELIGILKEPHFQSLLEAHDIVASKCYDSPPSSPEVN--SSGN
drosophila  ---------------------------------------------------------MPV
mosquito    AHDEIGETQLEKLKHPPTSATSSPGGNSQYHQASPSSETVDGSG-VELLTIN----DMPG
honeybee    IVDEVCHALRSSRDENARELVRLLRSSHLKALLETHDAVVERKE-APPSKPEPSLLAMPT

mousealpha  Q-LPVDAIRILGIHKKAGEPLGVTFRVEN-NDLVIARILHGGMIDRQGLLHVGDIIKEVN
mousebeta   Q-LPVDAIRILGIHKKAGEPLGVTFRVEN-NDLVIARILHGGMIDRQGLLHVGDIIKEVN
humanvam1   QLLPVDAIRILGIHKRAGEPLGVTFRVEN-NDLVIARILHGGMIDRQGLLHVGDIIKEVN
xenopus     QIAPVDAIRMVGIHKRTGEPLGVTFKVEN-NNLVIARILHGGMIDRQGLLHVGDIIKEIN
drosophila  -----ETIKMVGLRRDPSKPLGLTVELDEFKQLVVARILAGGVIDKQSMLHVGDVILEVN
mosquito    -----ETIKMVGIRRNPDEPLGLTVEVDEHNQLVVARIIAGGMIDRQGLLHPGDVILEVN
honeybee    N-ERMEAVRVVGLRRQPDEPLGLTVQVNESGNLIIARILGGSTAARQGLLRTGEVILEVN

mousealpha  GHEVGNNPKELQELLKNISGSVTLKILPSYRDTITPQQ--------------------
mousebeta   GHEVGNNPKELQELLKNISGSVTLKILPSYRDTITPQQSYVNMERHPAH---------VR
humanvam1   GHEVGNNPKELQELLKNISGSVTLKILPSYRDTITPQQ--------------------
xenopus     GHDVGNNPKELQELLKSISGSVTLKILPSYKDTVSPQQ--------------------
drosophila  GTPVRT-PDELQVEVSRAKENLTLKIGPNVDEEIKSGRYTVSGGQVKQNGIASLETGKKL
mosquito    GVPVTT-PEELQGEISVAKESVTLKIGPSVEEEMKSARITMAGGQVKNG--RNLDSGKKL
honeybee    GKEVHN-PEELQEAIHEAKENLSLKLAPGIATDGNRP-------------------VKS

mousealpha  -VFVKCHFDYNPFNDNLIPCKEAGLKFSKGEILQIVNREDPNWWQASHVKEGG--SAGLI
mousebeta   QVFVKCHFDYNPFNDNLIPCKEAGLKFSKGEILQIVNREDPNWWQASHVKEGG--SAGLI
humanvam1   -VFVKCHFDYNPYNDNLIPCKEAGLKFSKGEILQIVNREDPNWWQASHVKEGG--SAGLI
xenopus     -VFVKCHFDYNPFSDNLIPCKEAGLKFSKGEILHIVNREDPNWWQASHVKEGG--SAGLI
drosophila  TCYMRALFTYNPSEDSLLPCRDIGLPFKSGDILQIINVKDPNWWQAKNITAES-DKIGLI
mosquito    TCYMRALFDYDPNEDNLLPCKEIGLSFLRGDILQIINVKDPNWWQAK-HAGED-GPTGLI
honeybee    TCYMRALFDYDPSEDTLLPCREIGLPFQKGDVLQIVDQADPNWWQARRVEGEGLGPPGLI

mousealpha  PSQFLEEKRKAFVR--RDWDNSGPFCGTISNKKKKKMMYLTTRNAEFDRHEIQIYEEVAK
mousebeta   PSQFLEEKRKAFVR--RDWDNSGPFCGTISNKKKKKMMYLTTRNAEFDRHEIQIYEEVAK
humanvam1   PSQFLEEKRKAFVR--RDWDNSGPFCGTISSKKKKKMMYLTTRNAEFDRHEIQIYEEVAK
xenopus     PSQFLEEKRKAFVR--RDWDGSGQFCGTVTSKKKKKMMYLTTRNAEFDRHEIQIYEEVAR
drosophila  PSQELEERRKAFVAPEADYVHKIGICGTRISKRKRKTMYRSVANCEFDKAELLLYEEVTR
```

```
mosquito      PSQELEERRQAYVPPEADFVHKIGICGTRISKKKRKILYKTKQNSEFDKADLMLYEEVTK
honeybee      PSLELEERRKAFVPPEADFVHKISICGTKISKKKRKMYQSKSNGEFDSAELLLYEEVAR

mousealpha    MPPFQRKTLVLIGAQGVGRRSLKNRFIVLNPARFGTTVPFTSRKPREDEKDGQAYKFVSR
mousebeta     MPPFQRKTLVLIGAQGVGRRSLKNRFIVLNPARFGTTVPFTSRKPREDEKDGQAYKFVSR
humanvam1     MPPFQRKTLVLIGAQGVGRRSLKNRFIVLNPTRFGTTVPFTSRKPREDEKDGQAYKFVSR
xenopus       MPPFQRKTLVLIGAQGVGRRSLKNRLIVLNPTQFGTTIPFTSRKPKEDEKDGHAYRFVSR
drosophila    MPPFRRKTLVLIGVSGVGRRTLKNRLINSDVDKFGAVIPHTSRPKRALEENGSSYWFMDR
mosquito      MPPFKRKTLVLVGVAGVGRRTLKNRLINSDPDKFGSVLPHTSRQPRPLEESGKAYWFTDR
honeybee      MPPFRRKTLALVGARGVGRRTLKNRLINSDPEKFGTIVPYTSRPPRVLEEDGKSYWFIDR

mousealpha    SEMEADIKAGKYLEHGEYEGNLYGTKIDSILEVVQTGRTCILDVNPQALKVLRTS-EFMP
mousebeta     SEMEADIKAGKYLEHGEYEGNLYGTKIDSILEVVQTGRTCILDVNPQALKVLRTS-EFMP
humanvam1     SEMEADIKAGKYLEHGEYEGNLYGTKIDSILEVVQTGRTCILDVNPQALKVLRTS-EFMP
xenopus       VEMEADIKAGRYLEHGEYEGNLYGTKIDSIHEVVMAGRTCILDVNPQALKVLRTA-EFMP
drosophila    EEMEEAVRNNEFLEYGEHNGNLYGTHLQSIKDVINSGRMCILDCAPNALKILHNSQELMP
mosquito      EEMEQEIRENKFLEFGEHNGNLYGTHLDSIRDVIRQGKMCVLDCSPAALKTLHNSPEFMP
honeybee      ESMETDIREHRYLEYGEHGGHLYGTKLDSVRELIRAGKMCVLDCSPAALKILHNSTEFMP

mousealpha    YVVFIAAPELETLRAMHKAVVDAGITTKLLT----------------------DSDL
mousebeta     YVVFIAAPELETLRAMHKAVVDAGITTKLLT----------------------DSDL
humanvam1     YVVFIAAPELETLRAMHKAVVDAGITTKLLT----------------------DSDL
xenopus       YVVFIAAPEFETLRGMHKAVVDAGITTKLLT----------------------DSDL
drosophila    FVIFVAAPGMEQLKTIYADRR-ATGSNRNLSFDRQSSIRFSSRRARTLESLASLYEDDDL
mosquito      FVLFIAAPGMEQLKLLYSERRSASGSTRNL----------------------HDDDL
honeybee      YVIFIAAPGMEQLKWLYDLQRSTGTSSRNLTFDRQSSIRYSSRRARTLESLASLYEEDDL

mousealpha    KKTVDESARIQRAYNHYFDLIIVNDNLDKAFEKLQTAIEKLRMEPQWVPISWVY
mousebeta     KKTVDESARIQRAYNHYFDLIIVNDNLDKAFEKLQTAIEKLRMEPQWVPISWVY
humanvam1     KKTVDESARIQRAYNHYFDLIIINDNLDKAFEKLQTAIEKLRMEPQWVPISWVY
xenopus       KKTVDESARVMRAYNHYFDVSIVNDNLDKAFEKLQAAVHRLRTEQQWVPISWVY
drosophila    VATVEESSFVQRKYEKYFDMVIVNEDFDETFRQVVETLDQMSHEEQWVPVNWIY
mosquito      ISAVEESALLQRKYDKYLDMVIVNEDFDDTFRQVTEALEQLSHEHQWVPVNWIY
honeybee      KATLEESAALQRAYEKYIDLVIVNEDFDNTFRQVIAALDALATEHQWVPVNWIY
```

Appendix 9: Amino acid and nucleotide sequence of the original predicted

Drosophila **PALS2.** Originally predicted by flybase, *dPALS2* was 1728 base pairs in

length, corresponding to 576 amino acids. The cDNA and protein product contained an

L27, PDZ, SH3 and Guk domain. Domains are colour coded. Green = L27 domain,

Purple = PDZ domain, Blue = SH3 domain, and Red = GuK domain.

```
atgacatatgtacatctgaatccgaccgagccgcagccagtgccactattcctgcccgctcat
 M  T  Y  V  H  L  N  P  T  E  P  Q  P  V  P  L  F  L  P  A  H

ctgaacaacaagcccatctgcgacgacatcatccgcaagttctcaccctcgaggcgcctggag
 L  N  N  K  P  I  C  D  D  I  I  R  K  F  S  P  S  R  R  L  E

tcgcgagagctcgccaagcttctcgcccagccgcattttcgtgctcttttgcgcgcccatgat
 S  R  E  L  A  K  L  L  A  Q  P  H  F  R  A  L  L  R  A  H  D

gagataggagcactctacgagcaacgacttaaagctgccggggggatccaccagccaactagag
 E  I  G  A  L  Y  E  Q  R  L  K  A  A  G  G  S  T  S  Q  L  E

atcgccagccaacgccaaacgggaggctacctgttcaccgaggacgtcctcaacaccaagatg
 I  A  S  Q  R  Q  T  G  G  Y  L  F  T  E  D  V  L  N  T  K  M

ccagtggagaccatcaagatggtgggtctgcgccgagatcccagcaagccgctgggcctgacc
 P  V  E  T  I  K  M  V  G  L  R  R  D  P  S  K  P  L  G  L  T

gtcgaactggacgaattcaagcagctggtcgtggccaggattctggcgggcggggtgatcgac
 V  E  L  D  E  F  K  Q  L  V  V  A  R  I  L  A  G  G  V  I  D

aaacagagcatgctgcacgttggcgatgtcatcctagaggtgaacggtacgcccgttcgcact
 K  Q  S  M  L  H  V  G  D  V  I  L  E  V  N  G  T  P  V  R  T

cccgatgagctgcaggtggaggtgtcgcgggccaaggagaatctcaccctcaagatcgggccg
 P  D  E  L  Q  V  E  V  S  R  A  K  E  N  L  T  L  K  I  G  P

aacgtggacgaggagatcaagagcggtcgctatactgtgagtgggggtcaggtaaaacagaat
 N  V  D  E  E  I  K  S  G  R  Y  T  V  S  G  G  Q  V  K  Q  N

ggcatcgcgagtctcgagacgggcaagaaactgacgtgctacatgcgtgccctttttcacatac
 G  I  A  S  L  E  T  G  K  K  L  T  C  Y  M  R  A  L  F  T  Y

aatccatccgaagattccttgctgcccatgcagggatattggattgcccttcaagtcgggcgac
 N  P  S  E  D  S  L  L  P  C  R  D  I  G  L  P  F  K  S  G  D

attttgcagatcatcaacgtaaaggatcccaactggtggcaggccaagaacattactgccgaa
 I  L  Q  I  I  N  V  K  D  P  N  W  W  Q  A  K  N  I  T  A  E

tctgacaaaattggactcataccatcccaagagctggaggagcggcgcaaagccttcgtggcc
 S  D  K  I  G  L  I  P  S  Q  E  L  E  E  R  R  K  A  F  V  A

cccgaggcggactacgttcacaagattggcatttgcggaacaagaatctcgaagcgaaagcgc
 P  E  A  D  Y  V  H  K  I  G  I  C  G  T  R  I  S  K  R  K  R

aagaccatgtaccgatctgtggccaattgcgagttcgacaaggcggagctgctgctctatgag
 K  T  M  Y  R  S  V  A  N  C  E  F  D  K  A  E  L  L  L  Y  E

gaggtcacgcggatgccgccgttccgcaggaaaaccctggttctcatcggtgtttccggggtg
 E  V  T  R  M  P  P  F  R  R  K  T  L  V  L  I  G  V  S  G  V

ggaagacgtacgctcaagaatcggctgatcaacagcgatgtggacaagttcggagcggtcatt
 G  R  R  T  L  K  N  R  L  I  N  S  D  V  D  K  F  G  A  V  I

ccacataccagtcgccccaagcgcgccttggaggagaacggcagtagctactggttcatggac
 P  H  T  S  R  P  K  R  A  L  E  E  N  G  S  S  Y  W  F  M  D

cgcgaggaaatggaagaggccgtgcggaacaacgagttcctggagtacggcgagcacaatggc
 R  E  E  M  E  E  A  V  R  N  N  E  F  L  E  Y  G  E  H  N  G

aatctgtacggcacgcatttgcagtccatcaaggatgtgatcaacagtgggcgcatgtgcatc
 N  L  Y  G  T  H  L  Q  S  I  K  D  V  I  N  S  G  R  M  C  I

ctggattgtgcaccgaatgccctgaagatcctgcacaacagccaggaactgatgccttttgtc
 L  D  C  A  P  N  A  L  K  I  L  H  N  S  Q  E  L  M  P  F  V

atctttgtggcagcgcctggcatggagcagctcaagaccatctatgcggatcgcaggccacg
 I  F  V  A  A  P  G  M  E  Q  L  K  T  I  Y  A  D  R  R  A  T

ggctccaaccggaatttatcttttgatcgccagagttccataagattcagctccagacgcgcc
 G  S  N  R  N  L  S  F  D  R  Q  S  S  I  R  F  S  S  R  R  A

cgtacgctcgagtccctagcgtcgttgtatgaggacgacgaccttgtcgccaccgtcgaggag
 R  T  L  E  S  L  A  S  L  Y  E  D  D  D  L  V  A  T  V  E  E

agcagcttcgtccagcgcaagtatgagaaatacttcgacatggtcatcgtgaacgaggacttc
 S  S  F  V  Q  R  K  Y  E  K  Y  F  D  M  V  I  V  N  E  D  F

gatgagacgttccgccaggtggtggaaacgctggatcagatgagccacgaggagcagtgggtg
 D  E  T  F  R  Q  V  V  E  T  L  D  Q  M  S  H  E  E  Q  W  V

cctgtcaactggatctac
 P  V  N  W  I  Y
```

Appendix 10: A map of Potential Protein binding partners of dPALS2. A group of researchers from CuraGen Corporations, Wayne State University and Yale University proposed a map of protein interactions in *Drosophila melanogaster*. A mass Yeast two-hybrid screen using all known or predicted transcripts (10 623) against DNA libraries produced a draft map of 7048 proteins and 20 405 interactions (Giot *et al.*, 2003). dPALS2 (CG9326) interacts with dVELI (CG7662), a predicted protein binding partner along with putative genes CG9097, CG8288, CG5273, CG12019. Interaction confidences, indicated in the chart, have been rated on a scale of 0 (lowest) to 1 (highest). High confidence is 0.8+, moderate is 0.5-0.8, low is below 0.5. CG9097, also known as bric-a-brac1 and VELI show fairly high to moderate confidence of protein interaction with dPALS2, while the other proteins show moderate –low confidence of protein interaction.
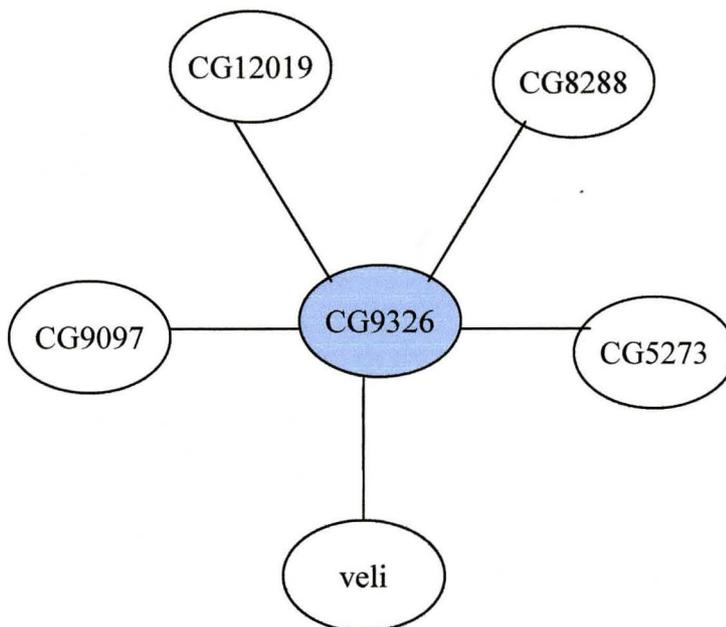
Name - CG - FBGN	Confidence
bab1 - CG9097 - FBgn0004870	0.867
veli - CG7662 - FBgn0039269	0.720
mRpL3 - CG8288 - FBgn0030686	0.542
CG5273 - FBgn0040382	0.542
Cdc37 - CG12019 – FBgn0011573	0.542