Summer 2016

Identification of Collagen IV Associated Proteins in Drsophila Using Genetics and Mass Spectrometry

Mayank S. Kapadia
Western Kentucky University, mayank.kapadia692@topper.wku.edu

Follow this and additional works at: http://digitalcommons.wku.edu/theses

Part of the Biochemistry Commons, Cell and Developmental Biology Commons, and the Molecular Biology Commons

IDENTIFICATION OF COLLAGEN IV ASSOCIATED PROTEINS IN *DROSOPHILA*
USING GENETICS AND MASS SPECTROMETRY

A Thesis
Presented to
The Faculty of the Department of Biology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Mayank S Kapadia

August 2016

IDENTIFICATION OF COLLAGEN IV ASSOCIATED PROTEINS IN *DROSOPHILA*
USING GENETICS AND MASS SPECTROMETRY

Date Recommended 07/21/16
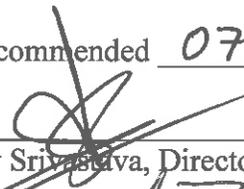
_____
Dr. Ajay Srivastava, Director of Thesis
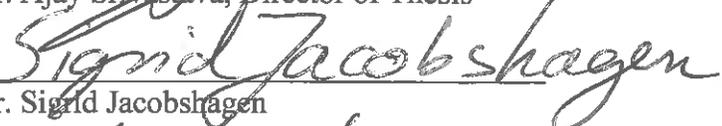
_____
Dr. Sigrid Jacobshagen

_____
Dr. Michael Smith

_____   7/28/16
Dean, Graduate Studies and Research   Date
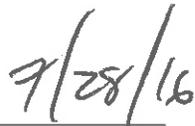
ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

IDENTIFICATION OF COLLAGEN IV ASSOCIATED PROTEINS IN *DROSOPHILA*
USING GENETICS AND MASS SPECTROMETRY

Mayank S Kapadia   August 2016   73 Pages

Directed by: Dr. Ajay Srivastava, Dr. Sigrid Jacobshagen, and Dr. Michael Smith

Department of Biology   Western Kentucky University

Metastatic cancer cells invade and spread to other locations by disrupting the basement membrane (BM). The membrane plays a major role during the normal development of an organism as well. In order to understand the invasion mechanism it is important to know about the interactions occurring between the proteins of the BM during normal development. This study concentrates on isolating and identifying the major factors associated with collagen IV, a major component of BM, during the third instar larval development of *Drosophila*. Western blot and mass spectrometry analysis revealed that collagen IV associates with various growth factors, signaling molecules, and proteins that may play a role during the development of *Drosophila*. Co-localization and knockdown studies performed on a single protein found through mass spectrometry suggested a possible role of this protein in the development of *Drosophila*. Further analysis of this proteins' function will provide new insights into its developmental role and its potential role in collagen IV transport.

1. INTRODUCTION

1.1 Overview of Basement Membrane

Basement membrane (BM) is a thin membrane-like structure located at the basal side of endothelial or epithelial cells [1, 2]. It acts as a substratum for the cells to adhere on and perform various cellular functions such as migration, proliferation, and differentiation [2]. It is 50-100nm thick and composed of large insoluble materials connected with each other to form a dense mesh-like structure [3]. Major proteins that are involved in the formation of this dense structure include type IV collagen, perlecan (a heparan sulphate proteoglycan), laminin, and entactin/nidogen [1, 2] (Figure 1). Type IV collagen accounts for approximately 50% of the overall molecular mass of the BM [4]. Laminin and type IV collagen each form their own suprastructure, which are then bridged together by entactin/nidogen and perlecan to form a stable and dense sheet-like structure. Other components that are also associated with BM include collagen XV, agrin, collagen XVIII, BM90, and secreted protein acidic and rich in cysteine (SPARC) [4]. Though BM found in various tissues consists of similar structure, molecular composition of minor components makes it unique at various locations within an organism [4, 5].
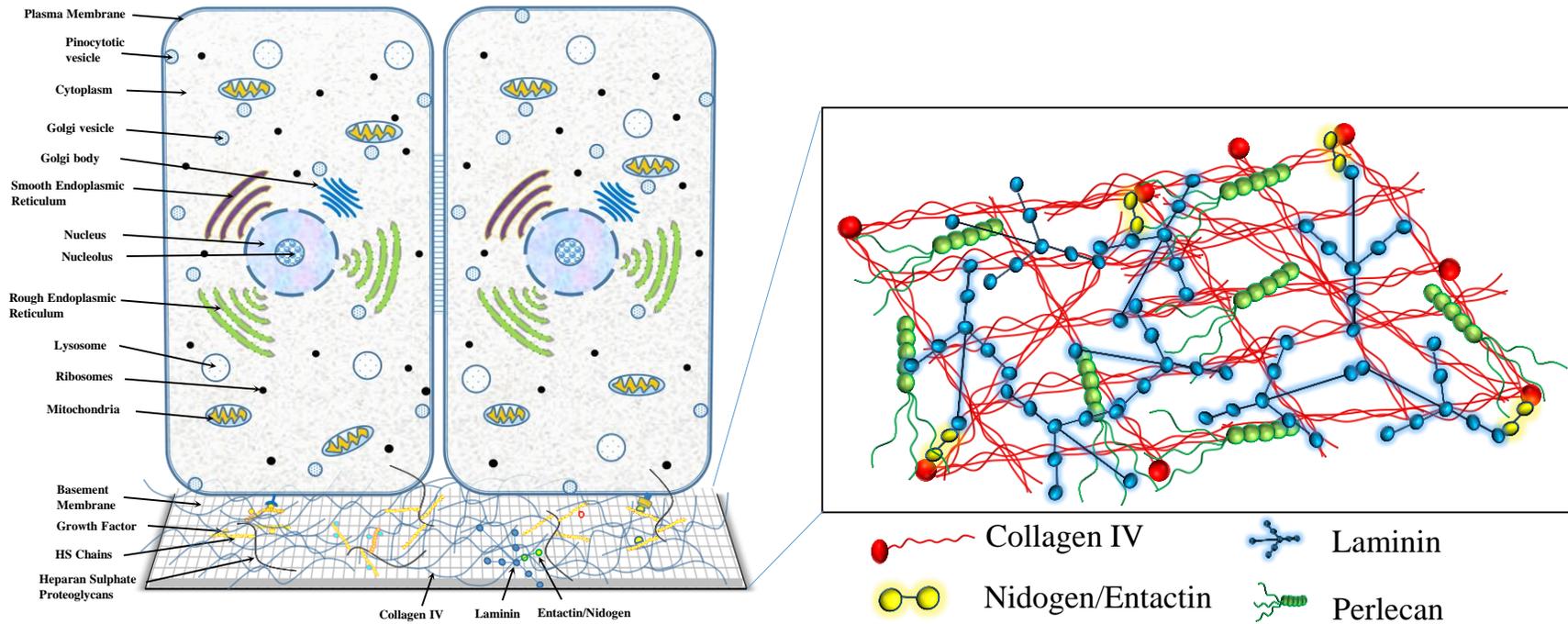
Figure 1. Basement membrane structure and composition The figure provides a graphical view of the interaction of the BM and cells. The magnified view shows the mesh-like network formation with major components of the BM. (Modified from http://www.bioon.com/book/biology/mboc/mboc.cgi@action=figure&fig=19-56.htm).

1.2 BM Major Components

Type IV collagen: Type IV collagen is a non-fibrillar 540-kDa protein with unique repeating sequence of Gly-X-Y; where X and Y represents hydroxylysine and hydroxyproline, respectively. The protein self-assembles to form a network-like structure providing BM its stability [4, 6]. Its structure consists of an α-chain polypeptide with three distinct parts - an amino terminus called 7S domain, a triple helical domain with Gly-X-Y sequence, and a carboxyl terminus called NC1 (non-collagenous) domain. The central triple helical domain consists of approximately 1,400 amino acids with 22 classical Gly-X-Y sequences while the NC1 domain is made up of 230 amino acids. The formation of network-like structure initiates with three monomer α-chains coiled to form a trimer called a Protomer (Figure 2). In mammals, six different types of α-chains (α1-α6) have been identified so far, which can form trimers in 56 different combinations [3]. These combinations have similar domain structure and show 50–70% homology in their amino acid sequence [7]. Two protomers then associate with each other through their NC1 domain to form dimers, which in turn form tetramers through linking of the 7S domain. These interactions form the nucleus for a type IV collagen scaffold. Through end-to-end and lateral associations between tetramers a type IV collagen suprastructure is formed [3].

Monomer (Single α-chain)

NC1 domain

Triple-helical domain (Gly-X-Y motif)

7S   ~1,400 aa   ~230 aa
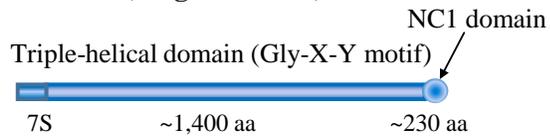
NC1 interactions promotes type IV protomer formation

Protomer (a trimer of α-chains)

Dimer   NC1 hexamer

Type IV collagen tetramer

7S domain →

NC1 hexamer
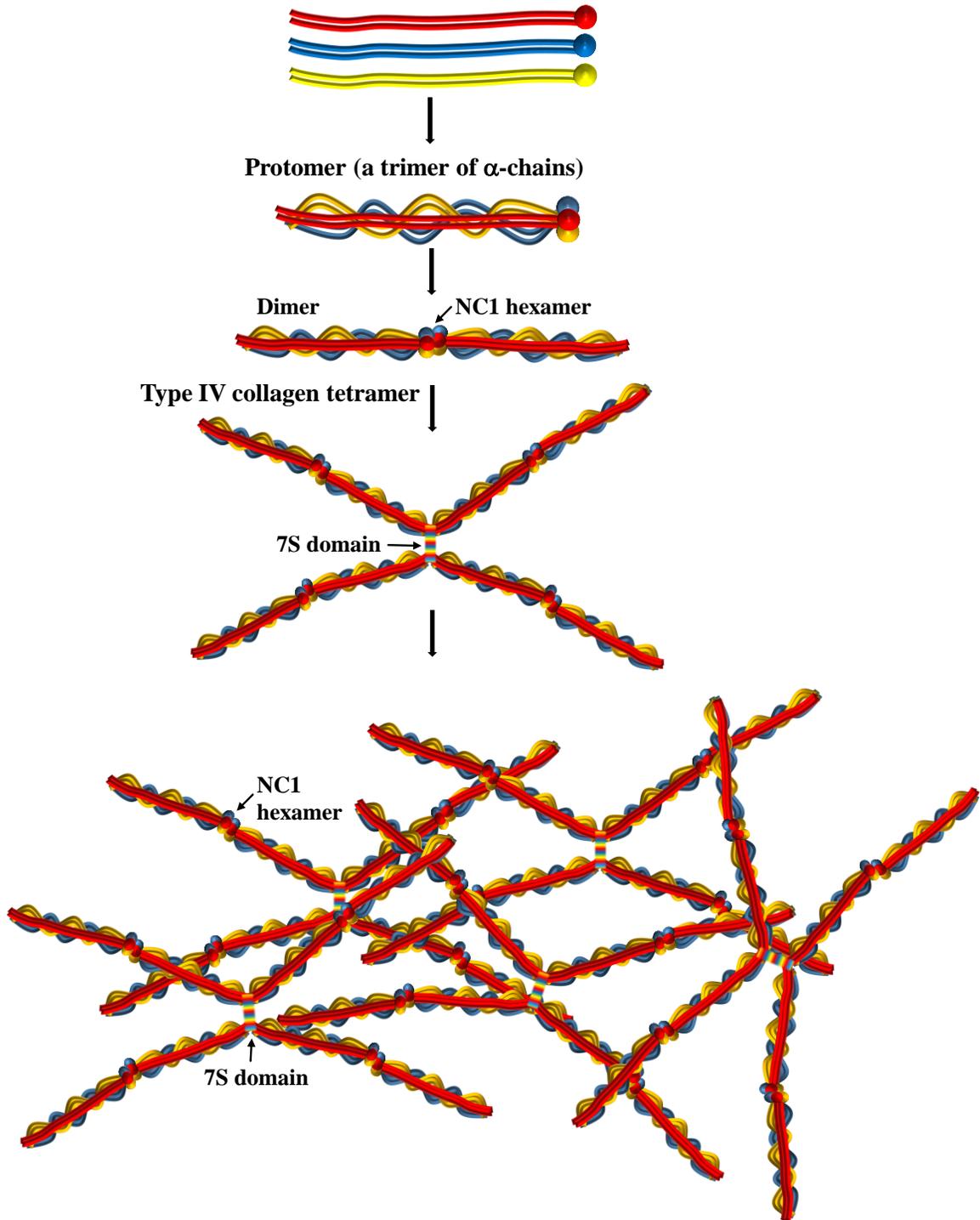
7S domain

Figure 2. Type IV collagen self-assembly. The figure shows a graphical representation of Type IV collagen protein. The protein is made up of α-chain polypeptide consisting of three domains – an amino terminal 7S domain, a triple helical domain with Gly-X-Y repeats, and a carboxyl-terminal non-collagenous (NC)-1 domain. The formation of collagen IV network initiates with three α-chains forming a trimer, also called Protomer, by connecting their NC1 domains. Two type IV collagen protomers then connect with each other through their C-terminal to form Dimers (NC1 hexamers). The next step involves the interaction of the glycosylated amino-terminal 7S region of the four protomers to form tetramers. These interactions form the nucleus of a type IV collagen, which then evolves into a type IV collagen suprastructure by end-to-end and lateral associations of the protomers. (Modified from [3]).

Laminin: Laminin is a non-collagenous protein with a size of 850-kDa [8]. Its structure consists of three polypeptide chains forming three short arms and one long arm. The arms are labeled α, β, and γ based on the sequence and protein domain organization. A total of eleven genes code for the eleven chains of laminin (α1-5, β1-3, and γ1-3). The α-chain is 400-kDa in size, while β- and γ-chains are 200-kDa in size each [5, 8, 9]. The laminin structure resembles a three-pronged fork, with C-termini of all three arms/chains acting as a handle for the fork [4, 10] (Figure 3). The α-chain C-terminus of laminin is 865-900 residues longer than the β- and γ-chains due to the presence of the laminin globular domains, also called G domains [9]. The presence of G-domain at the C-terminus helps laminin to interact with the proteins of the plasma membrane of cells; while the short arms at the N-termini are involved in interactions with other proteins of the BM [9]. Figure 3 shows various interacting sites on laminin for various proteins. *In vitro* analysis has shown that the arms interact in a domain-specific manner to form a network structure. It has also been found that collagen IV and laminins contain specific information within their amino acid sequence which guide them to initiate the self-assembly process [6]. Aumailley and colleagues showed that laminin can also interact directly with collagen IV via its short and long arms but these interactions are of low-affinity [9]. During the formation of BM, collagen IV acts as a scaffold while laminin forms the centerpiece with entactin/nidogen acting as a bridge connecting the two proteins [11, 12].

Entactin/Nidogen: Entactin is a 150-kDa tyrosine-sulphate glycoprotein. Studies of recombinant entactin/nidogen have revealed that the protein consists of three G domains (G1-G3). The G1 and G3 domains represent the N-terminus and C-terminus,

respectively; while G2 domain separates the two termini by two rod-like structures on each side [10, 12] (Figure 3). G2 domain consists of five cysteine-rich repeats with the presence of an RGD (Arg-Gly-Asp) sequence that promotes cell attachment. The C-terminal globular domain of entactin/nidogen strongly binds to laminin at the vicinity of the two arms and also interacts with collagen IV [1, 13].
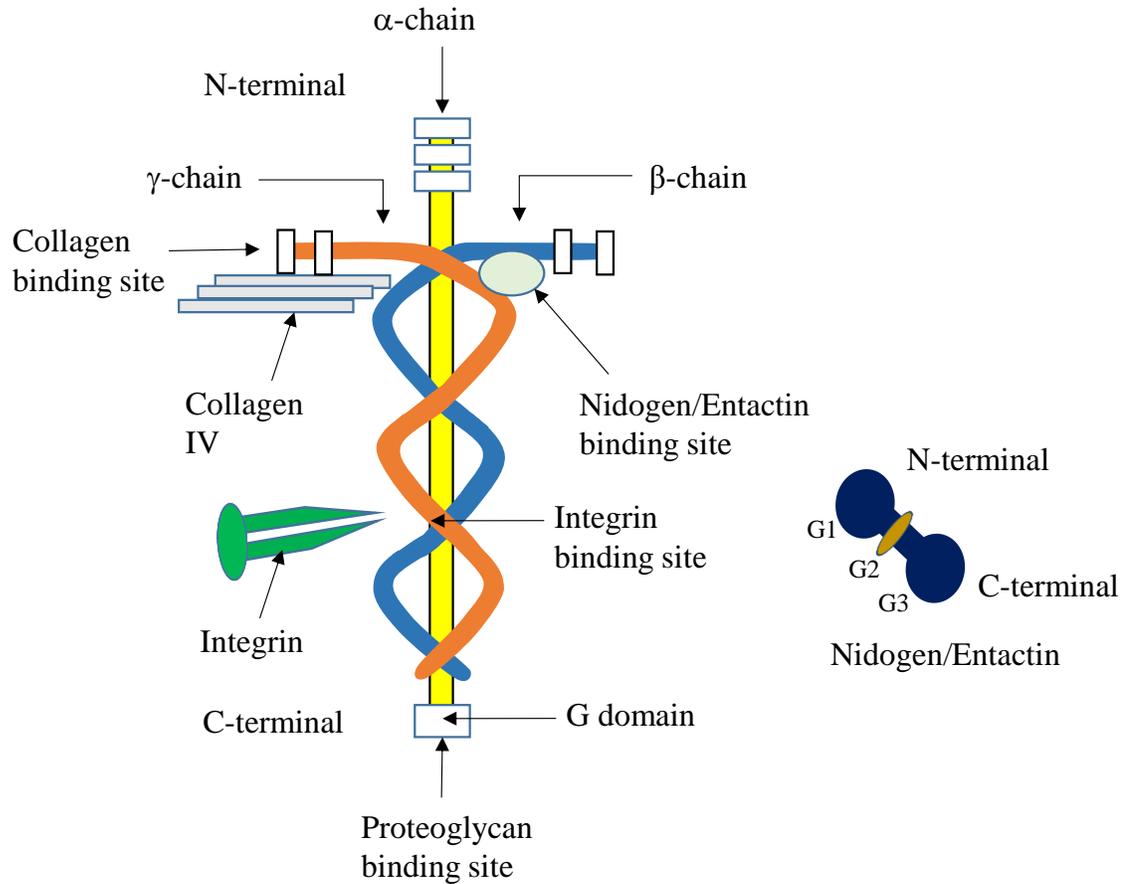
Figure 3. Structure of Laminin and Nidogen/Entactin. The laminin short arm (N-terminus) is involved in interaction with various BM proteins; while the long arm (C-terminus) is involved in interaction with cellular receptors. The C-terminus of nidogen/entactin interacts with the γ-chain of laminin (via G3) and also with collagen IV (via G2), forming a bridge between the two network proteins.

Heparan Sulphate Proteoglycans: Heparan sulphate proteoglycans (HSPGs), 400-450kDa in size, are glycoproteins with one or more covalently attached heparan sulphate (HS) glycosaminoglycan (GAG) chains [10]. There are three subfamilies of HSPGs, which include membrane-spanning proteoglycans (syndecans, betaglycan, and CD44v3), glycophosphatidylinositol-linked proteoglycans (glypicans), and extracellular matrix secreted proteoglycans (agrin, collagen XVIII, and perlecan) [14]. The presence of HS chains makes these glycoproteins very negatively charged and facilitates binding to a large number of proteins such as growth factors, receptor tyrosine kinases, chemokines and interleukins, enzymes and inhibitors, and ECM and plasma proteins [14]. HSPGs function by interaction of their HS/GAG chains with the signaling molecules or directly with different core receptor proteins. For instance, HS interacts with the fibroblast growth factors (FGFs) and their receptors resulting in a complex formation. This complex helps in lowering the concentration of FGFs required to initiate signaling and also in extending the duration of receptor response. In addition, HSPGs interact with cell-surface receptors like integrins to facilitate cell attachment, movement, and spreading. For example, fibronectin of ECM interacts with HS chains of syndecans and integrins to initiate cell movement and formation of focal adhesion [14–18]. Studies showed that all ECM proteins have binding domains for HS chains. For instance, HSPGs interact with laminin through nidogen/entactin to form a ternary complex [19]. Figure 4 provides a schematic representation of multiple interacting partners of HSPGs.
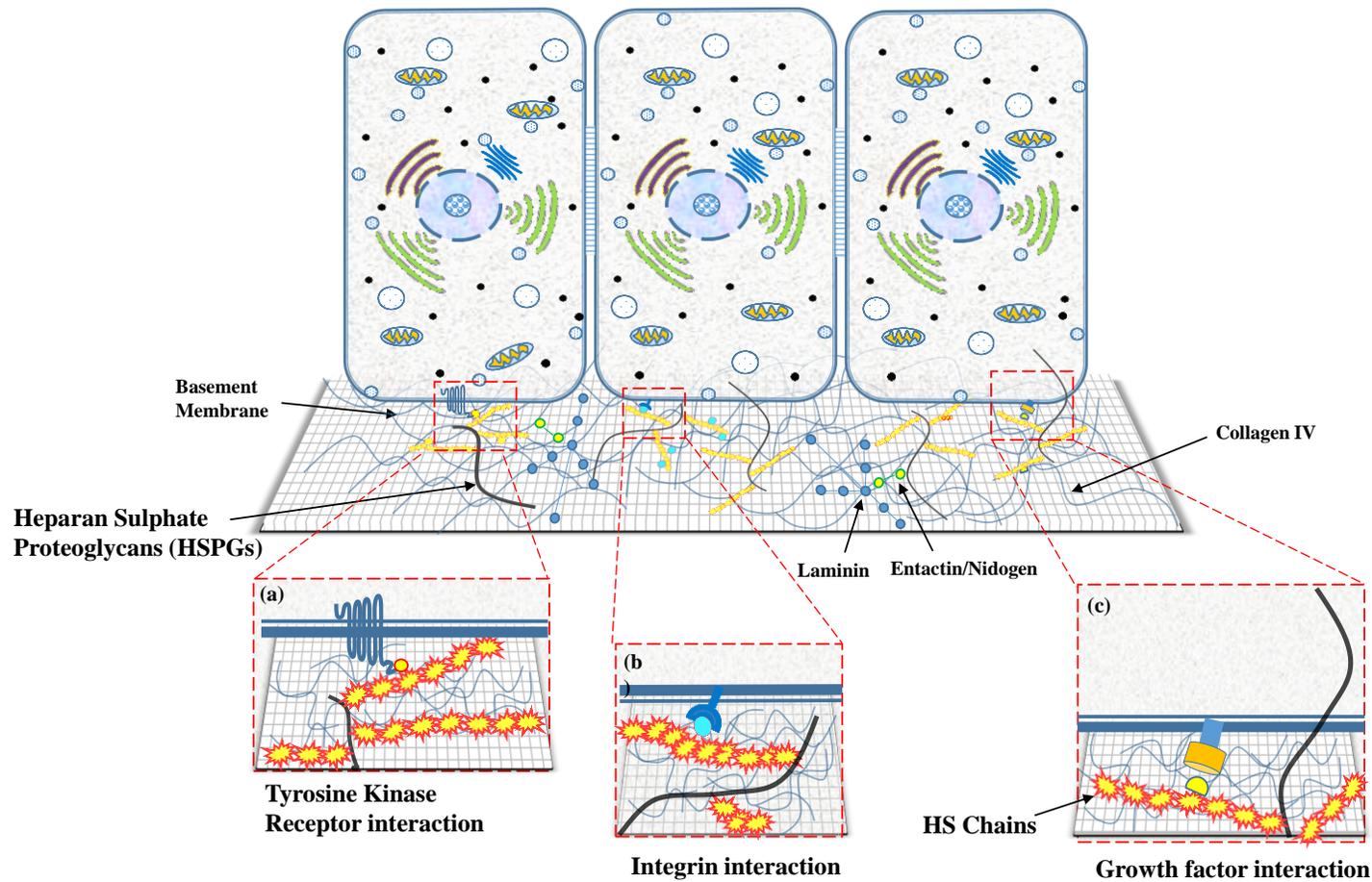
Figure 4. Multiple interacting partners of HSPGs. The figure shows HSPGs (Black S-shaped strings) interacting with laminin, entactin/nidogen or collagen IV through their HS chains (Yellow). It also interacts with receptor tyrosine kinase (a), integrins (b), and growth factors (c).

1.3 Components Associated with BM during Development

In addition to the conserved core components, BM is made up of various molecular components that differ between organisms, between tissues within an organism, and also with developmental age. These components perform various activities such as fill space between cells, act as a barrier between tissues, navigate migratory cells, provide signals to alter cell behavior, and sequester biologically active compounds such as growth factors [20].

A cell makes contact with the BM components through specialized receptor molecules present on its membrane. This contact allows the cell to adhere to the BM and in turn allows the BM to control the behavior of the cell [20]. It is also known that BM is constantly remodeled by the cells within and around it, by proteolytic enzymes, and through deposition or degradation of the BM components at various developmental stages. As a result, important changes in the cell–cell and cell–BM interactions occur generating new signals from the cell surface. This, in turn, affects gene expression and influences critical cellular behaviors such as proliferation, survival, differentiation, and motility [20, 21].

Integrins: Studies have found that cells bind to specific BM components using a variety of receptors present on their plasma membrane. These receptors include proteoglycans, lectins, and integrins [20]. For instance, α6β4 is an integrin–laminin receptor that anchors epithelium to the BM, forming a rigid structure called hemidesmosome [22, 23]. Integrins, when associated with the BM and a cell, bind with various signal-transducing molecules, including focal adhesion kinases, which on

activation phosphorylate various targets that control cell survival, differentiation, and proliferation [20].

Proteoglycans: Proteoglycans are found to be involved in promoting cell–BM adhesion. Most proteoglycans form hydrated gels by consuming water to fill space in the BM, and some, specifically the proteoglycan heparan sulfate, bind to a variety of growth factors, concentrating them in the BM and preventing their diffusion to other parts of the body [20]. Heparan sulfate chains also interact with proteins such as chemokines, morphogens, and enzymes [24]. Studies in *Drosophila*, mouse, and zebrafish have shown that the heparan sulfate chains on the HSPGs bind with various signaling molecules such as Fibroblast Growth Factors (FGFs), Hedgehogs (*Hh*), Wingless (*Wnt*) [25], Transforming Growth Factors β (TGFβs) [26], Hepatocyte Growth Factor (HGF) [27], and Heparin Binding-Epidermal Growth Factor (HB-EGF) [28]. It was also found that *Drosophila* contains genes coding for two glypican members, division abnormally delayed (*Dally*) and Dally-like (*Dly*) [29, 30]; one syndecan [31, 32]; and one perlecan [33]. Over-expression studies of glypican and syndecan in tissue culture cells restrict the response of cells to FGF [34, 35]. Perlecan, one of the largest members of the HSPG family, has been linked with signaling by heparin-dependent growth factors FGF2, Vascular Endothelial Growth Factor (VEGF), and Sonic Hedgehog (SHH) in the mammalian system [36, 37]. Perlecan is also involved in the maintenance of epithelial cell polarity by interacting with the BM receptor dystroglycan [38]. During the first instar brain development in *Drosophila*, perlecan's homolog *trol* (terribly reduced optic lobes) regulates the activity of Hedgehog and Branchless (an FGF homolog) to control the onset of stem cell proliferation [24]. HSPGs are also linked with growth factor signaling

pathways such as wingless (*wg*/*Wnt*) and decapentaplegic (*Dpp*/TGF-β). Both of these pathways take actively part along with Hh and Ras-MAPK pathways in the development of *Drosophila* eye disc and/or second instar brain [39, 40]. In addition, HSPGs are responsible for regulating Dpp movement in *Drosophila* wings [37]. Studies by Jackson and his co-workers showed that Dally is required for normal Dpp signaling during imaginal disc development. Using a Dally mutant, they also showed that Dally is responsible for altering the response of a cell to Dpp [41].

Collagen IV: Collagen IV is found to be associated with Dpp and plays a vital role during *Drosophila's* early embryonic development [42]. Dpp is a bone morphogenetic protein (BMP) signaling molecule that belongs to the TGF-β superfamily of growth factors. It is responsible for cell fates at different developmental stages of *Drosophila* and regulating the levels of Dpp signaling is vital during the development of *Drosophila*. For instance, gradients of Dpp form the anterior-posterior axis of the wing and proximal-distal axis of the leg [43, 44]. Similarly, alterations in the Dpp signaling may lead to several human diseases, including skeletal disorders, vascular diseases, and cancer [45, 46]. Collagen IV binds Dpp to inhibit the signaling of morphogen to the distant cells, thereby inhibiting the number of germline stem cells (GSC) [42, 47]. In an embryo, collagen IV promotes interaction of Dpp with heteromeric receptor complexes to form Dpp/Scw–Tsg-Sog complex. Disruption of this complex leads to reduced target gene expression and positive feedback, further decreasing subsequent signaling [42].

All these data indicate that the components of the BM play vital roles during *Drosophila* development.

1.4   Biogenesis of Collagen IV

The complete process of collagen IV, from production to its incorporation into the BM, is complex. The process starts with α chains undergoing extensive posttranslational modifications in the endoplasmic reticulum. Initially, prolines and lysines in the X and Y positions of the tripeptide repeats are hydroxylated to form hydroxyproline and hydroxylysines, respectively [48]. Hydroxylation of prolines at 3' or 4' position provides stability to the triple helix; while hydroxylation of lysine is followed by O-linked glycosylation in the ER. The triple helix is then properly folded and trimerized by chaperones and enzymes [49]. The post-translational modifications cause collagen IV protomers to be about 300nm long, which makes the transport of protein from ER to Golgi apparatus impossible into the COPII vesicles of 60-80nm in diameter. Studies performed in *Drosophila* showed that collagen IV is initially produced in soluble form and packaged into the COPII vesicle with the help of COPII cargo adaptor protein called TANGO1 (Transport And Golgi Organization) [50]. Once the protein reaches the Golgi apparatus it is further modified and transported out of the plasma membrane to form BM. Collagen IV accounts for 50% of the total BM mass and it is important that the protein is continuously supplied in the appropriate form. Hence, during the embryonic development of *Drosophila*, the requirement for collagen IV is fulfilled by hemocytes; however, the increasing need for the protein is fulfilled by the fat body during the later stages of development [51]. Literature review also showed that surf-4, an early secretory protein is involved in the transport of soluble proteins. This study provides with the insights on the role of surf-4 during the development of *Drosophila.*

1.5 Surf-4 and its Role

Three membrane organelles, namely, Endoplasmic Reticulum (ER), Endoplasmic Reticulum Golgi Intermediate Compartment (ERGIC), and Golgi apparatus constitute the early secretory pathway in higher eukaryotic cells [52]. Cycling proteins play an important role in the maintenance of these organelles by participating in proper trafficking of proteins in anterograde and retrograde directions. COPII vesicles are involved in anterograde mode of protein transport from ER to ERGIC and pleomorphic vesicles transport the cargo from ERGIC to Golgi apparatus. COPI vesicles perform the retrograde mode of transport from ERGIC or Golgi apparatus to the ER and they are mostly involved in the recycling of membrane proteins [52, 53]. Transport between these membrane organelles during the retrograde and anterograde modes is facilitated by transmembrane cargo receptors and inactivation of cargo receptors have been linked to several human diseases. For example, inactivation of ERGIC-53 leads to inefficient secretion of blood coagulation factors V and VIII resulting in provoking bleeding disorder. One of the cargo receptor that was identified in yeast to perform ER-to-Golgi apparatus transport is Erv29p. This receptor is required for efficient packaging of the glycosylated α-factor pheromone precursor into the COPII vesicles at the ER exit site [53]. The receptor is conserved among eukaryotes and the mammalian orthologue was found to be surf-4, a type of housekeeping protein. Surf-4 is part of the Surfeit locus which encodes six surf genes (*Surf-1* to *Surf-6*). All six genes differ in their amino acid sequence and have unique features such as overlapping genes, bidirectional transcriptional promoter, and CpG islands at the 5' end [54–56]. *Drosophila melanogaster Surfeit* genes are located on chromosome 3R. However, unlike in mouse

the *Surf-4* gene in *Drosophila* is not in close proximity to other *Surfeit* genes but located upstream from a gene encoding a homolog of a yeast seryl-tRNA synthetase protein [57, 58].The *Surf-4* gene encodes an integral membrane protein of 30-kDa mass with dilysine motifs at its C-terminal making it an ER resident protein [59]. The protein consists of 270 amino acids in *Drosophila* and between 250-270 amino acids in mouse, human, yeast, and *C. elegans* and is conserved between these species [58]. Knockdown studies performed by Mitrovic and her colleagues [53] showed no major effect on total protein secretion suggesting that the protein acts as a cargo receptor for a specific set of proteins.

1.6   Diseases Associated with BM Components

According to the literature, each individual component of the BM plays a vital role during the development of an organism (*Drosophila*). Hence defects in the genes expressing these components might result in severe abnormalities in an organism. Some of the diseases due to defects in the BM components are as follows:

a) Alport syndrome: It is an inherited disorder due to mutations in the *COL4A4*, *COL4A4*, or *COL4A5* gene. Mutation in any of these genes prevents the proper production/assembly of collagen IV network in the BM of kidney. This results in improper filtration of waste products from blood causing renal failure [60].

b) Knobloch syndrome: It is a rare autosomal recessive developmental disorder due to a mutation in the *COL18A1* gene. The mutation results in defects in collagen XVIII protein, which plays an important role in determining the retinal structure as well as closure of the neural tube. Mutations in the gene can also lead to occipital encephalocele and severe ocular alterations [61, 62].

c) Congenital muscular dystrophy: It is an autosomal recessive disorder due to a defect in the *LAMA2* gene. Mutation in this gene results in defects in the α2 chain of laminin. The complete and near-complete deficiency of α2 laminin chain results in severe hypotonia at birth or within the first few months of birth [63].

d) Schwartz-Jampel syndrome: It is a rare autosomal recessive skeletal dysplasia associated with myotonia. It is caused by a mutation in the HSPG2 gene which encodes perlecan. The disorder results in short stature, osteochondrodysplasia, myotonia, and a characteristic face with a fixed facial expression, blepharophimosis, pursed lips, low-set ears, and myopia [64].

1.7 *Drosophila* **as a Model Organism**

The overall goal of this research is to understand the mechanisms that cancer cells use to metastasize to different parts of the body. During the metastasis process, the Matrix Metalloproteases(MMPs) released by the cancer cells break the basement membrane at three different locations – one around its localized mass, another at the entry site into the blood stream, and finally at the exit site of the blood stream. As the current study is concentrated only on one protein of the BM, collagen IV, *Drosophila melanogaster* proves to be a perfect model organism. *Drosophila* has only two genes which code for collagen IV – *Viking* (*Vkg*) and *Collagen at 25C* (*Cg25C*) when compared to other model organisms like rats and mouse which consists of multiple genes coding for the same protein [65–67]. Knockdown studies of these genes in *Drosophila* have displayed embryonic lethality and absence of the protein has resulted in aberrant shapes in several organs during the larval stage [65, 66]. This study uses transgenic flies consisting of *Viking* gene tagged to *GFP* gene which upon expression produces a GFP-

tagged collagen IV protein. In addition, *Drosophila* has a completely sequenced genome and significant homology with the human genome exists, making it an ideal organism for understanding human biology and disease processes. Further, nearly 75% of the disease-related genes in humans have functional orthologs in the fly [68, 69]. Though the overall identity of nucleotide sequence between the fly and mammal is only approximately 40%, functionally they are 80-90% identical [68]. The fly also has a rapid life cycle producing hundreds of genetically identical offspring within 10 to 12 days at 25°C (Figure 5). Each stage of the fly growth cycle; i.e., from embryo to the adult; can be used as a model system to understand the effects caused due to mutations. The study uses late third instar stage as it helps in studying the developmental and physiological processes. In addition, during the late third instar stage the wing imaginal disc undergoes immense morphological changes which at the later stage develops into an adult fly wing. [69, 70].
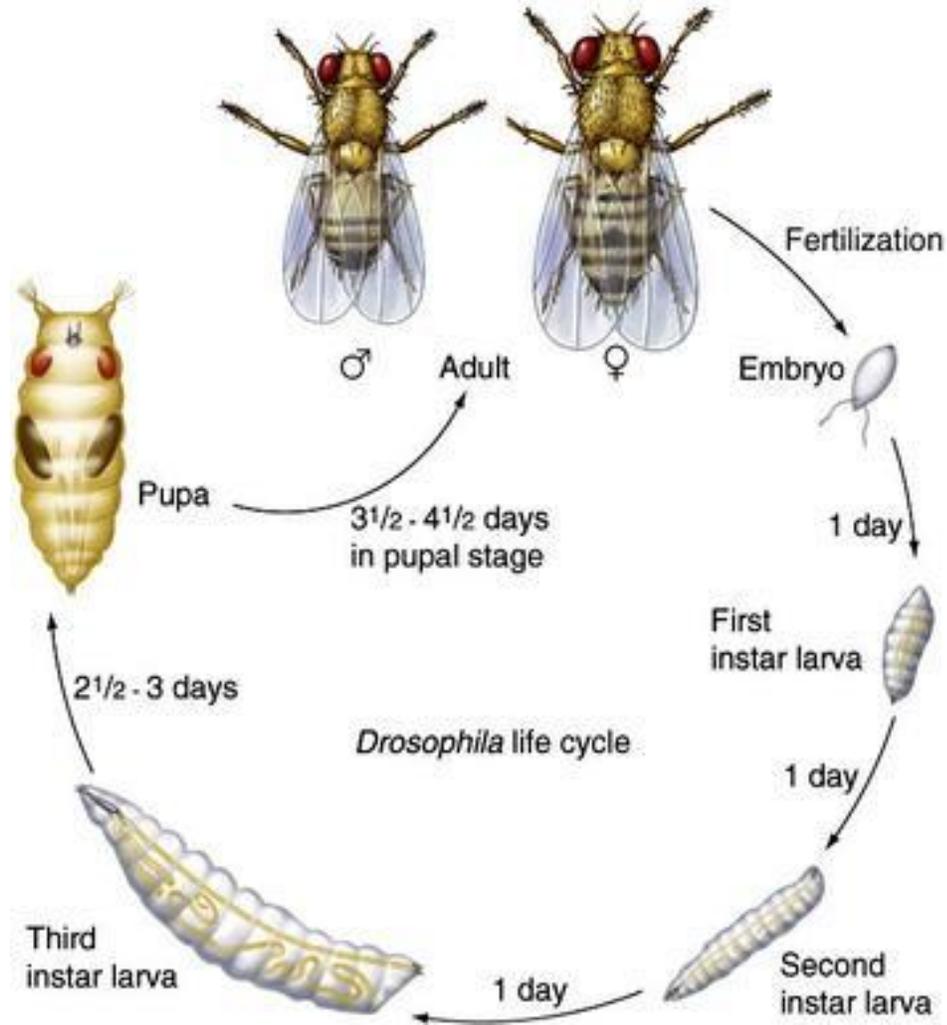
Figure 5. Lifecycle of *Drosophila melanogaster*. The lifecycle starts with adult flies mating and female flies laying eggs in the culture medium. The eggs undergo embryogenesis and hatch to form the first instar larvae in a day. The larvae continuously feed on the medium and molt to form second instar larvae in 24 hours. The process repeats for another 24 hours and again the larvae molt to produce third instar larvae. At this stage, larvae eat voraciously and finally enter the pupal stage where they undergo pupation for another 3-4 days. The pupae then metamorphose to produce adult flies on the 4th day of entering pupation and the lifecycle continues. Adopted from http://morphologicallydisturbed.weebly.com/the-biology.html.

1.8   Purpose of the Study

Malignant tumors spread to different parts of the body by disrupting the basement membrane. To understand the mechanism involved in this disruption process, it is important to identify the makeup of BM. However, BM is a complex structure with continuous modification, remodeling, and degradation involving interaction of several components during the development of an organism [71]. Therefore to understand the invasion mechanism, it is important to first know what components interact to assist in the function of BM during normal development. This study uses *Drosophila melanogaster's* late third instar larval stage to identify components/proteins that interact with type IV collagen of the BM through mass spectrometry analysis. Once these interacting partners and their relationship with collagen IV is defined, then similar studies would be performed in the cancer-bearing *Drosophila* to identify the interactive partners at the third instar larval stage. This has the potential to identify new basement membrane partners which may have a role in the cancer metastasis process.

2. MATERIALS & METHODS

2.1 Fly Stocks

Transgenic flies expressing *Actin-GFP* and *Viking-GFP* were obtained from the Srivastava lab stock collection and cultured at 25°C in the standard *Drosophila* cultured medium (LabExpress). *Viking-GFP* is a fusion protein which produces a GFP tagged collagen IV; whereas *Actin-GFP* produces a GFP protein wherever actin promoter is expressed. The study used transgenic *Actin*-GFP and *Vkg-GFP* third instar larvae for the isolation experiments. We also used transgenic flies for the *Surf-4* gene tagged with GFP for protein trap studies. For knockdown and overexpression studies of the *Surf-4* gene we used various drivers, RNAi lines, and over-expression lines. These lines were obtained from the Bloomington Drosophila Stock Center. Details of these fly stocks and their genotypes can be found in Tables 3 and 5.

2.2 Isolation Buffers

The following isolation buffers were used:

1. RIPA (50mM Tris HCl, pH 8.0; 150mM Sodium chloride; 1% NP40; 0.5% Sodium deoxycholate; 0.1% SDS; 1mM EDTA; and 1 tablet Roche's Protease inhibitor cocktail/10ml solution) [72]

2. Lysis buffer (0.5M Urea, 0.01% SDS, 2% Triton-X 100, 2mM PMSF (Phenylmethanesulfonyl fluoride), and 1 tablet Roche's Protease inhibitor cocktail (Catalogue no. 05892970001) in 1X PBS/10ml solution) [73]

3. 2X Laemmli buffer (65.8mM Tris-HCl, pH 6.8; 26.3% (w/v) Glycerol; 2.1% SDS; and 0.01% Bromophenol blue) [74]

4. MBL's IP lysis buffer (50mM Tris-HCl pH 7.2, 250mM NaCl, 0.1% NP-40, 2mM EDTA, 10% Glycerol, and 1 tablet Roche's Protease inhibitor cocktail (per 10ml) in 1X PBS; store at 4°C)

2.3   Stock Buffers

All buffer recipes mentioned below are for 1 liter, unless specified.

1. 10X PBS (8mM Na_2HPO_4, 2mM KH_2PO_4, 137mM Sodium Chloride, 27mM Potassium Chloride when diluted to 1X working concentration; pH 7.4) – Applied Biosystems (Catalogue no. 70011069)

2. 10X Running buffer (SDS-PAGE) (25mM Tris-HCl, 192mM Glycine, 0.1% w/v SDS when diluted to 1X working concentration; pH8.3) – Bio-Rad (Catalogue no. 1610734)

3. 10X Transfer buffer (Western Blotting) (25mM Tris-HCl, 192mM Glycine, 10% w/v Methanol, 20% w/v SDS when diluted to 1X working concentration)

4. 10X TBS (Western Blotting) (50mM Tris-HCl, 150mM Sodium chloride when diluted to 1X working concentration; pH 7.6)

2.4   Freshly-made Buffers

The concentrations of freshly made buffers used in this study are provided below.

1. 0.0-2.0% Formaldehyde in 1X PBS (Freshly prepared, keep at RT in dark)

2. 1.25M Glycine in 1X PBS (store at 4°C)

3. 1X PBS (pH 7.4)

4. 1X Running buffer

5. 1X Transfer buffer with 20% SDS and 10% methanol (store at 4°C – can be re-

   used)

6. 1X TBST (Tris-buffered Saline with Tween® 20)

7. 5% dried skim milk in 1X TBST (store at 4°C)

8. 2X Laemmli buffer with β-mercaptoethanol (950μl 2X Laemmli with 50μl BME

   – 1ml solution)

9. Clarity™ Western ECL Substrate (Bio-Rad Catalogue no. 1705061)

2.5   Gels and Membranes Used in the Study

1. 4–20% Mini-PROTEAN® TGX™ Precast Protein Gels (Bio-Rad catalogue no.

   4561093)

2. 7.5% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad catalogue no. 4561023)

3. Immun-Blot® PVDF membrane (Bio-Rad catalogue no. 1620239)

4. Precision Plus Protein™ Dual Color Standards (Bio-Rad catalogue no. 1610394)

2.6   Antibodies Used in the Study

1. Primary antibody (anti-GFP rabbit polyclonal antibody – Life Technologies A-

   6455) (1:2,000 dilution was prepared by mixing 1μl of antibody with 2,000μl of

   5% milk in 1X TBST)

2. Secondary antibody (Goat anti-rabbit polyclonal HRP conjugate – Jackson

   ImmunoResearch Laboratories, Inc. Catalogue No. 111-035-047) (1:20,000

   dilution was prepared by mixing 1μl of antibody with 20,000μl of 5% milk in 1X

   TBST)

3. MBL's anti-GFP mAb agarose beads (Catalogue no. D153-8)

2.7 Protein Staining Solutions

All solution recipes are for 1 liter, unless specified.

1. Gel-fixing solution – (50% Ethanol in deionized water, 10% acetic acid)

2. Coomassie stain – (Coomassie Brilliant Blue R250 2.5gm; Glacial acetic acid 100ml, Methanol: Deionized water (1: 1 v/v) 900ml)

3. Destaining solution – (Glacial acetic acid 100ml, Methanol: Deionized Water (1: 1 v/v) 900ml)

4. Ponceau S stain – 30ml (0.3ml glacial acetic acid, 0.033gm Ponceau S, 30ml Deionized water)

2.8 Sample Preparation

Viking-GFP-tagged flies were cultured using standard culture medium (LabExpress Fly Food B) and incubated at 25°C. The late third instar larvae were collected on the 5th/6th day of culture. *Actin-GFP*-tagged flies were cultured in the same way and used as control. The use of *Actin-GFP* larvae as a control aided in eliminating any non-specific detection of proteins, which might have bound to GFP during the normal development.

2.9 Protein Purification

2.9.1 Isolation and detection of proteins

Three different buffers (RIPA, Lysis buffer, and 2X Laemmli) were tested for efficient isolation of collagen IV and GFP proteins from the third instar larvae (Figure 6). I was interested in detection of a band at 250-kDa for collagen IV and at 27-kDa for GFP proteins in Western blot analysis.

2.9.1.1 RIPA buffer extraction process

Ten wandering third instar larvae of *Actin-GFP* and *Vkg-GFP* each were collected in separate microcentrifuge tubes and 500µl of freshly made RIPA buffer was added. Lysate was prepared by homogenizing the larvae using pestle and centrifuged at 12,000g for 10 minutes at 4°C. The supernatant was collected in a fresh microcentrifuge tube and stored at -20°C until used. SDS-PAGE analysis was performed by mixing 60µl of sample with 60µl of sample buffer (2X Laemmli with β-mercaptoethanol). After mixing, the samples were boiled at 95°C for 10 minutes and 30µl was loaded to the 4-20% Mini-Protean TGX precast polyacrylamide gradient gels for optimal resolution. Three microliters of Bio-Rad's Precision Plus Protein Dual Color Standard was used as ladder. The gels were run at 120 Volts in 1X running buffer for approximately 1 hour 20 minutes. The proteins were then transferred to the PVDF membrane using Western blot wet transfer method with transfer carried out at 40mAmp at 4°C for 18 hours. Once transfer was complete, the PVDF membrane was air dried and presence of protein was confirmed by staining with Ponceau S stain. After successful detection of proteins bands the membrane was washed with Deionized water until stained was completely washed. The membrane was incubated with 5% milk TBST solution for an hour on a rotary shaker at room temperature. The membrane was then incubated with primary antibodies (1:2000) overnight on a shaker at 4°C. Subsequently, the membrane was washed 4 times with 1X TBST for 10 minutes each. After washing, the membrane was incubated with secondary antibodies (1:20,000) for 1 hour at room temperature. The membrane was washed 4 times with 1X TBST for 10 minutes each and then incubated with ECL reagent (1:1 ratio of Clarity Western Peroxide and Western Luminol/Enhancer Reagents kit) for a

few second and observed for chemiluminescence in imager. The desired image was captured using attached camera.

2.9.1.2 *Lysis buffer extraction process*

The isolation of proteins using this buffer followed the same procedure as mentioned in Section 2.9.1.1; however, there is only one modification that was performed for this isolation process. The boiling step before loading samples to the SDS-PAGE was avoided as the isolation buffer contains urea. Boiling samples at elevated temperature (95°C) causes carbamylation and results in protein precipitation.

2.9.1.3 *2X Laemmli buffer extraction process*

The isolation of proteins using this buffer followed the same procedure as mentioned in Section 2.9.1.1.

2.10 Cross-linking

The cross-linking step was performed only with 2X Laemmli buffer as RIPA and Lysis buffers did not provide workable amounts of protein after extraction. To perform cross-linking, different concentrations (0.0 to 2.0% at increments of 0.2%) of formaldehyde were used to find a correct concentration for protein cross-linking. Four larvae each of *Actin-GFP* and *Vkg-GFP* were collected in 1.5μl microcentrifuge tubes and incubated with 500μl of different formaldehyde concentrations for 7 and 17 minutes on a vortex mixer. The samples were then centrifuged at 20,000g for 3 minutes at room temperature. This made a total of 10 and 20 minutes exposure to formaldehyde. The formaldehyde was quickly removed and 400μl of ice-cold 1.25M glycine was added to quench the reaction. The tubes were then centrifuged at 20,000g for 5 minutes at room

temperature. The supernatant was discarded and the quenching step was repeated with centrifugation carried out at 4°C. The glycine was removed [Protocol modified from Ref. 72] and 200μl of appropriate buffer (2X Laemmli) were added and larvae homogenization was performed using a pestle. The samples were centrifuged at 20,000g for 10 minutes at 4°C and supernatant was collected and stored at -20°C until used. SDS-PAGE and Western blot analysis were performed as mentioned in Section 2.9.1.1 to detect the presence of protein bands at or above 250kDa for the experimental sample and 27KDa for the Actin-GFP sample.

2.11 Co-immunoprecipitation

Once the optimal concentration of formaldehyde (0.4%) was determined, the co-immunoprecipitation step was performed. Two-hundred microliters of cross-linked sample was placed in a 0.5ml microcentrifuge tube and 20μl of anti-GFP monoclonal agarose beads were added. The tubes were then incubated on a rotary shaker for 2 hours at 4°C. All centrifugation steps were performed at 4°C, unless otherwise specified. The beads were pelleted by centrifugation at 2,500g for 10 seconds. The supernatant was discarded and to the pellet, 200μl of freshly-made ice cold IP lysis buffer was added. Tubes were gently inverted to mix the contents and then centrifuged at 2,500g for 10 seconds. This washing step was repeated twice and supernatant was collected in a fresh tube, which was later used to confirm the absence of high molecular weight protein complex in a Western blot analysis. After washing, the pellet was resuspended in 20μl of 2X Laemmli buffer containing β-mercaptoethanol and heated at 65°C for 5 minutes. The samples were quickly vortexed and centrifuged at 2,500g for 5 minutes. The supernatant was transferred to fresh tubes. This step was repeated one more time to ensure complete

protein complex isolation. SDS-PAGE and Western blot analysis were performed on the extracted and supernatant wash samples as mentioned in Section 2.9.1.1.

2.12 Coomassie Staining

Our interest was to find out the interactive partners of collagen IV and one of the best ways to identify these partners is to perform mass spectrometry analysis. The co-immunoprecipitated supernatant samples were loaded to a 7% SDS-PAGE gel and run at 120 Volts for 1 hour. The gels were carefully removed from the cast plates and kept in gel fixing solution for 30 minutes at room temperature with gentle agitation. The gels were incubated in Coomassie stain for 1 hour and then destained until the bands were clearly visible against a light background. The bands were carefully cut individually, collected in clean microcentrifuge tubes, labeled, and stored at -80°C.

2.13 Mass Spectrometry and Data Analysis

The individual bands obtained from the Coomassie staining step were sent for LC-MS/MS spectrometry analysis at the University of Kentucky core facility. Upon completion of the mass spectrometry analysis, data comparison between the genes from *Actin-GFP* and *Vkg-GFP* was performed to find the presence of unique genes. Flybase database was used to search the possible functions of these unique genes in the development of *Drosophila* and their role in the basement membrane. Details on the role of these unique genes during the *Drosophila* development are tabulated in the Results section 3.4.

2.14 Protein Trap Studies

The protein trap studies were performed on *Viking GFP* and *Surf-4 GFP* transgenic larvae. Ten third instar larvae of each were collected in ice-cold 1X PBS solution and

individual larva was dissected using no. 5 Dumont forceps. The dissected larvae were

inverted inside-out according to the dissection protocol described in Ref. [75]. The

individual larval tissues like wing disc, trachea, salivary gland, and fat body were

carefully dissected and mounted as described in the paper. The mounted slides were

viewed under Zeiss fluorescence microscope for GFP expression in mentioned tissues.
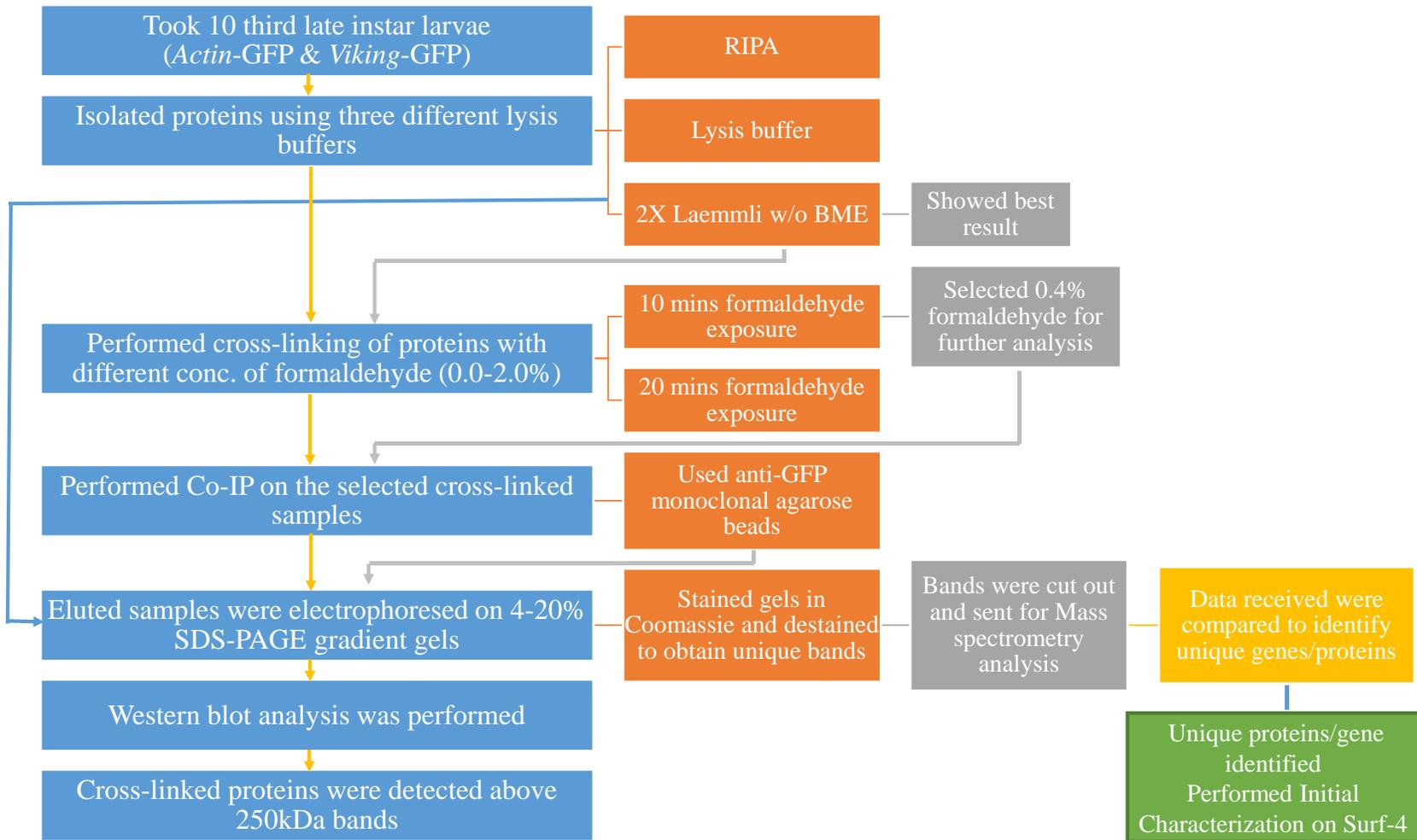
Figure 6. Schematic representation of the experimental setup. The figure provides a brief design of the overall experimental workflow.

3. RESULTS

3.1 Selection of Isolation Buffer

From the literature, we know that collagen IV accounts for 50% of the total BM mass [4]. Hence, we decided to concentrate the entire project on a single protein, collagen IV. To isolate it, it was first necessary to determine the best isolation buffer. We performed a literature search on buffers for protein isolation and found three commonly used buffers, namely RIPA, Lysis, and 2X Laemmli [72–74]. After proteins were extracted, Western blot analysis was performed using an anti-GFP antibody (Figure 7). As collagen IV (*Vkg*) is a high molecular weight protein, we expected a band at ~250kDa, while GFP being a small size protein a band is expected at ~25kDa. When using RIPA or lysis buffers, we did not observe protein bands detection at 250kDa and 25kDa, respectively; while clear bands were detected with 2X Laemmli buffer samples (Figure 7, Lysis buffer data not shown). Hence, further experiments were performed using 2X Laemmli buffer.

Figure 7. Comparison of proteins isolated using RIPA and Laemmli buffers. Vkg and GFP represent the samples isolated from *Viking-GFP* and *Actin-GFP* transgenic larvae, respectively. Collagen IV is detected at ~250kDa (green arrows), while GFP is detected at ~25kDa (red arrows) in a Western blot using an anti-GFP rabbit polyclonal antibody.

3.2   Standardizing the Formaldehyde Cross-linking Conditions

The main purpose of this study was to identify various interactive partners of collagen IV during the third larval instar stage. One of the best ways to perform interaction studies *in vivo* is to arrest the interacting partners using formaldehyde as a cross-linker. However, the use of formaldehyde as a cross-linker needs to be evaluated to obtain an optimal cross-linking complex. Hence, three parameters were considered 1) The reaction temperature, 2) formaldehyde concentration, and 3) reaction time. We started with different concentrations of formaldehyde ranging from 0.2% to 2.0%, at an incremental rate of 0.2%. The reaction was set up at room temperature (25°C) for 10 and 20 minutes, respectively, with vigorous vortexing. After formaldehyde incubation, proteins were isolated as mentioned earlier in Section 2.9.1.3. Before performing SDS-PAGE, the cross-linked samples were heated with 2X Laemmli buffer containing BME (β-mercaptoethanol) at 65°C for 5 minutes instead of 95°C for 10 minutes. At 95°C with 10 minutes of heat exposure proteins tend to lose their quaternary structure which might also break the formaldehyde bridge between the cross-linked proteins. SDS-PAGE and Western blot analysis showed the presence of high molecular weight bands (above 250kDa) in *Vkg-GFP* samples and at ~25kDa for *Actin-GFP* samples. We compared the Western blot results of *Vkg-GFP* and *Actin-GFP* which led us to choose the lowest concentration of formaldehyde (0.4%) as a standard cross-linker with an exposure time of 10 mins for further experiments (Figure 8).
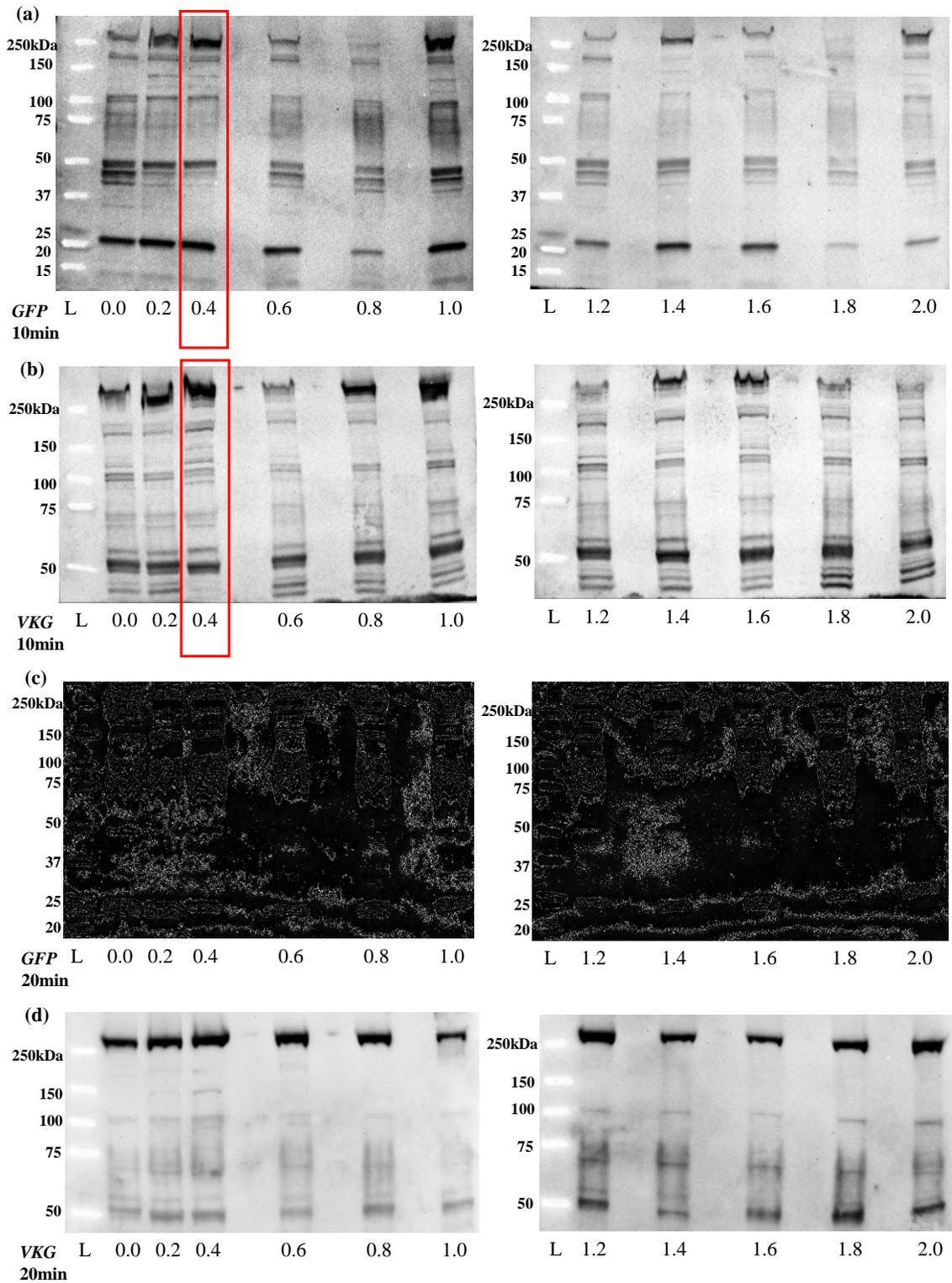
Figure 8. Cross-linking using different concentrations of formaldehyde. Western blot

analysis results using an anti-GFP rabbit polyclonal antibody. (a) *Actin-GFP* samples

with 10 minutes of increasing concentrations of formaldehyde exposure; (b) *Vkg-GFP* samples with 10 minutes of increasing concentrations of formaldehyde exposure; (c) *Actin-GFP* samples with 20 minutes of increasing concentrations of formaldehyde exposure; and (d) *Vkg-GFP* samples with 20 minutes of increasing concentrations of formaldehyde exposure. "L" represents ladder, X-axis represents formaldehyde concentrations, and Y-axis represents molecular weight in kDa. The lanes corresponding to the selected concentration is enclosed by a red box.

3.3   Co-immunoprecipitation of Cross-linked Protein Complex

As we were interested in the protein or factors associated with collagen IV, our

next step was to isolate only the collagen IV cross-linked complex and rule out the other

proteins from the sample. We accomplished this by utilizing the co-immunoprecipitation

technique. The process uses a bait protein linked to agarose beads to pull out the target

protein from the sample. From Figure 8 it is evident that there is a lot of non-specific

binding that may result in a false positive. As we were interested only in the collagen IV

associated complex and the protein from *Viking-GFP* and from the *Actin-GFP* control

was GFP tagged, an anti-GFP bait antibody linked to an agarose bead was judged to work

best for our experiments. We used MBL's anti-GFP mAb agarose beads as a bait protein

to extract the collagen IV cross-linked protein complex from the sample. The extracted

protein samples (*Actin-GFP* and *Vkg-GFP*) were resolved on SDS-PAGE gels and

Western blot analysis was performed as previously described. The Western blot analysis

revealed five distinct bands in the *Vkg-GFP* sample and a single band in the *Actin-GFP*

sample (Figure 9). The detection of five bands in the *Vkg-GFP* sample could be due to

the heating step at the SDS-PAGE analysis stage where proteins might have lost the

formaldehyde spacer arm separating the complex. We also ran the supernatant from the

IP washes. Detection of no high molecular weight band (250kDa) in the *Vkg-GFP*

supernatant confirmed the successful pull out of the collagen IV associated protein

complex.
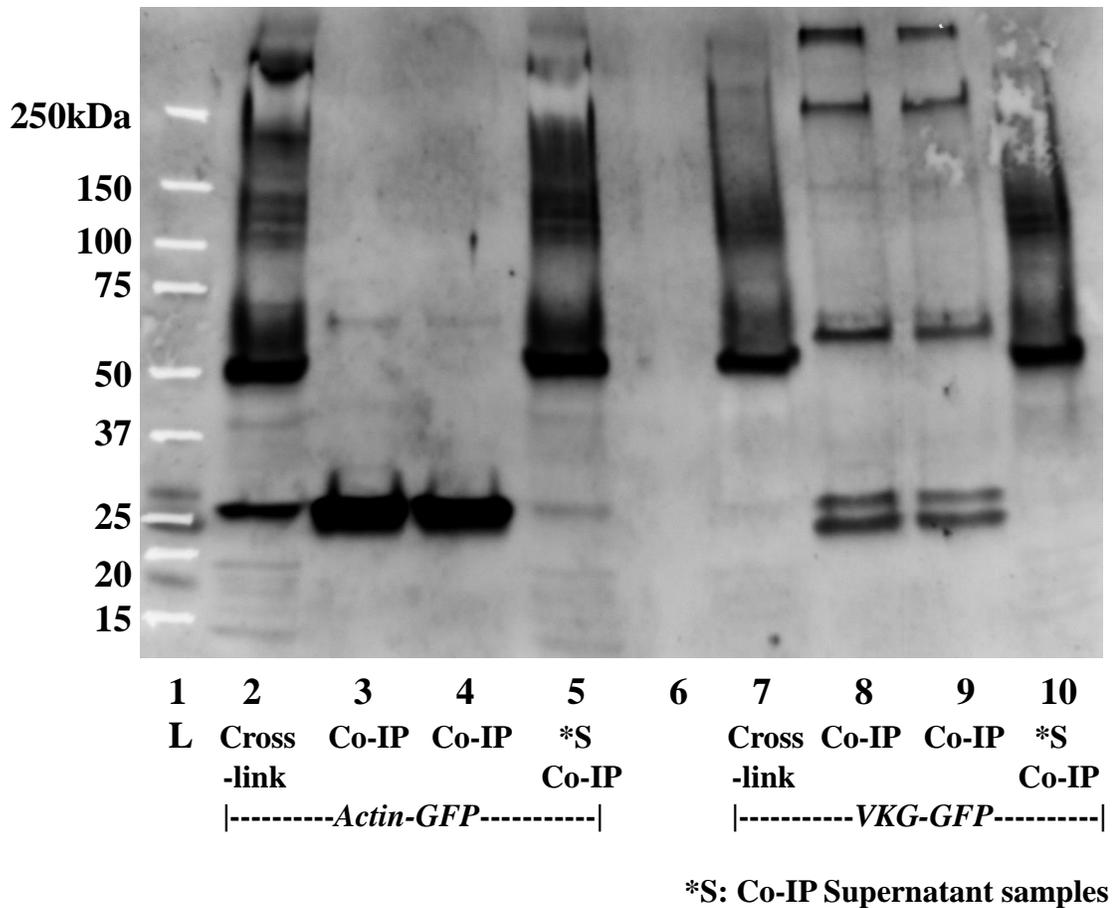
Figure 9. Co-immunoprecipitation results of the *Actin-GFP* and *Viking-GFP* cross-linked samples. Western blot analysis on Co-IP samples using an anti-GFP rabbit polyclonal as primary antibody. Lane 2 and 7 shows 0.4% formaldehyde cross-linked samples of *Actin-GFP* and *Vkg-GFP*, respectively. Lanes 3 & 4 and 8 & 9 show cross-linked co-immunoprecipitated samples of *Actin-GFP* and *Vkg-GFP*. Lanes 5 and 10 show the supernatant flow-through from the co-immunoprecipitated samples. Absence of 250kDa and 25kDa bands in the co-immunoprecipitated supernatant samples confirms the successful isolation of cross-linked complex. "L" represents ladder, X-axis represents formaldehyde concentrations, and Y-axis represents molecular weight in kDa.

3.4 Mass Spectrometric Analysis of Complexes Containing Collagen IV

Western blot analysis of the co-immunoprecipitated samples provided us with the confirmation of isolation of collagen IV and its associated complex. However, we still did not know which proteins or factors were present in this complex. To identify these collagen IV associated proteins and factors we had mass spectrometry analysis (LC-MS/MS) performed on the co-immunoprecipitated cross-linked samples. The collagen IV complex was precipitated using anti-GFP agarose beads, separated by SDS-PAGE, and visualized using Coomassie stain (Figure 10). The stained bands were excised carefully and sent to the collaborators for mass spectrometry analysis. The results from experimental (*Viking-GFP*) and control (*Actin-GFP*) samples were then compared and those proteins or factors that were unique to *Vkg* were selected. Literature reviews were performed on these unique proteins for their possible role in the development of *Drosophila* (Table 1). We again performed mass spectrometry analysis on the fresh Co-IP coomassie stained samples which resulted in several new candidates. Comparison of the unique *Viking* candidates between the two mass spectrometry results provided us with six candidates that were common (Table 2). Being able to find these six candidates in both rounds of mass spectrometry analysis we were confident of their interaction with collagen IV and decided to perform initial characterization on one of the candidate.
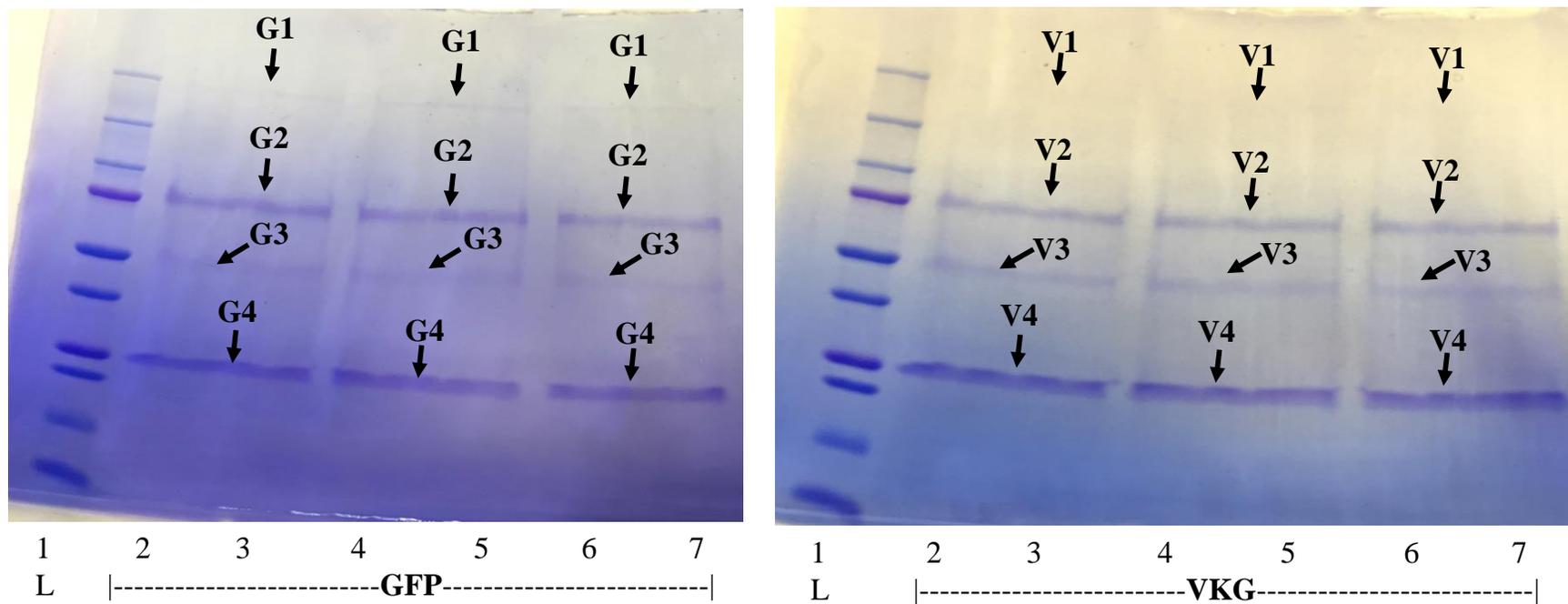
Figure 10. Coomassie stain on co-immunoprecipitated samples. Co-immunoprecipitated samples electrophoresed on 7% SDS-PAGE gel. "L" represents ladder, "G" represents Co-IP *Actin-GFP* sample, and "V" represents Co-IP *Viking-GFP* sample. Individual bands were excised carefully and sent for mass spectrometry analysis.

Table 1. Comparison of the mass spectrometry result between *Actin-GFP* and *Viking-GFP*. Only proteins unique to Vkg-GFP are listed.

Annotation Symbol	Gene Name	Gene Role	Functional Category	Ref.
CG1528	*gammaCop*	Responsible for secreting luminal components and assembles luminal chitinous during embryonic stage.	Development - Trachea	[76, 77]
CG15792	*zip*	Associated with ECM through PS2 integrin. Acts as an actin cross-linker and maintain the structural integrity of sarcomeric muscle cytoskeleton. It is also involved in the D-V compartmentalization.	Development - BM	[78, 79]
CG5210	*CG5210*	No data found.	-	-
CG1483	*Map205*	Interacts with polo, a kinase involved in regulating cell cycle division.	Cancer - Cell cycle	[80]
CG8193	*PPO2*	Found in the crystal cells in the hemolymph and responsible for protecting the injury site by producing melanin.	Protection	[81]

CG6705	*tsl*	A member of the membrane attack complex/perforin-like protein superfamily (MACPF), involved in pore-forming and immune defense roles and in influencing the developmental timing by localizing in the prothoracic gland.	Protection and Development	[82]
CG4314	*st*	Responsible for the biosynthesis of xanthommatin.	Eye pigmentation	[83]
CG5020	*CLIP-190*	Regulates microtubule dynamics and links microtubule plus-ends with other cellular structures.	Development - Cell cycle	[84]
CG3910	*mtTFB2*	Has a role in cell proliferation and differentiation.	Development	[85]
CG31623	*dtr*	No data found.	-	-
CG17291	*Pp2A-29B*	Regulates starvation-induced autophagy. It also acts as a brain tumor-suppressor that controls the self-renewal and differentiation of neural stem cells.	Cell Death and cancer	[86, 87]
CG1848	*LIMK1*	Involved in Rho signaling during metamorphosis.	Development	[88]
CG42338	*Ten-a*	Involved in the selection of specific synaptic partners in the olfactory circuit of Drosophila.	Others	[89]

CG4376	*Actn*	alpha-Actinin is an actin filament cross-linking protein found in muscle and non-muscle cells. Muscle alpha-actinin is found to be associated with the Z-disc involved in cross-linking of actinin filaments to the adjacent sarcomeres. Non-muscle alpha-actinin is found to be present in stress-fibers, lamellipodia, cell-cell and cell-matrix adhesion sites.	Development	[90]
CG17420	*RpL15*	Defect in the gene is associated with minute syndrome.	Disease	[91]
CG3395	*RpS9*	Associated with transcription of genes.	Others	[92]
CG10944	*RpS6*	Associated with the maintenance of the hematopoietic organ where upon mutation results in cell enlargement and over-proliferation of the hemocytes.	Cancer	[93, 94]
CG3203	*RpL17*	Mutation results in minute syndrome.	Disease	[91]
CG8922	*RpS5a*	Mutation results in minute syndrome.	Disease	[91]
CG4533	*l(2)efl*	Has role in clearing of poly-glutamine proteins.	Disease	[95]

CG4800	*Tctp*	Associates with ATM kinase and functions in growth regulation and cancer.	Cancer	[96, 97]
CG34410	*Rab26*	Interacts with different effector proteins and facilitates the transport of vesicles to the membrane.	Others	[98]
CG7576	*Rab3*	Localized at the presynaptic boutons and facilitates synaptic vesicle cycle.	Others	[99]
CG6601	*Rab6*	Regulates the N-cadherin trafficking in association with the Rich protein.	Others	[100]
CG1088	*Vha26*	Responsible for transporting ions in exchange of ATP. Maintain acidic environment of the lysosomes and pericellular spaces.	Others	[101, 102]
CG34090	*mt:Cyt-b*	No data found.	-	-
CG6202	*Surf-4*	Being a cargo receptor protein, surf-4 interacts with ERGIC-53 and proteins of the p25 family. Responsible for proper recruitment of COPI in the early secretory pathway.	Development	[54–59]

CG1404	*ran*	Involved in regulating several cellular functions of cell cycle which includes nucleocytoplasmic transport, nuclear membrane assembly, and spindle assembly	Others		[103]
CG6510	*RpL18A*	Mutation results in minute syndrome.	Disease		[91]
CG31034	*Jon99Cii*	Found abundantly in larval gut.	Others		[104]
CG7360	*Nup58*	No data found.	-		-
CG8987	*tam*	Mutation results in developmental defects of adult fly visual system.	Development Eye	-	[105]
CG12530	*Cdc42*	Regulator of axon outgrowth, branching, and guidance.	Others		[106, 107]
CG8274	*Mtor*	Regulator of spindle assembly checkpoint. Detaches kinetochore on the onset of mitosis.	Development Cell cycle	-	[108]
CG33180	*Ranbp16*	No data found.	-		-
CG3167	*aub*	Involved in guiding the siRNA for performing degradation activity.	Programmed cell death		[109]

CG33114	*Gyc32E*	No data found.	-	-
CG31196	*14-3-3Epsilon*	Involved in Ras1 signaling.	Cancer	[110]
CG30263	*stum*	Has role in proprioception feedback resulting in proper locomotion of Drosophila.	Others	[111]
CG8416	*Rho1*	Required for neuroblast proliferation.	Others	[112]

Table 2. Proteins found common in both rounds of mass spectrometry analysis and unique to *Viking-GFP* sample.

Gene Name	Description	Refs.
Actn	alpha-Actinin is an actin filament cross-linking protein found in muscle and non-muscle cells. In Drosophila, three alpha-actinin isoforms – non-muscle, larval muscle-specific, and adult muscle-specific – have been identified. Muscle alpha-actinin is found to be associated with the Z-disc involved in cross-linking of actinin filaments to adjacent sarcomeres. While non-muscle alpha-actinin is found to be present in stress-fibers, lamellipodia, cell-cell, and cell-matrix adhesion sites.	[90]
RpL18A	60S ribosomal protein L18a is also known as minute gene. Though studies could not find any direct correlation of it causing the syndrome as it lies in a gap in deletion coverage of the genome, but with point or transposon insertion mutations it can cause minute phenotype.	[91]
RpS9	40S Ribosomal protein S9 is a type of non-ribosomal protein associated with the transcription of genes. Studies carried out in Drosophila showed its relation with the transcription factors and found to localize at the transcription site.	[92]

RpL17	The mutation in the 60S Ribosomal protein L17 gene was found to cause minute syndrome in *Drosophila*, which results in prolonged development, short and thin bristles, and poor fertility and viability.	[91]
Surf-4	Co-localization studies showed that surf-4 localizes to ER, ERGIC (Endoplasmic Reticulum-Golgi Intermediate Compartment), and Golgi apparatus depending on the signals present at its N- and C-terminal. Being a cargo receptor protein, surf-4 interacts with ERGIC-53 and proteins of the p25 family. Double knockdown of surf-4 and ERGIC-53 showed that the complex is responsible for proper recruitment of COPI in the early secretory pathway. Studies have also shown that it is responsible for transport of soluble proteins.	[54–59]
ran	Studies showed that Ran is involved in regulating several cellular functions of the cell cycle, which includes nucleocytoplasmic transport, nuclear membrane assembly, and spindle assembly. One of the studies performed on Drosophila's neuroblast had shown that Ran associates with Canoe, a protein that regulates spindle orientation and cell polarity. Canoe/Ran-GTP complex then interact with Pins (Partner of Inscuteable) and then form complex with Mud (Mushroom body defect) resulting in the activation of Pins/Mud/dynein spindle orientation pathway.	[103]

3.5 Surf-4 and its Localization in *Drosophila* Third Instar Larvae

From all the unique proteins that were found associated with the collagen IV we decided to test the expression of surf-4 protein *in vivo*. Literature reviews indicated that surf-4 acts as a membrane receptor for the transport of soluble proteins between Endoplasmic Reticulum and Golgi apparatus [54–59]. As from the literature review on the biogenesis of collagen IV we found that the protein during its initial biosynthesis stage is expressed in the soluble form [50]. Therefore, we propose that surf-4 may be involved in the early secretory pathway and might play an important role in the transport of collagen IV. In order to understand the role of the surf-4 protein and its possible involvement in the early secretory pathway, we performed protein trap expression assay, over-expression, and knockdown studies. To better understand the function of *Surf-4* we utilized a transgenic third instar larvae expressing surf-4-GFP protein to locate the protein expression. We fixed various larval tissues from the surf-4 protein trap and analyzed the samples using fluorescence microscopy. Results from this experiment are provided in Figure 11, which displays the expression pattern of surf-4 protein in different *Drosophila* third instar larval tissues. It is evident from this figure that the protein is expressed ubiquitously throughout the organism. We also looked at the expression pattern for GFP-tagged collagen IV. From the figure it is evident that collagen IV is localized only at the basal side of the tissues.
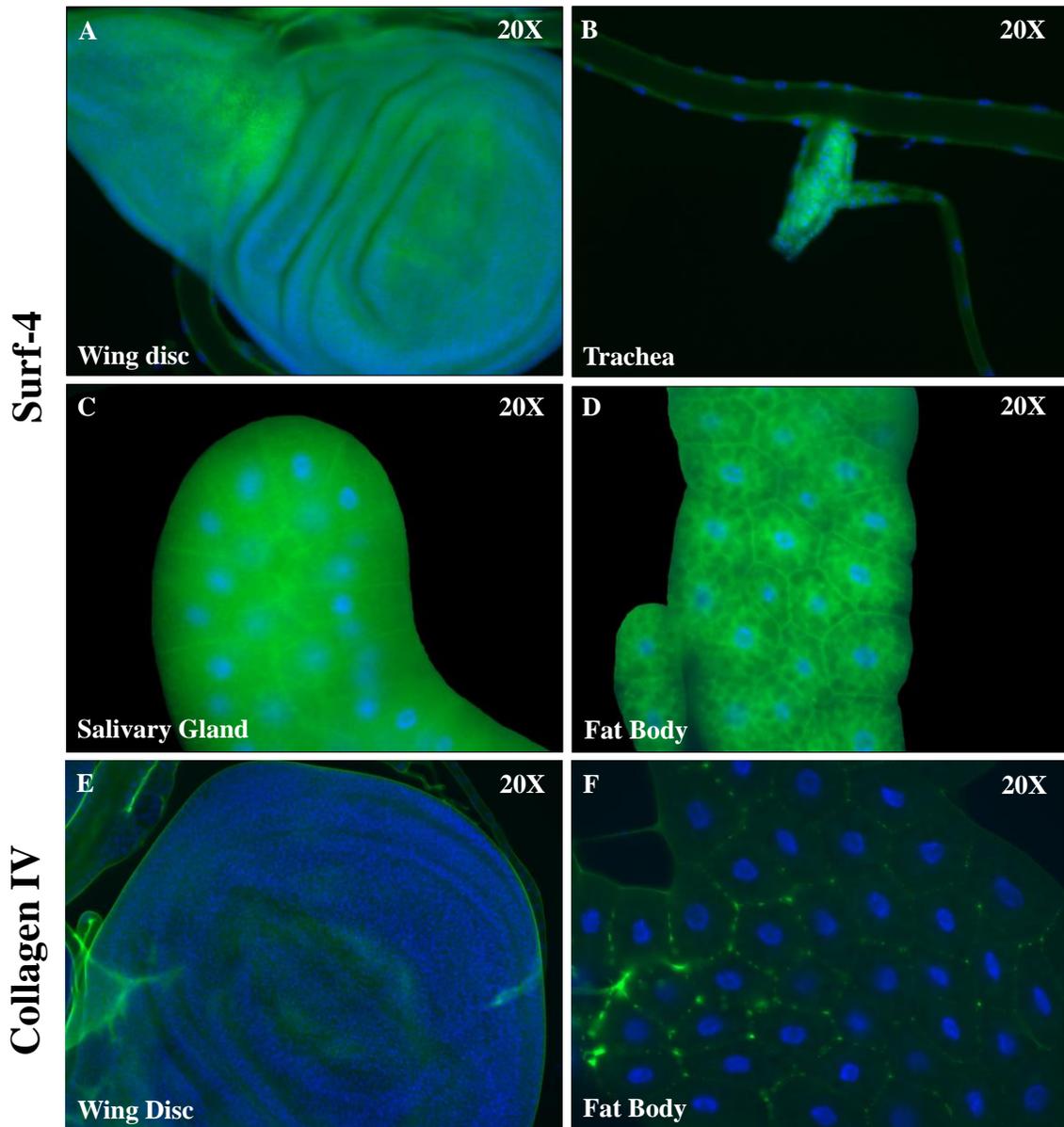
Figure 11. Expression pattern of Surf-4 and Collagen IV in different tissues of
Drosophila melanogaster. **A–D.** Expression of GFP-tagged Surf-4 in various structures
of third instar larvae. **E and F.** Expression of GFP-tagged Collagen IV in various
structures of third instar larvae. Green channel represents expression of GFP while blue
channel represents nucleus stained with DAPI.

3.6   *Surf-4* **RNAi Knockdown Study**

The surf-4 GFP protein trap expression study provided evidence of the protein being ubiquitously expressed throughout the larval tissues. Hence, we decided to perform knockdown study to see if there is any robust effect on the phenotype produced. We used two *Surf-4* RNAi lines obtained from Vienna Drosophila Research Center (VDRC - GD5883) and Bloomington Drosophila Stock Center (BDSC - 57471). We performed crosses between *Surf-4* RNAi lines and different RNAi drivers. The RNAi drivers contain Dicer-2 gene (a type of RNAse III enzyme) which enhances the RNAi effect. The crosses were performed between *Surf-4* RNAi virgin females and males of different RNAi drivers. The phenotypes produced using 57471 *Surf-4* RNAi line show defects in the wings (Figure 12C, D, G, H, K, and L and Figure 13G) and thorax bristles (Figure 13C and D). The phenotypic defects indicate that surf-4 might play some role during the development of *Drosophila* (Table 3, Figures 12 and 13). However, we could not observe any defects in the phenotypes produce in the GD5883 *Surf-4* RNAi line (Table 4). One of the possible reasons for this could be that the *Surf-4* RNAi line (GD5883) is weak to provide any robust phenotypic effect. In addition, we were still not sure if the phenotypic defects observed from the *Surf-4* RNAi line (57471) were due to the knockdown of surf-4. Therefore, performing RT-PCR on one of the *Surf-4* RNAi crosses would provide us with the indication of downregulation of *Surf-4* gene expression. In addition, the phenotypes from the knockdown of surf-4 were indicative of apoptosis, which could be mediated by the JNK pathway [113]. Hence, to explore the possibility that the downregulation of surf-4 resulted in induction of the JNK pathway we decided to assay for JNK pathway activation using a well-established LacZ reporter.

3.6.1 *Lac Z Staining*

We crossed *Surf-4* RNAi virgin females with the males of *Ptc-Gal4, UAS-GFP; Puc-lacZ/TM6Tb* genotype. The Puc-lacZ gene present in the driver helps to report the up-regulation of the JNK pathway, which will be indicated by a blue precipitate when stained for β-galactosidase reporter gene activity. However, when the cross was performed at 25°C we observed that most of the progenies growth was arrested at the first instar stage. We suspect that over-expression of *Surf-4* RNAi might have blocked the overall production of surf-4 mRNA which would have resulted in the cell death at the early growth stage. We also performed the cross at 18°C to observe the effect, as the Gal4 expression is reduced at lower temperature leading to less severe phenotype. However, we observed the same result as observed at 25°C. These observations are suggestive of a possible role for surf-4 in the early larval development stage.

Table 3. Generated phenotype description when *Surf-4* (BDSC - 57471) is knocked down using different RNAi drivers.

RNAi Drivers	Where Expressed	Progeny Phenotypes
25706 ♂ *w1118 P{GawB}BxMS1096; P{UAS-Dcr-2.D}2*	Wings	Wings found unfolded in ♀ progenies.
25708 ♂ *P{UAS-Dcr-2.D}1, w1118; P{Act5C-GAL4}25FO1/CyO*	Wings	No defects observed in progenies.
25750 ♂ *P{GawB}elavC155 w1118; P{UAS-Dcr-2.D}2*	Eye	Found oval-shaped eye in two ♂ progenies only.
25752 ♂ *P{UAS-Dcr-2.D}1, w1118; P{en2.4-GAL4}e16E, P{UAS-2xEGFP}AH2*	Wings	Wing defects were observed in progenies.
25753 ♂ *P{UAS-Dcr-2.D}1, w1118; P{bs-GAL4.Term}G1/CyO*	Wings	Wing defects were observed in progenies.
25757 ♂ *P{UAS-Dcr-2.D}1, w1118; P{GawB}bbgC96*	Wings	Wing defects were observed in ♀ progenies.
25758 ♂ *P{UAS-Dcr-2.D}1, w1118; P{GawB}pnrMD237/TM3, Ser1*	Thorax	Progenies showed less bristles on thorax.

Surf-4 RNAi line: y1 sc* v1; P{TRiP.HMC04782}attP40

P{UAS-Dcr-2.D}1, w1118; P{en2.4-GAL4}e16E, P{UAS-2xEGFP}AH2

y1 sc v1; P{TRiP.HMC04782}attP40*

A 25752♂

B 57471♀

C ♂

D ♀

P{UAS-Dcr-2.D}1, w1118; P{bs-GAL4.Term}G1/CyO

y1 sc v1; P{TRiP.HMC04782}attP40*

E 25753♂

F 57471♀

G ♂

H ♀
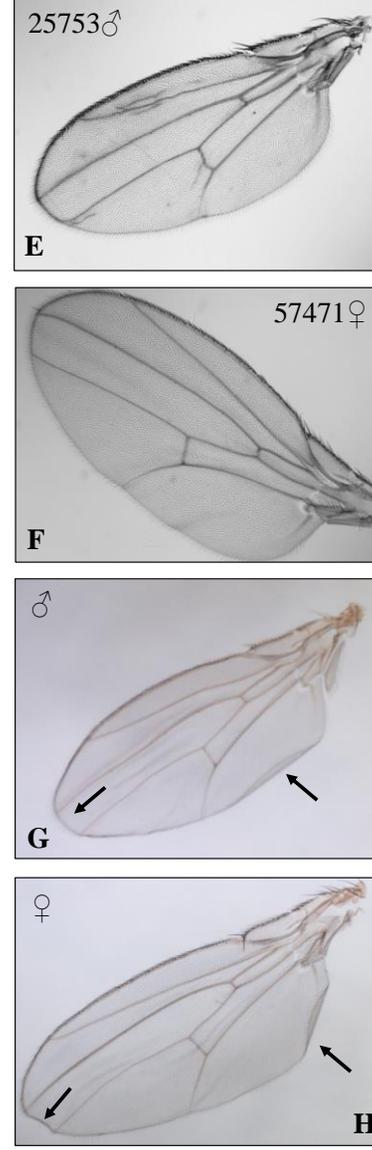
P{UAS-Dcr-2.D}1, w1118; P{GawB}bbgC96

y1 sc v1; P{TRiP.HMC04782}attP40*

I 25757♂

J 57471♀

K ♂

L ♀
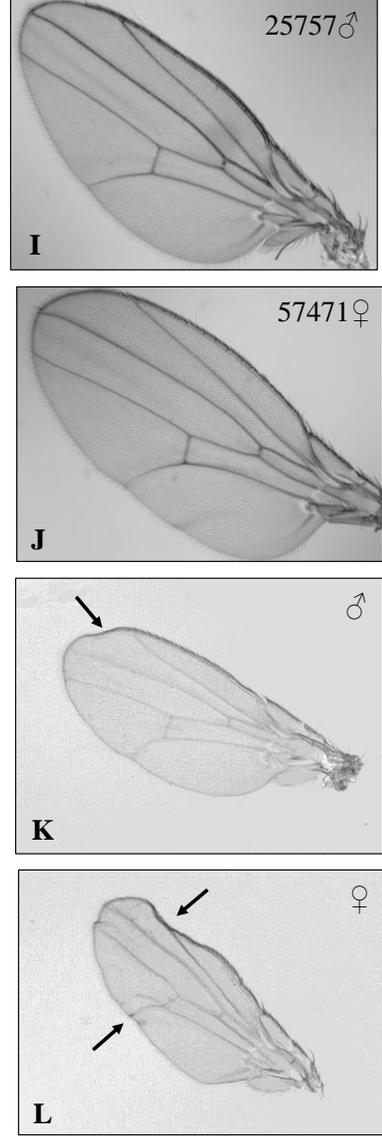
Figure 12. Defective phenotypes when *Surf-4* is knocked down. **A.** RNAi wild type driver: *P{UAS-Dcr-2.D}1, w1118; P{en2.4-GAL4}e16E, P{UAS-2xEGFP}AH2* (25752); **B, F, and J.** *Surf-4* RNAi wild type: *y1 sc* v1; P{TRiP.HMC04782}attP40* (57471); **C and D:** Progenies of crosses (A and B) showing defects in wings; **E.** RNAi Wild type driver: *P{UAS-Dcr-2.D}1, w1118; P{bs-GAL4.Term}G1/CyO* (25753); **G and H.** Progenies of crosses (E and F) showing defects in wings; **I.** RNAi wild type driver: *P{UAS-Dcr-2.D}1, w1118; P{GawB}bbgC96* (25757); **K and L.** Progenies of crosses (I and J) showing defects in wings. Arrows show the defect locations.
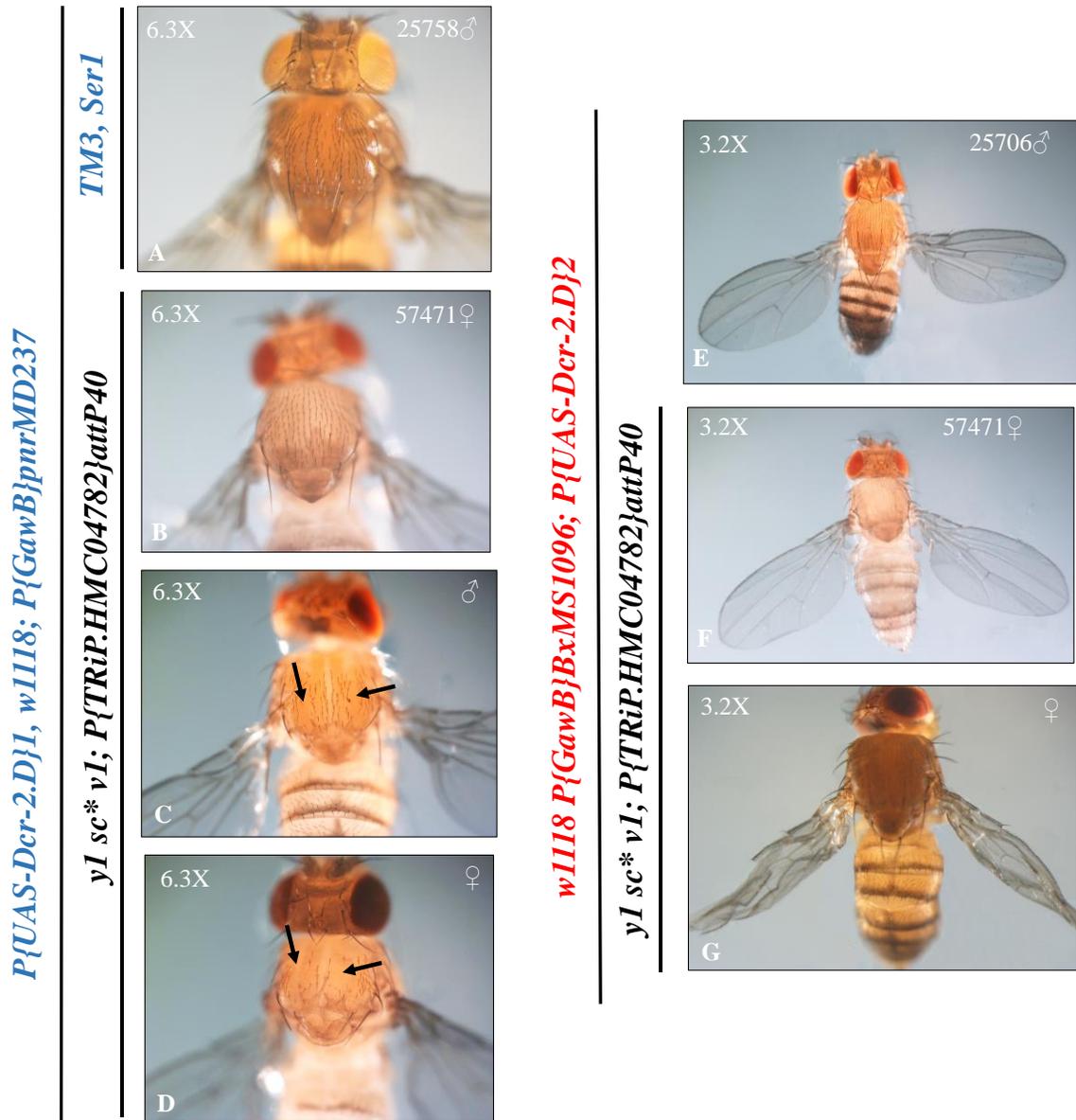
Figure 13. Defective phenotypes when *Surf-4* is knocked down. A. RNAi wild type

driver: *P{UAS-Dcr-2.D}1, w1118; P{GawB}pnrMD237/TM3, Ser1* (25758); **B and F.**

Surf-4 RNAi wild type: *y1 sc* v1; P{TRiP.HMC04782}attP40* (57471); **C and D.**

Progenies of crosses (A and B) showing less bristles on the thorax, indicated by arrows.

E. RNAi wild type driver: *w1118 P{GawB}BxMS1096; P{UAS-Dcr-2.D}2* (25706); **G.**

Female progenies of crosses (E and F) showing defects in wing unfolding.

Table 4. Phenotype description when *Surf-4* (VDRC – GD5883) is knocked down using different RNAi drivers.

RNAi Drivers	Where Expressed	Result
25706 ♂ *w1118 P{GawB}BxMS1096; P{UAS-Dcr-2.D}2*	Wings	No defects found.
25708 ♂ *P{UAS-Dcr-2.D}1, w1118; P{Act5C-GAL4}25FO1/CyO*	Wings	No defects found.
25750 ♂ *P{GawB}elavC155 w1118; P{UAS-Dcr-2.D}2*	Eye	No defects found.
25752 ♂ *P{UAS-Dcr-2.D}1, w1118; P{en2.4-GAL4}e16E, P{UAS-2xEGFP}AH2*	Wings	No defects found.
25753 ♂ *P{UAS-Dcr-2.D}1, w1118; P{bs-GAL4.Term}G1/CyO*	Wings	No defects found.
25757 ♂ *P{UAS-Dcr-2.D}1, w1118; P{GawB}bbgC96*	Wings	No defects found.
25758 ♂ *P{UAS-Dcr-2.D}1, w1118; P{GawB}pnrMD237/TM3, Ser1*	Thorax	No defects found.

Surf-4 RNAi line: w1118 P{GD2999}v5883

3.7 *Surf-4* **Over-expression Study**

Knockdown studies provided us with an indication of the possible effect of surf-4 in development. So we hypothesized that over-expression of surf-4 would result in an increase in its transport activity leading to aberrant phenotypes as well. We used UAS-Gal4 system to perform the over-expression of surf-4 protein using different UAS-Gal4 drivers. However, we did not find any specific difference in the phenotypes produced with different drivers (Table 5). The possible explanation for this observation could be that *Surf-4* over-expression line used was too weakly over-expressed to provide a robust phenotypic effect. This could be tested by performing RT-PCR on these flies in the future.

Table 5. Phenotype description when *Surf-4* (19202) is overexpressed using Gal4 drivers.

UAS-Gal4 Driver Lines	<u>Where</u> <u>Expressed</u>	Phenotype Description
Cg-gal4,UAS-GFP/CyO	Hemocytes	No phenotypic change observed.
w; Ey-gal4/CyO	Eyes	No phenotypic change observed.
GMR-gal4	Eyes	No phenotypic change observed.
w; LSP2-gal4	Fat Body	No phenotypic change observed.
Pnr-gal4/Tm6Tb	Thorax	No phenotypic change observed.
Ptc-gal4, UAS-GFP	Wings and other tissues	No phenotypic change observed.
Sd-gal4	Wings and other tissues	No phenotypic change observed.
yw; ; Tub-gal4/Tm3Sb	Ubiquitous	No phenotypic change observed.
w; ; Ubx-gal4/Tb	Haltere and other tissues	No phenotypic change observed.
Vg-gal4	Wings	No phenotypic change observed.

Surf-4 over-expression line: *w1118; P{XP}Surf4d04274/TM6B, Tb1*

4. **DISCUSSION**

Basement membrane (BM) plays a very important role during the development of an organism. Continuous remodeling and degradation of the BM occurs during the development of an organism, which involves interaction with various components like growth factors, enzymes, and proteins [14]. Invasion of cancer to other parts of the body involves breakage of the BM at various places, which makes it important to understand the BM and its interacting components. This thesis focuses on the isolation of these interacting components with one of the major BM protein, collagen IV, during the third instar larval stage of *Drosophila*. Because not many studies have been performed on the isolation of collagen IV in *Drosophila*, we began our study with the selection of a suitable isolation buffer. We used third instar larvae expressing *Actin-GFP* and *Viking-GFP* for this study. One of the reasons we chose *Actin-GFP* larvae as our control was because it enabled us to optimize GFP protein detection during Western blot analysis. In addition, as we were interested in finding the interactive partners of collagen IV, use of *Actin-GFP* as a control offered a way to rule out any interactive partners that non-specifically associated with GFP under our experimental conditions. We chose three different isolation buffers, namely, RIPA, Lysis, and 2X Laemmli to isolate collagen IV. After several repeated attempts and troubleshooting with the sample size, buffers, working conditions, and antibody concentrations, we were able to detect a collagen IV band at approximately 250kDa in the samples isolated using 2X Laemmli buffer. We repeatedly performed the isolation steps and optimized the protocol for isolation of collagen IV from the third instar larvae (Figure 7).

Once we confirmed the isolation of protein, our next step was to isolate the protein and its associated interactive partners. To perform protein interaction studies we used formaldehyde as a cross-linker. Formaldehyde has a spacer arm of 2.3–2.7 Å that helps to cross-link only those interactive partners that are in close vicinity of collagen IV. Three important criteria such as exposure times, temperature, and formaldehyde concentrations are required to be optimized when using formaldehyde as a cross-linker. We used different concentrations of formaldehyde (0.0-2.0% at an increment of 0.2%) with 10 and 20 minutes exposure at room temperature to determine the best cross-linking conditions. As we were working with the third instar larval stage we also performed the mentioned conditions on live larvae and dissected larvae. Figure 8 provides the results of the live larvae samples incubated with different formaldehyde concentrations and exposure time of 10 and 20 minutes, respectively at room temperature. As the concentration and exposure time of formaldehyde increased, it resulted in creating non-specific cross-links, which is clearly evident by the presence of less bands (Figure 8C and D). As we were interested in the interactions that are in close proximity, using higher concentration and longer exposure time could provide us with a false positive result. In addition, similar experiment on the dissected larvae samples did not yield good results (data not shown). One of the possible reasons could be the release of protease enzymes during the larval dissection that might have degraded the proteins. Therefore, we decided to use 0.4% of formaldehyde concentration with exposure time of 10 minutes on live larval samples as optimal condition for our further experiments.

Once we optimized the condition for cross-linking our next step was to pull out only the collagen IV and its associated complex and rule out other protein complexes that

might have formed. Hence, we decided to use co-immunoprecipitation technique on the cross-linked samples to pull out our protein of interest. The technique uses an agarose bead linked with a bait protein antibody and when incubated with the sample is capable of pulling out only the protein of interest. As our target protein was tagged with GFP, we used anti-GFP protein linked agarose beads as a bait to pull out collagen IV associated complex. Figure 9 provides the co-immunoprecipitation results on the cross-linked samples. Lanes 5 and 10 contain the samples of flow-through after incubation with the agarose beads. The absence of high molecular weight band (~250kDa) in Lane 10 and 25kDa band in Lane 5 indicated the successful pull out of collagen IV and its associated complex and GFP protein in the control sample, respectively. The presence of multiple bands in Lanes 8 and 9 is probably due to the addition of β-mercaptoethanol and boiling at 65°C before running the samples in the SDS-PAGE, which might have resulted in breakage of cross-links. Another possibility could be the GFP positive bands on the western blots are the degradation products from *Viking-GFP*.

Finally, we used mass spectrometry analysis to identify the interacting partners of collagen IV from the cross-link samples. Before the mass spectrometry analysis was performed by our collaborators at the University of Kentucky, the cross-linked Co-IP samples were electrophoresed on 7% SDS-PAGE gel. Upon coomassie staining of the gel similar bands appeared in both control *Actin-GFP* samples and experimental *Viking-GFP* samples. The reason for these similar bands could be due to the boiling step and addition of β-mercaptoethanol before the samples were electrophoresed. The presence of β-mercaptoethanol along with boiling would break the covalent bonds between the cross-links as well as the disulfide bonds in the protein structure resulting in similar band

61

detection. Once the data analysis was received we compared the genes detected in both the samples and ruled out all those genes from the *Viking-GFP* sample that were also found in the *Actin-GFP* control sample. The first round of the mass spectrometry analysis provided us with a total of 42 unique candidates while second round provided 51 unique candidates. The detection of more candidates in the second round could be due to the dynamic nature of BM in third instar larval samples. If this was the case then future experiment should be focused on timed-protein isolation. Another possibility could be the variability in protein isolation and cross-linking. While the difference in isolated proteins from the first and second round may be perplexing, it is important to note that we were able to identify six that were common in both rounds. Hence, we decided to perform further characterization studies on one of them. Initial literature reviews on these six candidates revealed that they play important roles during development. Out of these six candidates, we decided to perform initial characterization studies on *Surf-4* to understand about its possible role and effect through association with collagen IV.

Literature review showed that *Surf-4* plays an important role in the early secretory pathway and works as a cargo receptor protein between ER and Golgi apparatus [54]. It was also found that the protein might be responsible for the transport of soluble proteins [54–59]. In addition, biogenesis of collagen IV provided with the information that the protein during its initial biosynthesis stage is secreted in the soluble form. So we hypothesized that surf-4 might be involved in the transport of collagen IV from the ER to Golgi apparatus. In order to link the role of surf-4 with collagen IV we initially performed characterization studies on surf-4 and its role during the development of *Drosophila*. Protein trap studies using transgenic larvae expressing GFP-tagged surf-4

showed ubiquitous expression of protein within various tissues (Figure 11). This result

indicates the possible involvement of surf-4 protein in the development of various tissues.

We also performed knockdown of surf-4 protein using different RNAi drivers to observe

any defect in the progenies. The progenies showed defects in the wings and thorax

(Figures 12 and 13). However, when we over-expressed *Surf-4* we were unable to detect

any phenotypic change in the progenies. From this result it could be inferred that the

Surf-4 over-expression line we used was weak to provide an obvious phenotype. Another

possibility might be the existence of a feedback inhibition mechanism that might have

blocked the surf-4 from performing its activity. Additionally, the amount of surf-4 protein

may be present and in saturating quantities so that any additional protein has no effect.

RT-PCR on the progenies generated would be required to rule out the latter possibilities.

In addition, we noticed that when *Surf-4* RNAi crosses were driven using the Ptc-Gal4

driver, most of the progenies growth were arrested at the early development stage (1st

instar) when incubated at 25°C. Similar result was observed when the cross was

performed at 18°C, as the lower temperature reduces the Gal4 expression resulting in less

severe phenotype. These observations indicate that *Surf-4* might be involved during the

early growth stage of *Drosophila*.

4.1   Future Directions

4.1.1   *In-vivo Protein Knockdown Studies*

Protein trap, and lacZ staining allowed for initial characterization studies for *Surf-4*. However, we were not able to provide any connection between the surf-4 and collagen

IV. Hence, the next step would be to confirm these interactions *in vivo* and *in vitro*. In

order to perform this experiment, we would require to perform crosses between

transgenic lines expressing GFP-tagged collagen IV and *Surf-4* RNAi with different RNAi drivers. We hypothesize that if surf-4 is involved in the transport of collagen IV, performing this cross would result in either no or defective progenies. From the literature review of *Surf-4* [103–107] we know that it is involved in the transport of soluble proteins during the early secretory pathway. If *Surf-4* is knocked down there would not be any or only low level of collagen IV transport. As collagen IV is important in maintaining the structural integrity of various tissues, its irregular supply will result in absence or defective BM formation around the cells/tissues, causing early stage lethality.

4.1.2   *Characterization Studies for other Mass Spectrometry Candidates*

This study provides for the initial characterization of a single candidate found in the mass spectrometry analysis. However, there are other interactive partners of collagen IV that are yet to be explored. Initial literature reviews on these candidates showed that they are involved in performing critical functions during the cell cycle, for example, Ran [103]. One of these candidates is also involved in the transcription of genes, for example, Rps9 [92]. Thus, understanding their interaction mechanism with collagen IV could open-up new insights into the role they play during the development of *Drosophila*.

4.1.3   *Repetition of the Mass Spectrometry Analysis*

From the first two mass spectrometry analysis we observed that there were detections of several new candidates during each round. If the same pattern was repeated during the third round performing a time-based experiment would provide us with the pattern of the protein interaction at that particular time. Information gained through these timed experiments would indicate the dynamic nature of BM composition.

5. **REFERENCES**

1. Paulsson, M. (1992). Basement membrane proteins: structure, assembly, and cellular interaction. Crit. Rev. Biochem. Mol. Biol. 27, 93–127.

2. Schittny, J. C. and Yurchenco, P. D. (1989). Basement membranes: molecular organization and function in development and disease. Curr. Opin. Cell Biol. 1, 983–988.

3. Kalluri, R. (2003). Basement membranes: structure, assembly and role in tumour angiogenesis. Nat. Rev. Cancer, 3(6), 422-433.

4. LeBleu, V. S., MacDonald, B., and Kalluri, R. (2007). Structure and function of basement membranes.Exp. Biol. Med., 232(9), 1121–1129.

5. Yurchenco, P. D. and Schittny, J. C. (1990). Molecular architecture of basement membranes. FASEB J. 4, 1577–1590.

6. Yurchenco, P. D., Smirnov, S., and Mathus, T. (2002). Analysis of basement membrane self-assembly and cellular interactions with native and recombinant glycoproteins. Methods Cell Biol. 69, 111–144.

7. Kahsai, T.Z., Enders, G.C., Gunwar, S., Brunmark, C., Wieslander, J., Kalluri, R., Zhou, J., Noelken, M.E., and Hudson, B.G. (1997). Seminiferous tubule basement membrane. Composition and organization of type IV collagen chains, and the linkage of alpha3(IV) and alpha5(IV) chains. J. Biol. Chem. 272, 17023–17032.

8. Timpl, R., Rohde, H., Gehron-Robey, P., Rennard, S. I., Foidart, J. M., and Martin, G. R. (1979). Laminin-a glycoprotein from basement membrane. J. Biol. Chem. 254, 9933–9937.

9. Aumailley, M. (2013). The laminin family. Cell Adh. Migr., 7(1), 48-55.

10. Timpl, R., Tisi, D., Talts, J. F., Andac, Z., Sasaki, T., and Hohenester, E. (2000). Structure and function of laminin LG modules. Matrix Biol. 19, 309-17;

11. Aumailley, M., Battaglia, C., Mayer, U., Reinhardt, D., Nischt, R., Timpl, R., and Fox, J. W. (1993). Nidogen mediates the formation of ternary complexes of basement membrane components. Kidney Int. 43, 7–12

12. Aumailley, M., Wiedemann, H., Mann, K., and Timpl, R. (1989). Binding of nidogen and the laminin-nidogen complex to basement membrane collagen type IV. Eur. J. Biochem. 184, 241–248.

13. Mayer, U., Kohfeldt, E., and Timpl, R. (1998). Structural and genetic analysis of laminin-nidogen interaction. Ann. N. Y. Acad. Sci. 857, 130–142.

14. Bishop, J. R., Schuksz, M., and Esko, J. D. (2007). Heparan sulphate proteoglycans fine-tune mammalian physiology. Nature, 446(7139), 1030-1037.

15. Huntington, J. A. (2003). Mechanisms of glycosaminoglycan activation of the serpins in hemostasis. J. Thromb. Haemost. 1, 1535–1549.

16. Forsten-Williams, K., Chua, C. C., and Nugent, M. A. (2005). The kinetics of FGF-2 binding to heparan sulfate proteoglycans and MAP kinase signaling. J. Theor. Biol. 233, 483–499.

17. Couchman, J. R., Chen, L. G., and Woods, A. (2001). Syndecans and cell adhesion. Int. Rev. Cytol. 207, 113–150.

18. Kirkpatrick, C. A., Knox, S. M., Staatz, W. D., Fox, B., Lercher, D. M., and Selleck, S. B. (2006). The function of a Drosophila glypican does not depend entirely on heparan sulfate modification. Develop. Boil., 300(2), 570-582.

19. Aumailley, M., Battaglia, C., Mayer, U., Reinhardt, D., Nischt, R., Timpl, R., and Fox, J.W. (1993). Nidogen mediates the formation of ternary complexes of basement membrane components. Kidney Int. 43, 7–12

20. Davies, Jamie A. (2001). Extracellular matrix. In: eLS. John Wiley & Sons Ltd, Chichester. http://www.els.net [doi: 10.1038/npg.els.0001274]

21. DeClerck, Y.A., Mercurio, A.M., Stack, M.S., Chapman, H.A., Zutter, M.M., Muschel, R.J., Raz, A., Matrisian, L.M., Sloane, B.F., Noel, A., Hendrix, M.J., Coussens, L., and Padarathsingh, M. (2004). Porteases, extracellular matrix and cancer: a workshop of the path B study section. Am. J. Pathol. 164(4), 1131–1139.

22. Mercurio, A.M. and Rabinovitz, I. (2001). Towards a mechanistic understanding of tumor invasion: lessons from the α6β4 integrin. Semin. Cancer Biol. 11, 129–141.

23. Dajee, M., Lazarov, M., Zhang, J.Y., Cai, T., Green, C.L., Russell, A.J., Marinkovich, M.P., Tao, S., Lin, Q., Kubo, Y., and Khavari, P.A. (2003). NF-αβ blockade and oncogenic Ras trigger invasive human epidermal neoplasia. Nature 421, 639–643.

24. Dreyfuss, J.L., Regatieri, C.V., Jarrouge, T.R., Cavalheiro, R.P., Sampaio, L.O., and Nader, H.B. (2009). Heparan sulfate proteoglycans: structure, protein interactions and cell signaling. An. Acad. Bras. Cienc. 81(3), 409–429.

25. Lindner, J. R., Hillman, P. R., Barrett, A. L., Jackson, M. C., Perry, T. L., Park, Y., and Datta, S. (2007). The Drosophila perlecan gene trol regulates multiple signaling pathways in different developmental contexts. BMC Dev. Biol. 7(1): 1.

26. Lin X. (2004). Functions of heparan sulfate proteoglycans in cell signaling during development. Development 131(24), 6009–6021.

27. Zioncheck, T.F., Richardson, L., Liu, J., Chang, L., King, K.L., Bennett, G.L., Fugedi, P., Chamow, S.M., Schwall, R.H., and Stack, R.J. (1995). Sulfated oligosaccharides promote hepatocyte growth factor association and govern its mitogenic activity. J. Biol. Chem. 270, 16871–16878.

28. Aviezer, D. and Yayon, A. (1994). Heparin-dependent binding and autophosphorylation of epidermal growth factor (EGF) receptor by heparin-binding EGF-like growth factor but not EGF. Proc. Natl. Acad. Sci. USA 91, 12173–12177.

29. Lin, X. and Perrimon, N. (2003). Developmental roles of heparan sulfate proteoglycans in Drosophila. Glycoconj. J. 19, 363–368.

30. Nybakken, K. and Perrimon, N. (2002). Heparan sulfate proteoglycan modulation of developmental signaling in Drosophila. Biochim. Biophys. Acta 1573, 280–291.

31. Spring, J., Paine-Saunders, S.E., Hyne, R.O., and Bernfield, M. (1994). Drosophila syndecan: conservation of a cell-surface heparan sulfate proteoglycan. Proc. Natl. Acad. Sci. USA 91, 3334–3338.

32. Steigemann, P., Molitor, A., Fellert, S., Jackle, H., and Vorbruggen, G. (2004). Heparan sulfate proteoglycan syndecan promotes axonal and myotube guidance by slit/robo signaling. Curr. Biol. 14, 225–230.

33. Voigt, A., Pflanz, R., Schafer, U., and Jackle, H. (2002). Perlecan participates in proliferation activation of quiescent Drosophila neuroblasts. Dev. Dyn. 224, 403–412.

34. Mali, M., Elenius, K., Miettinen, H.M., and Jalkanen, M. (1993). Inhibition of basic fibroblast growth factor-induced growth promotion by overexpression of syndecan-1. J. Biol. Chem. 268, 24215–24222

35. Steinfeld, R., Van Den Berghe, H., and David, G. (1996). Stimulation of fibroblast growth factor receptor-1 occupancy and signaling by cell surface associated syndecans and glypican. J. Cell Biol. 133, 405–416.

36. Datta, S., Pierce, M., and Datta, M.W. (2006). Perlecan signaling: helping hedgehog stimulate prostate cancer growth. Int. J. Biochem. Cell. Biol. 38(11), 1855–1861.

37. Iozzo, R.V. (2005). Basement membrane proteoglycans: from cellar to ceiling. Nat Rev Mol. Cell Biol. 6(8), 646–656.

38. Schneider, M., Khalil, A.A., Poulton, J., Castillejo-Lopez, C., Egger-Adam, D., Wodarz, A., Deng, W.M., and Baumgartner, S. (2006). Perlecan and dystroglycan act at the basal side of the Drosophila follicular epithelium to maintain epithelial organization. Development, 133(19), 3805–3815.

39. Kaphingst, K. and Kunes, S. (1994). Pattern formation in the visual centers of the Drosophila brain: wingless acts via decapentaplegic to specify the dorsoventral axis. Cell 78(3), 437–448.

40. Silver, S.J. and Rebay, I. (2005). Signaling circuitries in development: insights from the retinal determination gene network. Development 132(1), 3–13.

41. Jackson, S.M., Nakato, H., Sugiura, M., Jannuzi, A., Oakes, R., Kaluza, V., Golden, C., and Selleck, S.B. (1997). dally, a Drosophila glypican, controls cellular responses to the TGF-beta-related morphogen, Dpp. Development 124, 4113–4120.

42. Wang, X., Harris, R.E., Bayston, L.J., and Ashe, H.L. (2008). Type IV collagens regulate BMP signaling in Drosophila. Nature 455, 72–77.

43. Campbell, G., Weaver, T., and Tomlinson, A. (1993). Axis specification in the developing Drosophila appendage: the role of wingless, decapentaplegic, and the homeobox gene aristaless. Cell 74, 1113–1123.

44. Jiang, J. and Struhl, G. (1996). Complementary and mutually exclusive activities of decapentaplegic and wingless organize axial patterning during Drosophila leg development. Cell 86, 401–409.

45. Chen, D., Zhao, M., and Mundy, G.R. (2004). Bone morphogenetic proteins. Growth Factors 22, 233–241.

46. Bobik A. (2006). Transforming growth factor-bs and vascular disorders. Arterioscler. Thromb. Vasc. Biol. 26, 1712–1720.

47. Kirilly, D. and Xie, T. (2007). The Drosophila ovary: an active stem cell community. Cell Res. 17, 15–25.

48. Hudson, D. M. and Eyre, D. R. (2013). Collagen prolyl 3-hydroxylation: a major role for a minor post-translational modification? Connect. Tis. Res. 54(4–5), 245–251.

49. Lamande´, S. R. and Bateman, J. F. (1999). Procollagen folding and assembly: the role of endoplasmic reticulum enzymes and molecular chaperones. Semin. Cell Dev. Biol. 10, 455–464.

50. Malhotra, V. and Erlmann, P. (2015). The pathway of collagen secretion. Ann. Rev. Cell Develop. Biol. 31, 109–124.

51. Pastor-Pareja, J. C., and Xu, T. (2011). Shaping cells and organs in Drosophila by opposing roles of fat body-secreted Collagen IV and perlecan. Dev Cell 21(2), 245–56.

52 Mitrovic, S., Ben-Tekaya, H., Koegler, E., Gruenberg, J., and Hauri, H. P. (2008). The cargo receptors Surf4, endoplasmic reticulum-Golgi intermediate compartment (ERGIC)-53, and p25 are required to maintain the architecture of ERGIC and Golgi. Mol. Biol. Cell 19(5), 1976–1990.

53. Breuza, L., Halbeisen, R., Jenö, P., Otte, S., Barlowe, C., Hong, W., and Hauri, H. P. (2004). Proteomics of endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. J Biol Chem 279(45), 47242–47253.

54. Duhig, T., Ruhrberg, C., Mor, O., and Fried, M. (1998). The human Surfeit locus. Genomics 52(1), 72–78.

55. Williams, T. and Fried, M. (1986). A mouse locus at which transcription from both DNA strands produces mRNAs complementary at their 3 ends. Nature 322: 275–279.

56. Lennard, A.C. and Fried, M. (1991). The bidirectional promoter of the divergently transcribed mouse Surf-1 and Surf-2 genes. Mol. Cell. Biol. 11, 1281–1294.

57. Huxley, C. L. and Fried, M. I. (1990). The mouse surfeit locus contains a cluster of six genes associated with four CpG-rich islands in 32 kilobases of genomic DNA. Mol. Cell Biol. 10(2), 605–614.

58. Armes, N. and Fried, M. (1996). Surfeit locus gene homologs are widely distributed in invertebrate genomes. Mol. Cell Biol. 16(10), 5591–5596.

59. Magoulas, C. and Fried, M. (2000). Isolation and genomic analysis of the human surf-6 gene: a member of the Surfeit locus. Gene, 243(1), 115–123.

60. Abreu-Velez, A. M. and Howard, M. S. (2012). Collagen IV in normal skin and in pathological processes. N. Am. J. Med. Sci., 4(1), 1.

61. Haghighi, A., Tiwari, A., Piri, N., Nürnberg, G., Saleh-Gohari, N., Haghighi, A., Neidhardt, J., Nürnberg, P., and Berger, W. (2014). Homozygosity mapping and whole exome sequencing reveal a novel homozygous COL18A1 mutation causing Knobloch syndrome. PloS one, 9(11), e112747.

62. Suzuki, O. T., Sertie, A. L., Der Kaloustian, V. M., Kok, F., Carpenter, M., Murray, J., Czeizel, A. E., Kliemann, S. E., Rosemberg, S., Monteiro, M. B. R. O., and Olsen, B. R. (2002). Molecular analysis of collagen XVIII reveals novel mutations, presence of a third isoform, and possible genetic heterogeneity in Knobloch syndrome. Am. J. Hum. Gen., 71(6), 1320–1329.

63. McGowan, K. A., Marinkovich, M. P. (2000). Laminins and human disease. Microsc. Res. Tech. 51, 262–279.

64. Arikawa-Hirasawa, E., Le, A.H., Nishino, I., Nonaka, I., Ho, N.C., Francomano, C.A., Govindraj, P., Hassell, J.R., Devaney, J.M., Spranger, J. and Stevenson, R.E., 2002. Structural and functional mutations of the perlecan gene cause Schwartz-Jampel syndrome, with myotonic myopathy and chondrodysplasia. Am. J. Hum. Gen., 70(5), 1368–1375.

65. Yasothornsrikul, S., Davis, W. J., Cramer, G., Kimbrell, D. A., and Dearolf, C. R. (1997). viking: identification and characterization of a second type IV collagen in Drosophila. Gene 198, 17–25

66. Le Parco, Y., Knibiehler, B., Cecchini, J. P., and Mirre, C. (1986). Stage and tissue-specific expression of a collagen gene during Drosophila melanogaster development. Exp. Cell Res. 163, 405–412

67. Natzle, J. E., Monson, J. M., and McCarthy, B. J. (1982). Cytogenetic location and expression of collagen-like genes in Drosophila. Nature 296, 368–371.

68. Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., and Bier, E. (2001). A systematic analysis of human disease-associated gene sequences in Drosophila melanogaster. Genome Res. 11, 1114–1125.

69. Pandey, U. B. and Nichols, C. D. (2011). Human disease models in Drosophila melanogaster and the role of the fly in therapeutic drug discovery. Pharmacol. Rev. 63, 411–436.

70. Bier E. (2005). Drosophila, the golden bug, emerges as a tool for human genetics. Nat. Rev. Genet. 6, 9–23.

71. Tanner, K. (2012). Regulation of the basement membrane by epithelia generated forces. Phy. Boil., 9(6), 065003.

72. Klockenbusch, C. and Kast, J. (2010). Optimization of Formaldehyde Cross-Linking for Protein Interaction Analysis of Non-Tagged Integrin β 1. J. Biomed. Biotech. 2010, Article ID 927585, 13 pages; doi:10.1155/2010/927585

73. Gao, M., McCluskey, P., Loganathan, S. N., & Arkov, A. L. (2014). An in vivo Crosslinking Approach to Isolate Protein Complexes From Drosophila Embryos. JoVE, 86, e51387-e51387.74. Stevens, L. J. and Page-McCaw, A. (2012). A secreted MMP is required for reepithelialization during wound healing. Mol. Biol. Cell, 23(6), 1068–1079.

75. Srivastava, A. (2013). A protocol for genetic induction and visualization of benign and invasive tumors in cephalic complexes of Drosophila melanogaster. JoVE 79, e50624-e50624.

76. Jayaram, S. A., Senti, K. A., Tiklová, K., Tsarouhas, V., Hemphälä, J., and Samakovlis, C. (2008). COPI vesicle transport is a common requirement for tube expansion in Drosophila. PLoS One. 3(4), e1964

77. Grieder, N. C., Caussinus, E., Parker, D. S., Cadigan, K., Affolter, M., and Luschnig, S. (2008). γCOP is required for apical protein secretion and epithelial morphogenesis in Drosophila melanogaster. PLoS ONE 3(9), e3241

78. Bloor, J. W. and Kiehart, D. P. (2001). Zipper nonmuscle myosin-II functions downstream of PS2 integrin in Drosophila myogenesis and is necessary for myofibril formation. Dev Biol 239(2), 215–228.

79. Major, R. J. and Irvine, K. D. (2006). Localization and requirement for Myosin II at the dorsal-ventral compartment boundary of the Drosophila wing. Dev Dynamics. 235:3051–3058.

80. Archambault, V., D'Avino, P. P., Deery, M. J., Lilley, K. S., and Glover, D. M. (2008). Sequestration of Polo kinase to microtubules by phosphopriming-independent binding to Map205 is relieved by phosphorylation at a CDK site in mitosis. Gene Develop. 22(19), 2707–2720.

81. Binggeli, O., Neyen, C., Poidevin, M., and Lemaitre, B. (2014). Prophenoloxidase activation is required for survival to microbial infections in Drosophila. PLoS Pathog. 10(5), e1004067.

82. Johnson, T. K., Crossman, T., Foote, K. A., Henstridge, M. A., Saligari, M. J., Beadle, L. F., Herr, A., Whisstock, J. C., and Warr, C. G. (2013). Torso-like functions independently of Torso to regulate Drosophila growth and developmental timing. Proc. Nat. Acad. Sci. 110(36), 14688–14692.

83. Tearle, R. G., Belote, J. M., McKeown, M., Baker, B. S., and Howells, A. J. (1989). Cloning and characterization of the scarlet gene of Drosophila melanogaster. Genetics 122(3), 595–606.

84. Dzhindzhev, N. S., Rogers, S. L., Vale, R. D., and Ohkura, H. (2005). Distinct mechanisms govern the localisation of Drosophila CLIP-190 to unattached kinetochores and microtubule plus-ends. J. Cell Sci. 118(16), 3781–3790.

85. Adán, C., Matsushima, Y., Hernández-Sierra, R., Marco-Ferreres, R., Fernández-Moreno, M. Á., González-Vioque, E., Calleja, M., Aragón, J. J., Kaguni, L. S., and Garesse, R. (2008). Mitochondrial transcription factor B2 is essential for metabolic function in Drosophila melanogaster development. J. Biol. Chem. 283(18), 12333–12342.

86. Bánréti, Á., Lukácsovich, T., Csikós, G., Erdélyi, M., and Sass, M. (2012). PP2A regulates autophagy in two alternative ways in Drosophila. Autophagy 8(4), 623–636.

87. Wang, C., Chang, K. C., Somers, G., Virshup, D., Ang, B. T., Tang, C., Yu, F., and Wang, H. (2009). Protein phosphatase 2A regulates self-renewal of Drosophila neural stem cells. Development 136(13), 2287–2296.

88. Chen, G. C., Gajowniczek, P., and Settleman, J. (2004). Rho-LIM kinase signaling regulates ecdysone-induced gene expression and morphogenesis during Drosophila metamorphosis. Current Biol. 14(4), 309–313.

89. Hong, W., Mosca, T. J., and Luo, L. (2012). Teneurins instruct synaptic partner matching in an olfactory map. Nature 484(7393), 201–207.

90. Wahlström, G., Norokorpi, H. L., and Heino, T. I. (2006). Drosophila α-actinin in ovarian follicle cells is regulated by EGFR and Dpp signalling and required for cytoskeletal remodelling. Mech. Develop. 123(11), 801–818.

91. Marygold, S. J., Roote, J., Reuter, G., Lambertsson, A., Ashburner, M., Millburn, G. H., Harrison, P. M., Yu, Z., Kenmochi, N., Kaufman, T. C., and Leevers, S. J. (2007). The ribosomal protein genes and Minute loci of Drosophila melanogaster. Genome Biol. 8(10), R216.

92. Rugjee, K. N., Chaudhury, S. R., Al-Jubran, K., Ramanathan, P., Matina, T., Wen, J., and Brogna, S. (2013). Fluorescent protein tagging confirms the presence of ribosomal proteins at Drosophila polytene chromosomes. Peer J. 1, e15.

93. Watson, K. L., Konrad, K. D., Woods, D. F., and Bryant, P. J. (1992). Drosophila homolog of the human S6 ribosomal protein is required for tumor suppression in the hematopoietic system. Proc. Nat. Acad. Sci. 89(23), 11302–11306.

94. Stewart, M. J. and Denell, R. (1993). Mutations in the Drosophila gene encoding ribosomal protein S6 cause tissue overgrowth. Mol. Cell. Biol. 13(4), 2524–2535.

95. Carra, S., Boncoraglio, A, Kanon, B., Brunsting, J. F., Minoia, M., Rana, A., Vos, M. J., Seidel, K., Sibon, O. C., and Kampinga, H. H. (2010). Identification of the Drosophila ortholog of HSPB8 - Implication of HSPB8 loss of function in protein folding diseases. J. Biol. Chem. 285(48), 37811–37822.

96. Hsu, Y. C., Chern, J. J., Cai, Y., Liu, M., and Choi, K. W. (2007). Drosophila TCTP is essential for growth and proliferation through regulation of dRheb GTPase. Nature 445(7129), 785–758.

97. Hong, S. T. and Choi, K. W. (2013). TCTP directly regulates ATM activity to control genome stability and organ development in Drosophila melanogaster. Nature Comm. 4(2986).

98. Chan, C. C., Scoggin, S., Wang, D., Cherry, S., Dembo, T., Greenberg, B., Jin, E. J., Kuey, C., Lopez, A., Mehta, S. Q., and Perkins, T. J. (2011). Systematic discovery of Rab GTPases with synaptic functions in Drosophila. Curr. Biol. 21(20), 1704–1715.

99. Graf, E. R., Daniels, R. W., Burgess, R. W., Schwarz, T. L., and DiAntonio, A. (2009). Rab3 dynamically controls protein composition at active zones. Neuron. 64(5), 663–677.

100. Tong, C., Ohyama, T., Tien, A. C., Rajan, A., Haueter, C. M., and Bellen, H. J. (2011). Rich regulates target specificity of photoreceptor cells and N-cadherin trafficking in the Drosophila visual system via Rab6. Neuron 71(3), 447–459.

101. Mauvezin, C., Nagy, P., Juhász, G., and Neufeld, T. P. (2015). Autophagosome-lysosome fusion is independent of V-ATPase-mediated acidification. Nature Comm. 6, 7007.

102. Allan, A. K., Du, J., Davies, S. A., and Dow, J. A. (2005). Genome-wide survey of V-ATPase genes in Drosophila reveals a conserved renal phenotype for lethal alleles. Physiol. Genom. 22(2), 128–138.

103. Wee, B., Johnston, C. A., Prehoda, K. E., and Doe, C. Q. (2011). Canoe binds RanGTP to promote PinsTPR/Mud-mediated spindle orientation. J. Cell. Biol. 195(3), 369–376.

104. Carlson, J. R. and Hogness, D. S. (1985). The Jonah genes: a new multigene family in Drosophila melanogaster. Develop. Biol. 108(2), 341–354.

105. Iyengar, B., Roote, J., and Campos, A. R. (1999). The tamas gene, identified as a mutation that disrupts larval behavior in Drosophila melanogaster, codes for the mitochondrial DNA polymerase catalytic subunit (DNApol-γ125). Genetics 153(4), 1809–1824.

106. Fritz, J. L. and VanBerkum, M. F. (2002). Regulation of rho family GTPases is required to prevent axons from crossing the midline. Develop. Biol. 252(1), 46–58.

107. Genova, J. L., Jong, S., Camp, J. T., and Fehon, R. G. (2000). Functional analysis of Cdc42 in actin filament assembly, epithelial morphogenesis, and cell signaling during Drosophila development. Develop. Biol. 221(1), 181–194.

108. Lince-Faria, M., Maffini, S., Orr, B., Ding, Y., Florindo, C., Sunkel, C. E., Tavares, Á., Johansen, J., Johansen, K. M., and Maiato, H. (2009). Spatiotemporal control of mitosis by the conserved spindle matrix protein Megator. J. Cell. Biol. 184(5), 647–657.

109. Simmons, M. J., Ryzek, D. F., Lamour, C., Goodman, J. W., Kummer, N. E., and Merriman, P. J. (2007). Cytotype regulation by telomeric P elements in Drosophila melanogaster: evidence for involvement of an RNA interference gene. Genetics 176(4), 1945–1955.

110. Chang, H. C. and Rubin, G. M. (1997). 14-3-3 epsilon positively regulates Ras-mediated signaling in Drosophila. Genes Develop. 11(9), 1132–1139.

111. Desai, B. S., Chadha, A., and Cook, B. (2014). The stum gene is essential for mechanical sensing in proprioceptive neurons. Science 343(6176), 1256–1259.

112. Lee, T., Winter, C., Marticke, S. S., Lee, A., and Luo, L. (2000). Essential roles of Drosophila RhoA in the regulation of neuroblast proliferation and dendritic but not axonal morphogenesis. Neuron 25(2), 307–316.

113. Kanda, H. and Miura, M. (2004). Regulatory roles of JNK in programmed cell death. J. Biochem., 136(1), 1-6.