Mucin-like proteins in *Drosophila* development

Zulfeqhar A. Syed

UNIVERSITY OF GOTHENBURG
Institute of Biomedicine
Department of Medical Genetics
Sahlgrenska Academy
2014

ISBN Print edition: 978-91-628-8885-5

ISBN Digital edition: 978-91-628-8884-8

http://hdl.handle.net/2077/33123

Institute of Biomedicine
Department of Medical Genetics
Sahlgrenska Academy
University of Gothenburg

Printed by Kompendiet, Göteborg 2014

Cover illustration: *Drosophila* dorsal vessel (heart) stained with anti-Tnc (red) and anti-α-Spectrin (green)

Abstract

Mucins are large and highly glycosylated proteins and major component of the mucus that coats the lining of epithelial organs. Mucins are characterized by the presence of extended regions rich in the amino acids Proline, Threonine and Serine (PTS domain), where the Serines and Threonines are *O*-glycosylated to form sugar-rich mucin domains. Mucins are classified into secreted gel-forming mucins and transmembrane mucins with possible signaling functions. The amino acid sequence of the PTS domains tends to be poorly conserved between species and different mucins. The goal of this thesis was to identify and study potential mucin-like proteins in *Drosophila melanogaster*. We devised a simple bioinformatic approach and developed a program that can identify PTS domains based on amino acid content. We thereby identified 36 mucins and mucin-related proteins. All proteins appear to be secreted, except for two that harbor a predicted transmembrane domain. Expression analysis at different stages of the *Drosophila* life cycle revealed that many mucins are expressed in the larval gut, consistent with a function in mucosal barrier formation. Interestingly, some of the mucins showed dynamic expression in different tubular organs during embryogenesis. Among these was Mur96B/Tenectin (Tnc) that was further studied to dissect its role in epithelial organ development. We found that Tnc is critical for diameter expansion of the developing hindgut. Tnc forms a transient matrix that fills the hindgut lumen and drives expansion in a dose-dependent manner, presumably by generating a luminal pressure. This study revealed a new mechanism in organ development, whereby the extent of lumen volume expansion can be regulated by the accumulation of single glycoprotein. In parallel to the bioinformatic approach, we identified a *Drosophila* protein that shares conserved domains with human SUSD2 and the non-mucin parts of human MUC4, called Mesh. We aimed to analyze Mesh function as a means to address the roles of these domains. Mesh was found to be is expressed in the digestive tract epithelium from mid-embryogenesis and throughout larval and adult life, localizing to the apical junction belt. Mesh is required for correct organization of the Scribble-complex, a main polarity complex conserved between fly and mammals, to prevent excess expansion of apical cell surface and for microvilli organization. The results demonstrate that mucin-like proteins, containing the PTS domains or other mucin-related domains, are essential for epithelial organ development in *Drosophila*.

Keywords: Mucins, PTS-domain, *Drosophila* development, Tube shape, Hindgut, Midgut, Malpighian tubules, Luminal matrix

PAPERS IN THIS THESIS

This thesis is based on the following papers, which will be referred to in the text by their roman number (I-III)

I. **Syed ZA**, Härd T, Uv A, van Dijk-Härd IF (2008) A Potential Role for *Drosophila* Mucins in Development and Physiology. PLoS ONE 3(8): e3041.

II. **Syed ZA**, Bougé A-L, Byri S, Chavoshi TM, Tång E, van Dijk-Härd, A Uv (2012) A Luminal Glycoprotein Drives Dose-Dependent Diameter Expansion of the *Drosophila* melanogaster Hindgut Tube. PLoS Genet 8(8): e1002850.

III. **Syed ZA**, Byri S, van Dijk-Härd, A Uv. Mesh is a lateral cell adhesion molecule required for apical cell membrane restriction in the *Drosophila* gut epithelium. (*Manuscript*)

Table of Contents

INTRODUCTION **1**
AIM 2
DROSOPHILA AS A MODEL ORGANISM 3
Transposons – a central tool in Drosophila *research* *4*
The UAS-GAL4 system for inducible gene expression *5*
Gene silencing by RNA interference using the UAS-GAL4 system *6*
EPITHELIAL ORGANIZATION 7
Marginal zone *9*
Adherens junctions *10*
Septate junctions *11*
Polarity regulators *12*
DROSOPHILA DIGESTIVE TRACT 14
MALPIGHIAN TUBULES 16
GLYCOSYLATION 17
MUCINS 18
Gel forming mucins *19*
Transmembrane mucins *19*
MUCIN-TYPE *O*-GLYCOSYLATION IN *DROSOPHILA* DEVELOPMENT 20
RESULTS AND DISCUSSION **23**
PAPER I 23
Identification of mucins and mucin-related proteins in Drosophila *23*
Drosophila *mucins are expressed at different stages of life cycle* *26*
Drosophila *mucins are expressed in developing epithelial organs* *27*
PAPER II 28
Tenectin is an intraluminal protein required for diameter expansion of the hindgut *28*
Tnc drives hindgut expansion in a dose-dependent manner *29*
Tnc is a component of O-glycosylated matrix in the hindgut lumen *30*
Model for Tnc-mediated tube dilation *30*
PAPER III 32
Mesh is expressed in the digestive tract and Malpighian tubules *32*
Loss of mesh *results in cysts in the Malpighian tubules* *32*
Loss of mesh causes enlarged apical cell surfaces *33*
Mesh affects localization of the Scribble-complex to the sSJ region *34*
CONCLUSIONS **36**
ACKNOWLEDGMENTS **37**
REFERENCES **41**

Abbreviations

ABP:	Apical basal polarity
AJ:	Adherens junction
AMOP:	Adhesion-associated domain present in MUC4 and other proteins
API:	Application Programming Interface
aPKC:	atypical protein kinase C
Baz:	Bazooka
BLAST:	Basic Local Alignment Search Tool
Bub:	Bubbles
Cdc42:	Cell division cycle 42
CK :	Cysteine Knot
Cont:	Contactin
Cora:	Coracle
Crb:	Crumbs
Dlg:	Discs Large
ECM:	Extracellular matrix
EGF:	Epidermal growth factor
ER:	Endoplasmic reticulum
FasIII:	Fasciclin III
FERM:	4.1/Ezrin/Radixin/Moesin
GalNAc:	*N*-acetylgalactosamine
GlcNAc:	*N*-acetylglucosamine
GlcNAcT:	*N*-acetylglucosaminyltransferase
Gli:	Gliotactin
GUI:	Graphical user interface

GUK:	Guanylate Kinase
Kune:	Kune-kune
Lac:	Lachesin
Lgl:	Lethal giant larva
LRR:	Leucine-rich repeat
MAGUK:	Membrane associated guanylate kinase
Mega:	Megatrachea
MZ:	Marginal zone
NIDO:	Extracellular domain of unknown function, found in nidogen (entactin) and hypothetical proteins
Nrg:	Neuroglian
Nrx-IV:	Neurexin IV
Par:	Partition defective
Par3:	Partition defective-3
Par6:	Partition defective-6
Patj:	Pals1-associated tight junction protein
PCP:	Planar cell polarity
PCR:	Polymerase chain reaction
PDZ:	Domain present in PSD-95, Dlg, and ZO-1/2.
PerA:	Peritrophin-A
ppGalNAcT:	polypeptide *N*-acetylgalactosaminyltransferase
Pro:	Proline
PTSPMiner:	Proline Threonine Serine Pattern –Miner
Scrib:	Scribble
Sdt:	Stardust
SEA:	Sea urchin sperm protein, enterokinase and agrin

Ser:	Serine
SH3:	SRC Homology 3
Sinous:	Sinu
SJ:	Septate junctions
Ssk:	Snake skin
SUSHI:	Complement control protein (CCP) modules, or short consensus repeats (SCR)
TEM:	Transmission electron microscopy
Thr:	Threonine
Tnc:	Tenectin
UAS:	Upstream activating sequence
UDP-GalNAc:	Uridine diphosphate *N*-acetyl-α-galactosamine
Vari:	Varicose
VWC:	Von Willebrand factor C
VWD:	Von Willebrand factor D
Yrt:	Yurt

Introduction

The defining characteristic of metazoans is the presence of epithelial cells that are organized into multicellular tissues and organs. Epithelial cells cover the outer surface of the body and line all our vital internal organs. Several types of epithelia exist to fulfil important cellular and physiological functions, such as control and delivery of gases, nutrient exchange, secretion of enzymes, secretion of hormones and excretion of waste products. Most epithelia also serve an important function in protecting the underlying tissues from mechanical injury, harmful chemicals, invading microorganisms and in preventing excess loss of water by acting as selective and dynamic barriers between the internal compartments and the external environment. In its simplest form, organ epithelia consist of a layer of structurally and functionally similar epithelial cells. These can be wrapped into complex three-dimensional hollow structures, thereby generating tubular organs with diverse shape and size. During development, tubular primordia can arise by various mechanisms [1]. Once formed, however, the rudimentary tube generally has a small lumen that must grow in size to acquire appropriate dimensions to satisfy physiological demands. Acquiring characteristic shape and size is critical, as an obstructed or misshapen tube leads to compromised organ function. Consequently, defects in tube size are implicated in human diseases, such as vessel aneurysms and polycystic kidney disease [1,2].

Mucins are large, highly *O*-glycosylated proteins, and are the main component of the protective mucosa that lines the luminal surface of epithelial organs. A few studies have suggested that mucins not only function to protect epithelia from the external environment, but also have roles in epithelial organ development. These studies reported expression of mucins in human fetal organs, such as the gastrointestinal tract, respiratory tract, kidneys and male genital ducts [3-7]. Parallel studies that employed various lectins and antisera against mucin-type *O*-glycosylation to label glycans during animal development have shown that the lumen of growing epithelial organs are lined, or sometimes filled with glycan-rich components in a dynamic pattern as development proceeds. Examples include the developing rabbit kidney, embryonic chick lung, and most epithelial organs of the fruit fly [8-12]. Thus, some of the *O*-glycans detected during development might represent mucin-like proteins. The presence of luminal components

rich in *O*-glycans in different developing organs and from different species is intriguing and suggests important roles for such components during development.

Aim

The aim of this thesis is to identify mucin and mucin-related genes in *Drosophila* and characterize their potential involvement in epithelial organ development.

Drosophila as a model organism

The fruit fly, *Drosophila melanogaster*, was first introduced as an invertebrate model organism to study classical genetics more than a century ago. In 1910, the discovery of the white mutation by Thomas Hunt Morgan and subsequent contribution by his graduate students kick-started the systematic use of *Drosophila* for genetic research. Since then, fly genetics has been successfully applied to study fields spanning from developmental biology to physiology, enriching our understanding of the genetic principles and molecular mechanisms underpinning biology [13]. Besides being a powerful model organism to study basic biology, *Drosophila* has over the years been widely used to study genetic components of various human pathologies, as it turns out that around 75% of the known human diseases genes have counterparts in fruit flies [14,15].

The enormous success of *Drosophila* as a model organism originates from the numerous practical advantages it has to offer. In addition to its powerful genetics and small size, fruit flies have a short generation time. *Drosophila* undergoes holometabolous development; this involves complete transformation of the immature larva, mostly lacking adult structures, into the adult fly (imago). The life cycle of *Drosophila* consists of six stages: embryo, 1st instar larva, 2nd instar larva, 3rd instar larva, pupa and adult. The duration of the life cycle varies with temperature. At 25°C, embryonic development takes approximately 21-22 hours, after which the embryo hatches into a larva. The larva grows continuously, undergoing two molts from 1st instar to 3rd instar larva and a final molt into an immobile pupa over a period of 4 days. During the pupal stages, larval tissues undergo histolysis while new adult body structures are built in a process called metamorphosis. After 5 days of pupariation, adult flies eclose from the pupal case, and it takes up to 8 hours for the newly eclosed flies to attain sexual maturity. Overall, the life cycle of *Drosophila* takes ten days at 25°C.

Drosophila is relatively easy and cost-effective to maintain in the lab, facilitating high throughput experiments involving large numbers of different fly stocks. The fruit fly has three pairs of autosomal chromosomes and two sex chromosomes (X and Y). Recombination is confined only to female flies, which is a major advantage when performing fly genetics. *Drosophila* has many genetic tools in its arsenal, of which the most unique are the balancer chromosomes, which supress recombination between

homologous chromosomes, and plethora of phenotypic markers. In addition, a multitude of transposon variants enable genetic strategies to manipulate a gene of interest at various temporal and spatial resolutions [16-18]. The constantly growing *Drosophila* research community and the collaborative effort to broaden the technical repertoire of *Drosophila* genetic tools make it a favourite organism for many researchers.

Transposons – a central tool in* Drosophila *research

Transposons are mobile genetic elements present in the genomes of most metazoans and have become an important tool in genome research [19]. Transposable elements provide powerful means to understand genome evolution and as tools for genetic manipulation. In general, transposable elements encode enzymes called transposases that mediate DNA cleavage and transposition of the transposable element in the genome. Transposable elements are commonly known as "jumping genes" and were discovered by geneticist Barbara McClintock in the 1940s while she was studying the pattern of pigmentation in maize. McClintock showed that the irregular pigmentation in maize was caused by genetic elements that transposed from one locus to another. For her ground breaking work on transposable elements, McClintock was awarded Nobel Prize in 1983 [20].

The *Drosophila P element* is one of the most widely used and best characterized eukaryotic transposons. *P elements* are thought to have entered *D. melanogaster* by horizontal transfer from another distantly related *Drosophila* species about 80 years ago [21]. They were first recognised as factors in P strains responsible for hybrid dysgenesis, and since then they have become widely used tools for studying gene function in *Drosophila* [22]. The *Drosophila P element* is a 2.9 kb DNA transposon that encodes a 87 kDa transposase protein, and transposition within the genome occurs by a cut-and-paste mechanism that requires approximately 150 bps of specific sequence at each end of the *P element* [23]. The sequences required for transposition includes 31 bps terminal inverted repeats, internal transposase-binding sites, and internal 11 bps inverted repeats [24-26]. *P elements* can either be autonomous, where they encode their own source of transposase needed for mobilisation, or non-autonomous, where an external source of transposase is required. Non-autonomous *P elements* in the form of engineered constructs were initially used for random gene disruptions [27]. Since then, *P elements* have been adapted and modified for different purposes of transgenesis, such as various types of gene tagging,

insertion of specific enzymatic target sites into the genome, and inducible gene expression [16,28-31]. Other well-studied DNA transposons that have now been adapted for use in *Drosophila* research are the piggyBac and Minos elements [32-35]. Genetic and molecular data of all *Drosophila* genes including different transposon-induced alleles and transgene information is made available by the FlyBase consortium [36].

The UAS-GAL4 system for inducible gene expression

Targeted gene expression is an important tool in the characterization of individual gene function. The UAS-GAL4 system is an extremely useful tool for selective expression of any cloned gene in a wide variety of cell and tissue specific patterns in *Drosophila*. This binary system is based on the *Saccharomyces cerevisiae* transcriptional activator called GAL4 that binds to specific DNA sequence called UAS_G (galactose upstream activating sequence) and activates the transcription of linked genes [16,37]. The key feature of this system is that the GAL4 gene and the UAS-target gene, both of which are introduced into the fly genome by transposon-mediated integration, are initially separated into two distinct transgenic lines. One strain expresses GAL4 under the control of a tissue specific enhancer or promoter and is generally referred to as the driver line. The other strain carries the gene of interest or reporter gene downstream of the UAS sequence. GAL4 has no detrimental consequences in the fly, even at elevated levels, and the UAS transgenes are largely silent in the absence of GAL4. When the fly strains are crossed to each other, the combination of the two transgenes in the progeny of the cross results in GAL4-binding to the UAS sequence and activation of the target gene (Figure 1). The progeny can thus be conveniently analyzed to study the effects of directed gene expression. There are wide selections of GAL4 driver lines available that express GAL4 in different cells and tissues at different stages of development, or upon conditional induction, allowing for targeted expression of a UAS transgene in a selected spatio-temporal manner.

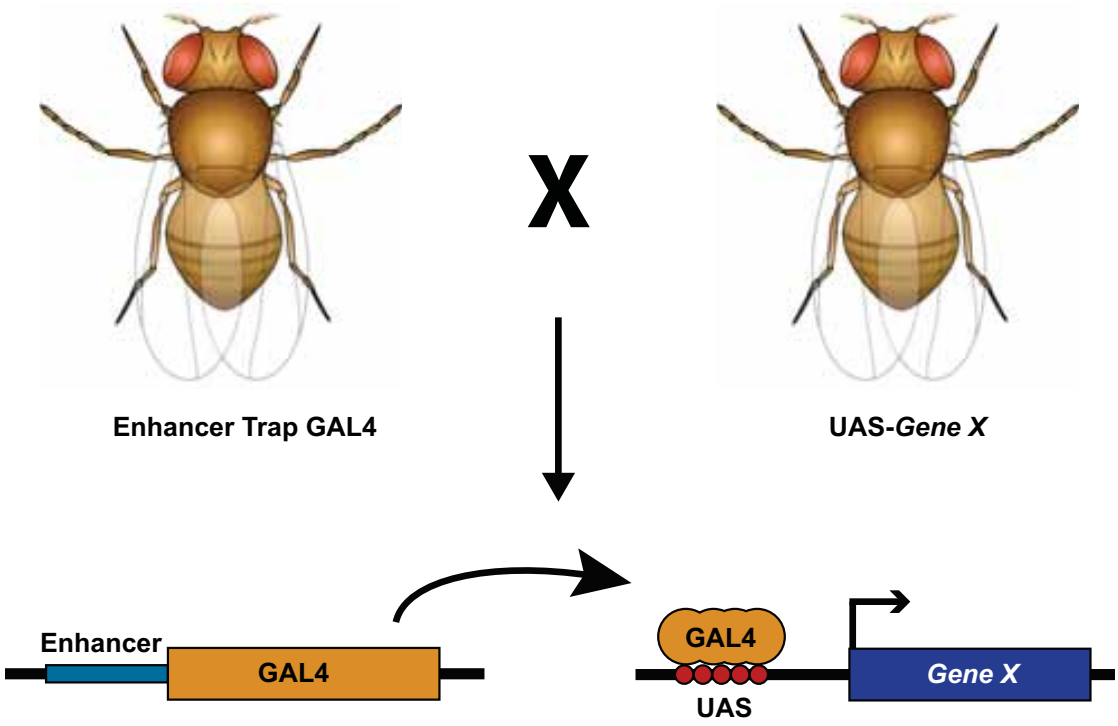

Figure 1: The UAS-GAL4 system allows targeted expression of any cloned gene in a tissue-specific manner. This bipartite system utilizes the yeast transcriptional activator GAL4 to activate expression of a target gene fused downstream of UAS. This system consists of two distinct transgenic strains: a GAL4-driver line and a UAS-line. The GAL4-driver carries the GAL4 gene inserted into the *Drosophila* genome and expresses GAL4 under the control of nearby enhancers. In the progeny of a cross between these transgenic fly strains, GAL4 binds to the UAS sequences to activate expression of the linked gene in the cells where GAL4 is expressed. *Figure adapted from [16]*

Gene silencing by RNA interference using the UAS-GAL4 system

Ribonucleic acid interference (RNAi), also referred to as post-transcriptional gene silencing, is an important biological pathway in which double-stranded RNA (dsRNA) molecules induce sequence-specific inactivation of gene function. The phenomenon of RNAi was first discovered in the nematode *Caenorhabditis elegans* [38] and is an endogenous cellular mechanism used by most eukaryotes to regulate gene expression [39]. When dsRNA molecules are present a cells, it triggers the RNAi machinery, wherein dsRNA is cleaved into smaller fragments that are used to target homologous mRNA sequences for degradation, resulting in inactivation of gene expression [40]. In recent years, RNAi has developed into a powerful tool to manipulate gene expression, and is

used for many invertebrate and vertebrate model organisms and cell culture systems to probe gene function [41].

In *Drosophila*, RNAi can be induced either by injecting embryos with *in vitro* transcribed dsRNA before cellularization or by germline transformation of a transgene that expresses a hairpin-forming RNA sequence under the control of the UAS promoter [41]. RNAi mediated by injection has its limitation in that studies of gene function are restricted to embryonic development, and sometimes, maternal contribution may alter embryonic phenotype. Targeted expression of an RNAi transgene, using the UAS-GAL4 system offers several advantages over the injection approach. Depending on the GAL4-driver line used, cell-type, tissue-specific or developmental stage specific probing of gene function can be achieved. This approach allows spatio-temporal control of gene knockdown and has been extensively used in reverse genetics for rapid investigation of gene function. Transgenic-based RNAi is simple, as it requires only one fly cross, and communal efforts have been made to generate RNAi libraries covering most of the *Drosophila* protein-encoding genes [42]. In *Drosophila* cell culture systems, RNAi can be activated by adding dsRNA to the cell culture [43].

Epithelial organization

Organ epithelia are polarized, such that the apical surface faces the luminal space of an organ or the exterior of an organism. The basal domain faces the basement membrane or the underlying extracellular matrix (ECM) mediating cell-matrix adhesion [1]. This apical to basal polarization is manifested in each epithelial cell, which displays well-defined apical, basal and lateral membrane domains with different protein and lipid composition and oriented organization of cytoskeletal components and cytoplasmic organelles (Figure 2). The lateral cell domain faces neighbouring epithelial cells and connects the cells by means of structurally defined intercellular junctions [1,44]. Such asymmetric partition of along the apical to basal axis of cells confers specific structural and functional properties to epithelia and is generally referred to as apical-basal polarity (ABP) [2]. In addition to ABP, many epithelial tissues are polarized along the plane of the epithelium; this is known as planar cell polarity (PCP) [2]. The ability of epithelial cells to maintain their ABP is essential for preserving tissue integrity and, consequently, loss of cell polarity due to infection, diseases or genetic predisposition underlie many human pathologies [45,46].

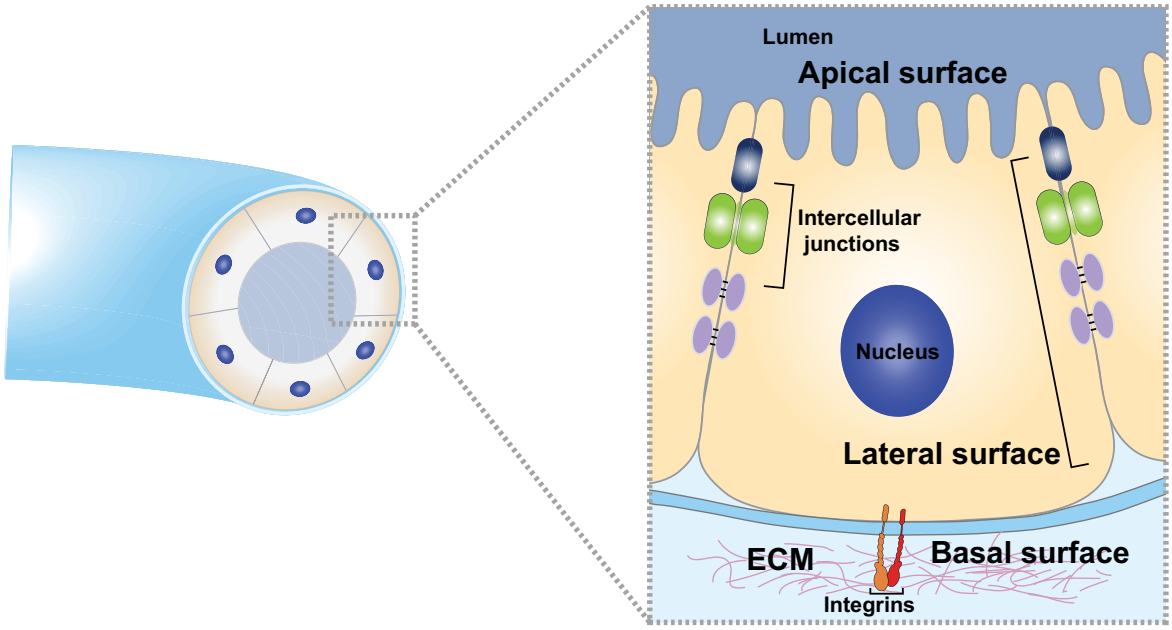

Figure 2: Typical architecture of a simple epithelial tube. The tube wall is composed of epithelial cells that display apical-basal polarity. The apical surface faces the lumen, while the basal surface is exposed to the underlying basement membrane and interacts with the ECM. The lateral membrane faces neighbouring cells and possesses structurally defined intercellular junctions that provide cell-cell adhesion and a diffusion barrier between the apical and basolateral surfaces. *Figure adapted from [2]*

In *Drosophila*, the apical membrane domain encompasses the free apical surface and a narrow region of cell-cell contact at the most apical part of the lateral domain known as the marginal zone (MZ) [44]. The MZ corresponds to the position of the vertebrate tight junctions (TJ), and several apical polarity regulators contained within the MZ have mammalian homologues that are found to localize to the TJs [47,48]. Unlike the MZ, the TJs also provide a permeability seal that restricts free diffusion of ions and solutes across the paracellular space and forms a "fence" that separates the apical and basolateral domains [49]. Basal to the MZ lies the adherens junction (AJ) that provides strong cell-cell adhesion by forming a circumferential belt around the epithelial cell [44]. The epithelial barrier functions in *Drosophila* are mediated by the septate junctions (SJ), which are found basal to the AJs (Figure 3) [44].

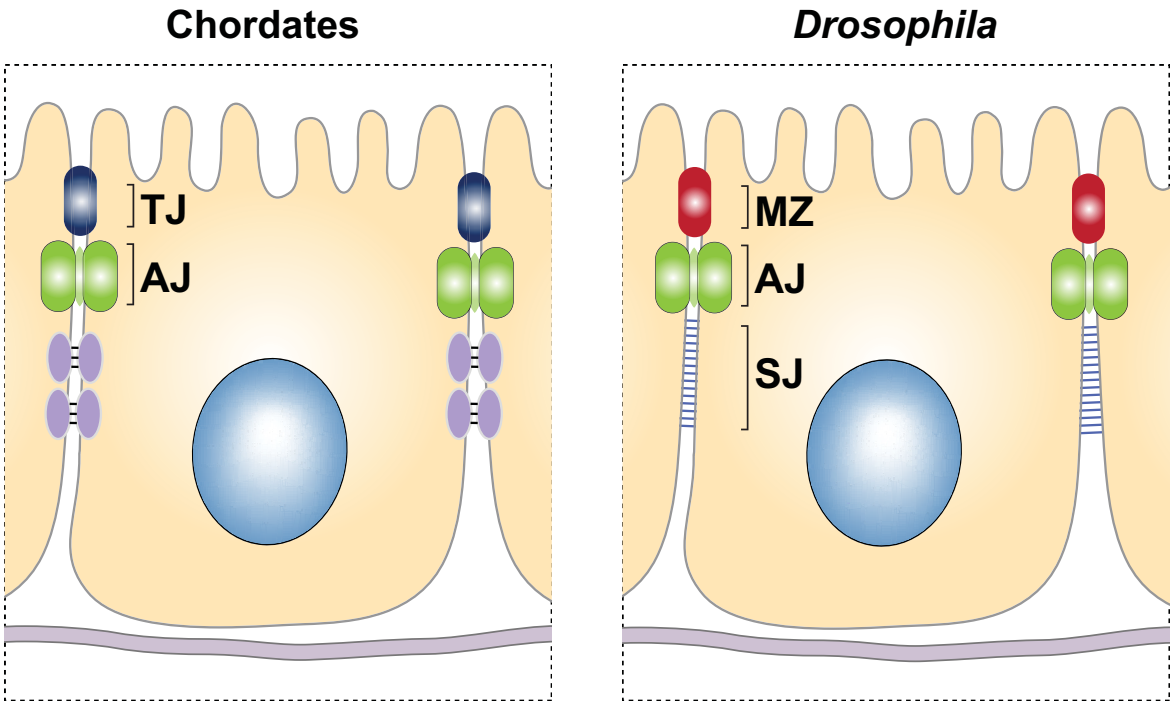

Figure 3: Schematic presentation of the apical junctional complex in chordates (left) and *Drosophila* (right). The apical-most region of cell-cell contact is represented by the tight junction (TJ) in chordates and the marginal zone (MZ) in *Drosophila*. Both *Drosophila* and chordates exhibit adherens junctions (AJs) basal to the MZ and TJ, respectively. Below the AJs in *Drosophila* are specialized junctional structures known as the septate junction (SJ). SJs are characterized by ladder-like bridges when observed using electron microscopy, unlike TJs that appear as anastomosing intramembranous strands. Despite the morphological differences, both junctions are responsible for maintaining the diffusion barrier and separating the apical and basolateral domains. Desmosomes are indicated below the AJs in chordates, and are absent in *Drosophila*. *Figure adapted from [44].*

Marginal zone

Several apical polarity proteins localizes to the MZ. Traditionally, these have been defined into the Crumbs and Par (partitioning defective) polarity complexes, which function as apical determinants and regulators of epithelial cell polarity in *Drosophila* [50]. The Crumbs complex consists of Crumbs (Crb), Stardust (Sdt), Pals1-associated tight junction protein (Patj) and Lin-7 [50]. Crb is a large transmembrane protein with twenty-nine epidermal growth factor (EGF)-like domains and four Laminin A G-domains in the extracellular region [51]. The extracellular domain of Crb is involved in homophilic Crb-Crb interactions [52]. The transmembrane region is followed by a small highly conserved cytoplasmic region, which contains a functionally important binding site for the

4.1/Ezrin/Radixin/Moesin (FERM) domain and a C-terminal Postsynaptic density 95/Discs large/Zonula occludens-1 (PDZ) binding motif [51]. Crb is detected in all ectodermally derived epithelia from the time of gastrulation and confers apical membrane characteristics and promotes apical membrane growth [53-55]. Sdt is a membrane-associated guanylate kinase (MAGUK) protein belonging to the MPP/P55 (membrane protein palmitoylated) subfamily of MAGUKs [50]. It contains a PDZ domain, an SH3 (SRC homology 3) domain, a GUK (guanylate kinase) domain, a HOOK domain and two L27 domains. In addition, it contains the evolutionary conserved ECR1 and ECR2 domains at the N-terminus. The PDZ domain of Std binds to the cytoplasmic C-terminus of Crb [56], and the two L27 domains bind to Patj and Lin-7, respectively [57]. Loss of Sdt results in similar epithelial defects as seen with loss of Crb [51]. Patj has a single L27 domain and four PDZ domains and plays a minor role in epithelial polarization [58].

The Par complex consists of Bazooka (Baz, Partitioning defective-3 (Par3) in *C. elegans*), Partitioning defective-6 (Par6), atypical protein kinase C (aPKC) and the small GTPase Cdc42 [50]. Baz contains three PDZ domains and together with Par6, which has a single PDZ domain and a semi-CRIB domain, forms a complex with aPKC [59,60]. Baz has an early role in the formation of AJs [61]. aPKC is a serine/threonine kinase with several important targets that contribute to its role as an evolutionary conserved epithelial polarity determinant [50,62,63]. There are multiple interactions among the members of the Par complex and the Crb complex, and both these complexes are required for the establishment of ABP [50,63,64].

Adherens junctions

The adherens junctions (AJs) are cell-cell adhesion complexes located below the marginal zone [44]. They define the boundary between the apical and basolateral domains, and have multiple roles in development and homeostasis, including cell-cell adhesion, anchoring of the cytoskeleton to the plasmamembrane, signal transduction and transcriptional control [65]. Classical cadherins are the core components of AJs, and the assembly of AJs typically begins by homophilic cis- and trans-clustering of these transmembrane proteins [61]. The basic features of cadherins are the presence of extracellular cadherin repeats that mediate calcium dependent cell-cell adhesion and a highly conserved cytoplasmic tail that interacts with cytoplasmic proteins called catenins

[61]. In *Drosophila*, members of the classical cadherin family include: DE-cadherin/Shotgun (DE-Cad), DN-cadherin (DN-Cad) and DN-cadherin2 (DN-Cad2) [61]. DE-Cad is the major epithelial cadherin. Its cytoplasmic tail binds to p120 catenin and beta-catenin (Armadillo), and beta-catenin binds to alpha-catenin that subsequently links to the actin filament, forming a circumferential actin belt around the cells [61].

Septate junctions

Septate junctions (SJs) are cell-cell junction complexes found basal to the adherens junctions [44]. After the establishment of ABP and AJs, these junctions can be detected by electron microscopy as ladder-like septa that span the intermembrane space [66]. The SJ strands meander along the lateral membrane forming a labyrinth-like structure that inhibits diffusion of molecules between the apical and basal domains [67,68]. They thereby provide a paracellular barrier and a fence function analogous to the vertebrate tight junctions, and mutational analyses of SJ proteins has revealed disruption of the paracellular barrier in dye permeation experiments [68].

SJs are structurally and molecularly similar to vertebrate paranodal junctions, which are formed between axons and myelinating glial cells at the node of Ranvier [44]. SJs are classified into two types, pleated septate junctions (pSJs) and smooth septate junctions (sSJs). pSJs are found in most of the ectodermally derived epithelia and glial sheets, whereas sSJs are found in endodermally derived tissues [66]. The difference between the pSJs and sSJs is the arrangement of SJ-strands. The strands in pSJs form regular undulating “pleated” lines, whereas the strands in sSJs form linear bands [66]. pSJs are the most prominent junctions in *Drosophila,* and recent studies have identified several SJ components involved in epithelial barrier formation, such as Neurexin IV (Nrx-IV), Neuroglian (Nrg) , Na+/K+-ATPase α- and β- subunit (ATPα and Nrv2), Gliotactin (Gli), Lachesin (Lac), Contactin (Cont), Coracle (Cora), Yurt (Yrt), Varicose (Vari), Disc Large (Dlg), Lethal giant larva (Lgl), Scribble (Scrib), and Fasciclin III (FasIII) [69-79]. In addition, three homologues of Claudin, a major component of vertebrate TJs, are part of *Drosophila* SJs and are called Megatrachea (Mega), Sinous (Sinu), and Kune-kune (kune) [80-82]. Intact pSJs also appear to be required for correct apical secretion of chitin modifying enzymes [83]. In a recent study, it was found that many pSJ components (Nrx-IV, Nrg, ATPα, Nrv2, Sinu, Mega, and Vari) are highly mobile on the lateral membrane at

embryonic stage 12, when the SJs begin to form, but become rather immobile at stage 13, suggesting that they form a structural core of the SJs [84]. Loss of any of the core SJ components dramatically affects the mobility of the others, indicating interdependence for stable SJ complex formation. On the other hand, Lgl, Dlg and Scrib remain relatively mobile, also after stage 13, and are not considered to be part of the core complex [84].

Scrib is a membrane-associated scaffolding protein and belongs to the LAP (LRR and PDZ) protein family. It contains sixteen N-terminal Leucine-rich repeats (LRRs) and four PDZ domains [78]. Dlg is a member of the MAGUK super family of proteins, which has three PDZ domains, a SH3 domain and a GUK domain and acts as a scaffolding protein [77]. PDZ domains bind to short PDZ-binding motifs that are located in the C-terminus of target proteins, and help to anchor the target protein to the correct membrane domain. Both transmembrane and cytosolic proteins can be targeted to membrane complexes through PDZ interactions [85]. Lgl is a WD40 repeat protein and, unlike Dlg and Scrib, is not a scaffolding protein. Lgl has conserved phosphorylation sites that are critical for its localization and function [86]. Mutations in any of these components result in disruption of epithelial organization and expansion of the apical membrane [64].

In contrast to pSJs, relatively few components of the sSJs have been characterized. Some studies have reported the localization of αβ spectrin, Ankyrin and FasIII to the sSJs region in the midgut [87,88]. A better understanding of the components of the sSJs, and how the polarization machinery works in endoderm-derived organs is only now emerging. Two molecular components specific for the sSJs have recently been identified, called Snake skin (Ssk) and Mesh [89,90]. These were found to be expressed in endoderm-derived tissues like the midgut, proventriculus (out layer) and Malpighian tubules. Loss of either of these components results in a compromised paracellular barrier in the midgut and mis-localization of other sSJ-associated proteins, such as FasIII, Cora and Lgl, accompanied by larval lethality [89,90].

Polarity regulators

In an epithelial tissue, the shape and function of the constituting cells are completely dependent on their polarization. The separation of the distinct surfaces of the plasmamembrane prevents mixing of receptors, channels and transporters between the

domains. It also retains key protein complexes that are involved in protein sorting, recycling, trafficking, and signalling to distinct compartments [50]. Genetic studies in *Drosophila* and *C.elegans* have identified a set of conserved proteins that are involved in the establishment and maintenance of cell polarity [55,64,78,91-93]. These include, in addition to the members of the Crb, Par and Scrib complexes described above, components of the recently identified Yrt/Cora complex (Yurt, Cora, Nrx-IV and Na+/K+-ATPase) [94] and Partitioning defective-1 (Par-1), which localize to the basolateral cell domain [95]. These complexes act in a mutually antagonistic relationship to define the apical and basolateral domains.

The traditional description of the apical Crb and Par polarity complexes is now changing, as the components of the two complexes appear to interact in an interdependent apical protein network. For example, Crb can bring together Sdt, PatJ, PAR-6 and aPKC though its intracellular domain [50,96]. Polarization of different epithelial cell types can require different components of the polarization network and can occur by slightly different mechanisms, but a simplified mechanism of polarity establishment has been put forward [96]. In this model, the transmembrane protein Crb defines apical cell identity by localizing Sdt, PatJ, PAR-6 and aPKC to the MZ through protein interactions mediated by its intracellular domain. This network excludes the presence of Baz in the apical cell domain via aPKC-mediated phosphorylation of Baz. Baz interacts with components of the AJs, recruiting these to the region where Baz is enriched. The apical polarity regulators therefore restrict AJ formation in the apical direction. The basolateral polarity regulators prevent Baz from moving in the basal direction by Par-1-mediated phosphorylation of Baz. The apical and basal polarity proteins also antagonize each other. The apical polarity proteins exclude the Scrib complex from the apical cortex through phosphorylation of Lgl by aPKC. The Scrib complex in turn antagonizes the apical regulators, partly through interaction of Lgl with aPKC, to maintain the basolateral domain. The second basolateral polarity complex Yurt/Cora negatively regulates the activity of the Crb complex and stabilizes the basolateral membrane [94]. Thus, the apical and basal polarity regulators establish ABP by creating an equilibrium through mutually modulating each other activity (Figure 4) [96].

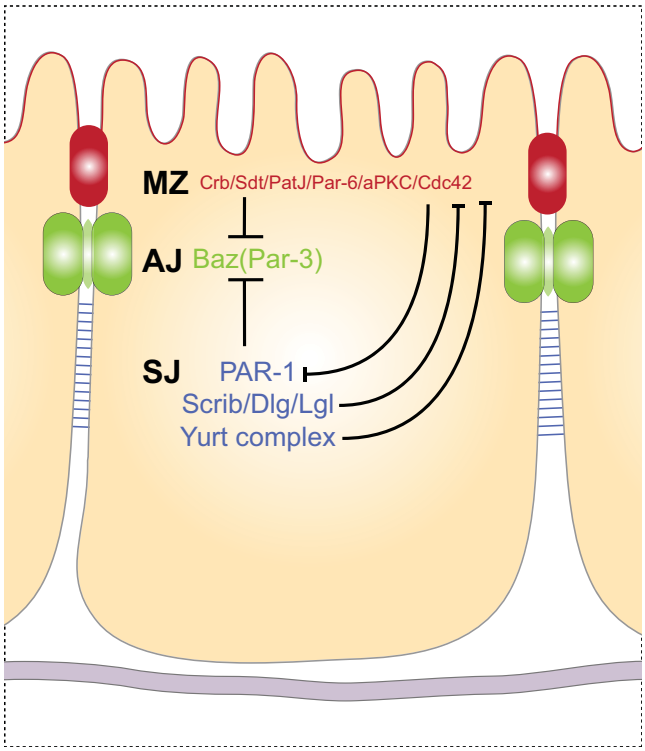

Figure 4: Epithelial polarity factors and their interactions. Positive feedback among members of the apical polarity determinants and mutual antagonism between the apical and basolateral determinants are required for the formation and maintenance of apical-basal polarity. *Figure adapted from [96]*

Drosophila digestive tract

The *Drosophila* larval digestive tract is divided into three distinct anatomical regions: the foregut, midgut and hindgut. The foregut and hindgut are ectodermal in origin, while the midgut is derived from the endoderm and forms a secondary epithelium by undergoing a mesenchymal-to-epithelial transition during late embryogenesis [97,98]. The foregut arises from invagination of stomodeal cells from the anterior region of the blastoderm embryo. The posterior part of the foregut makes contacts with the anterior midgut, and through a series of cellular events, including cell division and cell shape changes, the posterior part of the foregut tube makes an inward movement and becomes surrounded by the anterior-most part of the midgut. This gives rise to a bulb-like three-layered structure called the proventriculus, where the outer layer is derived from the endoderm and the middle and inner layers are ectodermal in origin [99,100]. Thus, the proventriculus develops at the boundary of the foregut and midgut and serves as a valve regulating the

passage of food into the midgut. The fully developed foregut consists of the atrium, pharynx, oesophagus and proventriculus [98-100].

The larval midgut is composed of two cell layers [97]. The outer visceral muscle layer is organized into circular and longitudinal muscles and is derived from visceral mesoderm, while the inner epithelial layer is derived from the endoderm [97,101]. The midgut arises from two spatially separated primordia at the anterior and posterior ends of the blastoderm embryo, called the anterior midut primordium (amp) and the posterior midgut primordium (pmg) [102]. The visceral mesoderm derives from clusters of mesodermal cells in parasegment 2-13, which join together to form a continuous band of cells on each side of the embryo [102]. The formation of the midgut begins when the midgut primordia (amg and pmg) invaginate, lose their epithelial properties through epithelial-to-mesenchymal transition and start migrating along the bands of visceral mesoderm that serve as tracks [97,101]. The migrating primordia meet in the middle of the embryo, undergo a mesenchymal-to-epithelial change and fuse to form two bands of cells, that along with visceral mesoderm, extend ventrally and dorsally to wrap around the yolk to form a midgut tube. Depending on the interaction of the visceral mesoderm with the endoderm, the midgut tube generates three constrictions that subdivide the midgut into four lobes. The first lobe gives rise to the outer layer of the proventriculus and the four gastric caeca. The second, third and the fourth lobe develop into the anterior midgut, middle midgut and posterior midgut respectively [97,102].

The hindgut arises by invagination of a group cells, called the proctodeal primordium, at the posterior end of the blastoderm embryo [98]. The invagination elongates through convergent cell extension, accompanied by changes in cell size and cell shape to form a narrow, left-right asymmetric, shepherd's crook shaped tube. A transcriptional hierarchy consisting of Drumstick (Drm), Lines (Lin) and Bowl controls hindgut patterning during tube elongation [103,104]. The elongated hindgut tube is divided into morphologically distinct sub domains. The anterior-most domain, called the small intestine, lies just posterior to the midgut and is followed by the large intestine and the posterior-most rectum. The large intestine is partitioned into ventral and dorsal domains by two lines of cells at each lateral side of the tube, referred to as 'border cells' [105]. The border cells also form circumferential rings at the border between the small and the large intestine and

between the large intestine and the rectum. During later stages of development, the hindgut grows by tube elongation and by luminal diameter expansion [104].

Malpighian tubules

The Malpighian tubules are the excretory organs of *Drosophila*, and are functionally equivalent to the vertebrate kidney. The larval Malpighian tubules consist of two pairs of single-cell layered epithelial tubes that originate from the hindgut during embryogenesis [98,106]. The development of the Malpighian tubules in *Drosophila* involves successive morphogenetic events including (a) cell specification and eversion of the tubule primordia (b) cell proliferation (c) cell rearrangement and tube elongation and (d) cell differentiation [106].

Malpighian tubule cells are specified by interactions between the midgut and the hindgut and depends on the zinc-finger transcription factor Krüppel (Kr) and the homeodomain-containing protein Cut [107]. In the presence of the Kr and Cut transcriptional regulators, four clusters of Malpighian tubule primordial cells start to bud from the hindgut. One cell from each bud is selected to become a tip cell by lateral inhibition, specified by the Notch pathway. This tip cell secretes EGF and promotes cell division in its neighbouring cells [107,108]. When cell division ceases, the tubules are short with 8-12 cells encircling the lumen. These cells make up the main cell type of the Malpighian tubules, known as the principal cells (PC) [106]. Subsequent tubule growth and elongation occurs largely by cell rearrangements and cell intercalation that decreases the number of cells at the circumference of the tubules. As the tubules extend, they undergo stereotypic path-finding through the body cavity with the anterior tubules moving forward towards the thorax region and the posterior tubule protruding along either side of the hindgut [106]. During this phase of tube elongation, a population of cells from the caudal mesoderm incorporate into the tubule, by undergoing a mesenchymal-to-epithelial transition, and differentiate into a physiologically distinct cell type known as stellate cells (SC) [109]. As the tubules develop, SC are progressively integrated into the epithelium between the PC [106]. The SCs become apical-basal polarized once they are incorporated into the tubules. At the end of embryogenesis, the Malpighian tubules have attained their final architecture with an extensive increase in length and a narrow lumen with two cells at the circumference. Before hatching, precipitates of uric acid are visible in the tubule lumen, indicating the

onset of excretory activity. The excretory function of the Malpighian tubules relies on the combined function of PCs and SCs [106].

Glycosylation

Post-translational modifications are of critical importance to the function of an expressed protein [110]. Two of the most abundant forms of posttranslational modifications that involve carbohydrates are *N*- and *O*-linked glycosylation, distinguished by their glycosidic linkages to amino acid side chains [111]. Glycosylation results in the addition of sugar groups to the protein, and takes place in the lumen of the endoplasmic reticulum (ER) and the Golgi complex [111] .

N-linked glycosylation is initiated in the ER, and further processing takes place in the Golgi complex. *N*-linked glycans are characterized by being linked to the amide nitrogen atom in the side chain of Asparagine (Asn). An asparagine residue can accept an oligosaccharide only if the residue is part of an Asn-X-Ser or Asn-X-Thr consensus sequence, where X can be any amino acid except proline [112]. *N*-linked glycoproteins acquire their initial sugars from Dolichol donors in the ER [112]. All *N*-linked oligosaccharides have in common a oligosaccharide core, consisting of three glucoses, nine mannoses and two *N*-acetylglucosamine (GlcNAc) residues ($Glc_3Man_9GlcNAc_2$), which serve as the foundation for a wide variety of *N*-linked oligosaccharides that are categorized into High-mannose type, Complex type and Hybrid type [112]. The final oligosaccharide structure acquired on the mature glycoprotein is dictated by the action of different glycosyltransferases and glycosidases residing in the ER and Golgi complex [112].

O-linked glycosylation takes place in the Golgi complex [113] . *O*-linked glycans are linked to the oxygen atom in the side chain of Serine (Ser) or Threonine (Thr). Unlike *N*-glycosylation, *O*-glycosylation does not begin with the transfer of an oligosaccharide from a Dolichol precursor, but with the addition of a single monosaccharide [113]. Mucin type *O*-glycosylation is initiated by the enzymatic addition of a *N*-acetylgalactosamine (GalNAc) residue to the side chain of Ser or Thr by the UDP-GalNAcT:polypeptide *N*-acetyl-galactosaminyltransferase (ppGalNAcTs), referred to as GalNAc transferases in mammals and PGANTs in *Drosophila* (EC 2.4.1.41) to generate the Tn-antigen (GalNAc-

α-1-O-Ser/Thr) [114,115]. Subsequent elongation by transferases yields eight distinct core structures, which can be further elongated or modified by Sialylation, Sulfatation, Acetylation, Fucosylation, and Polylactosamine-extension to build hundreds of different *O*-glycan chains [113,114]. *O*-glycans with *O*-linkages to Ser or Thr other than GalNacA includes *O*-linked fucose, glucose, mannose, xylose and GlcNAc. A large family of ppGalNAcTs exists, which indicates redundancy in the activity of these enzymes and spatio-temporal expression and substrate preference [114-116].

Mucins

Mucins are large and highly glycosylated multifunctional proteins found on the surface of epithelial tissues lining the respiratory, digestive and urinogenital tracts [117,118]. Mucins are the major component of the mucus that protects underlying epithelial cells from infection, dehydration and physiological or chemical injury [119]. A common structure in mucins is a protein backbone termed "apomucin", which is decked with a large number of *O*-linked oligosaccharides and a few *N*-glycan chains [118]. Apomucins contain variable numbers of tandem repeats that are particularly rich in amino acids Ser and Thr whose hydroxyl groups will become *O*-linked with oligosaccharides. These tandem repeat regions are called PTS domains (Proline, Threonine and Serine) or mucin-domains. The *O*-linked oligosaccharides account for up to 80% of the molecular mass of the mucin and results in a highly extended and rigid structure of the mucin [118]. They have high water holding capacity, and are therefore largely responsible for the viscous nature of mucus [119]. The PTS domains are not conserved between species and can vary from one mucin to another [120,121]. The heavily glycosylated mucin domains adapt an outstretched conformation, best described as a "bottle brush", where the stalk represents the protein backbone and the bristles are represented by oligosaccharide chains [111].

Mucins have been subdivided into gel-forming and membrane-bound forms. In humans, there are nine membrane-bound mucins (MUC1, MUC3, MUC4, MUC12, MUC13, MUC15, MUC16, MUC17 and MUC20) and five secreted gel-forming mucins (MUC2, MUC5B, MUC5AC, MUC6 and MUC 19 [118].

Gel forming mucins

A characteristic of gel-forming mucins is the capacity of monomers to form polymeric structures. Secreted mucins are produced by specialized cells, generally referred to as "goblet cells" [119]. The secreted mucins contribute to the formation of a physical barrier that protects epithelial cells lining the respiratory, urinogenital and gastrointestinal tracts [118]. Gel-forming mucins have several VWD (Von Willebrand factor-D) and VWC (Von Willebrand factor-C) domains flanking the mucin domains. They also harbour cysteine-rich regions named 'CK' domains (Cystine Knot) at their C-terminal ends [118]. MUC2, a major gel-forming mucin of the colon forms dimers via its C-terminal and trimmers via its N-terminal that leads to a polymeric structure [122-124].

Transmembrane mucins

Transmembrane mucins are present along the apical surface of epithelial cells. The human transmembrane mucins are characterized by either a SEA (sea urchin sperm protein, enterokinase and agrin) domain or a special variant of the VWD (Von Willebrand factor D) domain that is lacking cysteines. From amino- to carboxyl ends, the overall structure of membrane-bound mucins exhibits three main regions: (I) An extracellular domain, which carries the mucin domain and extends far from the surface of the cell, (II) a type I transmembrane domain that spans the lipid bilayer layer and (III) a short cytoplasmic tail (Figure 5B) [118]. Several of the human transmembrane mucins are known or predicted to be cleaved in their SEA or in VWD domain to yield two peptides that remain attached by non-covalent forces [125-127]. The cytoplasmic tails of some of the transmembrane mucins have been implicated in different cell signalling events [128,129].

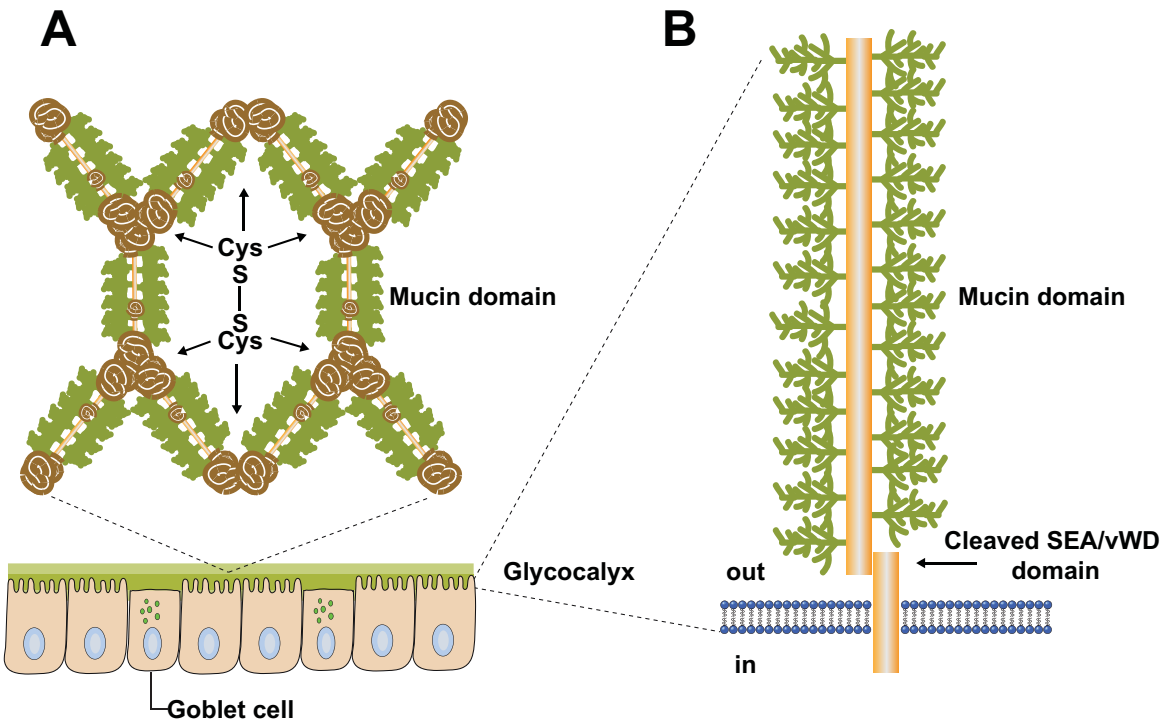

Figure 5: Typical example of a gel-forming mucin (A) and a transmembrane mucin (B). Gel-forming mucins are synthesized in specialized cells, known as goblets cells that are characterized by large mucin-packed secretory granules. Upon regulatory signals or stimulation, mucin granules are released and can expand up to 1000-fold on hydration. Gel-forming mucins form large polymers through oligomeration/multimerization, which are held together with numerous disulphide bonds. Transmembrane mucins are expressed at the apical cell surface and appear cleaved at SEA/vWD domains into amino- and carboxy-terminal subunits that are held together by non-covalent forces. The N-terminal subunit harbours highly glycosylated mucin-domains that are tethered to the C-terminal transmembrane subunit. The mucin-domains extend far from the cell membrane into the glycocalyx. (Protein backbones are shown in brown and oligosaccharides in green). *Figure adapted from [130]*

Mucin-type *O*-glycosylation in *Drosophila* development

In *Drosophila,* mucin-type *O*-glycosylation is initiated by PGANTs. There are at least twelve putative genes in the *Drosophila* genome encoding PGANTs, out of which nine have been demonstrated to have enzymatic activity *in vitro* [131,132]. Structurally, members of this family are type II transmembrane proteins, consisting of a short cytoplasmic tail at the N-terminus that is tethered to the Golgi membrane by means of a transmembrane domain and a highly conserved catalytic domain at the C-terminus that lies within the Golgi lumen [131,132]. Biochemical studies have shown that members of the PGANT family have a hierarchy of enzymatic activity [132]. Similar to mammalian

GalNAc transferases, PGANTs have been categorized into two groups based on their activity. One group consists of enzymes that catalyse the initial addition of GalNAc to unmodified peptides (peptide transferases), while the other group of enzymes act on previously glycosylated substrates that contain GalNAc residues (glycopeptide transferases) [132]. The initial addition of GalNAc to selected Ser/Thr residues on the protein backbone by PGANTs results in the so-called Tn-antigen, which can be further extended by addition of galactose by core 1 β1,3-galactosyltransferase (C1GalT1) to form a core 1 structure, called the T-antigen [133]. Unlike mammalian mucin-type *O*-glycans, which have several arrays of high-order *O*-glycan structures, *Drosophila O*-glycans tend to be shorter and less extended, and they mainly consist of Tn-antigens and T-antigens [134,135]. Expression analysis of individual PGANTs has revealed highly dynamic spatio-temporal and frequently overlapping patterns of expression during *Drosophila* embryogenesis [136]. This dynamic expression of PGANTs indicates a specific requirement of *O*-glycans in diverse tissues and at various stages of development. Indeed, labelling of *Drosophila* embryos with lectins and an antibody against the Tn-antigen has shown the presence of mucin-type *O*-glycans in most of the developing embryonic tissues [8-10]. In particular, *O*-glycans were predominantly found along the luminal and apical surfaces of epithelial tubes of the salivary glands, developing gut and the tracheal system [8].

The first evidence implicating mucin-type *O*-glycosylation in *Drosophila* development was demonstrated by the observation that one of the members of the PGANT family, *pgant35A,* is recessive lethal [131,137]. Subsequently, it was found that loss of *pgant35A* was associated with an altered tracheal tube morphology, accompanied by mislocalization of SJ proteins and a compromised paracellular barrier [138]. Loss of *pgant35A* also resulted in reduced levels of Crb and tracheal luminal components (the 2A12 antigen and *O*-glycans), suggesting a role for *pgant35A* in trafficking of apical and luminal components during tracheal development [138]. *pgant35A* is also expressed in the developing salivary glands and hindgut [136], but the irregular tube morphology and cell polarity defects seen in *pgant35A* mutants were restricted to the tracheal system. This could be due to functional redundancy among PGANT isoforms [138].

Specific roles of *O*-glycosylation in *Drosophila* have also been demonstrated for integrin mediated cell adhesion during wing development [139]. Loss of *pgant3* results in

reduction of *O*-glycans along the basal surface of the larval wing imaginal discs, causing irregular adhesion of the two epithelial cell layers that will ultimately form the adult wing blade [139]. This aberrant adhesion was evident in localized blisters in the adult wing soon after eclosion. A combination of bioinformatics and *in vitro* glycosylation assays showed that PGANT3 glycosylates the integrin-binding ECM protein Tiggrin that is normally secreted into the basement membrane [139]. This was further confirmed by immunoprecipitation and genetic interaction experiments. In *pgant3* mutants, reduced *O*-glycosylation of Tiggrin was observed, and it was proposed that *O*-glycans found on Tiggrin could affect some aspects of integrin-ECM adhesion and also could influence protein stability, secretion and binding interactions [139]. Additional roles of *pgant3* in secretion were further demonstrated in a *Drosophila* cell culture system, in which RNAi against *pgant3* resulted in altered Golgi structure and reduced secretion of a reporter construct [140].

In a recent study, tissue-specific knockdown of multiple members of the PGANT family using RNAi identified *pgant4, pgant5, pgant7* and *CG30463* to be essential in various developing organs [141]. Loss of *pgant5* was found to cause altered copper cell morphology (disorganized apical microvilli), reduced levels of *O*-glycans along the apical and luminal surfaces of the copper cells and defects in larval midgut acidification [141]. Copper cells are specialized cup-shaped cells found in the *Drosophila* midgut and are responsible for gut acidification [142]. Although no target substrate for PGANT5 was identified, this study suggested the possibility that PGANT5 could be responsible for glycosylation of components essential for localization of ion transporters or proteins involved in organizing apical polarity in copper cells [141]. Together, these studies point to important roles for mucin-type *O*-glycosylation in *Drosophila* development.

Results and Discussion

Paper I

Mucins are a large family of heavily *O*-glycosylated proteins and are the major components of the protective mucosal surfaces lining several vital organs of the body. Malfunction of this protective mucosal-barrier can lead to infections, acute or chronic inflammation and development of cancer [117,118]. Despite their importance in human pathologies, there is limited knowledge about the mechanisms regulating mucin expression and glycosylation. Moreover, it has been difficult to study mucins in relation to disease development, largely due to their physical and biochemical properties. One way to approach these questions is to address them in simpler invertebrate model systems, such as *Drosophila,* from which parallels to vertebrates can be drawn. The main aim of the study in Paper I was to identify mucin or mucin-like proteins in *Drosophila* and describe their expression pattern during development from embryo to adult.

Identification of mucins and mucin-related proteins in Drosophila

The PTS domains found in mucins tend to be poorly conserved, and there is no general consensus sequence defined to predict mucin-type *O*-glycosylation [143]. Identification of PTS domains using sequence similarity methods like BLAST is unreliable. However, statistical studies using experimentally verified *O*-glycosylation sites have led to development of sophisticated algorithms to predict mucin-type O-glycosylation [143-145].

To identify mucins in *Drosophila*, we devised a simple bioinformatic strategy that targets PTS repeats. To accomplish this, we developed a program called PTSPMiner that was used to find PTS domains in the predicted *Drosophila* proteome. PTSPMiner is developed in Java programming language and utilizes BioJava API. The first step in the program is implemented to calculate the total frequency of the amino acids Ser, Thr and Pro in a predicted protein, and the second step identifies the number of amino acid tandem repeats in the sequence. When applying cut-offs for Ser and Thr content > 25% and number of repeats > 4, forty-two proteins encoded by different genes were identified (Figure 6). PTSPMiner allows visualization of PTS repeats, by highlighting Pro, Thr and Ser in

different colours (Figure 7). By manual analysis, we found that nine of the proteins lacked PTS repeats and instead contained other types of repeats. These were excluded from further analysis. In addition, we included three mucin-like proteins that were identified through homology searches for mucin-associated domains. These were not picked by PTSPMiner because their Ser and Thr contents were below threshold (Figure 6).

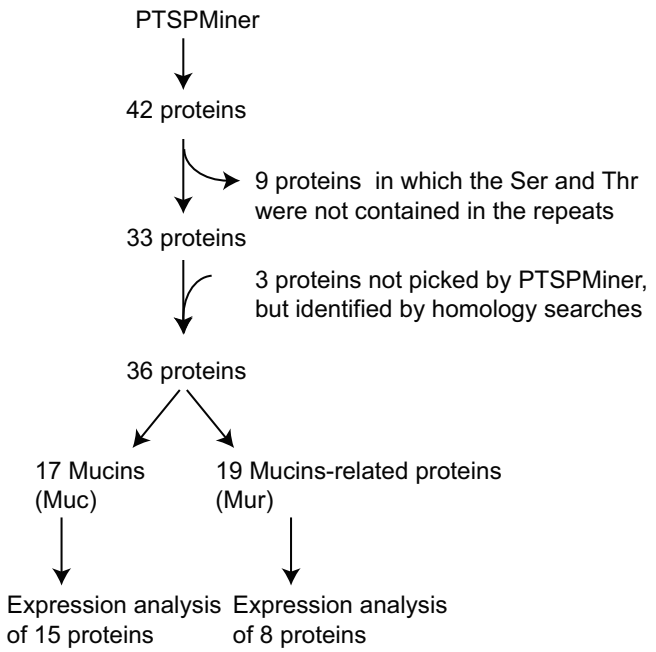

Figure 6: Schematic workflow of the devised strategy for the identification of mucins in *Drosophila*.

In order to name the thirty-six identified proteins, we adopted a simple nomenclature: Proteins in which the PTS domain(s) constitute more than 30% of the protein length were termed Mucins (Muc), and proteins with PTS domains constituting less than 30% of the protein, or where the Ser and Thr-rich regions contained no Pro, were termed mucin-related proteins (Mur) (Figure 6). This nomenclature was followed by the cytological position of the encoding gene.

In our further analysis, we focused on fifteen mucins and eight mucin-related proteins (Figure 7). Of these, two mucins and two mucin-related proteins had no predicted signal sequence or transmembrane domains, which might be due to inaccurate prediction of the encoding genes. None of the identified mucins contain a transmembrane domain, nor did they harbour a vWD, SEA or CK domain found in human mucins. However, other conserved protein domains involved in protein interactions, such as the vWC domain, chitin binding Per-A domains and EGF-like domains were found both in the identified mucins and mucin-related proteins. Moreover, three of the identified mucins contain cysteines within their PTS domains, and this is also observed in the PTS domains of gel-forming mucins in *Xenopus tropicalis* [121].

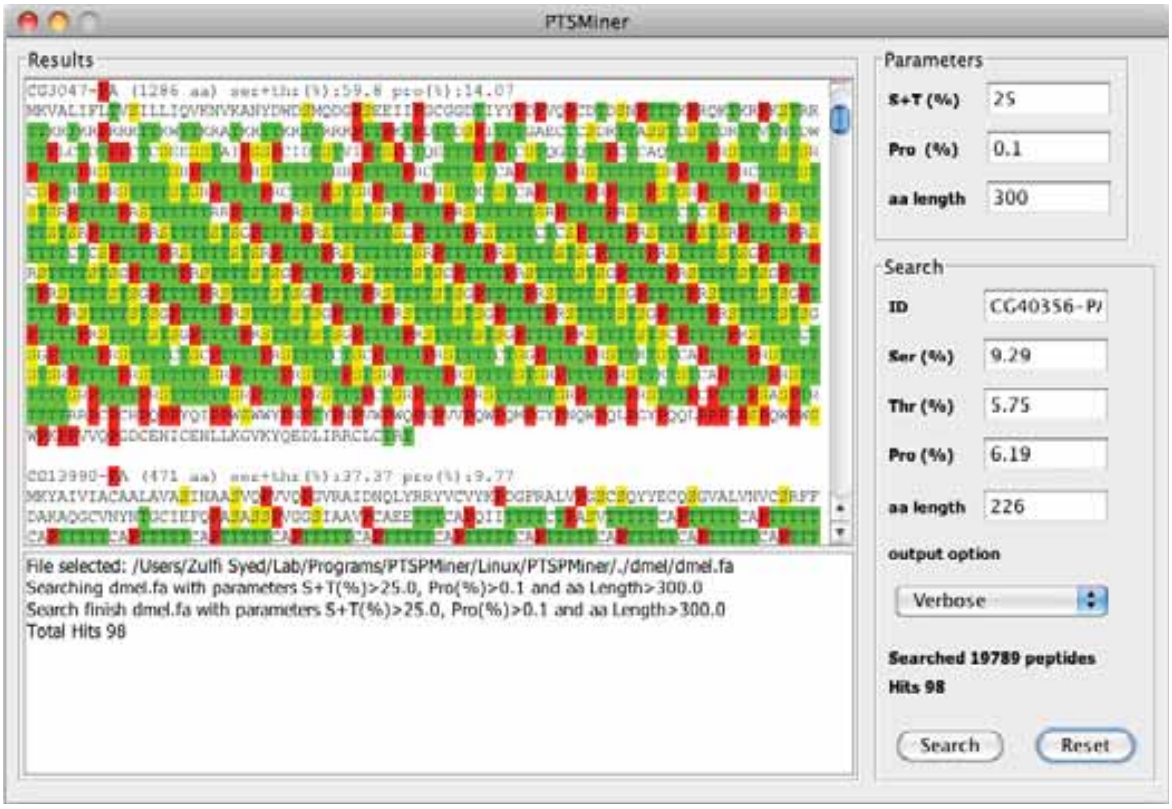

Figure 7: Screenshot of PTSMiner after scanning the entire protein database. The image shows the PTS repeats in CG3047-PA, where Pro, Thr and Ser are highlighted in different colours to visualize the repeats.

A previous bioinformatic study to identify mucins lead to the development of two approaches that were implemented in two programs called PTSPRED and MPRED [145]. PTSPRED identifies regions in a protein sequence that shows high content of Ser, Thr and Pro, while MPRED uses a statistical model called the hidden Markov model to make probabilistic predictions of whether an amino acid sequence conforms to a mucin domain. PTSPRED and MPRED were used to identify mucin domains in several different species, including *Fugu rubripes* [145], chicken [146] and *Drosophila* [121]. PTSPMiner was able to identify most of the mucin-domain containing *Drosophila* proteins identified by PTSPRED and MPRED [121], with the exception of a few that either lack repetitive nature or had low Ser and Thr-content. Like PTSPRED, PTSPMiner is based on amino acid compositional bias, but it differs in some architectural elements such as GUI, programming language and platform independence. The GUI allows for easy manual inspection for repeat regions with different colouring schemes and convenient file input

and output handling (Figure 7). PTSPMiner, being written in Java, makes it platform-independent, enabling it to run on multiple operating systems.

Drosophila ***mucins are expressed at different stages of life cycle***

To gain further insight into possible functions of the identified mucins and mucin-related proteins, we analysed their expression pattern at different stages of development from embryo to adult. We performed reverse transcription PCR on RNA extracts from different developmental stages and on dissected organs from third instar larva, and found that several of the mucins and mucin-related proteins were dynamically expressed during the fly life cycle and showed tissue-specific expression. Three of the mucins were exclusively expressed either at embryonic stage (Muc30E), larval stage (Muc68D) or adult stage (Mur11Da).

In *Drosophila*, many epithelia are protected by an apical chitinous cuticle, such as the epidermis, the tracheal system and parts of the digestive system (foregut and hindgut). If *Drosophila* mucins were to have similar physiological functions as vertebrate mucins, their primary site of expression should be within cuticle-free organs. We found that a majority of the mucins and mucin-related proteins were expressed in cuticle-free organs, including the salivary glands, digestive tract and Malpighian tubules of third instar larvae. The midgut is protected on the luminal side by a non-cellular apical matrix known as the Peritrophic matrix (PM) [147]. The PM plays the role of a physical barrier and consists of a scaffold of chitin fibres embedded with glycosylated and most often chitin-binding proteins (Peritrophins) [147], and is regarded to be functionally similar to vertebrate mucosa [148]. The mucins and mucin-related proteins detected in the digestive tract could potentially be components of the PM. Some of these have similar domains to Invertebrate Intestinal Mucin (IIM), a PM protein of *Trichoplusia ni* [149].

In addition to the digestive tract, the salivary glands showed prominent mucin expression. Two of the mucins expressed in salivary glands, Muc25B/Sgs1 and Muc68Cb/Sgs3, were previously reported to belong to the salivary gland secretion (Sgs) family of proteins. These are secreted towards the end of third instar larva to produce a sticky secretion by which the larvae attach themselves to a solid surface prior to pupa formation [150]. The other mucins that are expressed in the salivary glands might be glue proteins or have

protective roles, and further characterization has to be carried out to address their function.

Drosophila ***mucins are expressed in developing epithelial organs***

An interesting finding was that many of the mucins and mucin-related proteins were expressed during embryogenesis. In order to identify their expression pattern, we performed whole mount RNA *in situ* hybridization. This revealed that mucins and mucin-related proteins were expressed in the developing salivary glands, midgut, foregut, hindgut and trachea, as well as in the proventriculus and epidermis during late stages of embryogenesis. One of the mucins, Muc30E*,* was exclusively expressed in the extraembryonic amniosera and its expression was ceased after dorsal closure. This selective expression indicates a specific role for this mucin during development, and further characterization is required to understand its function. The observed expression patterns of mucins and mucin-related proteins parallel the previously reported expression of PGANTs [136] and detection of *O*-glycans in the lumen and along the apical surface of developing organs [8]. It is therefore possible that the identified mucins and mucin-related proteins represent novel components of the apical *O*-glycan-rich matrices seen in developing epithelial organs. During tracheal development, a transient chitin-based luminal matrix is required for shaping the tube [151,152], and earlier studies have shown that one of the mucin-related proteins, Mur96/Tnc, is present in the lumen and along the apical surface of the developing trachea, foregut and hindgut [153]. It is an exciting possibility that glycosylated proteins like mucins could form luminal matrices needed to shape epithelial organs.

Paper II

It was intriguing that many of the mucins and mucin-related proteins were expressed during embryogenesis. To study a possible involvement of mucin in embryonic epithelial organ development, we choose to analyse the function of Mur96/Tnc in epithelial tubes like the foregut, hindgut and trachea.

Tenectin is an intraluminal protein required for diameter expansion of the hindgut

Tnc is a large mucin-related protein consisting of 2788 amino acids. The deduced protein sequence reveals that Tnc is secreted and harbours two PTS domains flanked by five vWC domains. The domain organization of Tnc resembles that of secreted gel-forming mucins, in which dense *O*-glycosylated PTS domains are separated by cysteine-rich von Willebrand factor domains that mediate polymerization [119]. Confocal imaging of wild type embryos labelled with antisera against Tnc showed Tnc to be secreted at the apical surface in the developing hindgut, foregut and tracheal dorsal trunks, consistent with the previously reported expression of Tnc in these tissues [153]. In addition, Tnc was present in the proventriculus, salivary gland ducts and in the lumen of dorsal vessel. In all tissues analysed, Tnc localized to the apical surface and spanned the entire lumen, thus behaving like a secreted intraluminal component.

To address a possible function of Tnc in epithelial organ development, *tnc* mutants were generated and the loss of function allele called tnc^{13c} was analysed. Labelling of *tnc* mutant embryos with the apical marker Crb to visualize the organ lumens revealed that *tnc* mutants have an unusually narrow hindgut lumen. At the end of embryonic stage 13, the hindgut layout is established with the partitioning of the hindgut into distinct compartments (small intestine (Si), large intestine (Li) and rectum) [154]. From stage 14 onwards, the hindgut tube (Si and Li) grows by tube elongation and diameter expansion. Morphometric analysis of the hindgut (Si and Li) in wild type and *tnc* mutants at stage 14 and stage 16 indicated that the volume of the hindgut lumen in *tnc* mutants was at least 50% less than in the wild type, while tube length was slightly longer than normal. The narrow hindgut lumen in *tnc* mutants was accompanied by detectable cellular changes in the epithelium. Staining of wild type and *tnc* mutant embryos with DE-cad to highlight the apical cell circumference showed that *tnc* mutants have reduced apical cell domains,

and this was particularly evident in Si. Examination of the apical cell circumference in Li of *tnc* mutants further revealed aberrations in cell arrangement, coherent with the narrow diameter. A reduction in outer tube diameter was also observed in *tnc* mutants. These observations show that Tnc is required for hindgut tube dilations, both at the apical and basal surface, by mechanisms associated with changes in cell shape and slight cell rearrangements. We could not find that epithelial polarity or patterning of the hindgut tube was affected in the *tnc* mutants.

Tnc drives hindgut expansion in a dose-dependent manner

Expression analysis of Tnc in the hindgut showed that Tnc gradually accumulates in the lumen from stage 13 onwards, and by stage 16, Tnc was abundant in the lumen with stronger expression in Si than in Li. The period of accumulation of Tnc in the lumen, together with the severe reduction of Si diameter compared to Li and the stronger expression of *tnc* in Si versus Li suggested that the degree of diameter expansion might be connected to the levels of Tnc in the lumen. To test this possibility, *tnc* was overexpressed in the hindgut at varying levels using different GAL4-driver lines. Over-expression of *tnc* caused excessive dilation of the hindgut lumen, and over-expression of *tnc* at higher levels in Si than in Li resulted in an elevated increase in Si diameter compared to Li. The excessive lumen dilation associated with Tnc over-expression was accompanied by an increase in outer tube diameter and enlarged apical circumferences of the hindgut epithelial cell. These observations contrasted the defects seen in the hindgut of *tnc* mutants and indicate that Tnc drives hindgut lumen dilation depending on the pattern and levels of expression. Interestingly, ectopic *tnc* expression in other epithelial organs also lead to excessive tube dilation and increase in apical cell circumference, similar to the effects of Tnc over expression in the hindgut.

When Tnc was over-expressed in only the dorsal compartment of Li, this resulted in enlarged apical cell circumferences in both the dorsal and ventral compartments, suggesting that Tnc acts non-cell autonomously. Moreover, when *tnc* was expressed in the epidermis, to test its effect on epithelial organization in a non-luminal context, no anomalies or changes in cell shape were seen in epidermis. Together these results argue that Tnc promotes tube dilation by exerting an internal luminal pressure.

Tnc is a component of O-glycosylated matrix in the hindgut lumen

Tnc carries two large PTS domains, which could potentially serve as substrates for *O*-glycosylation. When protein extracts from embryos and larvae were probed on western blots, Tnc was detected as high molecular weight species substantially larger than the predicated molecular mass of 290 kDa. This migration pattern of Tnc would be consistent with Tnc carrying post-translational modifications. To test if Tnc might be glycosylated, embryonic extracts were treated with deglycosylation enzymes. Treatment with O-glycanase or with O-glycanase together with three other enzymes, Sialidase, β(1-4) Galactosidase and β-N-Acetylglucosaminidase, caused slightly faster migration of Tnc, while N-glycanase treatment did not affect the migration, indicating that Tnc carries mucin-type *O*-glycans. If Tnc carries mucin-type *O*-glycans, it should contribute to detectable *O*-glycans in the hindgut lumen. To address this, both wild type and *tnc* mutant embryos were labelled with an antibody that detects the Tn-antigen. Wild type embryos showed strong staining along the luminal surface and in intraluminal compartment of the hindgut, whereas *tnc* mutants displayed reduced intraluminal Tn-staining. Staining with *Vicia villosa* lectin (VVA) that also recognizes the Tn-antigen showed a similar reduced intraluminal VVA staining in the *tnc* mutant hindgut. Both Tnc and Tn-staining showed punctate, and sometimes discontinuous, staining in the hindgut lumen with formaldehyde fixation. It has earlier been demonstrated that ethanol-based fixation methods best preserve the texture of glycan-rich matrices [155], and when embryos were fixed with the ethanol-based Clark's fixative, Tnc appeared as a striated structure that filled the entire hindgut lumen. Tnc therefore appears to be an important carrier of mucin-type *O*-glycans, forming a glycan-rich matrix in the hindgut lumen.

Model for Tnc-mediated tube dilation

Our results indicated that Tnc has biochemical characteristics similar to secreted gel-forming mucins and forms a striated structure that fills the hindgut lumen during tube expansion. Overexpression of Tnc showed that Tnc is able to drive tube dilation in dose-dependent manner, and that this lumen dilation is associated with an increase in apical cell circumferences and an increase in both inner and outer tube diameter. Based on these observations, a model for Tnc can be presented in which Tnc drives tube dilation by generating a luminal pressure upon its secretion into the lumen, and thus the extent of

lumen volume expansion would directly relate to the amount of Tnc in the lumen. This represents a new mechanism of shaping epithelial tubes.

The proposed function of Tnc would be analogous to that of a luminal hydrostatic pressure, which can cause an increase in lumen volume [156,157]. This is believed to occur in the initial inflation of the brain ventricle lumen [158,159] and Kupffer's vesicle in zebrafish [160]. However, Tnc can also cause local tube dilation, possibly because Tnc is relatively immobile in the lumen and forms an expanding matrix at the site of secretion. This differs from the effect of hydrostatic pressure, where the force acts uniformly on the tube wall.

The function of Tnc differs from that of luminal chitin matrix in trachea. The chitinous matrix does not drive lumen dilation, but appears to provide a scaffold for uniform diameter expansion [151,152], while Tnc is needed to increase the luminal volume. Tnc function is also distinct from that Eye Shut in the fly retina [161] and CD34 sialomucin Podocalyxin [162] that are required for initiating lumen formation, as Tnc is not required for opening a lumen cavity, but for subsequent diameter expansion. It is intriguing how a luminal pressure generated by Tnc is perceived by the epithelium to cause epithelial remodelling. Recent studies examining tube elongation in salivary glands have shown that apical membrane dynamics and membrane-cytoskeleton interactions are critical for facilitating cell shape changes [163,164]. Another recent study proposed that epithelial relaxation, mediated by the activity of myosin phosphate facilitates inflation of the brain ventricular lumen during zebrafish development [165]. Further investigation of the *Drosophila* hindgut might provide insights into the molecular basis of tube wall plasticity in response to increase in luminal pressure.

Conclusions

In this thesis work potential *Drosophila* mucins and mucin-related proteins were identified in a genome wide screen. *Drosophila* mucins lack the conserved domains associated with vertebrate mucins such SEA, vWD and CK domains, but have potential protein- and chitin-interaction domains that could mediate multimerization and apical matrix formation.

Expression analysis of *Drosophila* mucin-like proteins revealed two interesting findings. First, many mucins were expressed in the digestive tract, suggesting that they contribute to a protective apical matrix analogous to vertebrate mucosa. Second, several of the mucins and mucin-related proteins showed dynamic and tissue-specific expression during embryogenesis, implicating a role for mucins in organ development. Characterisation of one of the mucin-related proteins, Tnc, shows that it fills the lumen of developing hindgut and drives tube diameter expansion in a dose-dependent manner. We believe that this represents a new mechanism of shaping epithelial tubes, where a glycan-rich luminal matrix generates a mechanical force on the surrounding epithelium.

Although conserved domains associated with vertebrate mucins are absent in *Drosophila* mucins, many are found in other *Drosophila* proteins. One of these is called Mesh and contains several protein domains found in human MUC4. As a first step to understand the possible roles of these domains, we have shown that Mesh is important for epithelial integrity.

Further studies of the identified mucins and mucin-related proteins should enrich our understanding of mucin biology, both in physiological settings and during development, and possibly reveal new mechanisms of mucin function.

Acknowledgments

I would like to take this opportunity to thank everybody who has contributed directly or indirectly to this thesis work, and without their kind help and support this would not have been possible.

Iris Härd for introducing me to the world of mucins and giving me this opportunity. You have given me enormous freedom to explore and develop my scientific skills. Discussing scientific and non-scientific matters with you has always been fun and interesting. Thank you so much for your unlimited support throughout my studies and your motivation, mentorship and friendship is invaluable to me.

Anne Uv for her unlimited support and taking me into her lab. Throughout my time in lab, your energy and kind nature have created relaxed and very inspiring scientific environment. I admire your commitment, and your endless discussion about science has always been interesting and fun. Above all you have been such great friend, mentor and sport. I thoroughly enjoyed my time in lab with you and all this made my PhD a nice experience.

I thank you both for your invaluable scientific guidance and constructive comments on my scientific writing.

A special thanks to Torleif Härd for the PhD position and support.

Thanks to Gunnar Hansson for being my co-supervisor and his inputs. Thanks to all the group members for letting me borrow reagents and equipment.

I would like to thank all the past and present members of the A.Uv and I. Härd lab.

Thank you Simin for sharing your extensive scientific experience, kind nature and friendship. It was lot of fun sharing lab bench with you.

Kate, thank so much for proofreading my thesis and all the laughs. I enjoy your sense of humor and conversations about gretchen and entities that grind our gears.

Hamid (hello jaaamess!!!), It was fun working with you and being part of akbar brothers tea club. Thank you so much for your friendship and all the help and good food. Good luck with your future ventures!!

Anna Tonning, for always smiling and helping me out during my early lab days, even though you were busy preparing for your thesis. Erika, thank you for the good time in lab and sharing jokes and laughs about the akbar brothers. Tina, thanks for the all talks and discussions during our early days in lab.

Fredrik and Erik (the cool dude), thanks for the good time and fun being part of the akbar brothers. I enjoyed our lunch conversations and thanks for the fishing advice!! Good luck in your future endeavors. Moin, Banaz, Jonathan and Mengistu thanks for your time and help and good luck with your studies. Jonathan, thank you for introducing me to one of my favorite pizzeria.

Thanks to Levent and his group members Sashi and Srujuna for letting me use their equipment. Sashi and Srujuna, thank you both for your friendship. Sashi thanks for the good time and for being my fishing buddy. I had lot fun with you during our fishing trips and you are good sport, keep it up. Good luck with your studies!!!

Cinna, thank you so much for the laughs, lunches to Hani and for sharing your interest in electronics. Good luck with your studies!!!

Arun, you are truly a nice guy!! Thank you for the good time and interesting conversations and sharing your views. Keep up the good nature and I wish you all the best for your studies.

Ali (Gotiya sahib), thank you so much for all the jokes and help. It was really great hanging out with you and thank you for being such a good friend. Good luck with your new job!!!

Bernard, it was great having you in the lab and thank you so much for all the interesting scientific discussions and collaboration. I have never seen someone with such a cool moves on the dance floor.

Michaela, Gosa and Mehrnaz from Homa's group and Katarina, thank you all for the good time in lab and interesting lab discussions.

Thanks to Chandra and all his group members (Tanmoy, Gaurav, Kankadeb, Sanhita, Arif, Shantilal, Mathieu) for all the laughs and discussions. Erik, Babak and Zhiyuan thanks for interesting lunch conversations and discussing cool stuff!!!

Thanks to Julia and Maria at the Cellular imaging core facility for such a fantastic course and answering all my annoying questions and help.

Bengt, Yvonne and Canita thank you so much for such an interesting EM course.

Thanks to the administration at Biomedicine, Carina Ejdeholm, Ulrica Lilja, Andrea Frateschi, Carina Petersson for the all their help.

I would thank all my friends in India, Vipul, Aman, Parvez, Tabrez, Hamed, Bushra, Ruby, Razi, Masood, Feroze, Sainath, Kiran, Murli, Arif Bhai, Rafi, Salman and Asjad. You all have been great and we had a good time during our schooling, college and university days. Good luck with your jobs and research.

To my parents for their endless love and support, Abba and Ammi you're the best and thank you so much for your prayers and motivation. I miss you Dada, Dadi and Nana, Nani.

To my brother, Jeelani, you have been such a great inspiration and role model for me. Thank you so much for introducing me to music, books and all the cool stuff a younger brother could wish for. Thanks a lot bhaiyya!!, I couldn't ask for anything more!!

My brother, Rabbani, you are such great sport; we had so much fun during our schooling. Thank you for supporting me in whatever I plan to do. Salma appa and Farha, my adorable sisters!!! Thank you for your love, caring and support and teaching me how to read and write urdu. Your Biryanis are delicious and thank you so much for teaching me how to cook!!!!

Thanks to my brother-in-law Mohiuddin and Malik and Sister-in-law Ruth and Sara for your caring, affection and motivation. A Special thanks to Sunita's parents, you have been amazing and thank you so much for your love and support.

My two lovely niece, Gousia, Nazia and newphews Rafic, Toufeeq, Levi, Faiz and Ryan, I love you all and promise to have loads of fun when I visit you.

My beloved Sunita, for her endless love, support and encouragement. I admire your great patience and sharing all the moments of happiness and difficulties. Thank you for sharing with me your passion for science. I love you!!!

References

1. Lubarsky B, Krasnow MA (2003) Tube morphogenesis: making and shaping biological tubes. Cell 112: 19-28.

2. Bryant DM, Mostov KE (2008) From cells to organs: building polarized tissue. Nat Rev Mol Cell Biol 9: 887-901.

3. Buisine MP, Devisme L, Copin MC, Durand-Reville M, Gosselin B, et al. (1999) Developmental mucin gene expression in the human respiratory tract. Am J Respir Cell Mol Biol 20: 209-218.

4. Leroy X, Devisme L, Buisine MP, Copin MC, Aubert S, et al. (2003) Expression of human mucin genes during normal and abnormal renal development. Am J Clin Pathol 120: 544-550.

5. Buisine MP, Devisme L, Savidge TC, Gespach C, Gosselin B, et al. (1998) Mucin gene expression in human embryonic and fetal intestine. Gut 43: 519-524.

6. Reid CJ, Harris A (1999) Expression of the MUC 6 mucin gene in development of the human kidney and male genital ducts. J Histochem Cytochem 47: 817-822.

7. Reid CJ, Harris A (1998) Developmental expression of mucin genes in the human gastrointestinal system. Gut 42: 220-226.

8. Tian E, Ten Hagen KG (2007) O-linked glycan expression during Drosophila development. Glycobiology 17: 820-827.

9. Fredieu JR, Mahowald AP (1994) Glycoconjugate expression during Drosophila embryogenesis. Acta Anat (Basel) 149: 89-99.

10. D'Amico P, Jacobs JR (1995) Lectin histochemistry of the Drosophila embryo. Tissue Cell 27: 23-30.

11. Gheri G, Sgambati E, Bryk SG (2000) Glycoconjugate sugar residues in the chick embryo developing lung: a lectin histochemical study. J Morphol 243: 257-264.

12. Schumacher K, Strehl R, Minuth WW (2002) Detection of glycosylated sites in embryonic rabbit kidney by lectin chemistry. Histochem Cell Biol 118: 79-87.

13. Arias AM (2008) Drosophila melanogaster and the development of biology in the 20th century. Methods Mol Biol 420: 1-25.

14. Fortini ME, Skupski MP, Boguski MS, Hariharan IK (2000) A survey of human disease gene counterparts in the Drosophila genome. J Cell Biol 150: F23-30.

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

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

17. St Johnston D (2002) The art and design of genetic screens: Drosophila melanogaster. Nat Rev Genet 3: 176-188.

18. del Valle Rodriguez A, Didiano D, Desplan C (2012) Power tools for gene expression and clonal analysis in Drosophila. Nat Methods 9: 47-55.

19. Kazazian HH, Jr. (2004) Mobile elements: drivers of genome evolution. Science 303: 1626-1632.

20. Ravindran S (2012) Barbara McClintock and the discovery of jumping genes. Proc Natl Acad Sci U S A 109: 20198-20199.

21. Clark JB, Kidwell MG (1997) A phylogenetic perspective on P transposable element evolution in Drosophila. Proc Natl Acad Sci U S A 94: 11428-11433.

22. Rubin GM, Kidwell MG, Bingham PM (1982) The molecular basis of P-M hybrid dysgenesis: the nature of induced mutations. Cell 29: 987-994.

23. Beall EL, Rio DC (1997) Drosophila P-element transposase is a novel site-specific endonuclease. Genes Dev 11: 2137-2151.

24. Kaufman PD, Doll RF, Rio DC (1989) Drosophila P element transposase recognizes internal P element DNA sequences. Cell 59: 359-371.

25. Kaufman PD, Rio DC (1992) P element transposition in vitro proceeds by a cut-and-paste mechanism and uses GTP as a cofactor. Cell 69: 27-39.

26. O'Hare K, Rubin GM (1983) Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. Cell 34: 25-35.

27. Spradling AC, Rubin GM (1982) Transposition of cloned P elements into Drosophila germ line chromosomes. Science 218: 341-347.

28. Wilson C, Pearson RK, Bellen HJ, O'Kane CJ, Grossniklaus U, et al. (1989) P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila. Genes Dev 3: 1301-1313.

29. Morin X, Daneman R, Zavortink M, Chia W (2001) A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in Drosophila. Proc Natl Acad Sci U S A 98: 15050-15055.

30. Xu T, Rubin GM (1993) Analysis of genetic mosaics in developing and adult Drosophila tissues. Development 117: 1223-1237.

31. Buszczak M, Paterno S, Lighthouse D, Bachman J, Planck J, et al. (2007) The carnegie protein trap library: a versatile tool for Drosophila developmental studies. Genetics 175: 1505-1531.

32. Ryder E, Russell S (2003) Transposable elements as tools for genomics and genetics in Drosophila. Brief Funct Genomic Proteomic 2: 57-71.

33. Venken KJ, Bellen HJ (2007) Transgenesis upgrades for Drosophila melanogaster. Development 134: 3571-3584.

34. Franz G, Savakis C (1991) Minos, a new transposable element from Drosophila hydei, is a member of the Tc1-like family of transposons. Nucleic Acids Res 19: 6646.

35. Metaxakis A, Oehler S, Klinakis A, Savakis C (2005) Minos as a genetic and genomic tool in Drosophila melanogaster. Genetics 171: 571-581.

36. FlyBase C (2002) The FlyBase database of the Drosophila genome projects and community literature. Nucleic Acids Res 30: 106-108.

37. Fischer JA, Giniger E, Maniatis T, Ptashne M (1988) GAL4 activates transcription in Drosophila. Nature 332: 853-856.

38. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, et al. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391: 806-811.

39. Shabalina SA, Koonin EV (2008) Origins and evolution of eukaryotic RNA interference. Trends Ecol Evol 23: 578-587.

40. Hannon GJ (2002) RNA interference. Nature 418: 244-251.

41. Perrimon N, Ni JQ, Perkins L (2010) In vivo RNAi: today and tomorrow. Cold Spring Harb Perspect Biol 2: a003640.

42. Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, et al. (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. Nature 448: 151-156.

43. Steinbrink S, Boutros M (2008) RNAi screening in cultured Drosophila cells. Methods Mol Biol 420: 139-153.

44. Tepass U, Tanentzapf G, Ward R, Fehon R (2001) Epithelial cell polarity and cell junctions in Drosophila. Annu Rev Genet 35: 747-784.

45. Tanos B, Rodriguez-Boulan E (2008) The epithelial polarity program: machineries involved and their hijacking by cancer. Oncogene 27: 6939-6957.

46. Marchiando AM, Graham WV, Turner JR (2010) Epithelial barriers in homeostasis and disease. Annu Rev Pathol 5: 119-144.

47. Izumi Y, Hirose T, Tamai Y, Hirai S, Nagashima Y, et al. (1998) An atypical PKC directly associates and colocalizes at the epithelial tight junction with ASIP, a mammalian homologue of Caenorhabditis elegans polarity protein PAR-3. J Cell Biol 143: 95-106.

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

49. Tsukita S, Furuse M, Itoh M (2001) Multifunctional strands in tight junctions. Nat Rev Mol Cell Biol 2: 285-293.

50. Tepass U (2012) The apical polarity protein network in Drosophila epithelial cells: regulation of polarity, junctions, morphogenesis, cell growth, and survival. Annu Rev Cell Dev Biol 28: 655-685.

51. Bulgakova NA, Knust E (2009) The Crumbs complex: from epithelial-cell polarity to retinal degeneration. J Cell Sci 122: 2587-2596.

52. Letizia A, Ricardo S, Moussian B, Martin N, Llimargas M (2013) A functional role of the extracellular domain of Crumbs in cell architecture and apicobasal polarity. J Cell Sci 126: 2157-2163.

53. Wodarz A, Grawe F, Knust E (1993) CRUMBS is involved in the control of apical protein targeting during Drosophila epithelial development. Mech Dev 44: 175-187.

54. Wodarz A, Hinz U, Engelbert M, Knust E (1995) Expression of crumbs confers apical character on plasma membrane domains of ectodermal epithelia of Drosophila. Cell 82: 67-76.

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

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

57. Bulgakova NA, Kempkens O, Knust E (2008) Multiple domains of Stardust differentially mediate localisation of the Crumbs-Stardust complex during photoreceptor development in Drosophila. J Cell Sci 121: 2018-2026.

58. Zhou W, Hong Y (2012) Drosophila Patj plays a supporting role in apical-basal polarity but is essential for viability. Development 139: 2891-2896.

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

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

61. Harris TJ, Tepass U (2010) Adherens junctions: from molecules to morphogenesis. Nat Rev Mol Cell Biol 11: 502-514.

62. St Johnston D, Ahringer J (2010) Cell polarity in eggs and epithelia: parallels and diversity. Cell 141: 757-774.

63. Morais-de-Sa E, Mirouse V, St Johnston D (2010) aPKC phosphorylation of Bazooka defines the apical/lateral border in Drosophila epithelial cells. Cell 141: 509-523.

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

65. Baum B, Georgiou M (2011) Dynamics of adherens junctions in epithelial establishment, maintenance, and remodeling. J Cell Biol 192: 907-917.

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

67. Flower NE, Filshie BK (1975) Junctional structures in the midgut cells of lepidopteran caterpillars. J Cell Sci 17: 221-239.

68. Lamb RS, Ward RE, Schweizer L, Fehon RG (1998) Drosophila coracle, a member of the protein 4.1 superfamily, has essential structural functions in the septate junctions and developmental functions in embryonic and adult epithelial cells. Mol Biol Cell 9: 3505-3519.

69. Baumgartner S, Littleton JT, Broadie K, Bhat MA, Harbecke R, et al. (1996) A Drosophila neurexin is required for septate junction and blood-nerve barrier formation and function. Cell 87: 1059-1068.

70. Genova JL, Fehon RG (2003) Neuroglian, Gliotactin, and the Na+/K+ ATPase are essential for septate junction function in Drosophila. J Cell Biol 161: 979-989.

71. Llimargas M, Strigini M, Katidou M, Karagogeos D, Casanova J (2004) Lachesin is a component of a septate junction-based mechanism that controls tube size and epithelial integrity in the Drosophila tracheal system. Development 131: 181-190.

72. Faivre-Sarrailh C, Banerjee S, Li J, Hortsch M, Laval M, et al. (2004) Drosophila contactin, a homolog of vertebrate contactin, is required for septate junction organization and paracellular barrier function. Development 131: 4931-4942.

73. Fehon RG, Dawson IA, Artavanis-Tsakonas S (1994) A Drosophila homologue of membrane-skeleton protein 4.1 is associated with septate junctions and is encoded by the coracle gene. Development 120: 545-557.

74. Hoover KB, Bryant PJ (2002) Drosophila Yurt is a new protein-4.1-like protein required for epithelial morphogenesis. Dev Genes Evol 212: 230-238.

75. Wu VM, Yu MH, Paik R, Banerjee S, Liang Z, et al. (2007) Drosophila Varicose, a member of a new subgroup of basolateral MAGUKs, is required for septate junctions and tracheal morphogenesis. Development 134: 999-1009.

76. Strand D, Jakobs R, Merdes G, Neumann B, Kalmes A, et al. (1994) The Drosophila lethal(2)giant larvae tumor suppressor protein forms homo-oligomers and is associated with nonmuscle myosin II heavy chain. J Cell Biol 127: 1361-1373.

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

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

79. Snow PM, Bieber AJ, Goodman CS (1989) Fasciclin III: a novel homophilic adhesion molecule in Drosophila. Cell 59: 313-323.

80. Behr M, Riedel D, Schuh R (2003) The claudin-like megatrachea is essential in septate junctions for the epithelial barrier function in Drosophila. Dev Cell 5: 611-620.

81. Wu VM, Schulte J, Hirschi A, Tepass U, Beitel GJ (2004) Sinuous is a Drosophila claudin required for septate junction organization and epithelial tube size control. J Cell Biol 164: 313-323.

82. Nelson KS, Furuse M, Beitel GJ (2010) The Drosophila Claudin Kune-kune is required for septate junction organization and tracheal tube size control. Genetics 185: 831-839.

83. Wang S, Jayaram SA, Hemphala J, Senti KA, Tsarouhas V, et al. (2006) Septate-junction-dependent luminal deposition of chitin deacetylases restricts tube elongation in the Drosophila trachea. Curr Biol 16: 180-185.

84. Oshima K, Fehon RG (2011) Analysis of protein dynamics within the septate junction reveals a highly stable core protein complex that does not include the basolateral polarity protein Discs large. J Cell Sci 124: 2861-2871.

85. Fanning AS, Anderson JM (1999) PDZ domains: fundamental building blocks in the organization of protein complexes at the plasma membrane. J Clin Invest 103: 767-772.

86. Betschinger J, Eisenhaber F, Knoblich JA (2005) Phosphorylation-induced autoinhibition regulates the cytoskeletal protein Lethal (2) giant larvae. Curr Biol 15: 276-282.

87. Baumann O (2001) Posterior midgut epithelial cells differ in their organization of the membrane skeleton from other drosophila epithelia. Exp Cell Res 270: 176-187.

88. Dubreuil RR, Wang P, Dahl S, Lee J, Goldstein LS (2000) Drosophila beta spectrin functions independently of alpha spectrin to polarize the Na,K ATPase in epithelial cells. J Cell Biol 149: 647-656.

89. Izumi Y, Yanagihashi Y, Furuse M (2012) A novel protein complex, Mesh-Ssk, is required for septate junction formation in the Drosophila midgut. J Cell Sci 125: 4923-4933.

90. Yanagihashi Y, Usui T, Izumi Y, Yonemura S, Sumida M, et al. (2012) Snakeskin, a membrane protein associated with smooth septate junctions, is required for intestinal barrier function in Drosophila. J Cell Sci 125: 1980-1990.

91. Tepass U, Knust E (1993) Crumbs and stardust act in a genetic pathway that controls the organization of epithelia in Drosophila melanogaster. Dev Biol 159: 311-326.

92. Knust E, Tepass U, Wodarz A (1993) crumbs and stardust, two genes of Drosophila required for the development of epithelial cell polarity. Dev Suppl: 261-268.

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

94. Laprise P, Lau KM, Harris KP, Silva-Gagliardi NF, Paul SM, et al. (2009) Yurt, Coracle, Neurexin IV and the Na(+),K(+)-ATPase form a novel group of epithelial polarity proteins. Nature 459: 1141-1145.

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

96. St Johnston D, Sanson B (2011) Epithelial polarity and morphogenesis. Curr Opin Cell Biol 23: 540-546.

97. Tepass U, Hartenstein V (1994) Epithelium formation in the Drosophila midgut depends on the interaction of endoderm and mesoderm. Development 120: 579-590.

98. Campos-Ortega JA, Hartenstein V (1997) The embryonic development of Drosophila melanogaster. Berlin ; New York: Springer. xvii, 405 p. p.

99. Myat MM (2005) Making tubes in the Drosophila embryo. Dev Dyn 232: 617-632.

100. Skaer H (1993) The alimentary canal; M Bate AMA, editor. New York: Cold Spring Harbor Laboratory Press.

101. Reuter R, Grunewald B, Leptin M (1993) A role for the mesoderm in endodermal migration and morphogenesis in Drosophila. Development 119: 1135-1145.

102. Nakagoshi H (2005) Functional specification in the Drosophila endoderm. Development, Growth & Differentiation 47: 383-392.

103. Iwaki DD, Johansen KA, Singer JB, Lengyel JA (2001) drumstick, bowl, and lines are required for patterning and cell rearrangement in the Drosophila embryonic hindgut. Dev Biol 240: 611-626.

104. Lengyel JA, Iwaki DD (2002) It takes guts: the Drosophila hindgut as a model system for organogenesis. Dev Biol 243: 1-19.

105. Murakami R, Takashima S, Hamaguchi T (1999) Developmental genetics of the Drosophila gut: specification of primordia, subdivision and overt-differentiation. Cell Mol Biol (Noisy-le-grand) 45: 661-676.

106. Beyenbach KW, Skaer H, Dow JA (2010) The developmental, molecular, and transport biology of Malpighian tubules. Annu Rev Entomol 55: 351-374.

107. Hatton-Ellis E, Ainsworth C, Sushama Y, Wan S, VijayRaghavan K, et al. (2007) Genetic regulation of patterned tubular branching in Drosophila. Proc Natl Acad Sci U S A 104: 169-174.

108. Skaer H (1989) Cell division in Malpighian tubule development in D. melanogaster is regulated by a single tip cell. Nature 342: 566-569.

109. Denholm B, Sudarsan V, Pasalodos-Sanchez S, Artero R, Lawrence P, et al. (2003) Dual origin of the renal tubules in Drosophila: mesodermal cells integrate and polarize to establish secretory function. Curr Biol 13: 1052-1057.

110. Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. Glycobiology 3: 97-130.

111. Varki A (2009) Essentials of glycobiology. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. xxix, 784 p. p.

112. Helenius A, Aebi M (2004) Roles of N-linked glycans in the endoplasmic reticulum. Annu Rev Biochem 73: 1019-1049.

113. Van den Steen P, Rudd PM, Dwek RA, Opdenakker G (1998) Concepts and principles of O-linked glycosylation. Crit Rev Biochem Mol Biol 33: 151-208.

114. Hang HC, Bertozzi CR (2005) The chemistry and biology of mucin-type O-linked glycosylation. Bioorg Med Chem 13: 5021-5034.

115. Ten Hagen KG, Fritz TA, Tabak LA (2003) All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. Glycobiology 13: 1R-16R.

116. Raman J, Fritz TA, Gerken TA, Jamison O, Live D, et al. (2008) The catalytic and lectin domains of UDP-GalNAc:polypeptide alpha-N-Acetylgalactosaminyltransferase function in concert to direct glycosylation site selection. J Biol Chem 283: 22942-22951.

117. Hollingsworth MA, Swanson BJ (2004) Mucins in cancer: protection and control of the cell surface. Nat Rev Cancer 4: 45-60.

118. Andrianifahanana M, Moniaux N, Batra SK (2006) Regulation of mucin expression: mechanistic aspects and implications for cancer and inflammatory diseases. Biochim Biophys Acta 1765: 189-222.

119. Perez-Vilar J, Hill RL (1999) The structure and assembly of secreted mucins. J Biol Chem 274: 31751-31754.

120. Gendler SJ, Spicer AP (1995) Epithelial mucin genes. Annu Rev Physiol 57: 607-634.

121. Lang T, Hansson GC, Samuelsson T (2007) Gel-forming mucins appeared early in metazoan evolution. Proc Natl Acad Sci U S A 104: 16209-16214.

122. Asker N, Axelsson MA, Olofsson SO, Hansson GC (1998) Dimerization of the human MUC2 mucin in the endoplasmic reticulum is followed by a N-glycosylation-dependent transfer of the mono- and dimers to the Golgi apparatus. J Biol Chem 273: 18857-18863.

123. Lidell ME, Johansson ME, Morgelin M, Asker N, Gum JR, Jr., et al. (2003) The recombinant C-terminus of the human MUC2 mucin forms dimers in Chinese-hamster ovary cells and heterodimers with full-length MUC2 in LS 174T cells. Biochem J 372: 335-345.

124. Godl K, Johansson ME, Lidell ME, Morgelin M, Karlsson H, et al. (2002) The N terminus of the MUC2 mucin forms trimers that are held together within a trypsin-resistant core fragment. J Biol Chem 277: 47248-47256.

125. Macao B, Johansson DG, Hansson GC, Hard T (2006) Autoproteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. Nat Struct Mol Biol 13: 71-76.

126. Palmai-Pallag T, Khodabukus N, Kinarsky L, Leir SH, Sherman S, et al. (2005) The role of the SEA (sea urchin sperm protein, enterokinase and agrin) module in cleavage of membrane-tethered mucins. FEBS J 272: 2901-2911.

127. Lidell ME, Johansson ME, Hansson GC (2003) An autocatalytic cleavage in the C terminus of the human MUC2 mucin occurs at the low pH of the late secretory pathway. J Biol Chem 278: 13944-13951.

128. Malmberg EK, Pelaseyed T, Petersson AC, Seidler UE, De Jonge H, et al. (2008) The C-terminus of the transmembrane mucin MUC17 binds to the scaffold protein PDZK1 that stably localizes it to the enterocyte apical membrane in the small intestine. Biochem J 410: 283-289.

129. Singh PK, Hollingsworth MA (2006) Cell surface-associated mucins in signal transduction. Trends Cell Biol 16: 467-476.

130. Johansson ME, Sjovall H, Hansson GC (2013) The gastrointestinal mucus system in health and disease. Nat Rev Gastroenterol Hepatol 10: 352-361.

131. Schwientek T, Bennett EP, Flores C, Thacker J, Hollmann M, et al. (2002) Functional conservation of subfamilies of putative UDP-N-acetylgalactosamine:polypeptide N-

acetylgalactosaminyltransferases in Drosophila, Caenorhabditis elegans, and mammals. One subfamily composed of l(2)35Aa is essential in Drosophila. J Biol Chem 277: 22623-22638.

132. Ten Hagen KG, Tran DT, Gerken TA, Stein DS, Zhang Z (2003) Functional characterization and expression analysis of members of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family from Drosophila melanogaster. J Biol Chem 278: 35039-35048.

133. Muller R, Hulsmeier AJ, Altmann F, Ten Hagen K, Tiemeyer M, et al. (2005) Characterization of mucin-type core-1 beta1-3 galactosyltransferase homologous enzymes in Drosophila melanogaster. FEBS J 272: 4295-4305.

134. Aoki K, Porterfield M, Lee SS, Dong B, Nguyen K, et al. (2008) The diversity of O-linked glycans expressed during Drosophila melanogaster development reflects stage- and tissue-specific requirements for cell signaling. J Biol Chem 283: 30385-30400.

135. North SJ, Koles K, Hembd C, Morris HR, Dell A, et al. (2006) Glycomic studies of Drosophila melanogaster embryos. Glycoconj J 23: 345-354.

136. Tian E, Ten Hagen KG (2006) Expression of the UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase family is spatially and temporally regulated during Drosophila development. Glycobiology 16: 83-95.

137. Ten Hagen KG, Tran DT (2002) A UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase is essential for viability in Drosophila melanogaster. J Biol Chem 277: 22616-22622.

138. Tian E, Ten Hagen KG (2007) A UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase is required for epithelial tube formation. J Biol Chem 282: 606-614.

139. Zhang L, Zhang Y, Hagen KG (2008) A mucin-type O-glycosyltransferase modulates cell adhesion during Drosophila development. J Biol Chem 283: 34076-34086.

140. Zhang L, Ten Hagen KG (2010) Dissecting the biological role of mucin-type O-glycosylation using RNA interference in Drosophila cell culture. J Biol Chem 285: 34477-34484.

141. Tran DT, Zhang L, Zhang Y, Tian E, Earl LA, et al. (2012) Multiple members of the UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase family are essential for viability in Drosophila. J Biol Chem 287: 5243-5252.

142. Dubreuil RR (2004) Copper cells and stomach acid secretion in the Drosophila midgut. Int J Biochem Cell Biol 36: 745-752.

143. Hansen JE, Lund O, Engelbrecht J, Bohr H, Nielsen JO, et al. (1995) Prediction of O-glycosylation of mammalian proteins: specificity patterns of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase. Biochem J 308 (Pt 3): 801-813.

144. Hansen JE, Lund O, Tolstrup N, Gooley AA, Williams KL, et al. (1998) NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. Glycoconj J 15: 115-130.

145. Lang T, Alexandersson M, Hansson GC, Samuelsson T (2004) Bioinformatic identification of polymerizing and transmembrane mucins in the puffer fish Fugu rubripes. Glycobiology 14: 521-527.

146. Lang T, Hansson GC, Samuelsson T (2006) An inventory of mucin genes in the chicken genome shows that the mucin domain of Muc13 is encoded by multiple exons and that ovomucin is part of a locus of related gel-forming mucins. BMC Genomics 7: 197.

147. Lehane MJ (1997) Peritrophic matrix structure and function. Annu Rev Entomol 42: 525-550.

148. Hegedus D, Erlandson M, Gillott C, Toprak U (2009) New insights into peritrophic matrix synthesis, architecture, and function. Annu Rev Entomol 54: 285-302.

149. Wang P, Granados RR (1997) Molecular cloning and sequencing of a novel invertebrate intestinal mucin cDNA. J Biol Chem 272: 16663-16669.

150. Roth GE, Wattler S, Bornschein H, Lehmann M, Korge G (1999) Structure and regulation of the salivary gland secretion protein gene Sgs-1 of Drosophila melanogaster. Genetics 153: 753-762.

151. Devine WP, Lubarsky B, Shaw K, Luschnig S, Messina L, et al. (2005) Requirement for chitin biosynthesis in epithelial tube morphogenesis. Proc Natl Acad Sci U S A 102: 17014-17019.

152. Tonning A, Hemphala J, Tang E, Nannmark U, Samakovlis C, et al. (2005) A transient luminal chitinous matrix is required to model epithelial tube diameter in the Drosophila trachea. Dev Cell 9: 423-430.

153. Fraichard S, Bouge AL, Chauvel I, Bouhin H (2006) Tenectin, a novel extracellular matrix protein expressed during Drosophila melanogaster embryonic development. Gene Expr Patterns 6: 772-776.

154. Takashima S, Murakami R (2001) Regulation of pattern formation in the Drosophila hindgut by wg, hh, dpp, and en. Mech Dev 101: 79-90.

155. Matsuo K, Ota H, Akamatsu T, Sugiyama A, Katsuyama T (1997) Histochemistry of the surface mucous gel layer of the human colon. Gut 40: 782-789.

156. Ferrari A, Veligodskiy A, Berge U, Lucas MS, Kroschewski R (2008) ROCK-mediated contractility, tight junctions and channels contribute to the conversion of a preapical patch into apical surface during isochoric lumen initiation. J Cell Sci 121: 3649-3663.

157. Gin E, Tanaka EM, Brusch L (2010) A model for cyst lumen expansion and size regulation via fluid secretion. J Theor Biol 264: 1077-1088.

158. Lowery LA, Sive H (2005) Initial formation of zebrafish brain ventricles occurs independently of circulation and requires the nagie oko and snakehead/atp1a1a.1 gene products. Development 132: 2057-2067.

159. Zhang J, Piontek J, Wolburg H, Piehl C, Liss M, et al. (2010) Establishment of a neuroepithelial barrier by Claudin5a is essential for zebrafish brain ventricular lumen expansion. Proc Natl Acad Sci U S A 107: 1425-1430.

160. Navis A, Marjoram L, Bagnat M (2013) Cftr controls lumen expansion and function of Kupffer's vesicle in zebrafish. Development 140: 1703-1712.

161. Husain N, Pellikka M, Hong H, Klimentova T, Choe KM, et al. (2006) The agrin/perlecan-related protein eyes shut is essential for epithelial lumen formation in the Drosophila retina. Dev Cell 11: 483-493.

162. Strilic B, Eglinger J, Krieg M, Zeeb M, Axnick J, et al. (2010) Electrostatic cell-surface repulsion initiates lumen formation in developing blood vessels. Curr Biol 20: 2003-2009.

163. Kerman BE, Cheshire AM, Myat MM, Andrew DJ (2008) Ribbon modulates apical membrane during tube elongation through Crumbs and Moesin. Dev Biol 320: 278-288.

164. Cheshire AM, Kerman BE, Zipfel WR, Spector AA, Andrew DJ (2008) Kinetic and mechanical analysis of live tube morphogenesis. Dev Dyn 237: 2874-2888.

165. Gutzman JH, Sive H (2010) Epithelial relaxation mediated by the myosin phosphatase regulator Mypt1 is required for brain ventricle lumen expansion and hindbrain morphogenesis. Development 137: 795-804.

166. Rizki MTM (1956) The secretory activity of the proventriculus of Drosophila melanogaster. Journal of Experimental Zoology 131: 203-221.