Regulated assembly of a prion-like phosphatase, *CG5830*, controls its phosphatase activity and embryonic development in *Drosophila melanogaster*

By

Zelha Nil

M.Sc., Middle East Technical University, 2012
B.Sc., Middle East Technical University, 2008

Submitted to the graduate degree program in Molecular and Integrative Physiology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Co-Chair: Dr. Kausik Si

Co-Chair: Dr. Michael Wolfe

Dr. Joan Conaway

Dr. Robb Krumlauf

Dr. Chad Slawson

Date Defended: June 19, 2018

The dissertation committee for Zelha Nil certifies that this is the approved version of the following dissertation:

Regulated assembly of a prion-like phosphatase, *CG5830*, controls its phosphatase activity and embryonic development in *Drosophila melanogaster*

Co-Chair: Kausik Si

Co-Chair: Dr. Michael Wolfe

Date Approved: June 19, 2018

Abstract

Prion-like proteins can assume distinct conformational states in the same cell, and often one of these states leads to a self-assembling oligomer that can vary in physical nature. The conformational state change from monomers to oligomers can act like a switch, and most importantly this switch also leads to a change in protein function, i.e. loss of the original function or gain of a new function. Such protein switches have been shown to be involved in transcriptional regulation, translational regulation and signaling, regulating many different physiological processes such as stress response, memory formation, and immune response. However, it is unclear how common such mechanism(s) are utilized in multicellular organisms and the diversity of the processes they regulate. Here, I report the characterization of CG5830, a candidate a prion-like protein identified through a systematic screen in *Drosophila melanogaster*. I find that in the *Drosophila* embryo, CG5830 changes its conformational state from monomers to amyloid like oligomers on the membrane during gastrulation. It functions as a membrane phosphatase required for embryonic patterning and interacts with major regulators of developmental pathways, such as TGF-β, EGF and FGF signaling pathways. Finally, a conformational state change of CG5830 leads to gain of enzymatic function. My data indicate that prion-like protein switches can have essential roles in animal development.

Acknowledgements

"Once you free your mind about a concept of harmony and of music being correct,

You can do whatever you want

So nobody told me what to do, and there was no preconception of what to do…" Giorgio

With all the unknowns and uncertainties, PhD was a challenging but an amazing journey for me. Looking back, I was not alone. I always had Kausik backing me up, guiding me, supporting me. When I did not have any preconception, he already had the outline ready for me. When I did not have the courage, he was the one teaching me how to be strong in front of difficulties. I cannot thank him enough for preparing me for my future adventures in science.

I feel very lucky to have my PhD career here, in a very special program of both the University of Kansas Medical Center and the Stowers Institute. I am very thankful to all my committee members for their contributions and guidance throughout my studies, especially Dr. Robb Krumlauf for being my second mentor both academically and professionally. As my third mentor throughout my PhD, Dr. Susan Abmayr, I will always carry a piece of her enthusiasm and trust in students with me in my future career.

I owe lots of thanks to all the members of Si lab for creating the most friendly and supportive environment! I will never forget the nights with Paulo and Matthew that we danced, listened to music, or listened to Mr. muffin ("why don't you let me die!") while collecting embryos… I cannot thank Paulo enough for being with me and for patting my back during the hardest times. It was such an amazing opportunity to work with a really good scientist, Ruben, throughout my last year. I will always remember Blake's crazy but right questions, and his hair. I thank you so much: Younshim for listening my presentations over and over again, and an outside member of Si lab, Zainab for making Stowers a better place to be. I also owe big thanks to all the past members who were here during my PhD.

This work would be impossible without the help of Stowers Institute core members that I have worked together. I owe a very special thanks to Jeffrey Lange for working with me, especially tolerating all my ‘sometimes’ unbearable jokes. I cannot forget helps of the best PostDocs in the institute, Narendra Singh, Takuya Akiyama and Stephanie Nowatarski, during learning how to work with *Drosophila*, which was a whole another world to me at the beginning.

All my friends here who made Kansas City somehow livable, I thank you all so much: Olcay, Damon, Melek, Enis, Ayse, Ural, Richard, Kim, Neil, Anju, Rick, Izzie, Petros, Kelly, Sirma, Ahmetcan, Molly, Sila, Banu.

And my lifetime friends, Tugba, Gulistan, Dincer, Aycan, Gokce, Beyaz, Ahmet A, Efsun, Emin baba, Aysegul, who I went to college with, who I spent my childhood with, who I cried, who I laughed together, who I talked about my experiments for endless hours, who I talked about academia, who I talked about life, I hope I will always have you with me.

Finally, I thank my family for their endless love, trust and support in every step of my life. My father always says, “keep your mind busy, always have some book you are reading”, my mother always says, “I know you will do it” … Thank you for believing in me. Without my sisters though, I would not have fun. Lastly, I thank the most to my dearest sister, Umran, for being with me throughout all this beautiful but challenging journey, and in life.

To all the academicians of Turkey, who lost their jobs and were found guilty of being "Academics for Peace"

Table of Contents

List of Figures

List of Tables

Chapter I: Introduction

Forming appropriate conformational states are integral to protein function and most proteins adopt one specific conformation in a cell dictated by their primary amino acid sequence and folding dynamics during synthesis (Anfinsen, 1973). This relationship between a protein's sequence and one definite structure was challenged with the discovery of prions and prion-like proteins (S. Prusiner, 1982; Ross, Minton, & Wickner, 2005; Staaf et al., 2008). Contrary to most proteins, these groups of proteins are found to exist in at least two different conformations in the cell: a monomeric state and a self-replicating oligomeric state. Once the oligomeric state is attained, it leads to conversion of monomers to the oligomeric state in a dominant manner and maintains this conformation over time as well as across cell divisions (S. B. Prusiner, 2013). For prions, this mechanism of self-maintenance is found to be the underlying cause of several neurodegenerative diseases (Horwich & Weissman, 1997; Kelly, 1998; S. Prusiner, 1991). However, non-pathogenic functional counterparts of prions, called prion-like proteins, were identified in yeast and they have been found to be the basis of different epigenetic heritable phenotypes, acting as regulatory switches in a controlled manner (Shorter & Lindquist, 2005; Reed B Wickner, Edskes, Shewmaker, & Nakayashiki, 2007). Recently, sporadic examples of such prion-like proteins have been shown in many different organisms, from bacteria to human, and the diversity of biological processes that they participate in is also remarkably broad (*e.g.*, memory, regulation of interferon expression, and transcriptional regulation (Cai & Chen, 2014; Fowler et al., 2005; Franklin et al., 2014; Hou et al., 2011b; Ishimaru et al., 2003; Majumdar et al., 2012; Si, Choi, White-Grindley, Majumdar, & Kandel, 2010; Tariq, Wegrzyn, Anwar, Bukau, & Paro, 2013; Yuan & Hochschild, 2017). The biological and functional diversity of prion-like proteins suggests that proteins with self-sustaining conformational states may be a part of an evolutionarily conserved regulatory mechanism involved in normal physiological function (Soto, 2012). However, since they are primarily characterized *in vitro*, it remains unclear how prevalent

these proteins are, what function they serve in different conformational states, and how they are regulated in normal physiology of organisms.

1.1 Characteristics of Prion-like Proteins

The most profound feature of prion-like proteins is their ability to exist in at least two different conformational states: a non-prion monomeric form and a self-maintaining prion-like oligomeric form (Harbi & Harrison, 2014). This oligomeric form can vary in physical nature from flexible liquid crystals to hydrogels to the most commonly found highly stable state: amyloid (H. Wu & Fuxreiter, 2016). Structural studies of prion-like proteins showed that amyloidogenic oligomeric state often forms fibrils which can have a cross β-sheet structure (D. S. Eisenberg & Sawaya, 2017). Moreover, these amyloidogenic oligomers bind to amyloid specific dyes such as thioflavin-t and congo red; they show resistance to detergents, chaotropic agents, proteases or heat; and they can propagate their conformation from cell to cell (Si, 2015). However, since these proteins can adapt multiple conformational states with different biophysical natures, there can also be variations among different prion-like proteins in terms of their biochemical properties (Harbi & Harrison, 2014; Si, 2015).

Another common feature of prion-like proteins is that they often contain a modular domain, called a prion domain (PrD), which is necessary and sufficient for self-propagation of the protein but can be dispensable for its biochemical function (Alberti, Halfmann, King, Kapila, & Lindquist, 2009). Even though the PrD is sufficient for protein state changes into prion-like form, this process is regulated by a physiological signal that is used to change protein function, i.e. loss of the original function or gain of a new function (Si, 2015). PrDs in different proteins share similar amino acid composition: enriched for polar residues such as glutamine (Q) and asparagine (N), and depleted of hydrophobic and charged residues (Batlle, Iglesias, Navarro, & Ventura, 2017). To a large extent, this composition determines the amyloid propensity of these proteins, leading to formation of disordered conformational ensembles which allows amyloid nucleating contacts to

be made (Alberti et al., 2009). Indeed, these domains are located in intrinsically disordered (ID) and/or low-complexity (LC) regions, which contains similarly charged amino acids and repetitive amino acids, respectively. These regions in proteins collectively lead to structural disorder causing low folding probability (Batlle et al., 2017). Even though these common characteristics of PrDs led to development of several computer tools to screen for such proteins *in-silico* at proteomic level, such as PLAAC, pWALTZ, and PrionW, these tools mainly utilize sequence features of yeast PrDs (Lancaster, Nutter-Upham, Lindquist, & King, 2014; Sabate, Rousseau, Schymkowitz, & Ventura, 2015; Zambrano et al., 2015).

Amyloid structure

Amyloids are among the most stable protein conformations which forms through folding of PrD, ID or LC region in the proteins upon binding to themselves (H. Wu & Fuxreiter, 2016). Amyloids often have non-branching fibril structure containing a central cross-β sheet core with solvent excluded, self-complementing steric zipper interactions (D. Eisenberg & Jucker, 2012). In fact, the structural disorder of PrD, ID, or LC regions of such proteins leads to exposure of short segments called steric zippers from these proteins (D. S. Eisenberg & Sawaya, 2017). These segments form cross-β spines of the fibrils via homotypic interdigitating interactions in parallel or antiparallel arrangements (Hughes et al., 2018) with a high barrier of dissociation leading to an irreversible assembly (H. Wu & Fuxreiter, 2016).

Although previous studies of amyloid structures of different proteins showed extensive similarity between their structural features, recent high-resolution data of these proteins showed that amyloids at molecular level have a high level of conformational heterogeneity (Toyama & Weissman, 2011). It has been shown that identical polypeptides can fold into multiple, distinct amyloid conformations with differences in distances between turns in the fibrils, or in the length of the fibrils (Chuang, Hori, Hesketh, & Shorter, 2018; Colby et al., 2009; Hiramatsu et al., 2010).

Moreover, one surprising finding revealed that the amyloid structure of phenol-soluble modulin α3 (PSMα3) peptide is formed of cross-α fibrils with similar biophysical features to cross-β fibrils such as binding to Thioflavin T (Tayeb-Fligelman et al., 2017). In addition to conformational heterogeneity, recent findings also suggested that amyloid structures can be dynamically regulated, e.g. Rim4 amyloidogenic aggregates form, and disappear upon phosphorylation within minutes (Carpenter, Bell, Yunus, Amon, & Berchowitz, 2018; Ford & Shorter, 2015).

1.2 Screen for Prion-like Proteins in *Drosophila melanogaster*

To identify novel prion-like proteins in the *Drosophila* proteome, members of the Si lab carried out an *in-silico* screen by using common sequence characteristics of known PrDs found in yeast. In addition to these characteristics, the *Drosophila* proteome was also screened for characteristics of prion-like proteins in multicellular eukaryotes (Hou et al., 2011b; Si et al., 2010) including: presence at either the N- or C-terminal end of the protein; enrichment in Ser/Gly/His residues; presence of a similar PrD in the homologs of these PrD containing proteins in other species (R B Wickner, Taylor, Edskes, & Maddelein, 2000). Proteins with more than ten continuous stretches of glutamine residues were excluded from this screen because the *Drosophila* genome is highly enriched in opa (CAA and CAG) repeats, which encode for glutamine residues (Wharton, Yedvobnick, Finnerty, Tsakonas, & Artavanis-Tsakonas, 1985). After performing computational analysis based on these criteria, results were compared with 470 previously identified putative prion domain containing proteins in the *Drosophila* genome (Harrison & Gerstein, 2003; Michelitsch & Weissman, 2000). From the shared list of 43 proteins, the highest-scoring 21 candidates were experimentally tested for common characteristics of prion-like aggregates by previous Si lab members (Alberti, Halfmann, & Lindquist, 2010). Their properties were also compared to a well-known prion-like protein of *Drosophila,* Orb2A, and control proteins with no known prion-forming capacity, such as calmodulin (Figure 1.1. Data and figures is kindly provided by previous Si lab members).

First, secondary structure analysis of purified recombinant candidate proteins was performed by infrared spectroscopy (FTIR) to determine if the types of molecular vibrations made by the molecule are similar to those made by amyloidogenic aggregates. An extensive representation of amyloid-type vibration, similar to Orb2, supported the idea that some of these proteins have the ability to form amyloidogenic aggregates (Figure 1.1A).

Second, the aggregation propensity of each candidate protein was detected as visible fluorescent puncta formation by confocal microscopy through the use of GFP tagged variants of the proteins in S2 cells (Figure 1.1B). The stability of these fluorescent aggregates for each candidate protein was tested by fluorescence recovery after photobleaching (FRAP), in which fluorescent aggregates of the protein were hit with very intense laser light, such that the fluorophore conjugated to the protein would bleach and become unable to fluoresce. Therefore, any fluorescence appearing after bleaching should be due to new molecules diffusing into the aggregate. The aggregated form of a prion-like protein is structurally very ordered which would lead to a very little exchange with new molecules. Hence, a lengthy FRAP time would indicate that the protein could be prion-like (Figure 1.1C).

Third, self-assembling propensity of each candidate protein was detected by measuring fluorescence energy transfer (FRET) efficiency between YFP and CFP tagged copies of the same protein in S2 cells. In this assay, a laser initially excites CFP leading it to emit a 477-nanometer photon which is in the absorption spectrum of YFP. If the two fluorescent proteins are closely located, YFP would be able to absorb the emitted photon of CFP. Therefore, in FRET assays, the YFP fluorescent signal serves as a readout for the self-assembled aggregates of a prion-like protein. The putative prion-like proteins were compared to Orb2 and a FRET efficiency that was similar to or greater than Orb2 was considered prion-like behavior (Figure 1.1D).

Another feature of prion-like proteins is that they are resistant to detergents such as SDS. This behavior can be assayed using semi denaturing detergent agarose gel electrophoresis (SDD-AGE) (Alberti et al., 2010). In this assay, most proteins will run through an agarose gel at the same pace. However, prion-like proteins in aggregates, which are resistant to denaturation by SDS, would migrate slowly and appear as smears on a western blot. To test this, GFP-tagged proteins were expressed in S2 cells, and run on an SDD-AGE, and subsequently probed with an anti-GFP antibody (Figure 1.1E). A smear or broad staining pattern was considered evidence that the protein may behave in a prion-like manner.

Last, a Sup35 based prion assay was performed by swapping prion domain of Sup35 with the predicted prion domain of each candidate protein and assaying for the ability of each domain to recapitulate prion like activity of Sup35 (Figure 1.1F). Sup35, a yeast translation termination factor, is one of the best characterized prion-like proteins (True, Berlin, & Lindquist, 2004). The Sup35 protein has become a very useful tool to identify novel prion-like domains since it has a modular prion domain which can be replaced without affecting its function, and it has different phenotypes in monomeric non-prion and aggregated prion-like state which can be easily scored. The phenotypic assay is based on a yeast strain with *ade1-14* mutation, which leads to a premature UGA stop codon in the *Ade1* gene in the adenine biosynthesis pathway. When Sup35 is in a monomeric state, it is active, so the translation of Ade1 gene stops at the premature stop codon. This makes cells unable to synthesize adenine and therefore unable to grow in media without adenine, accumulating a metabolic intermediate that turns cells red. When the Sup35 protein is in an aggregated prion state, it loses its translation termination activity, allowing for read-through translation and synthesis of adenine. This allows the cells to grow in adenine (-) media with a white color because they do not accumulate the metabolic intermediate. Therefore, Sup35 chimeric proteins can be scored for the non-prion and prion state based on color and growth in

adenine (-) media. The prion-like domain of Sup35 was replaced with each putative prion domain and their ability to confer the prion-like state was assayed.

Candidate proteins which show prion-like behavior in multiple assays, namely CG5830, Hr4, Pipsqueak, CG42540 and Murashka, are considered prion-like. Interestingly, out of these five candidates, CG5830, Hr4 and Psq, are required for developmental progression of *Drosophila.* CG5830, a strong candidate for a prion-like protein based on the aforementioned assays, is an uncharacterized putative phosphatase with homology to small CTD phosphatases (SCPs). It is indeed the only *Drosophila* ortholog of mammalian SCPs with a HAD phosphatase signature motif D*X*D*X*(T/V) where both Asp residues are essential for activity (Yeo, Lin, Dahmus, & Gill, 2003).

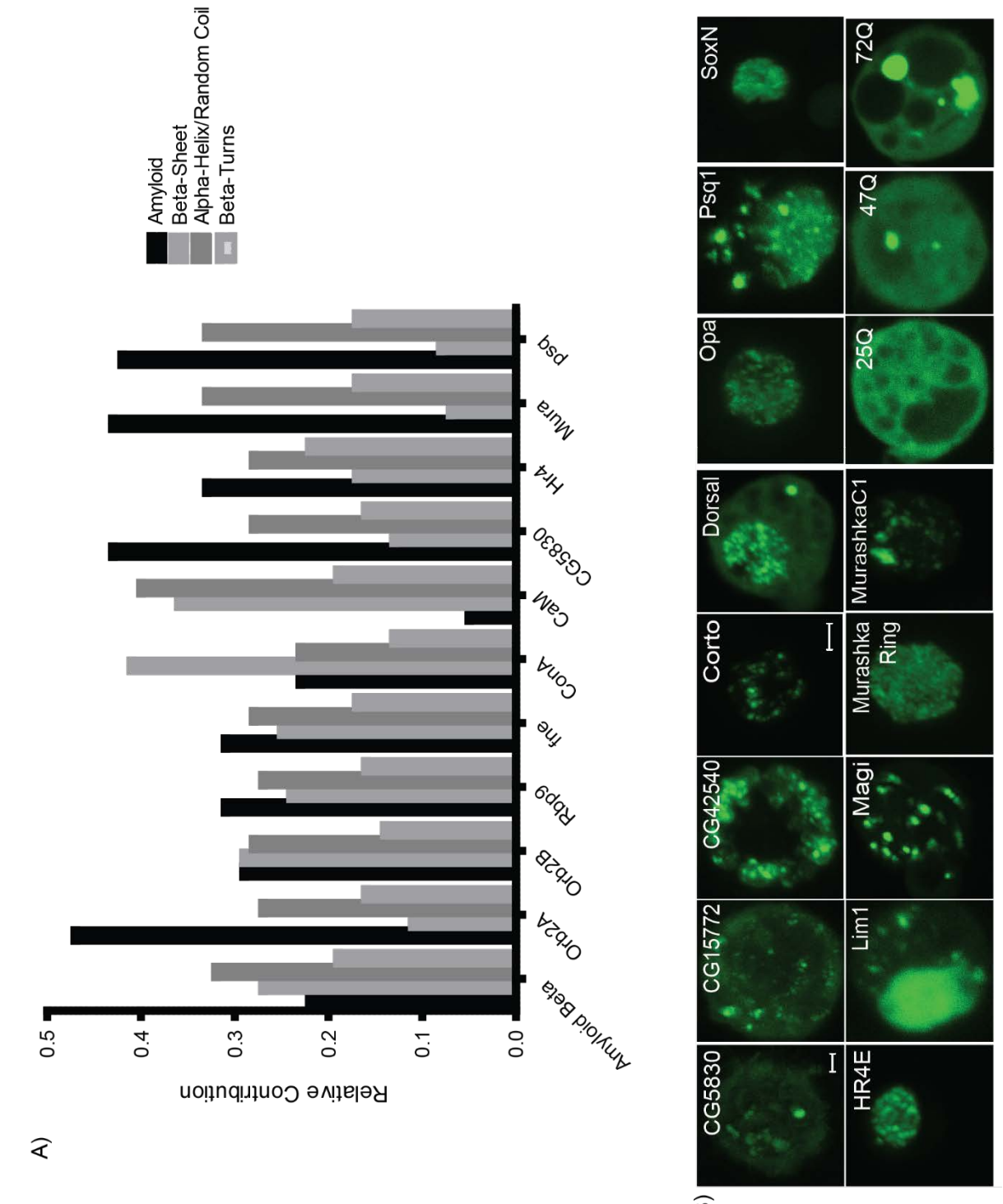


Figure 1.1 Screen for prion-like proteins.

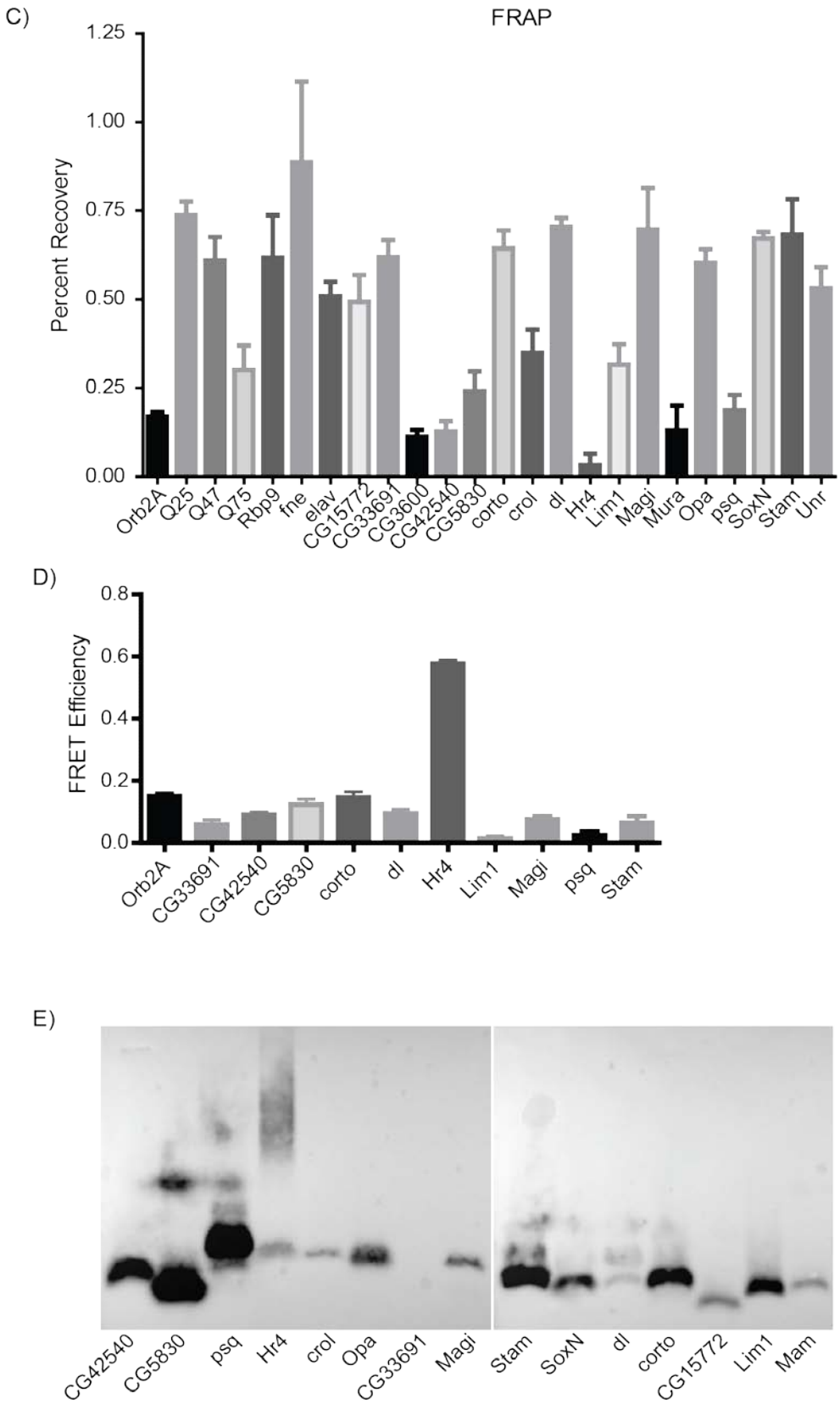


Figure 1.1 Continued.

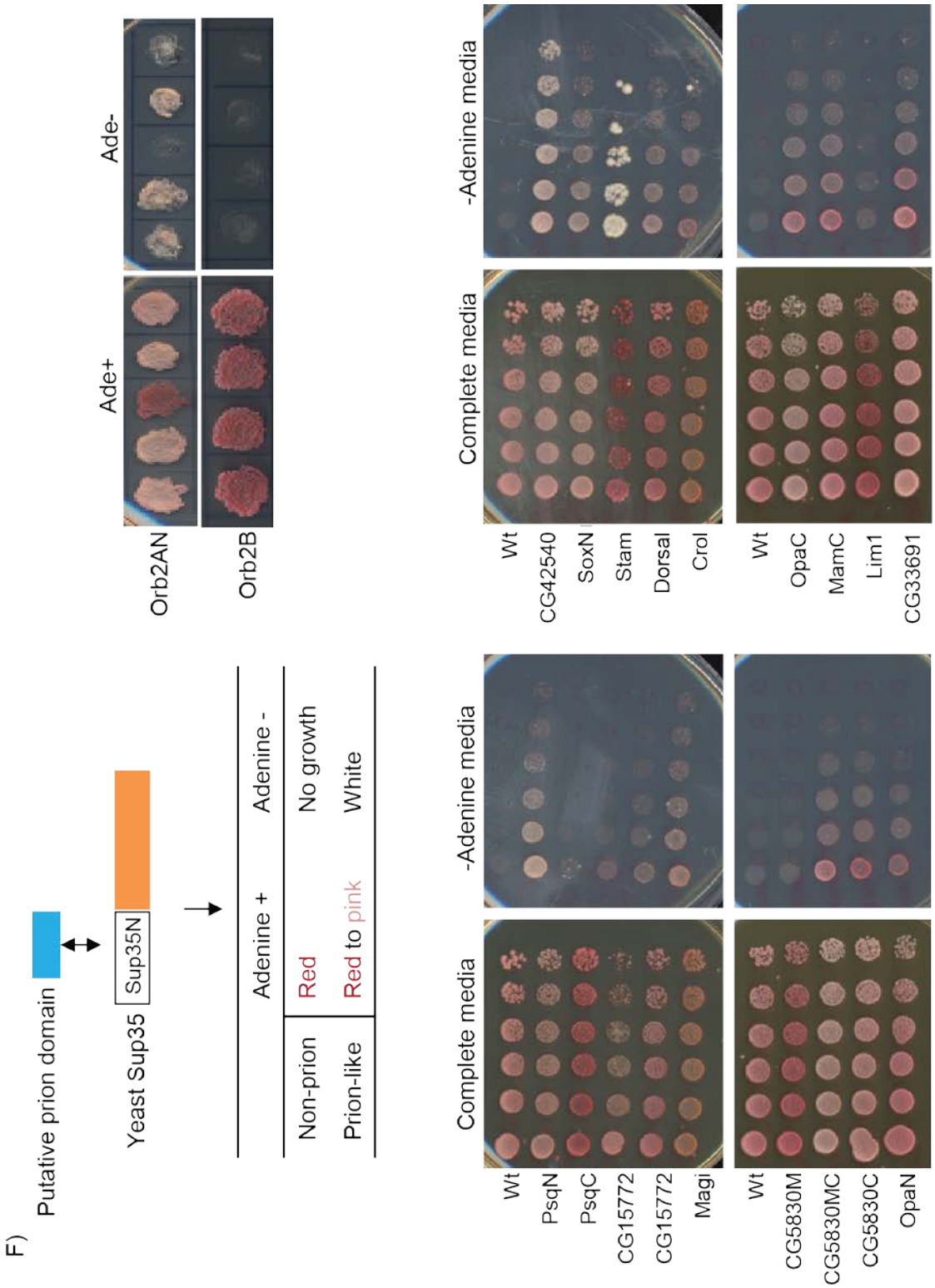


Figure 1.1 Continued.

Figure 1.1 Screen for prion-like proteins.

a. Secondary structure determination by FTIR. Relative integrated area under classic tertiary structure peaks from infrared spectrophotometry performed on a variety of purified potential prion-like domains and on positive and negative control proteins.
b. Fluorescent puncta detection of GFP tagged candidate proteins by confocal microscopy.
c. FRAP analysis of candidate proteins. Percent of fluorescence recovered after photobleaching of a variety of putative prion-like proteins compared to a well-known prion like protein Orb2A, liquid crystal-like aggregates of glutamine repeats or non-aggregating proteins (Elav, Fne, Rpb9).
d. FRET efficiency of candidate proteins. Efficiency of energy transfer between two copies of the putative prion-like proteins of interest, compared to Orb2A, a well-established prion-like protein.
e. SDS-resistant oligomer detection by SDD-AGE. (Purified putative prion-like proteins run through 1.5% SDS agarose gel.)
f. Yeast Sup35 domain substitution assay. Yeast Sup35 was replaced with putative prion-like domains and assayed for prion functionality, which is observable by a white colony on Ade- media. Orb2A (a well-defined prion-like protein) has white growth on Ade- media, indicating that the amyloid form has generated and therefore Adenine is being produced, whereas Orb2B, a non-prion-like splice variant of Orb2A, doesn't form the amyloid Sup35 structure and therefore does not have growth on Ade- media Putative prion-like protein domains were substituted into the Sup35 gene and ability to recapitulate Sup35 prion-like activity was assayed.

1.3 Phosphatases: Classification and Regulation

Protein phosphorylation is a dynamic post translational modification (PTM) that allows cells to translate extracellular signals into functional changes through intracellular signaling pathways in diverse processes, such as transcription, translation, cell cycle, cell motility, intercellular interactions, differentiation and basic metabolism (Almo et al., 2007; Shenolikar, 2007). Protein kinases and phosphatases contribute to regulation of these robust signaling pathways by modulating and setting the levels of cellular protein phosphorylation at specific times and locations. Complexity and specificity of kinases in these networks are generally associated with the number of genes encoding for the proteins. Indeed, almost 2% of the eukaryotic protein coding genes, i.e. yeast, fruit fly, mouse and human, encode for protein kinases. However, only one third as many genes encode for protein phosphatases, implying a low level of substrate specificity and regulation (Table 1.1) (Caenepeel, Charydczak, Sudarsanam, Hunter, & Manning, 2004; Morrison, Murakami, & Cleghon, 2000; Snyder et al., 2000). This assumed promiscuity and redundancy of phosphatases has been supported by their relatively indistinguishable catalytic activities in many *in vitro* experiments comparing their action on the same or multiple substrates and site preferences in a substrate. In contrast, *in vivo* genetic studies, mostly in fruit fly and to a lesser degree in mice, showed distinct phenotypes for specific phosphatases (Mustelin, 2007). This suggests that there are context-dependent mechanisms *in vivo* that allow them to exert specific activities, and that they may be precisely regulated non-redundant enzymes.

Table 1.1 Distribution of kinases and phosphatases in different organisms.

	Protein coding genes	Kinases	Phosphatases
Human	~24000	518 (0.022)	200 (0.0083)
Mouse	~22000	510 (0.023)	160 (0.0072)
Fruit fly	~13900	251 (0.018)	86 (0.006)
Yeast	~5400	122 (0.022)	38 (0.007)

Based on their catalytic domain structure, protein phosphatases are classified into four different families, namely protein tyrosine phosphatases (PTPs), phosphoprotein phosphatases (PPPs), protein phosphatases dependent on Mn2+/Mg2+ ions (PPMs), and haloacid dehalogenases (HADs) (Mustelin, 2007). However, this structural classification does not reflect substrate specificity, according to which these enzymes can be classified into three categories: Ser/Thr- specific, Tyr-specific, and dual-specific phosphatases. While PPPs and PPMs are specifically Ser/Thr phosphatases, PTPs include both tyrosine specific and dual specific enzymes. The HAD family phosphatases also have been shown to be specific for phospho-serine (Gohla, Birkenfeld, & Bokoch, 2005) and/or tyrosine (Rebay, Silver, & Tootle, 2005; Yeo et al., 2003).

Table 1.2 Phosphatase families (Adapted from (Mustelin, 2007))

Phosphatase families	Examples of members
1. *PTP family*	CD45, TCPTP, MKP7, VHZ, SSH1, PRL-1, CDC14A, PTEN, MTM1, CDC25A, LMPTP
2. *PPP family*	PP1, PP2A, calciunerin, PP5
3. *PPM family*	PP2C
4. *HAD family*	Eya, CTD, cronophin

Thus far, all phosphatases are regulated either by modulating direct interaction with their substrate or controlling their localization to a specific subcellular location/ compartment containing the substrate. These mechanisms are mediated by different structural elements in the phosphatases, in addition to the catalytic domain, namely i) targeting domains such as BRCT (breast cancer 1 C terminus) domain of FCPs (F-cell production, also called TFIIF-stimulated CTD phosphatase); ii) small binding motifs such as DNA binding motif of PNKP (polynucleotide kinase phosphatase); iii) regulatory or scaffold subunits such as ~200 different regulatory subunits of PP1 (Figure 1.2) (Sacco, Perfetto, Castagnoli, & Cesareni, 2012; Seifried, Schultz, & Gohla, 2013).

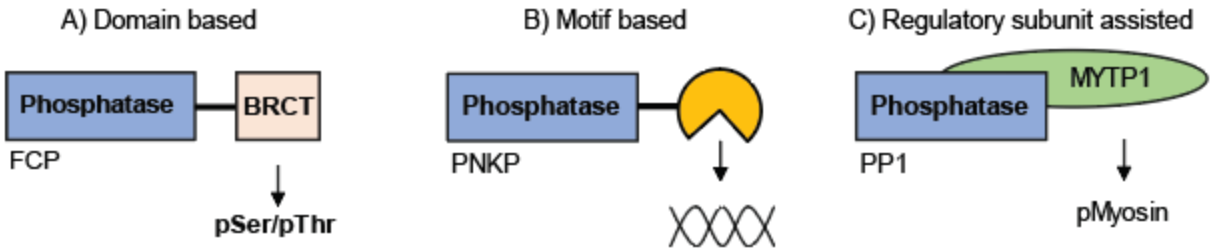


Figure 1.2 Regulatory mechanisms of phosphatases for substrate recognition. Adapted from (Sacco et al., 2012)

1.3.1 Small CTD Phosphatases: Regulation and Function

CG5830, a strong candidate for a prion-like protein based on screening done in Si lab, is an uncharacterized putative phosphatase with homology to small CTD phosphatases (SCPs). It is indeed the only *Drosophila* ortholog of mammalian SCPs with a HAD phosphatase signature motif (Yeo et al., 2003). Generalization of above-mentioned regulatory mechanisms for all phosphatases would be misleading, since HAD subfamily of SCPs, like prokaryotic HADs, contains only a single catalytic domain without any of the recognizable structural elements listed above (Seifried et al., 2013). They also do not have a cap module structure in their active site, which is used by other HAD family members for substrate selectivity (M. Zhang, Cho, Burstein, Siegel, & Zhang, 2011). But different isoforms of SCPs in diverse organisms have been shown to localize to distinct subcellular compartments, such as the nucleus and cell membrane, and to dephosphorylate a variety of substrates, including carboxy terminal domain of RNAPII, SMADs, REST, and Snail (Wani, Sugita, Ohkuma, & Hirose, 2016). Therefore, it is still elusive how these seemingly structurally primitive enzymes, resembling their prokaryotic counterparts, achieve specific substrate recognition.

SCPs are highly conserved among eukaryotes including, human, mouse, chicken and fruit fly. While human, mouse and chicken each has at least four orthologous genes, fruit fly has only one gene encoding an uncharacterized phosphatase orthologous to SCPs, called CG5830. They all

share a common catalytic CTD phosphatase domain characterized by HAD phosphatase signature motif DXDX(T/V) and a disordered N-terminal tail, with no known function (Jaramillo-Tatis, Bamm, Vassall, & Harauz, 2014; Wani et al., 2016). In chicken, downregulation of *SCP1* promotes neuronal differentiation during embryonic CNS development (Visvanathan, Lee, Lee, Lee, & Lee, 2007). In mouse embryonic cells, SCP1 inhibits neurogenesis (Yeo et al., 2005). In human cells, different SCPs have been found to interact and dephosphorylate diverse members of major developmental pathways, including Wnt signaling and TGFβ/BMP signaling, in addition to their common substrate RNAPII CTD tail (Knockaert, Sapkota, Alarcón, Massagué, & Brivanlou, 2006; Sapkota et al., 2006; Wrighton et al., 2006; Y. Wu, Mark Evers, & Zhou, 2009). Together, these observations suggest that SCPs play significant and diverse roles during early development.

1.4 *Drosophila melanogaster* as a model for early development

Drosophila has been an important model organism since it was first used by Thomas Hunt Morgan in early 1900s for studies of heritability of various traits (Morgan, 1909, 1910). Until 1937, when D. F. Poulson showed how chromosomal deficiencies affected embryonic development of *Drosophila* (Poulson, 1937), genetics studies were done only on adult flies due to ease in tracking phenotypes. These early studies established both the concept of the genetic control of embryonic development and the framework of how to study this concept. Even though Poulson and others continued to use this approach to study embryogenesis, *Drosophila* embryology remained as a niche field for several years due to technical limitations of working with small embryos. It was not until 1970s-1980s that it became possible to do histological analysis on embryos without breaking them apart. This made it possible to study early embryogenesis in much greater detail than before (Turner & Mahowald, 1976). These technical improvements allowed Christiane Nüsslein-Volhard and Eric Wieschaus to perform the “Heidelberg screen” for mutations that changed the segmentation pattern of the *Drosophila* embryo (Nüsslein-volhard & Wieschaus, 1980). In this

screen, they identified 15 loci, mutations of which altered the segmental pattern of larva and were scored by the changes in its cuticle pattern. Based on the effects on segmentation patterns they grouped the mutants into three, categories; i) gap mutants (deletions of a whole stretch of segments), ii) pair-rule mutants (deletions in alternating segments) and iii) segment polarity mutants (deletions of parts of each segment that are replaced by a mirror-image of the remainder) (Detailed in sections 1.5 and 1.6). This pioneering work led the way for uncovering a large number of other genes and pathways involved in key developmental processes in embryogenesis. Furthermore, due to genomics and genome sequencing, the appreciation that a wide range of animals from invertebrates to vertebrates have a shared gene toolkit has helped to translate this knowledge to show that it is highly relevant to human biology and disease, owing to the functional conservation of *Drosophila* genes across evolution. In fact, 75% of disease causing genes in humans have functional counterparts in fly (Pandey & Nichols, 2011). For that reason, *Drosophila* has enormous potential as a model system to dissect a specific gene function and translate its relevance to other organisms.

In addition to these evolutionary merits, use of *Drosophila* as a model organism has many practical advantages too. These include low-cost to culture, short generation time (10 days), high fecundity of females, easy *in vivo* manipulation due to well-annotated genome, and external development of embryo making it visible to study. Under normal conditions (25°C), it takes approximately a day for the *Drosophila* embryo to develop into a larva from fertilization. During this time, we can observe cells dividing and differentiating into the precursors of the organs and appendages of the adult fly. After hatching into larva, it develops over 4–5 days until it pupates, and then it metamorphoses into adult fly over the course of 4 days (Kölling, 2015; Weigmann et al., 2003).

The combination of these genetic and practical advantages makes the *Drosophila* embryo a great system to characterize functions of a gene and its protein product both temporally and spatially.

1.5 A General Overview of Embryonic Development in *Drosophila melanogaster*

In the following sections, an overview of embryonic development of *Drosophila melanogaster* is given based on the morphological staging described by Campos-Ortega and Hartenstein (Campos-Ortega & Hartenstein, 1997). (All the embryo images in this section are adapted from the Atlas of Drosophila Development (Hartenstein, 1993) and used with permission.)

0–2h50: Fertilization, the syncytium and cellularization (morphological stages 1–5)

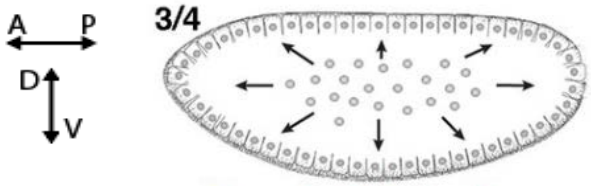


Figure 1.3 *Drosophila melanogaster* embryo, lateral view, morphological stages3/4: syncytial blastoderm.
A-P: Anteroposterior D-V: Dorsoventral axis. This orientation is the same for the following images.

During the first 2h of development, the nuclei go through a series of 13 mitotic cell divisions in a single shared cytoplasm. The first five divisions occur in the central region and after the fifth division, the nuclei migrate outwards while continuing to divide, and they form the syncytial blastoderm (Figure 1.3). At 10^{th} division, 15 nuclei move towards the posterior end of the embryo and form pole cells which are the progenitors of adult germ cells. After 13^{th} nuclear division when maternal gene products are depleted and zygotic genes are activated (maternal-to-zygotic transition) (Benoit et al., 2009; Edgar & Schubiger, 1986), all the nuclei are enclosed by a plasma membrane in a process called cellularization, and embryo turns into a cellular blastoderm.

2h 50min–4h after fertilization: Gastrulation (morphological stages 6–9)

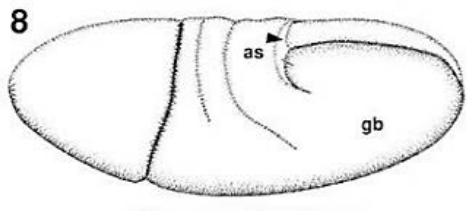


Figure 1.4 *Drosophila melanogaster* embryo, lateral view, morphological stage 8.
as: amnioserosa and gb: germ band.

Gastrulation begins at 2h 50min and causes major cell shape changes and movements that separates the embryo into three germ layers: the endoderm, mesoderm and ectoderm (Leptin, 1999). Around the 3h 10min, the germ band starts elongating, folding back upon itself on the dorsal side (Figure 1.4).

4–7h20min after fertilization: Germ band elongation (morphological stages 9–11)

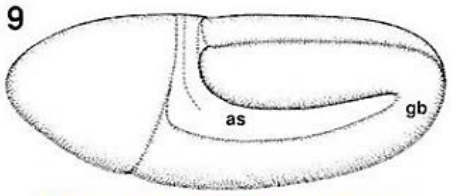


Figure 1.5 *Drosophila melanogaster* embryo, lateral view, morphological stage 9.
as: amnioserosa and gb: germ band.

Between 4-6h, the germ band slowly elongates further, reaching its maximum length around 5h (Figure 1.5).

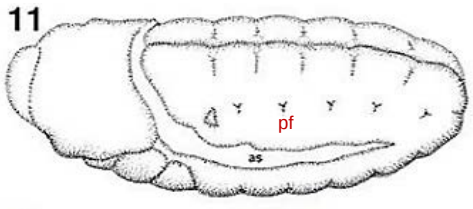


Figure 1.6 *Drosophila melanogaster* embryo, lateral view, morphological stage 11.
as: amnioserosa, pf: para-segmental furrow.

When the germ band is fully extended, the cells arrange to form clearly visible segments. This stage is marked by para-segmental furrows, which divides the germ band into distinct segments.

7h20min–10h after fertilization: Germ band retraction (morphological stages 12–13)

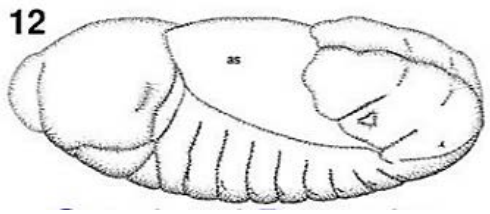


Figure 1.7 *Drosophila melanogaster* embryo, lateral view, morphological stage 12.
as: amnioserosa.

At 7h 20min, germ band retraction is initiated (Figure 1.7). Around this time, with coordinated removal of cells in specific locations through programmed cell death, the embryo continues the process of segmentation (Abrams, White, Fessler, & Steller, 1993). Germ band retraction ends at around 9h 20min, and most organ primordia start to differentiate. Segments become more visible, and anterior and posterior midgut meet in the middle of the embryo.

10–12 h after fertilization: Tissue differentiation and dorsal closure (morphological stages 13–15)

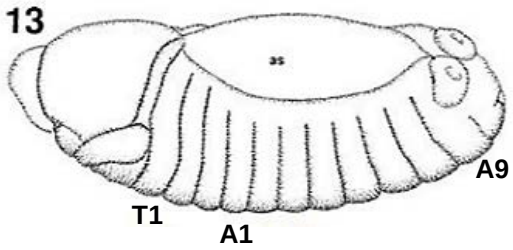


Figure 1.8 *Drosophila melanogaster* embryo, lateral view, morphological stage 13.
as: amnioserosa, T1: thoracic segment 1, A1-9: abdominal segments 1 and 9.

Segments of the germ band separate into 12 domains, with T1–3 making up the future thorax and A1–9 forming the abdomen (Figure 1.8). Head involution (10h 20min) and dorsal closure

(11h) happens at this stage, representing the last morphogenetic movements of *Drosophila* embryogenesis.

12–22 h after fertilization: End of embryogenesis (morphological stages 15–17)

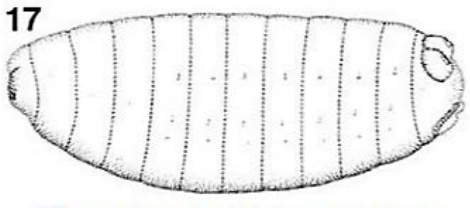


Figure 1.9 *Drosophila melanogaster* embryo, lateral view, morphological stage 17: final larval morphology.

After closure of the dorsal opening and completion of head involution, an outer layer of larva, called the cuticle, begins to form. This is the protective coating of larva required for structural integrity (Ostrowski, Dierick, & Bejsovec, 2002). At around 21–22 h, the larva hatches.

The entire process of embryonic development described above, from egg to larva, is tightly controlled by the coordinated temporal and spatial expression of specific genes detailed below.

1.6 Genetic Control of Patterning During Embryonic Development in *Drosophila melanogaster*

During development, pattern formation describes the process by which equivalent cells in a developing embryo progressively adopt distinctly different complex forms and functions, generating the overall body plan of the organism in three major axes: anteroposterior (A-P), dorsoventral (D-V) and left-right (L-R). In early embryos, each cell senses and responds to its position along a morphogen gradient setting the very first pattern. This is further refined by short distance cell to cell communications through cell signaling pathways, resulting in fields of cells with the same properties. Many animal body plans are composed of these morphologically similar metameric units of cells called segments. Molecular control of this process is highly conserved in invertebrates and very well characterized in *Drosophila* (Alberts et al., 2002).

The *Drosophila* body plan consists of 15 segments: three head (Mn, Mx, Lab), three thoracic (T1-T3), and nine abdominal (A1-A9) (Hartenstein, 1993). These are formed by two distinct segmental register events: an early para-segmental register during gastrulation, and a late segmental register during germ band retraction. This process is controlled by the coordinated expression of five classes of genes which are organized in a hierarchy and grouped based on the phenotype of loss of function alleles in these genes (S. DiNardo & O'Farrell, 1987; Martinez-Arias & Lawrence, 1985; Nüsslein-volhard & Wieschaus, 1980; Nüsslein-Volhard, Wieschaus, & Kluding, 1984; Schüpbach & Wieschaus, 1986). The first four classes of genes establish the patterning and numbers of segments, and the last class determines the identity of each segment (Figure 1.10).

First class of genes is called maternal coordinate genes, mRNA products of which are maternally deposited in the egg and produce maternal morphogens such as *bicoid* (bcd). bcd diffuses from the anterior (future location of head) region activating *hunchback* (Hb) and inhibiting *caudal* (Cad). Conversely, *nanos* (nos) diffuses from the posterior end (future location of tail) and inhibits Hb. The gradients of these morphogens allow establishment of the A-P axis by differential regulation of transcription factors such as *giant* (Gt), *krüppel* (Kr) and *knirps* (Kni) (Johnston & Nüsslein-Volhard, 1992). Patterning of the dorsal-ventral axis in early embryo is controlled by the Rel/NFκB-family transcription factor *dorsal*, which activates genes in a concentration dependent manner. The loss of function phenotype of this first class of genes produce eggs which do not develop correctly. For example, *bcd* mutant mothers produce embryos with abdominal segments replacing head and thoracic segments, and a duplication of posterior most structures, while *dorsal* mutants have the entire cuticle at the dorsal side. (Driever & Nüsslein-Volhard, 1988).

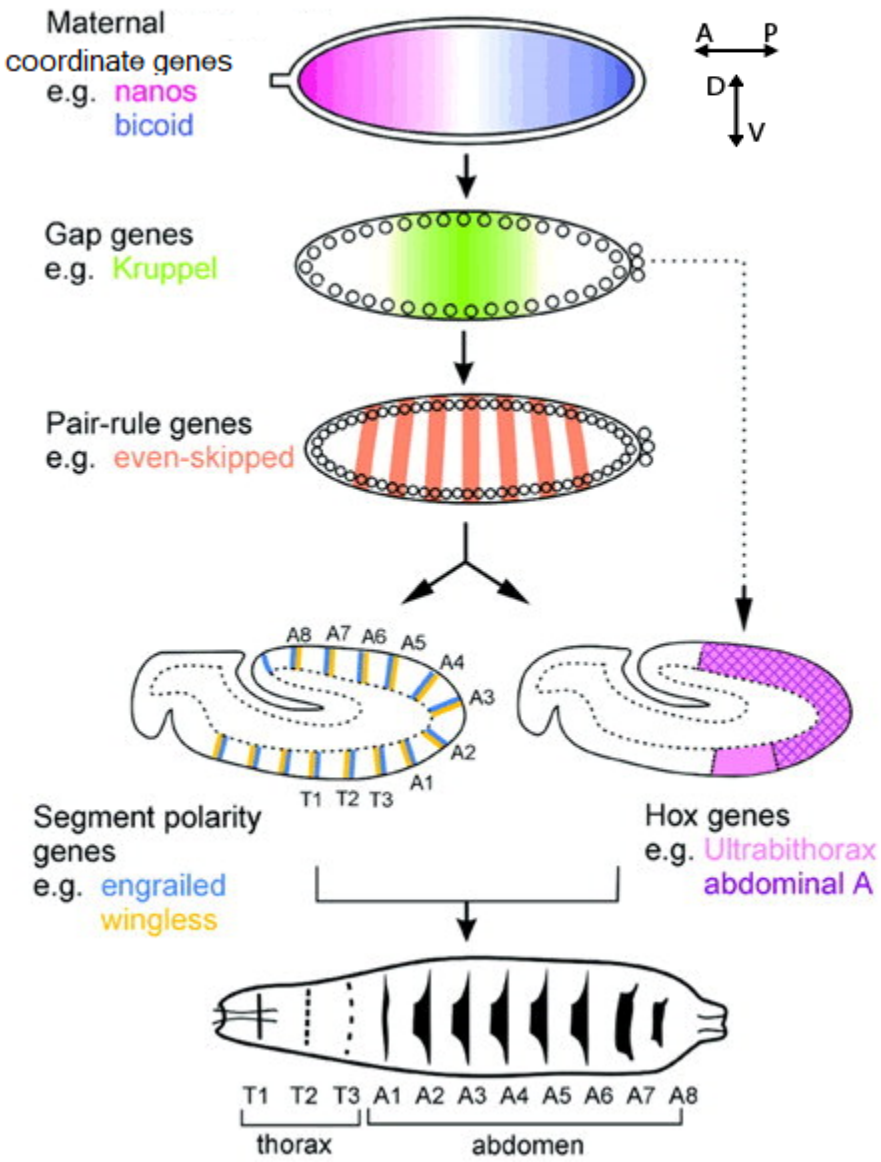


Figure 1.10 Patterning of A-P axis in *Drosophila* embryonic development and expression pattern of segmentation genes.

Image adapted from Sanson, 2001 and used with permission.

The maternal coordinate genes regulate the expression of second class of genes called gap genes. They are among the first zygotically transcribed genes being expressed in large contiguous regions of up to 7 units along the A-P axis in syncytium. Mutants of these genes causes deletions of several segments generating gaps in larval cuticle (Nüsslein-Volhard et al., 1984).

The gap genes in turn regulate the expression of the third class of segmentation genes called pair-rule genes, such as *even skipped* (eve), *odd skipped* (odd) and *fushi tarazu* (ftz). Expression of these genes is precisely limited to 7 defined stripes, called para-segmental domains, perpendicular to the A-P axis which are clearly visible in cellular blastoderm. Loss of function mutations of these genes result in deletions in alternating segments. For example, the odd numbered para-segments are deleted in *eve* mutants while even numbered segments are deleted in *ftz* mutants (Jürgens, Wieschaus, Nüsslein-Volhard, & Kluding, 1984; Nüsslein-Volhard et al., 1984).

The pair-rule genes in turn regulate the expression pattern of the segment polarity genes, *wingless* (wg) and *engrailed* (en) which establish cell fates and AP polarity within each parasegment through cell to cell signaling. These genes are expressed in 14 stripes which are the precursors of final body segments. During gastrulation and early germ band extension, *en* and *wg* are expressed at the anterior and posterior ends of each parasegment, respectively, and they set the boundaries between parasegments. Right after the completion of germ band extension and during germ band retraction, parasegmental furrows disappear. Next very deep grooves, which will be the segmental boundaries, start forming halfway down the length of each parasegment (Gilbert, 2000), leaving *en* and *wg* expressing cells next to each other at the posterior end of every segment. The loss of function phenotype for these genes is the mirror-image transformation of all segments.

Once segmentation is established, the fifth group of genes, selector or segment identity genes (*Hox* genes), determine the identity of each segment which results in formation of particular anatomical structures from each segment. These genes are also activated by pair-rule genes (Sanson, 2001). *Hox* mutants, either gain of function or loss of function, show homeotic transformation of one segment into the likeness of another, causing development of body structures in wrong place (Mallo & Alonso, 2013).

All these patterning and segmentation events of embryogenesis eventually lead to an alternating pattern of denticle formation on the ventral epidermis of larva, separated by regions of naked cuticle (A. Bejsovec & Martinez Arias, 1991). This specific pattern of denticles is indeed determined by the members of different signaling pathways expressed in underlying epidermal cells (Figure 1.12) (François Payre, 2004). The connection between these structures, and signaling pathways controlling embryonic patterning and polarity made it possible to understand the roles of above-mentioned five group of genes identified in the Heidelberg screen for pattern mutants (Wieschaus & Nüsslein-Volhard, 2016) (Figure 1.11).

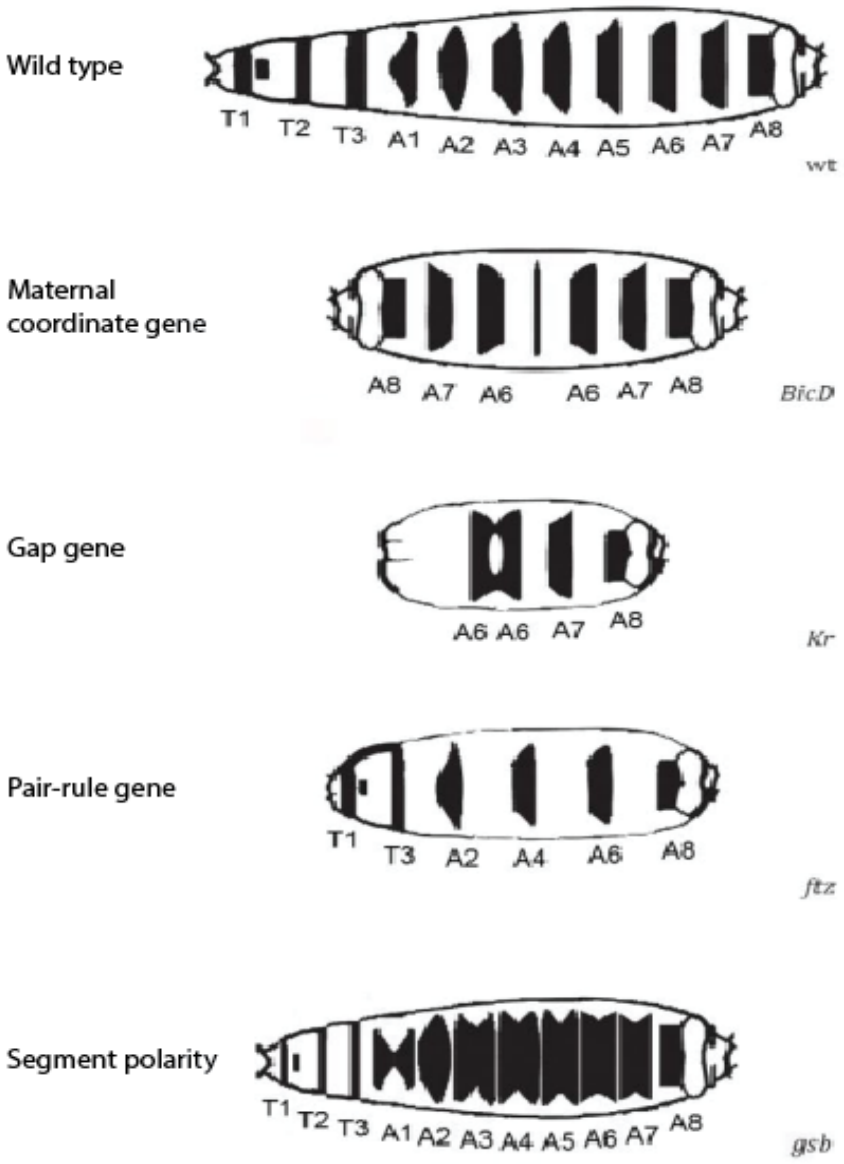


Figure 1.11 Graphic representations of wild-type and mutant first instar larva cuticular phenotypes observed in loss of function mutants for each class of segmentation genes.

Bicaudal (Bic-D), Kruppel (Kr), fushi tarazu (ftz), gooseberry (gsb). Images adapted from Wieschaus & Nusslein-Volhard (2016) and Sivanantharajah (2013) and used with permission.

1.7 Signaling Pathways in Patterning of *Drosophila* Embryo

Cells translate extracellular signals into functional changes through a cascade of molecular events called signaling pathways, which regulate processes like apoptosis, cell proliferation, cytoskeletal reorganization, cell polarity, adhesion, and cell migration (Shenolikar, 2007). During

development, these pathways usually act as developmental switches and push development forward by regulating transcriptional programs, which determine cell fates (Perrimon, Pitsouli, & Shilo, 2012). There are 11 main classes of developmental signaling pathways, which are listed in Table 1.3 (Perrimon et al., 2012). Development of embryonic patterning is regulated by the interplay of these pathways with each other. This interplay can happen in three ways: activation of one pathway can induce another pathway, e.g. hedgehog and wingless pathways during patterning (Sanson, 2001), or they can work in concert to control the activity of same target, e.g. EGF and notch pathways controlling the expression of the same transcription factor (Flores et al., 2000), or one pathway can interfere with the activity of a component of the other, e.g. hedgehog pathway leading to degradation of BMP receptors in TGFβ signaling pathway (L. Xia et al., 2010) and decapentaplegic induced inhibition of EGFR signaling during dorsal closure (Lusk, Lam, & Tolwinski, 2017).

Table 1.3 Developmental signaling pathways, their receptors-ligands-outputs.

Signaling pathway	**Receptor**	**Ligand**	**Output**
Wnt/Wg	Frizzled, dFrizzled2	Wg/Wnt	Patterning, growth, PCP (β-catenin independent)
Hh	Patched	Hh	Patterning, growth
TGFβ	Thickveins Saxophone	Dpp/TGFβ	Patterning, growth
RTK	EGFR	Spitz, Gurken, Keren, Vein	Patterning, morphogenesis
	FGFR (Breathless, Heartless)	Branchless, Thisbe, Pyramus	Patterning, morphogenesis, migration
	InR	dIlp1-dIlp7	Growth, metabolism, aging
	PDGF/VEGF receptor (PVR)	Pvf1-3	Morphogenesis, migration
	Torso	Trunk, PTTH	Patterning, metamorphosis
	dALK	Jelly belly	Growth on starvation (CNS)

	Sevenless	Boss	Patterning, cell-fate specification
Notch	Notch	Delta, Serrate	Patterning, lateral inhibition, cell-fate specification
Hippo	Fat	Dachsous	Growth, PCP
NF-κB	Toll	Spatzle	Patterning, innate immunity
JAK/STAT	Domeless	Unpaired1-3	Patterning, innate immunity
JNK	Eiger/TNF	Wengen	Migration, patterning, innate immunity
Nuclear receptors	EcRA, EcRB	Ecdysone	Patterning, growth, metabolism

Adapted from Perrimon (2012) and used with permission.

In the following sections, I have briefly discussed the mechanistic details of four of these signaling pathways that are implicated in this thesis: Transforming growth factor beta (TGF-β) signaling pathway, Epidermal growth factor (EGF) signaling pathway, Fibroblast growth factor (FGF) signaling pathway, and Wnt/Wingless signaling pathway. In addition, I have briefly discussed the activity of the gene *discontinuous actin hexagon* (dah) and its roles in cellularization.

TGF-β signaling pathway

The transforming growth factor beta (TGFβ) pathway is one of the most conserved signaling pathways in multicellular eukaryotes (De Robertis, 2008). It is known to be involved in many cellular processes in both the adult organism and the developing embryo, including cell proliferation, differentiation, apoptosis, and cellular homeostasis (De Robertis, 2008). Despite the wide range of processes that this pathway controls, the signaling mechanism is relatively simple. In the canonical signaling cascade, an extracellular ligand binds to a type II receptor which is a serine/threonine kinase. The ligand bound type II receptor phosphorylates and activates a type I receptor, which in turn phosphorylates response-Smads (R-Smad). After binding of phosphorylated R-Smad to a Co-Smad, this complex translocates to the nucleus and alters the

expression of genes containing TGF-β signaling specific response elements (Parker, Stathakis, & Arora, 2003).

In *Drosophila,* TGF-β signaling family can be divided into two branches based on the type of ligands: Bone morphogenic protein (BMP) and activin signaling pathways. BMP family ligands include *decapentaplegic* (dpp), *screw* (scw), and *glass bottom boat* (gbb), which bind to one of the two BMP-specific Type I receptors, *thickveins* (Tkv) or *saxophone* (Sax), and lead to activation of BMP-specific R-Smad called *mothers against decapentaplegic* (Mad). Activin family ligands include *activin-P* (Act-P) and *dawdle* (Daw), which bind to Activin-specific Type I receptor *baboon* (Babo) and activates Activin R-Smad called *smad on X* (Smox). Type II receptors, *pun*t (Put) and *wishful thinking* (Wit), and Co-Smad *medea* (Med) are common components of both pathways.

BMP signaling is involved in dorso-ventral patterning of early embryo through a morphogen gradient based mechanism of its ligands, *dpp* and *scw*, which interact with *tkv* and *sax*, respectively (O'Connor, 2005). Activin signaling through type I receptor *babo* has been implicated in neuronal remodeling during metamorphosis of larva (Parker et al., 2003). Both pathways are regulated by different mechanisms such as degradation of type-I receptors and R-Smads, dephosphorylation of type I receptors by phosphatases like PP1c, and complex formation of inhibitory Smads, i.e. Dad (daughters against dpp), with R-Smads (Itoh & ten Dijke, 2007). It has also been shown that *inwardly rectifying potassium channel 2* (irk2) protein is involved in wing-disc patterning through *dpp* signaling (Dahal et al., 2012).

EGF signaling pathway

EGF is another well-known conserved pathway involved in cell growth, proliferation and regulation of apoptosis, and it is no surprise that it has fundamental roles in *Drosophila* embryogenesis. There are several different EGF ligands which has differential tissue expression patterns (Malartre, 2016). In addition, it has been suggested that each ligand can activate signal

transduction to a different degree, which leads tissues receiving the EGF signal to differentiate according to the strength of signaling (Lusk et al., 2017). A major outcome of EGF signaling is activation of Ras, mediated through a phosphorylation cascade. EGF receptor (EGFR) is a receptor tyrosine kinase which auto phosphorylates itself upon ligand binding. The phosphorylated tyrosine residues on EGFR in turn interacts with the SH2 domain of *downstream of receptor kinase* (drk) localizing to the plasma membrane. Activated *drk* then binds to *son of Sevenless* (sos) which is a guanine exchange factor. *sos* triggers activation of Ras by promoting GTP binding. Activated Ras localizes *polehole* (D-Raf) to the plasma membrane and this serine-threonine kinase activates mitogen activated kinase (MAPK) pathway. Eventually, *rolled* (ERK kinase homolog in *Drosophila*) phosphorylates and activates transcription factors such as *pointed* and *yan* (Rebay, 2002).

EGF signaling has major roles in differentiation of neuroectoderm (Kim, Jung, Kim, & Jeon, 2008). Moreover, through its interplay with *wg* and *hh* signaling pathways, it regulates ventral epidermal patterning. There, it induces denticle formation through transcription factor *shaven baby* (svb) while inhibition of EGF signaling by *wg* signaling leads cells to develop a smooth cuticle fate (FranÇois Payre, Vincent, & Carreno, 1999; Urban, 2004). The EGF pathway has also been implicated in specification of muscle precursors and tracheal invagination, and dorsal midline patterning during embryogenesis (Lusk et al., 2017)

The EGF signaling pathway can be inactivated by multiple mechanisms which include ligand dissociation, ligand mediated degradation, sequestration of downstream effectors and dephosphorylation of the receptor and MAPKs (Ceresa & Vanlandingham, 2008). The transcriptional targets of the pathway itself can also work as positive and negative feedback loops to amplify and dampen the signal (Rebay, 2002).

FGF signaling pathway

FGF signaling is another conserved pathway which controls cell differentiation, proliferation and movements during the development of multicellular animals (De Robertis, 2008). There are two different FGF receptors in *Drosophila*, *heartless* (htl) and *breathless* (btl) which signal through a similar intracellular cascade. Upon ligand binding and stabilization by heparan sulfate proteoglycans, these activated receptor tyrosine kinases auto- and trans phosphorylate their kinase domains, and in turn phosphorylate the adaptor protein *stumps/downstream of FGF* (dof) (Lin, Buff, Perrimon, & Michelson, 1999; Petit, Nussbaumer, Dossenbach, & Affolter, 2004). Through three different cascades, including Csw/Shp2, Grb2/Drk and Src64B proteins, *dof* leads to activation of MAPK pathway. It has been shown that Src64B mediates activation of D-Raf in response to FGF signaling (F. Xia et al., 2008). The end result of these cascades is induction of gene expression changes to execute proliferative and anti-apoptotic responses (Csiszar, Vogelsang, Beug, & Leptin, 2010; F. Xia et al., 2008). FGF signaling is negatively regulated by the classical negative feedback loop of RTK signaling which is mediated by *sprouty* and dual specificity phosphatases that targets MAPKs (C. Li, Scott, Hatch, Tian, & Mansour, 2007).

FGF signaling is implicated in regulation of cell migration and morphogenetic movements during gastrulation in *Drosophila* embryogenesis. It controls gene expression that leads to differentiation of cells, it regulates cell shape and cell–cell interactions during mesoderm layer formation, caudal visceral muscle (CVM) formation, tracheal morphogenesis and glia differentiation (Muha & Müller, 2013).

Wnt/Wingless signaling pathway

Wnt/Wingless signaling pathway is one of the most conserved pathways among bilaterians, and regulates patterning, cell fate specification and morphogenesis of tissue layers (De Robertis, 2008; Graba, Aragnol, Rothbächer, & Pradel, 2013). There is only one ligand of this pathway in

Drosophila, called wingless (wg), which is a segment polarity regulator. Wingless is a glycoprotein that acts as a morphogen over a long distance, especially in wing disc, and target cell responds to this morphogen in a concentration dependent manner (Perrimon et al., 2012).

Canonical wg pathway is the pathway that leads to accumulation of *armadillo* (arm), the β-catenin homolog in *Drosophila*, in cytoplasm and its eventual translocation into the nucleus where it serves as a transcriptional co-activator of the transcription factor TCF/pangolin. Without wg signaling, *arm* is degraded by a destruction complex comprised of Axin, adenomatosis polyposis coli (Apc), zeste white 3/glycogen synthase kinase (Zw3/GSK) and casein kinase 1 (CK1) in wg responsive cells (Stamos & Weis, 2013). Phosphorylation of *arm* by CK1 and Zw3/GSK targets it for ubiquitination and subsequent degradation by the proteasomal machinery (Verheyen & Gottardi, 2010). In the absence of wg signaling, wg target genes are repressed by a corepressor g*roucho*, which binds to TCF/pangolin transcription factor (Levanon et al., 1998). When wg is present, the destruction complex function becomes disrupted. Wg first concentrates on the cell membrane through interaction with the modulator glypican dally. Then it binds to its receptor frizzled (fzd) and this leads the co-receptor arrow to cluster with fzd. Ligand binding and activation of fzd leads to polymerization of cytoplasmic phosphoproteins axin and dishevelled (dsh) at the membrane (Amy Bejsovec, 2018). First, phosphorylated axin binds to the cytoplasmic tail of arrow, and there it becomes dephosphorylated leading to a decrease in its stability and levels (Tacchelly-Benites et al., 2018). Then, dsh becomes activated via phosphorylation and inhibits Zw3/GSK activity of the destruction complex, which no longer targets *arm* to destruction (Gao & Chen, 2010). This allows *arm* to accumulate and eventually localize to the nucleus. There, it interacts with Tcf and directly displaces groucho from Tcf (Daniels & Weis, 2005). This recruits the transcription activation complex involving legless and pygopus to the corresponding DNA elements (Städeli & Basler, 2005). The transcriptional targets of this pathway include other

important segment polarity regulators *engrailed* (en) and *hedgehog* (hh) (Ramakrishnan & Cadigan, 2017).

The most well characterized interplay between different developmental pathways is the one between wg and hh signaling, through which segments of the embryo are established and patterned, generating the specific arrangement of denticles and naked cuticle on larval epidermis (Gilbert, 2000) (Figure 1.12). This patterning starts with the activation of *en* by pair-rule genes at the anterior boundary of each para-segment where concentrations of both eve and ftz is high. At the posterior boundary, where the pair-rule genes *sloppy paired* and *odd paired* are highly concentrated and *eve* and *ftz* are low, *en* is inhibited while *wg* gets activated (Swarup & Verheyen, 2012). The para-segmental borders are sharpened and maintained by the intercellular signaling between *en* and *wg* expressing cells at the boundaries (Stephen DiNardo, Heemskerk, Dougan, & O'Farrell, 1994). *En* cells secrete a morphogen called *hedgehog* which binds to the *patched* receptors on the nearby cells and releases the inhibitor, *smoothened*. This results in *wg* expression in these cells. Secreted wg morphogen then binds fzd and activates *wg* pathway in *en* cells, which results in *en* and *hh* expression in these cells. This positive feedback loop maintains the identity of *wg* and *hh (en)* secreting cells. Then, morphogen gradients of *wg* and *hh* across each para-segment specifies the anterior-posterior boundaries of future segments (Amy Bejsovec, 2018; Gilbert, 2000). The cells, which receives *wg* morphogen, produce naked cuticle in larva because *wg* causes repression of *shaven baby* (svb) gene. On the other hand, denticle producing cells express svb which is required to direct denticle formation cell autonomously (A. Bejsovec & Martinez Arias, 1991; FranÇois Payre et al., 1999; Swarup & Verheyen, 2012).

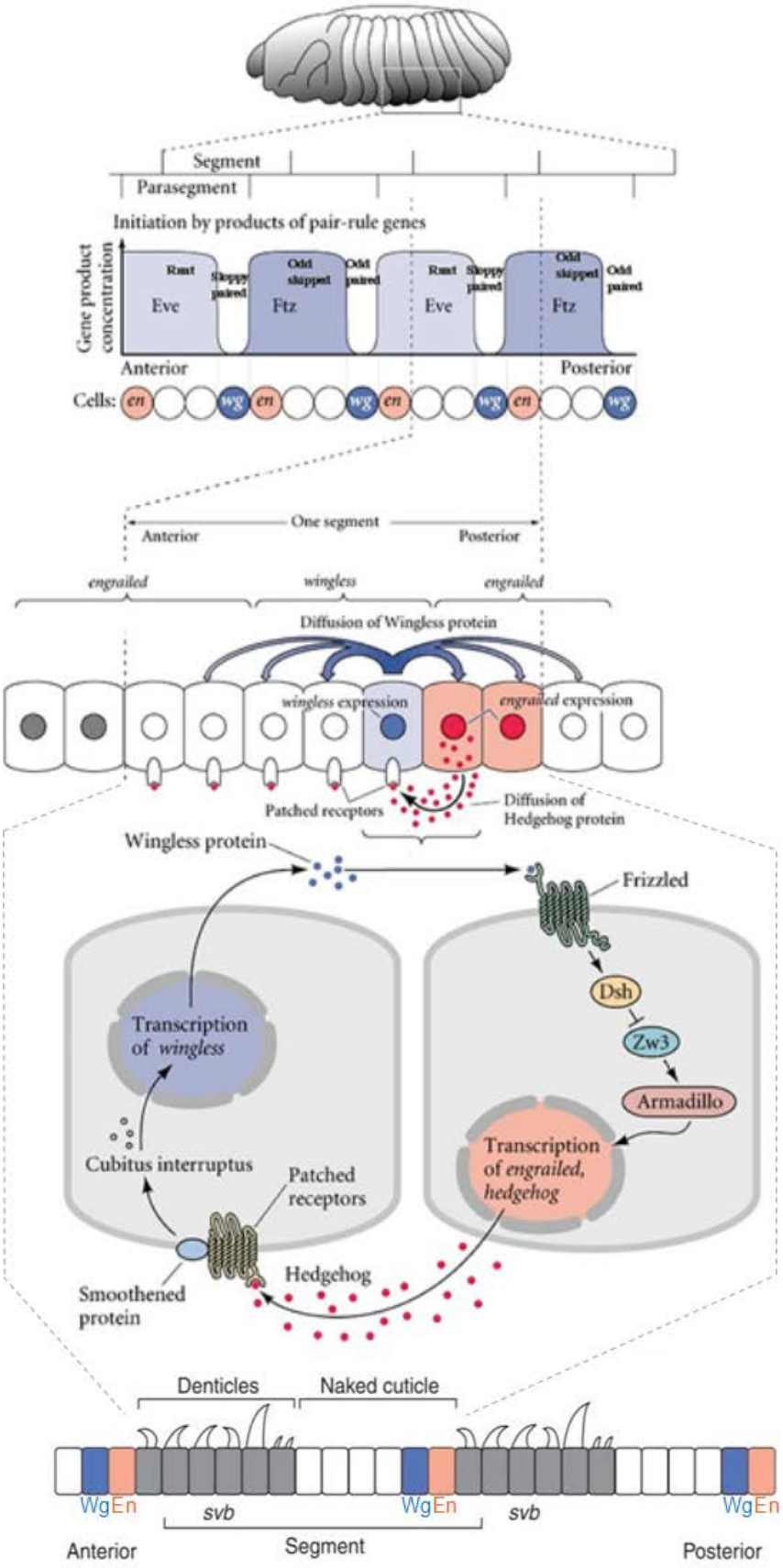

Segment
Parasegment
Initiation by products of pair-rule genes
Gene product concentration
Rnt
Sloppy paired
Odd skipped
Odd paired
Eve
Ftz
Anterior
Posterior
Cells:
en
wg
One segment
engrailed
wingless
Diffusion of Wingless protein
wingless expression
engrailed expression
Patched receptors
Diffusion of Hedgehog protein
Wingless protein
Frizzled
Dsh
Zw3
Armadillo
Transcription of wingless
Transcription of engrailed, hedgehog
Cubitus interruptus
Patched receptors
Hedgehog
Smoothened protein
Denticles
Naked cuticle
WgEn
svb
Segment

Figure 1.12 Segment polarity and denticle formation in embryo though *wg* and *hh* pathways.
The interplay between the w*g* and h*h* (en) signaling pathways first sets the parasegment boundaries and then directs the segmental borders, establishing the denticle pattern and naked cuticle at the end of embryogenesis. Eve: even skipped, Ftz: fushi tarazu, Dsh: disheveled, Zw3: Zeste-white3, Svb: Shavenbaby. Adapted from Gilbert (2000) and Swarup & Verheyen (2012).

The regulation of *wg* signaling upstream of the pathway itself is achieved by regulation of *arm* through targeted proteosomal degradation mentioned above. In addition, the *wg* pathway regulates itself in a negative feedback loop by inducing the expression of an antagonist, called naked cuticle (nkd). In absence of *nkd*, embryos produce more *wg,* which in turn inhibits denticle formation (W. Zeng et al., 2000).

Discontinuous actin hexagon (*Dah*) and cellularization

Dah is a maternally deposited protein. It gets recruited to the plasma membrane with the help of nuclear fallout (nuf) and colocalizes with actin on the membrane (Rothwell, Zhang, Zelano, Hsieh, & Sullivan, 1999). Dah is essential for cortical furrow formation during cellularization of embryos. Dah is incrementally phosphorylated during the nuclear divisions of the early embryo and the degree of phosphorylation peaks throughout cellularization, where Dah plays an essential functional role. *In vitro* phosphorylation experiments showed that casein kinase 1 can hyperphosphorylate dah. After cellularization, Dah is dephosphorylated in the embryo by an unknown mechanism (C. X. Zhang, Rothwell, Sullivan, & Hsieh, 2000).

1.8 Phospho-regulation of Developmental Pathways

Reversible phosphorylation of proteins is one of the mechanisms that cells utilize to precisely control the state and the amplitude of signaling pathways. Many of the components of the developmental signaling pathways mentioned above are either inhibited, activated or stabilized through phosphorylation and dephosphorylation events. Indeed, several components of these pathways themselves are kinases, i.e. Sax, Babo, EGFRs, FGFRs, MAPKs, Src64B, CK1, and

Zw3/GSK, which phosphorylate either themselves or downstream components of the pathways. There are a few phosphatases that has been identified as part of these pathways, including PP1, PP2A, PP4 and MAPK phosphatases (Hall, Pradhan-Sundd, Samnani, & Verheyen, 2017; McEwen, 2005; Swarup, Pradhan-Sundd, & Verheyen, 2015; Zhao, Wang, Chi, Lan, & Su, 2017). In most cases, there are loss of function phenotypes of these phosphatases, but we still do not know the basis of substrate specificity, when and how they are activated or inhibited, and the relevance of specific dephosphorylation events on the phenotype.

1.9 Prion-like proteins as regulators normal cellular processes

As mentioned in section 1.1, a conformational change of prion-like proteins from monomers to self-sustaining oligomers is often accompanied by a change in the functional state of the protein, i.e. loss of the original function or gain of a new function. Because of the dominant nature of the oligomeric form, this process acts like a conformational-functional switch. It has been previously shown that these types of protein switches are important regulators of several physiological processes, including adaptation to changing environmental conditions (Halfmann et al., 2012), immune response (Hou et al., 2011b), and memory formation (Majumdar et al., 2012). In all these processes, they either act as transcriptional regulators, translational regulators or signaling components at molecular level, leading to a global change in cellular response.

In the following sections, I briefly discuss at the molecular level some examples of such proteins and implications of their properties.

Prion-like proteins in transcriptional regulation

Many prion-like proteins, including Ure2, Swi1, Sfp1, Cyc8, and Mot3 in yeast (Alberti et al., 2009; Du, Park, Yu, Fan, & Li, 2008; Patel, Gavin-Smyth, & Liebman, 2009; Rogoza et al., 2010; R. Wickner, 1994) and GAGA factor in *Drosophila* (Tariq et al., 2013), are transcriptional regulators. URE2 represses nitrogen catabolism enzymes, when grown on a rich nitrogen source,

by sequestering Gln3p in the cytoplasm (Drillien, Aigle, & Lacroute, 1973). URE2 is also capable of aggregating into a prion form under limited nitrogen conditions and aggregates lose function, releasing catabolite repression. This creates a selective and epigenetically heritable advantage on poor nitrogen food sources (R. Wickner, 1994). Swi1 is a component of the ATP-dependent nucleosome remodeling complex SWI/SNF which regulates transcription of almost 6% of all genes and it can switch into a prion-like form, resulting in partial loss of SWI/SNF function (Du et al., 2008). Despite the growth limitations exhibited by this partial loss of function, the Swi1 prion-like form confers greater cell resistance to the fungicide benomyl, suggesting that the switch into a prion-like state may offer a necessary trade-off in desperate conditions (Alberti et al., 2009). Sfp1 is a global transcriptional regulator of ribosomal proteins and ribosome biogenesis genes in response to stress and nutrients. Sfp1 prion form accumulates in nucleus resulting in larger cell sizes and drug resistance (Rogoza et al., 2010). This form of the protein has a different transcriptional response than deletion of the gene itself, suggesting differential activity of separate states of the protein (Drozdova, Rogoza, Radchenko, Lipaeva, & Mironova, 2014). Cyc8, another yeast transcription factor, can also form a prion (Patel et al., 2009). Cyc8 functions in the Cyc8-Tup1 corepressor complex, which modifies expression of 7% of yeast genes in a wide variety of physiological processes including heme-regulated and catabolite repressed genes (Wong & Struhl, 2011). Mot3 is prion forming transcription factor of yeast responsible for repression of hypoxic genes under aerobic conditions (Alberti et al., 2009). The prion form of Mot3 leads to both loss and gain of function phenotypes, regulating multicellularity in response to nutrient deprivation (Holmes, Lancaster, Lindquist, & Halfmann, 2013a). *Drosophila* GAGA factor (GAF) is a transcription factor with roles in multiple processes, including regulation of gene expression, the structural organization of heterochromatin and chromatin remodeling (Adkins, Hagerman, & Georgel, 2006). It has been previously shown that GAF can form prion-like assemblies *in vitro* through its C-terminal Q rich domain (Tariq et al., 2013). These diverse functions require GAF to

interact with different partners in distinct multiprotein complexes and its conserved prion domain is implicated in complex formation with these partners (Tariq et al., 2013).

Many of these proteins operate at pivotal positions in determining how genotype is translated into phenotype by regulating both transcription and chromatin remodeling. It is suggested that by virtue of their prion-like conformational functional switching, once engaged such transcriptional regulators can maintain the altered transcriptional state following cell division in an epigenetic manner (Halfmann et al., 2012; Si, 2015; Soto, 2012). This protein-based inheritance can happen in several ways: creating diffusion barriers, as in the case of Ure2 restricting Gln3 localization to cytoplasm in non-prion form; or interacting with different protein complexes in non-prion and prion forms, as suggested in the case of Swi1 and GAF; or acting as hubs on DNA for reconstitution of repressor or activator complexes after cell division (Si, 2015). In yeast, the hypothesis that prions act as epigenetic determinants has been extensively studied and shown to be valid in several cases, as mentioned above. However, in multicellular eukaryotes, the situation in much less clear due to a lack of evidence at *in vivo* level.

Prion-like proteins in translational regulation

Two of the well-known examples of prion-like proteins, Sup35 and Orb2 (Drosophila homolog of CPEB), are translational regulators (Khan et al., 2015; Reed B Wickner et al., 2010). Sup35p is a subunit of the translation termination factor, and the normal function of its prion domain is in shortening the 3' polyA of mRNAs and thus in mRNA turnover (Cox, 1965). In its prion form, Sup35p loses its translation termination activity, resulting in increased readthrough of termination codons (R. Wickner, 1994). This readthrough of termination codons allows different set of proteins to be translated and provides a survival advantage under diverse conditions with an increase in phenotypic diversity (True et al., 2004).

Orb2 is Drosophila homolog of Aplysia and mammalian CPEB (cytoplasmic polyadenylation element-binding protein) which is an activity-dependent regulator of synaptic protein synthesis (Richter, 2007). Both the neuronal CPEB of Aplysia, and Orb2 exist in two different states in their respective neuronal contexts: a monomeric state and an amyloidogenic oligomeric state (Majumdar et al., 2012; Si et al., 2003, 2010; White-Grindley et al., 2014). It has been found that the conversion of monomers to oligomers is regulated by repeated stimulation with neurotransmitters (Majumdar et al., 2012; Si et al.,2010; White-Grindley et al., 2014). In their oligomeric states, both Aplysia CPEB and Orb2 retain their mRNA binding ability (Khan et al., 2015; Majumdar et al., 2012; Raveendra et al., 2013). More importantly, Orb2 oligomerization transforms it from a translational repressor to an activator and during this process there are also changes in its interacting partners (Khan et al., 2015). It is also suggested that the dominant nature of oligomeric Orb2 creates a self-sustaining active state of protein synthesis at the activated synapse and maintains memory over time (Khan et al., 2015). This has been supported by the evidence that facilitation of Orb2 aggregation allowed *Drosophila* to form memory in suboptimal conditions (L. Li et al., 2016).

In both examples, Sup35 and Orb2, cells either maintain their phenotypic or functional identity through a stable protein-based mechanism. As seen with the other examples, yeast provides an opportunity to follow functional outputs of prion-like protein switches both quantitatively (molecular) and qualitatively (phenotype). However, because of its unicellular nature, it does not give us information at a system level. *Drosophila* Orb2 is an excellent example of studying such proteins in *in vivo* context at a system level.

Prion-like proteins in signaling

Signaling pathways are dynamic information networks of the cell. While dynamic in nature they can carry the information in a stable manner based on interactions of many different proteins in

the cascade with each other. In view of the stable nature of prion-like switches and the diversity of protein-protein interactions of prion-like proteins in different states, it is interesting to see how such switches are used in signaling pathways. One of the examples of prion-like phenotype in yeast, [GAR+], is important for glucose signaling pathway in yeast. It is an unusual prion strain because it has at least two components: the plasma membrane proton pump Pma1, and the glucose signaling factor Std1. They form an oligomeric complex that is non-amyloid. In prion form, Pma1 and Std1 complex activates transcription at *HXT3* locus and allows yeast cells to circumvent glucose repression of alternative carbon substrates (Brown & Lindquist, 2009).

Another example of a prion-like protein in signaling is an immune response protein in humans, called the mitochondrial transmembrane antiviral signaling (MAVS) protein (Hou et al., 2011b). In response to viral infection, the RIG-1 protein, a cytoplasmic RNA helicase, recognizes viral double-stranded RNA (dsRNA). MAVS, present in the mitochondrial membrane, link RIG-1 to downstream signaling molecules. This in turn activates the transcription factors eRF3 and NF-kB, leading to expression of interferons (Hou et al., 2011b; Jacobs & Coyne, 2013; Xu et al., 2014). It has been shown that *in vitro* activation of RIG-1 causes its interaction with MAVS through CARD domains, and this leads aggregation of MAVS into functional detergent- and protease-resistant form which is capable of signal transduction (Xu et al., 2014).

1.10 Hypothesis and Objectives

As discussed in the previous section, in yeast and more recently in multicellular eukaryotes proteins with prion domains (PrDs), called prion-like proteins, act as regulatory switches controlling non-pathogenic heritable phenotypes. The fact that non-pathogenic prion-like proteins are distributed across multiple phyla raises the possibility that proteins with self-sustaining conformational states are part of an evolutionarily conserved regulatory mechanism involved in normal physiological functions. Although this is a testable hypothesis, functional prion-like proteins in multicellular eukaryotes have thus far been discovered entirely by chance. Therefore,

to determine how broadly these prion-like-protein-based biochemical switches are employed, we carried out a systematic screen in *Drosophila*, which led to the identification 5 potential prion-like-proteins. One of these candidate proteins is CG5830, which is an uncharacterized phosphatase in *Drosophila* with a homology to HAD subfamily SCPs.

The *in vitro* assays mentioned in section 1.2 suggested that CG5830 has prion-like properties, but most of these assays used a modified version of the protein in heterologous conditions. Each of these factors can influence protein behavior, including conformational state and aggregation. More importantly, none of these assays revealed whether its prion-like properties have any functional significance. Therefore, my two main objectives in this thesis are:

1: Determine the function of CG5830 in *Drosophila melanogaster*.

a) Determine if there is any functional orthology of CG5830 to mammalian SCPs

b) Determine the phenotypic effect of CG5830 null-mutation

c) Determine expression pattern of CG5830 protein across at different developmental stages

d) Determine what protein complexes CG5830 forms in the cell and its putative substrates

e) Determine the enzymatic activity of CG5830 on its putative substrates

2: Determine whether the prion-like properties of CG5830 contribute to the biochemical, cellular, and/or physiological functions.

a) Determine the prion-domain of CG5830

b) Determine the effect of prion-domain on enzymatic activity of CG5830

c) Characterize non-prion and prion form of endogenous CG5830 at structural and functional levels.

Chapter II: Materials and Methods

Antibodies: The anti-CG5830 polyclonal guinea pig IgG (#2677) was raised against full-length recombinant CG5830 protein and affinity purified against the purified recombinant antigens. The following antibodies were obtained from commercial sources; anti-HA HRP-linked monoclonal (3F10, Roche), anti-Myc 9B11 monoclonal (#2276, Cell signaling Technology), anti-GFP polyclonal (#598, MBL International), RNA pol II CTD phospho-Ser5 antibody monoclonal (3E8, Active motif), anti-Rbp1 polyclonal (N20, Santa Cruz), anti-Tubulin monoclonal (mouse, Molecular Probes), anti-Oligomer A11 antibody (AHB0052, Invitrogen), Anti-fiber OC (AB2286, EMD Milipore), anti-mouse-HRP linked (#7076S, Cell signaling Technology), anti-guinea pig HRP linked secondary antibody (#7076, Cell signaling Technology), anti-Rat HRP linked (AB97057, Abcam) and anti-rabbit HRP linked secondary antibody (A7289, Sigma).

For immunostaining experiments, the following antibodies were used: Anti-HA monoclonal (3F10, Sigma), anti-Wingless monoclonal (4D4, DSHB), anti-Engrailed monoclonal (4D9, DSHB), anti-Even-skipped monoclonal (2B8, DSHB), anti-Dorsal monoclonal (7A4, DSHB), anti-Discs-large monoclonal (4F3, DSHB), anti-Mouse 488 (A11001, Invitrogen), anti-mouse 594 (A11032, Invitrogen).

Plasmids and clones: For all protein-protein interaction studies in S2 cells the open reading frames were cloned into TopoD vector (Invitrogen) and transferred into the following vectors obtained from *Drosophila* Genome Research Center (DGRC) using the Gateway system; PAC 5.1 C-GFP, PAC 5.1 C-6XMyc, PMT C-3XHA-FLAG. For *in vitro* protein purification the genes were cloned in pDEST42-6xHis vector (Invitrogen) using the Gateway cloning system. For genomic tagging experiments, homology arms were cloned in pHD-dsRed-Attp, and guide RNAs were cloned into pBFv-U6.2 and pBFv-U6.2B.

Generation of *CG5830* mutant flies: gRNAs for *CG5830* (target sequence "GACAACACCGCCACCATTACCGG" and "GCAAAAGACGTAACATCAGTGGG") were cloned into vector pU6.2 and vector pU6.2B, respectively. They were then injected into attp1 line (BDSC#8621). gRNA flies and nos-cas9 flies were crossed to obtain founders. KO were screened and confirmed by PCR and sequencing. Then KO flies were backcrossed to Canton-S background.

Generation of *CG5830-3xHA* and *CG5830-eGFP* flies: Tagged flies were prepared as detailed in FlyCrispr (Bier, Harrison, O'connor-Giles, & Wildonger, 2018). Briefly, gRNA for *CG5830* (target sequence "TGCAAAAGACGTAACATCAGTGG") were cloned into vector pU6.2 and homology arms (5'arm: 1kb region before the stop codon and 3'arm: 1kb region after the stop codon of the gene) were cloned into pHD-dsRed-3xHA-Attp or pHD-dsRed-eGFP-Attp. A mixture of gRNA and homology arm clones (100ng/ul:500ng/ul) were then injected into nos-cas9 flies. dsRed positive flies collected and balanced for the third chromosome. After confirmation of the tag insertion into the correct position, flies were crossed to transposase line (BDSC#32072) to remove dsRed marker. After second round of screening, flies carrying the transgene were confirmed by PCR and sequencing, and homozygous stocks are prepared.

Embryo counts: For embryo-hatching counts, homozygous mutant fly crosses and control wild-type (CantonS) crosses were set up in cages (60mm). Hatching of thousand embryos was counted for each cross. Embryo hatching rate was calculated by normalization of hatch counts of each cross with wild type cross and represented as percentages.

For sterility experiments, homozygous mutant fly crosses with each other and with wild type (as 5 females and 5 males of same age) were set up in cages (30mm) in triplicates. Embryos laid by each cross were counted each day in a total of 7 days. Embryos laid by per female per day

calculated for each set of cages. Statistical analysis was performed by one-way ANOVA for multiple groups and data was plotted as mean ± SEM.

Immunohistochemistry: Embryos were prepared as in (Narasimha & Brown, 2006). Briefly, flies were allowed to lay eggs on apple juice/agar plates with yeast paste for up to 2 hours and aged to 4h, 7h, 10, 13h, 16h. Embryos were collected in water using a paintbrush, then dechorionated for 5 minutes in 50% bleach in 1X PBS. Embryos were fixed for 20 minutes in 1:1 heptane to 4% formaldehyde in 1X PBS. Embryos were then devitellinized by vortexing in 1:1 heptane to methanol or ethanol to preserve actin. Embryos were then washed in methanol or ethanol followed by 0.1% Triton-X or 0.1% Tween20 in 1X PBS, followed by incubation in blocking buffer (5% normal goat serum diluted in 0.1% Triton-X or 0.1% Tween20 in 1X PBS) for 2 hours. Embryos were incubated with primary antibody overnight at 4°C, washed in 0.1% Triton-X or 0.1% Tween20 in 1X PBS, then incubated with secondary antibody overnight at 4°C. Embryos were mounted in Vectashield mounting medium (Vector Laboratories). Images were acquired on a Zeiss LSM 700 confocal microscope and processed using Fiji (https://fiji.sc/) as follows: maximum intensity projections were created, filtered with gaussian blur (sigma: 1). For engrailed, wingless, even-skipped and dorsal staining, images were also thresholded to highlight cells expressing them, and eliminate background noise.

Live imaging of *CG5830-GFP* embryos: Flies were allowed to lay eggs on apple juice/agar plates with yeast paste for up to 2 hours. Embryos were collected in water using a paintbrush, then dechorionated for 5 minutes in 50% bleach in 1X PBS. After washing in water, they were collected into 1X PBS. Timelapses of embryo development were acquired on a Perkin Elmer Ultraview Vox Spinning Disc microscope with a variety of conditions. Embryos were immobilized in a custom made microfluidic device which allowed for trapping and hydration of the embryos with 1X PBS. Images were acquired with a 40x C-Apochromat (NA 1.2, Zeiss), a 63x Plan-Apochromat (NA 1.4, Zeiss), or a 20x Plan-Apochromat (NA 0.8, Zeiss) objective onto a C9100-

23B EMCCD (Hamamatsu) camera. CG5830-GFP was excited at 488nm and its fluorescence was filtered through a 525-50 bandpass filter, while H2A-RFP was excited at 561nm and its fluorescence was filtered through a 615-70 bandpass filter. Z-stacks were acquired with 2um steps for 20x and 0.3 or 0.5um steps for 40x and 63x images. The time between frames also varied from 30ms to several minutes depending on the experiment. Integration times were adjusted for image brightness and sample preservation but were typically 30-100ms. For high magnification movies, image tiling and stitching was also employed. An overlap of 10% between tiles was used. Images were stitched manually in Fiji (https://fiji.sc/) using the grid/collection stitching plugin (Preibisch, Saalfeld, & Tomancak, 2009).

Fluorescence Recovery after Photobleaching of full-length and truncated CG5830-GFP: For FRAP experiments, the data was acquired on the same system as the timelapses mentioned above. Here, images of full-length and truncated CG5830-GFP transfected S2 cells were acquired with an alpha-Plan Apochromat (NA 1.4, Zeiss) objective. The cells were plated in Mat-tek dishes with No. 1.5 coverslip bottoms. Five images were acquired at maximum speed (30ms integration) before bleaching. Bleaching of GFP occurred over user selected ROIs with the 488nm laser. The power and number of iterations for bleaching were set manually per cell to ensure proper bleaching. Recovery images were acquired for a rate and duration that was appropriate for the recovery. Typically, quickly recovering species were imaged at 30ms per frame (for 10s total), and slow species were imaged once every 1s (for 2min total).

CG5830-His protein purification and RNAPII phosphatase assay: Protein purification was performed in native conditions following IPTG induction. Both wild type and phosphatase dead CG5830 were cloned into a pDEST42 vector (Invitrogen) with a C-terminal histidine-tag and expressed in BL21 *E.coli* cells (Invitrogen). Bacteria were grown in 50ml 1X LB with 0.1mg/ml Ampicillin overnight at 37ºC. The next day 20ml was used to inoculate a liter of 1X LB with 0.1mg/ml Ampicillin and bacteria were grown at 25ºC to an OD^{600} of 0.5 to 0.6. Protein synthesis

was induced with 1mM IPTG and cells were grown an additional 4hours. Cells were subsequently harvested, washed with cold phosphate buffer saline (PBS) and lysed in cold native buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole and pH adjusted to 8.0 using NaOH). The 6xhis-tagged proteins were purified in Ni-NTA resin (Qiagen), eluted with 300 mM imidazole (Sigma) and dialyzed against phosphatase reaction buffer without detergents (50mM TrisHCl pH7.9, 10mM $MgCl_2$, 20% glycerol, 0.1mM EDTA, 5mM DTT, 20mM KCl) at 4°C for subsequent assays.

Purified RNAPII holoenzyme and TFIIH was kindly provided by the Conaway lab (Stowers Institute). Kinase reactions were performed as described (Serizawa, Conaway, & Conaway, 1992), and phosphorylated RNAPII was separated from the kinase with a centricon filter system (YM10, Sigma). Phosphorylated RNAPII was incubated with 20ug purified wild-type or phosphatase dead CG5830 in phosphatase reaction buffer containing 0.025% Tween80 for 45min at 30ºC. Reaction was stopped by the addition of laemmli sample buffer (Bio-Rad) and incubated at 70ºC for 10min. Samples were loaded onto a 4-12% SDS-Page gel (Invitrogen) and immunoblotted for phosphorylated anti-RNAPII phospho-Ser5 and anti-Rbp1 antibodies.

Western blot detection of tagged CG5830 proteins from embryos: Flies were allowed to lay eggs on apple juice/agar plates with yeast paste for time-points 0-30min, 30min-1, 1-2, 2-3, 3-4, 4-5, 5-7, 7-10 and 10-13-hour. Embryos were collected in water using a paintbrush, then dechorionated for 5 minutes in 50% bleach in 1X PBS. After washing in water, they were collected into 1X PBS and pelleted down at high speed. Embryos were lysed using a pestle with a cordless motor (VWR) in a buffer containing 150 mM NaCl, 1.0% NP-40, 50 mM Tris-Cl pH 8.0, 0.1% Triton X 100, 0.5% Na-Deoxycholate, 0.1% SDS in the presence of EDTA-free protease inhibitor (Roche), followed by rotating at 4ºC for 20 min. The lysates were centrifuged at 4500 rpm, 4ºC and the cleared supernatant was collected in fresh 1.5 mL Eppendorf tube. Pellet was dissolved in the same buffer by rotating at 4ºC for 20 min. Total of 20ug protein for each sample was boiled

with 2X laemmli sample buffer, resovled on a 4-12% SDS-polyacrylamide gel (Invitrogen), and immunoblotted with anti-HA, anti-GFP and anti-CG5830 antibodies.

Generation of stable S2 lines: S2 cells were maintained at 25°C in SFX insect culture medium (Fisher), containing 1X penicillin/streptomycin (Invitrogen). Cells were transfected with Effectene (QIAGEN), according to the manufacturer's protocol. Stable cell lines were generated by cotransfection of pMT promoter driven C-terminally 3XHA/3XFlag tagged wild-type or phosphatase dead CG5830 with pCoHygro (Invitrogen). Selection was carried out for a month in medium containing 0.25 mg/ml hygromycin (Invitrogen).

Protein purification from stable S2 cell lines and MudPit analysis: Three independent purification experiments were performed on parental S2 cell line, stable cell lines of WT-CG5830-3XHA/3XFlag and PD-CG5830-3XHA/3XFlag with a protocol kindly provided by Tamaki Suganuma (Stowers Institute). All three cell lines were grown in two-liters of SFX insect culture medium with 1X penicillin/streptomycin and induced with 0.7uM $CuSO_4$ for 36h. Cells were pelleted by centrifuging at 2000 rpm for 10min at 4ºC. After washing with ice cold 1X PBS, cells were lysed in 5X volumes of extraction buffer I (1% NP40,10mM HEPES pH 7.9, 1.5mM $MgCl_2$, 10mM KCl, 1mM PMSF, 1mM DTT, EDTA-free protease inhibitor (Roche)). Nuclei were pelleted by centrifuging at 5000 rpm for 5min and cytoplasmic supernatant was moved to new tubes. Nuclei were lysed in 3X volumes of extraction buffer II (20mM HEPES pH 7.9, 25% (v/v) glycerol, 420 mM NaCl, 1.5mM $MgCl_2$, 0.2mM EDTA, 1mM PMSF, 1mM DTT, EDTA-free protease inhibitor (Roche)) for an hour, vortexed for 15sec, and centrifuged at 14000 rpm at 4ºC. Both nuclear and cytoplasmic fractions were ultra-centrifuged at 40K for an hour to remove DNA. Supernatants were collected, and salt concentration was adjusted to 300mM NaCl before immunoprecipitation. For anti-FLAG purifications, approximately 200 mg of protein extract were bound to 120 µl M2-agarose beads (Sigma) overnight at 4°C. The beads were washed three times for 10 min in wash buffer (10mM HEPES pH 7.9, 300mM NaCl, 1.5mM $MgCl_2$, 0.1% Triton X-100, 10mM KCl). The

complexes were eluted off the beads by competing with 120 µl of 1 mg/ml of triple FLAG peptide (Sigma) in elution buffer (10mM HEPES pH 7.9, 150mM NaCl, 1.5mM $MgCl_2$, 0.1% Triton X-100) for 1h 30 min at 4ºC. Affinity-purified complexes were resolved on denaturing polyacrylamide gels and visualized by silver staining. Following overnight TCA precipitation, CG5830 protein complexes were analyzed by MudPIT as previously described (Lee, Florens, Swanson, Washburn, & Workman, 2005). Precipitated pellets were solubilized in TRIS-HCl pH 8.5 and 8M Urea. TCEP (Tris(2-Carboxylethyl)-Phosphine Hydrochloride, Pierce) and CAM (Chloroacetamide, Sigma) were added to a final concentration of 5mM and 10mM, respectively. Protein suspensions were digested overnight at 37°C using Endoproteinase Lys-C at 1:50 wt/wt (Roche). Samples were brought to a final concentration of 2M urea and 2mM CaCl2 before performing a second overnight digestion at 37°C using Trypsin (Promega) at 1:100 wt/wt. Formic acid (5% final) was added to stop the reactions. Samples were loaded on split-triple-phase fused-silica micro-capillary columns (McDonald, et al. 2002) and placed in-line with linear ion trap mass spectrometers (LTQ, Thermo Scientific), coupled with quaternary Agilent 1100 HPLCs. Fully automated 10-step chromatography run (for a total of 20 hours) was carried out for each sample, as described in (Florens et al., 2006), enabling dynamic exclusion for 120 sec. The MS/MS datasets were searched using SEQUEST (Link et al., 1999) against a database of 37466 sequences, consisting of 18556 *D. melanogaster* non-redundant proteins (downloaded from NCBI on 2012-03-08), 177 usual contaminants (such as human keratins, IgGs, and proteolytic enzymes), and 18733 randomized amino acid sequences derived from each non-redundant protein entry to estimate false discovery rates. Peptide/spectrum matches were sorted, selected and compared using DTASelect/CONTRAST (Tabb et al., 2002).

Two types of statistical analyses were performed on label-free quantitative proteomics values to define significantly enriched proteins in the WT CG5830 FLAG-affinity purifications (APs) compared to FLAG negative controls and FLAG-APs with PD CG5830. Proteins were deemed

significantly enriched if their STN (signal-to-noise) values were greater than 0.9 with Benjamini and Hochberg adjusted p-values (Benjamini & Hochberg, 1995) lower than 0.05 using PLGEM (Pavelka et al., 2008); and/or if their log2 Fold Change values were greater than 2 with p-values lower than 0.05 using QPROT (Choi, Fermin, & Nesvizhskii, 2008). Proteins also had to be detected in at least 2 of the three replicate analyses.

Co-Immunoprecipitation and Western blot with S2 cell lysates: To study pair wise protein-protein interaction, S2 cells were transfected with 0.3 µg of CG5830-Myc and/or HA/FLAG tagged target DNA constructs using Effectene transfection reagent (Qiagen). After 48 hrs of transfection cells were harvested, washed with PBS and lysed in PBS based homogenization/lysis buffer (150 mM NaCl, 10 mM Tris pH 7.5, 1% NP-40, 0.1% Triton X-100) for 20 mins. The lysates were centrifuged at 10,000 rpm, 4ºC and clear supernatant was collected in fresh 1.5 mL Eppendorf tube. For immunoprecipitation 1.5-2.0 mg of total S2 cell lysate was incubated in each IP with 10µL pre-washed anti-FLAG and anti-Myc agarose beads (Sigma) for 2 h at 4ºC with continuous rotation. After washing 3 times and boiling with 2X laemmli sample buffer, the immunoprecipitates were run on a 4-12% SDS-polyacrylamide gel (Invitrogen). The IPs were immunoblotted with either anti-HA or anti-CG5830 antibody. Total S2 cell lysates were also immunoblotted for CG5830 and CG5830 partners as loading controls. Pairwise interaction experiments of Chn-3XHA/3XFlag and CG5830-Myc; truncated CG5830 constructs and full-length CG5830-Myc or Dah-3XHA/3XFlag were performed similarly.

Phos-tag IP with S2 cell lysates: To study phosphatase activity of CG5830 and its truncated versions, S2 cells were transfected with 0.3 µg of CG5830-Myc (PD, full length, N, M, C, NM, MC, and mutants) and/or Dah-HA/FLAG tagged target DNA constructs using Effectene transfection reagent (Qiagen). After 48 hrs of transfection cells were harvested, washed with PBS and lysed in lysis buffer (150 mM NaCl, 1.0% NP-40, 50 mM Tris-Cl pH 8.0, 0.1% Triton X 100, 0.5% Na-

Deoxycholate, 0.1% SDS) in the presence of phosphatase inhibitors and 1mM $MgCl_2$ for 20 mins. The lysates were centrifuged at 10,000 rpm, 4ºC and clear supernatant was collected in fresh 1.5 mL Eppendorf tube. For immunoprecipitation total S2 cell lysates were incubated in each IP with 10µL pre-washed Phos-tag agarose beads (Wako) for 2 h at 4ºC with continuous rotation. After washing 3 times and boiling with 2X laemmli sample buffer, the immunoprecipitates were run on a 4-12% SDS-polyacrylamide gel (Invitrogen). The IPs were immunoblotted with anti-HA antibody. Total S2 cell lysates were also immunoblotted for CG5830 and Dah as loading controls. Phosphorylation status of other substrates were detected in absence of CG5830 with the same IP system. To detect dephosphorylation with nonspecific phosphatases, cell lysates were incubated with 2U calf intestine phosphatase (NEB) and 2U lambda phosphatase (NEB) in absence of phosphatase inhibitors and followed with IP and western blotting.

***Drosophila* embryo extract preparation for chromatography:** Embryo extracts were prepared according to the protocol described by Wodarz (Wodarz, 2008). Briefly, the flies, CG5830-HA and $CG5830^{C91}$ were housed at 25ºC and 0-to-20hr old embryos were collected on large petri-dish containing apple juice agar medium and yeast paste. After 20-h the embryos were collected into a sieve by scraping them off the agar surface with a soft paint brush and rinsing with a stream of cold tap water. Approximately 5 gm of embryos was obtained per 0-20-h collection. Embryos were dechorionated at room temperature by submerging them in a beaker containing 50% sodium hypochlorite solution in 1X PBS with periodic stirring. After hypochlorite treatment for approximately 5 minutes, the embryos were washed extensively with cold tap water to remove any chlorine from the preparation. The embryos were put into 50ml falcon tubes containing 1X PBS and centrifuged to remove all the liquid. Approximately, 5 gm of embryos were lysed in 10 ml of freshly prepared embryo lysis buffer (150 mM NaCl, 50 mM Tris-Cl pH 7.5, 1% NP-40 (Sigma), 0.1% Triton X-100 (Sigma), 0.1% SDS) containing EDTA-free protease inhibitor (Roche). The embryos were homogenized using Dounce homogenizer (KontesTMDounce) on ice

and rotated in 4ºC for 30min. Lysates were sonicated at 20 kHz for 10 cycles of 10 secs “on” and 10 sec “off”. Whole lysates were treated with 5ug Benzonase (EMD Milipore) and 50ug RNAse A (Qiagen) for 45min at 4ºC. Following centrifugation at 4500 rpm at 4ºC for 10min, the soluble cytoplasmic phase was filtered (0.45um) into a new tube. Insoluble pellet was solubilized in 10ml of lysis buffer by homogenizing and processing with the same protocol and filtered into the cytoplasmic fraction. Whole lysate was processed by chromatography (see next section).

Purification and separation of CG5830-HA and CG5830^{C91} monomers and polymers: 30ml of total cell lysates were loaded, at 0.2 ml/min, onto a Mono Q 4.6/100 PE AEX column (GE healthcare) equilibrated with a buffer containing 100mM NaCl and 50mM Tris-HCl pH 8.0. Elution was performed using a 30 minutes salt gradient ranging from 0.1M to 1M. Most CG5830-HA protein was eluted at around 350-450 mM salt concentration. All CG5830-HA-containing fractions collected were pooled together and the sample was dialyzed overnight using CEX binding buffer (50 mM MOPS pH 6.5 with 100 mM NaCl). The following day, sample was loaded into Mono S 4.6/100 PE CEX column (GE healthcare) equilibrated with the same buffer. Elution was performed as described for the AEX column. Most CG5830-HA protein was eluted at around 200-300 mM salt concentration. Fractions were analyzed by western blotting with anti-HA and anti-CG5830 antibodies. One fraction containing either CG5830-HA or CG5830^{C91} protein were dialyzed against 2000ml of 150mM NaCl, 10mM $MgCl_2$ and 50mM Tris-HCl pH 7.5 for 12 hours. Following dialysis, each fraction was loaded onto a Superose 6 10/300 column (GE healthcare) equilibrated with a buffer containing 150mM NaCl, 10mM $MgCl_2$ and 50mM Tris-HCl pH 7.5 and fractionated at 0.2 ml/min using a HPLC (AKta Purifier) system. Altogether 36 fractions (500µL each) were collected for each lysate and alternate fractions were analyzed by 4-12% SDS-PAGE and 1.5% Semi-Denaturing Detergent Agarose Gel (SDD-AGE) Electrophoresis. Blots were probed with anti-HA and anti-CG5830 antibodies to detect different size classes of CG5830-HA and CG5830^{C91} proteins, respectively.

DifMUP phosphatase assay: Fractions of oligomers and monomers were dialyzed against 1X reaction buffer (100mM Tris acetate pH 5.5 0.1% Triton X-100 1mM $MgCl_2$) at 4C overnight. DiFMUP substrate working solution was prepared to a final concentration of 200uM in 1X reaction buffer and 50ul substrate solution was distributed to each microplate well (96well). Dialyzed fractions containing 20ug of phosphatase were distributed to corresponding wells. As a positive control, 1U of potato acid phosphatase activity was measured. As a negative control, only reaction buffer was added to one well. Sample activity was measured every 10min over 1h at room temperature in the dark with a fluorimeter using excitation at ~360 nm and emission detection at ~460 nm. Activity of CG5830-HA monomers and oligomers was normalized with the activity of corresponding fractions from similarly purified $CG5830^{N91}$ sample. Statistical analysis was performed by one-way ANOVA for multiple groups and data of activity at 25min was plotted as mean ± SEM.

ThioflavinT (ThT) binding assay with CG5830-HA and $CG5830^{C91}$ monomers and oligomers: The ThT binding assay was performed essentially as described previously (LeVine, 1999). Briefly, 100ul of a 500*µ*l eluate from size exclusion chromatography was mixed with 50 *µ*l of 25 *µ*M ThT (Sigma) and 50 ul of 50 mM Glycine buffer (pH 8.5). The assay was carried out at 25ºC in 96-well fluorescence plate reader (442nm excitation and 485 nm emission) using an infinite M200 PRO spectro fluorimeter (TECAN). 10 independent measurements were taken for each sample using Magellan software every min with 10 sec shaking between the measurements. To obtain the ThT enhancement fluorescence due to CG5830 alone, the fluorescence of ThT in elution buffer was subtracted. Binding data after 5 minutes are represented in the figure.

Proteinase K digestion: Fractions containing CG5830-HA and $CG5830^{C91}$ monomers and polymers (0.6 µg of total protein) were digested with 1, 2 and 5 ng of proteinase K for two minutes at 37ºC. All reactions were assembled in ice and then incubated in a 37ºC water bath for 2 mins. The enzyme activity was stopped by immediately heating the samples to 75ºC. The enzyme

treated reaction mixture was passed through nitrocellulose membrane. The membrane was blocked in 5% milk in TBS buffer and probed with anti-HA and anti-CG5830 antibodies.

A11 and OC Western blot assay: For A11 and OC western analysis, different amounts of total protein (2, 4 and 6 µg) of the corresponding CG5830-HA and $CG5830^{C91}$ monomeric and oligomeric fractions (with a blank fraction) were dot blotted on nitrocellulose transfer membrane (Whatman). The membranes were blocked in 5% milk and incubated with either rabbit A11 anti-oligomer antibody (Invitrogen) or rabbit OC anti-fibril antibody overnight at 4ºC. The same amount of total protein was dot-blotted and probed with either anti-HA or anti-CG5830 antibodies for the detection of CG5830-HA and $CG5830^{C91}$ proteins, respectively.

Chapter III: Functional characterization of *CG5830* gene in *Drosophila melanogaster*

3.1 Objectives and Rationale

CG5830 is a strong candidate for a prion-like protein based on *in silico* and experimental screening performed in the Si lab (Section 1.2). However, it is a functionally uncharacterized gene in *Drosophila melanogaster*. In Table 3.1 are listed the characteristics of the CG5830 *gene* and protein based on flybase.org, which is a database and the primary repository of genetic and molecular data for the insect family *Drosophilidae*. Most of the functional information listed below is based on sequence homologies and high-throughput assays.

Table 3.1 Characteristics of CG5830 gene and protein based on FlyBase.

Gene	FlyBase ID	FBgn0036556
	Genomic location	3L: 15,992,669...16,004,736[-]
	Gene group	Haloacid dehalogenases
	Putative transcripts	RA and RB
	High throughput tissue expression	Low expression in adult ovary and testis Moderate expression in rest of all tissues
	High throughput developmental expression	Low to moderate expression in all stages embryo, larva, pupa and adult
	Orthologs	Small CTD phosphatases (CTDSP) in human and model organisms: mice, chicken, xenopus, zebra fish, *C. elegans*
Protein	Putative isoforms	RA and RB
	Protein domain	FCP1 homology domain
	Molecular function	Phosphoprotein phosphatase activity
	Location	Cell membrane
	Interaction partner	*Staufen*

The *CG5830* gene is located on the 3rd chromosome, extending over 12.1 kb of genomic DNA with six annotated exons. It has two putative transcripts and two putative protein isoforms, RA and RB. In a high-throughput developmental proteome analysis of *Drosophila melanogaster*, it has been shown that CG5830-RA protein is expressed constitutively throughout whole life cycle of flies (Figure 3.1) (Casas-Vila et al., 2017).

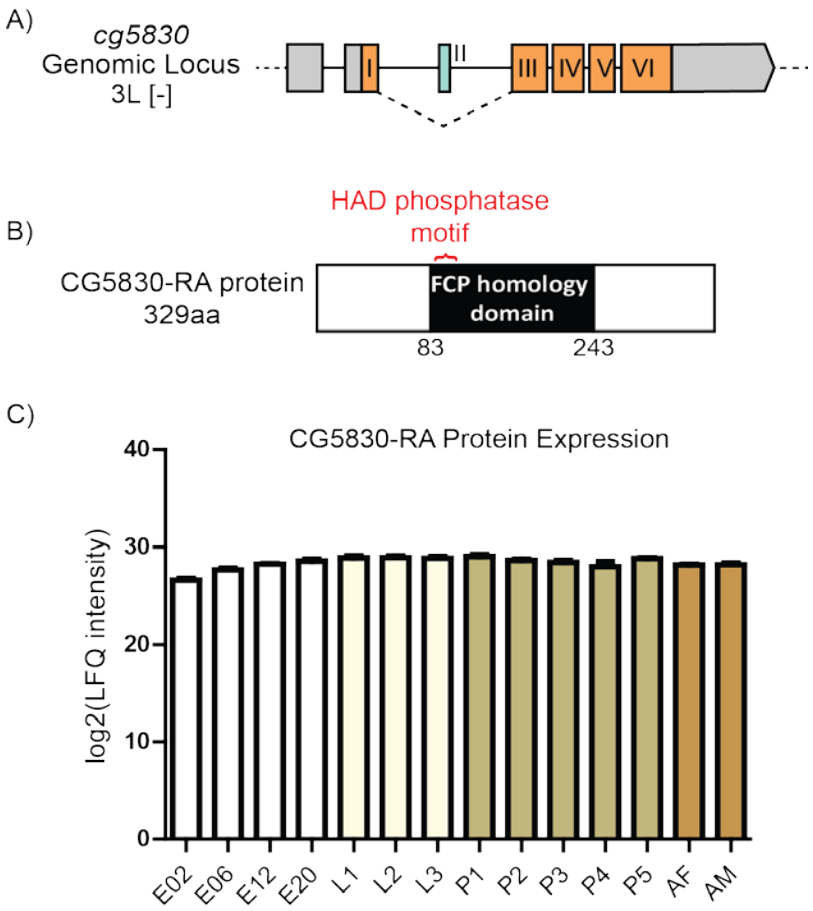


Figure 3.1 CG5830 gene and protein.

A) Genomic locus of CG5830. Gray: untranslated regions, Orange: Exons. Dashed line: Splicing of RA isoform.
B) CG5830-RA protein with sequence homology domain and phosphatase motif indicated.
C) CG5830-RA protein expression pattern across different developmental stages adapted from Casas-Vila (2017). LFQ: label-free quantitation based on unique peptide intensities for each protein. E: embryo, L: Larva, P: Pupa, AF: Adult female, AM: Adult male.

CG5830 has sequence homology to HAD phosphatases and specifically HAD subfamily od Small C-terminal domain phosphatases (SCPs). SCPs are conserved in human and several model organisms and has an FCP (F-cell production, also called TFIIF-stimulated CTD phosphatase) homology domain. Previously, it has been shown that CG5830 knockdown in *Drosophila* embryonic cell line, Schneider Line 2 (S2), enhances the expression of neuronal and glial genes, including *embryonic lethal abnormal vision* (*elav*), similar to SCP1 knockdown in mouse embryonic stem cells (Yeo et al., 2005). Moreover, *in vitro* mammalian SCP1

dephosphorylates the RNA polymerase II (RNAPII) CTD. These observations implied that CG5830 may function as a CTD phosphatase to control transcription of neuronal genes, however these experiments did not rule out the possibility that the observed transcriptional changes are not due to alteration of some other cellular processes.

To understand the function of CG5830 at different levels - organismal, cellular-molecular and biochemical-, I set the following objectives:

a) Determine if there is any functional similarity between *Drosophila* CG5830 and mammalian SCP1.

b) Determine the phenotype of CG5830 mutation.

c) Determine expression pattern of CG5830 protein across different developmental stages.

d) Determine what protein complexes CG5830 forms in the cell and its putative substrates.

e) Determine the enzymatic activity of CG5830 on its putative substrates.

3.2 Results

3.2.1 *Drosophila* CG5830 is not a functional ortholog of mammalian SCP1

Based on protein sequence homology, CG5830 is the only *Drosophila* ortholog of mammalian SCP1. In mammals SCP1 belongs to a family of phosphatases that dephosphorylate the RNA polymerase II (RNAPII) CTD specifically at Serine-5 (Yeo et al., 2003) at the end of the transcription cycle and regulates subsequent transcription initiation (Hajheidari, Koncz, & Eick, 2013). It has been previously shown that overexpression of SCP1 represses transcription, while the phosphatase dead (PD) mutant, acts dominant negatively and enhances transcription (Yeo et al., 2003). SCP1 is also known to interact with Repressor Element 1 (RE1) Silencing Transcription factor or REST (Yeo et al., 2005), which represses neuronal gene expression in embryonic stem (ES) cells and differentiated non-neuronal cells (Ballas & Mandel, 2005). REST binds to a 23-bp

RE1-DNA element found in numerous neuronal genes (Schoenherr, Paquette, & Anderson, 1996) and recruits a silencing complex comprised of histone deacetylases and DNA methyl transferases (Hakimi et al., 2002; Naruse, Aoki, Kojima, & Mori, 1999). REST is a phospho-protein that is stabilized by SCP1-dependent dephosphorylation (Nesti, Corson, McCleskey, Oyer, & Mandel, 2014). Perturbation of SCP1 function promotes expression of neuronal genes in cultured non-neuronal cells (Yeo et al., 2005). Taken together these observations lead to the model that SCP1 controls neuronal gene expression via the REST transcription factor, which serves as a master regulator of neuronal gene expression. To determine whether CG5830 is truly an orthologue of SCP1, I investigated its function at three levels.

First, I performed a CTD phosphatase assay. For this, recombinant His-tag wild type (WT) and PD (D89E, D91N) CG5830 were purified using Ni-NTA columns under native conditions. RNAPII was *in vitro* phosphorylated by TFIIH at Ser-5 residues at its CTD tail and subsequently incubated with either wild type $CG5830^{WT}$ or $CG5830^{PD}$ variants of the protein Reactions were stopped with SDS-PAGE sample buffer and the phosphorylation status of RNAPII was determined with Serine-5-phosphate (Ser5P) specific antibodies. I observed that only $CG5830^{WT}$ but not $CG5830^{PD}$ dephosphorylates RNAPII (Figure 3.2A).

Next I asked whether CG5830 interacts with charlatan (Chn), the *Drosophila* orthologue of mammalian REST (Tsuda et al., 2006). To this end, I did co-immunoprecipitation experiments using S2 cells transfected with CG5830-Myc, or Chn-HA/Flag, or CG5830-Myc and Chn-HA/Flag together. Immunoprecipitation of Chn did not pull down CG5830, suggesting either they don't interact or do not form a stable complex (Figure 3.2B).

Lastly, I tested the effect of overexpression of CG5830 on transcriptional output from a neuronal gene promoter (elav) using a dual luciferase reporter based assay. In this assay, the *elav* promoter drives the expression of *firefly luciferase* and an *actin* promoter drives *renilla*

luciferase expression and the output of *elav* promoter was measured as ratio of firefly/renilla. Following co-transfection of *CG830*^{WT} or *CG5830*^{PD} with the reporter plasmids, I observed that unlike SCP1, overexpression of CG5830^{WT} significantly enhanced transcription from neuron specific promoter ($p<0.001$), while CG5830^{PD} did not have any effect on promoter activity compared to control (Figure 3.2C).

Taken together these results suggest that in isolation, although CG5830 can act as a CTD phosphatase, the effect on the transcriptional output is different from SCP1. Therefore, I concluded that CG5830 phosphatase is not a functional ortholog of mammalian SCPs.

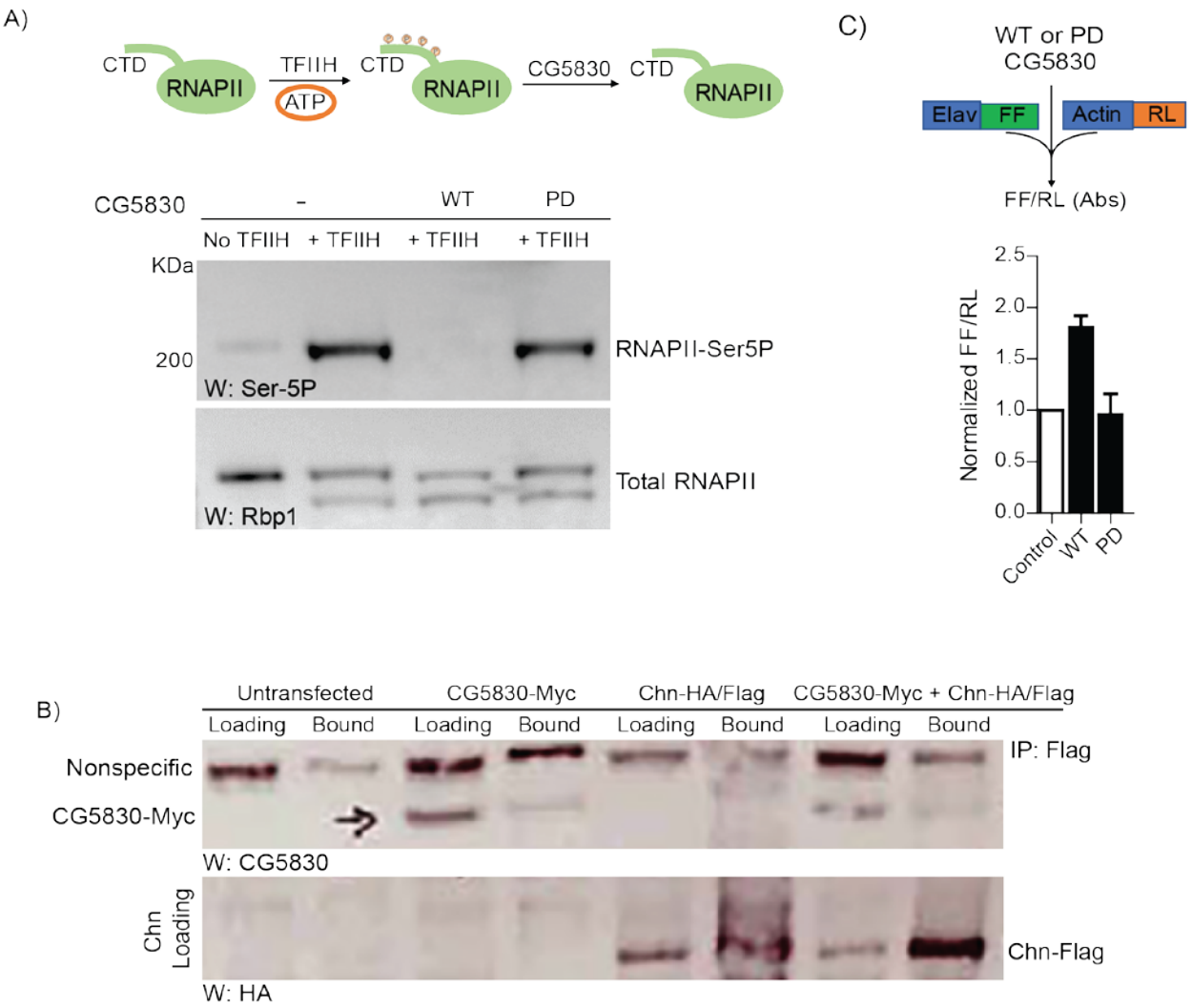


Figure 3.2 CG5830 is not a functional ortholog of SCP1.

A) RNAPII CTD phosphatase assay. Top: experimental scheme. Bottom: Western blot of phosphorylated RNAPII. Total RNAPII serves as loading control. Addition of WT CG5830 abolishes Ser5 specific phosphorylation of RNAPII, while PD mutant has no effect.

B) Co-immunoprecipitation of CG5830 with Chn. Immunoprecipitation for Chn and blotting is done for CG5830. Faint band in CG5830-Myc and Chn-HA/Flag co-transfected bound fraction is nonspecific binding of CG5830-Myc protein to Flag beads as also seen in only Cg5830-Myc transfected bound fraction.

C) Left panel: Experimental scheme for luciferase assay. Right: WT-CG5830 enhances expression from *elav* promoter compared to PD mutant. The statistical significance was measured by unpaired two-tailed t-test, n=7, $p<0.001$.

3.2.2 CG5830 is required for embryonic development

As mentioned above, the *CG5830* gene has two predicted transcripts, CG5830-RA and CG5830-RB. However, only the RA isoform was detected with global gene expression analysis

using RNA-Seq in flies and in S2 cells (Celniker et al., 2009; Roy et al., 2010). This transcript encodes a 329-aa protein (37kDa) with a HAD phosphatase signature motif (ψψψD*X*D*X*(T/V)ψψ) between residues 91 to 101 in the conserved FCP1 homology domain (83-243) (Figure 3.1A and Figure 3.3A).

To determine the *in vivo* function of CG5830 in *Drosophila*, I used CRISPR/Cas9 mediated gene deletion to generate *CG5830* deficient flies and systematically analyzed the associated phenotypes (Figure 3.3). I obtained three deletion lines that express only the N-terminal 69, 91 or 169 amino acids, and I therefore named them $CG5830^{P69}$, $CG5830^{C91}$, and $CG5830^{N169}$. I confirmed these deletions at both mRNA (Figure 3.3B) and protein levels (Figure 3.3C). Among these, only $CG5830^{N169}$ contains the signature HAD phosphatase motif.

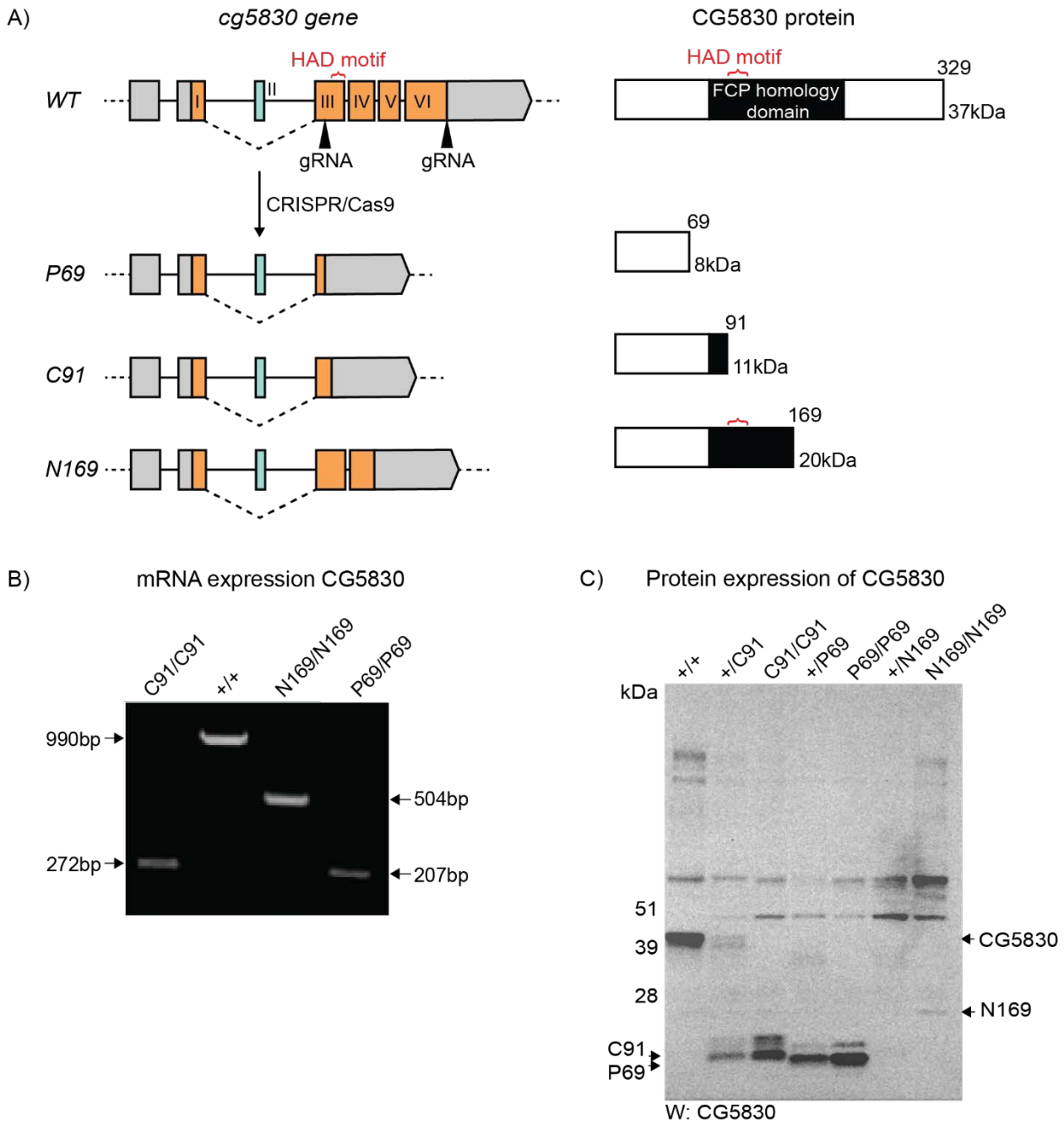


Figure 3.3 Generation of *CG5830* deletion lines.

A) *CG5830* genomic locus and schematic of the generation of mutant lines by CRISPR/Cas9. Guide RNAs (gRNA) targeting indicated regions in the genomic locus were cloned and injected into a *vasa-Cas9* fly line. Resulting fly lines was screened and balanced with a 3rd chromosome balancer. WT: Wild type, Gray: untranslated regions, Orange and blue: Exons. Dashed line indicates splicing into RA isoform.

B) mRNA expression of selected mutant lines and wild-type flies was detected using primers for 5'- and 3'-UTR regions of *CG5830*. Corresponding bands for each deletion line and wild-type is indicated by arrows.

C) Protein expression of CG5830 in wild-type and heterozygous and homozygous mutant larvae was detected by SDS-PAGE. Corresponding bands for each deletion line and wild-type is indicated by arrows.

Next, I systematically screened phenotypes of homozygous mutant flies as shown in the crossing scheme in Figure 3.4. I observed that all homozygous mutant lines are viable in the F1 generation, with no obvious developmental phenotype. However, homozygous mutants of *CG5830*^{P69} and *CG5830*^{C91} show female fertility defects laying significantly less eggs (eggs laid/ female/day) compared to wild type control and *CG5830*^{N169} (Figure 3.4B). To understand the contribution of possible male fertility defects in this phenotype, I performed sterility tests by crossing homozygous *CG5830* mutant males to wild type females or vice versa (Figure 3.4A and B). Mutant females crossed to wild type males showed the same fertility defect, however, mutant males crossed to wild type females did not cause a fertility defect (Figure 3.4B). To understand the reason behind this reduced fertility phenotype, I looked at the size of the ovaries, the number of ovarioles and the number of eggs per ovary in *CG5830*^{P69} and *CG5830*^{C91} mutant females. I did not detect any significant difference in these phenotypic assays compared to wild type and *CG5830*^{N169} (Figure 3.4C). With these experiments, I could not reach a clear conclusion about the reason of the female fertility phenotype of *CG5830* mutant flies. In order to understand the function of *CG5830*, I focused on embryonic phenotype discussed below.

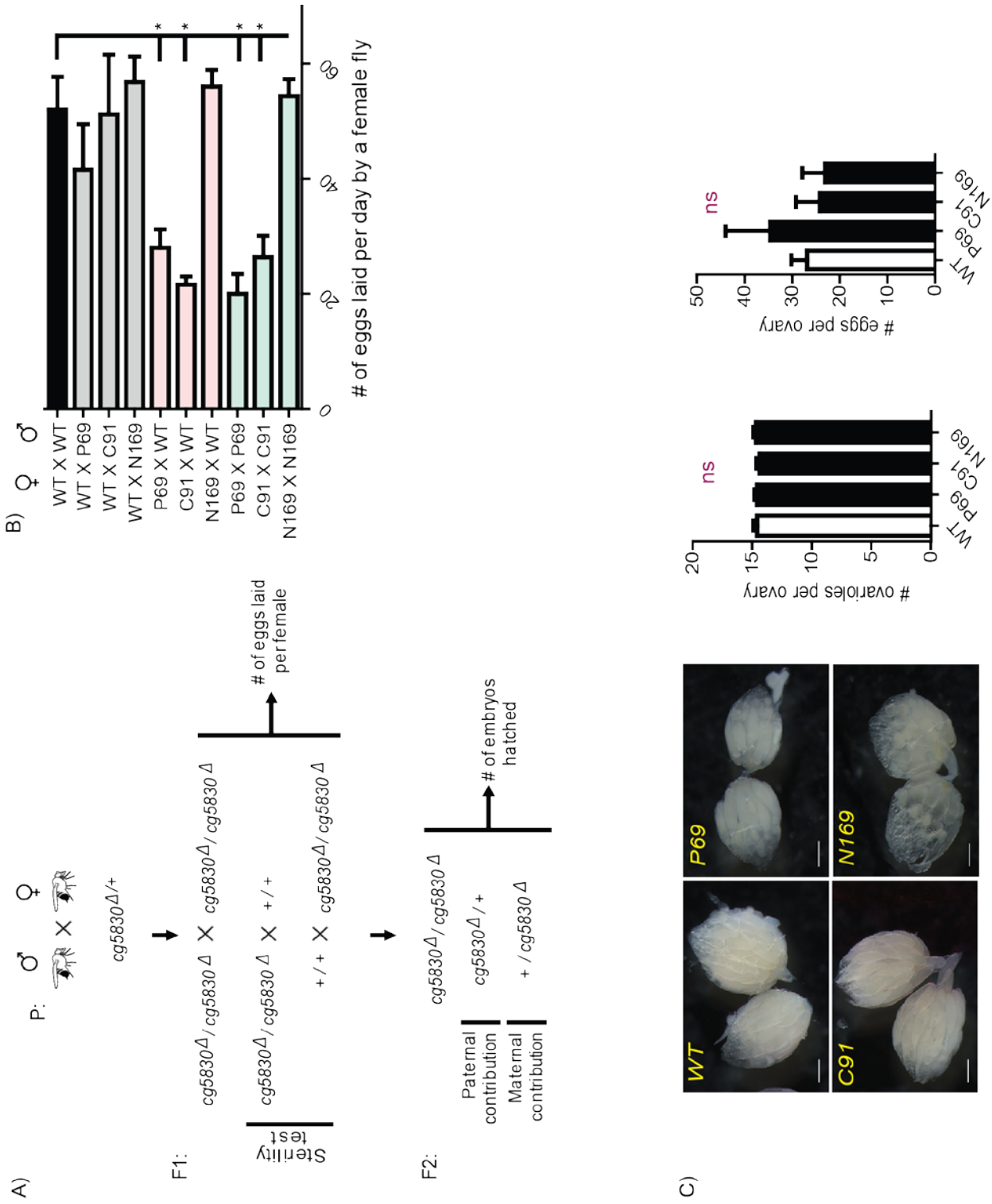


Figure 3.4 Characterization of *CG5830* deletion mutants.

A) Crossing scheme. Heterozygous parental lines were crossed to obtain homozygous F1 progeny. Homozygous mutant flies were either self-crossed or crossed with wild type flies. Number of embryos laid by F1 progeny was counted.

B) *CG5830* deletion causes female sterility. Number of embryos laid by F1 female flies for each group of crosses was plotted as mean ± SEM, and statistical significance was determined by one-way ANOVA for multiple groups.
C) Female fertility defects of $CG5830^{P69}$ and $CG5830^{C91}$ mutants are not caused by ovary morphogenesis and egg storage defects. Top: Representative images of a pair of ovary from wild type and mutant flies. Bottom: Ovaries were dissected from a-week old mutant females (n=10), number of ovarioles and eggs per/ovary counted under dissection microscope.

In the F2 generation, embryos laid by homozygous F1 $CG5830^{P69}$ and $CG5830^{C91}$ mutants die during embryogenesis. This is a classical maternal effect phenotype, indicating that the maternal deposition of the CG5830 is enough for the organism to survive to adulthood (Figure 3.5A). To characterize F2 generation mutant embryos, I took a deductive approach: first find what kind a cuticular phenotype they present, which would also show at what stage development stopped, and based on the cuticular phenotype focus on that particular developmental point. Cuticle preparations of mutant embryos showed that they do not form any denticles, and they die with germ band retraction, head involution and dorsal closure defects (Figure 3.5B) before going into differentiation (Stage 15 and on). A similar segment polarity phenotype has been previously seen in embryos exposed to excess *wg* (Simcox & Roberts, 1989) and embryos lacking zygotic copies of *nkd* (W. Zeng et al., 2000). Therefore, I focused on *wingless* signaling components (Section 1.7) by immunostaining mutant embryos with *en* and *wg* antibodies. Indeed, $CG5830^{P69}$ and $CG5830^{C91}$ mutant embryos show severe segment polarity phenotypes with diffused *en* expression in each parasegment at stage 10-11. Later at stage 13-14, embryos, even though they have normal number of segments, show an ectopic stripe of *wg* at the anterior side of each segment. This suggests that the diffused expression pattern of *en* led cells to gain *wg* identity, causing segments to lose their polarity (Figure 3.5C). Such abnormal patterns of *en* and *wg* have been previously seen in embryos lacking both maternal and zygotic contributions from any of the three genes that antagonize *wg* signaling: *zw3/GSK* (Siegfried et al., 1992), *D-axin* (Hamada et al., 1999), and *D-Apc2* (McCartney et al., 1999) in addition to *nkd*. To understand, if this defect

arises at an earlier time point, I also checked early embryonic patterning of these embryos by immunostaining with *eve* and *dl*. *CG5830*^{P69} and *CG5830*^{C91} embryos do not have any defect in dorsoventral and anteroposterior axis patterning shown by *dl* gradient and *eve* stripe patterns, respectively, indicating that *CG5830* mutant phenotype is not caused by an earlier developmental defect (Figure 3.5D). Moreover, rescue of embryonic lethality in the F2 generation, to nearly the level of wild type control by introduction of a wild-type copy of *CG5830* through crossing mutant females with wild-type males, suggests that zygotic expression of *CG5830* is sufficient and required for the progress of embryonic development (Figure 3.5A). Contrary to *CG5830*^{P69} and *CG5830*^{C91}, *CG5830*^{N169} mutant flies develop normally without any detectable phenotype and can be maintained over multiple generations.

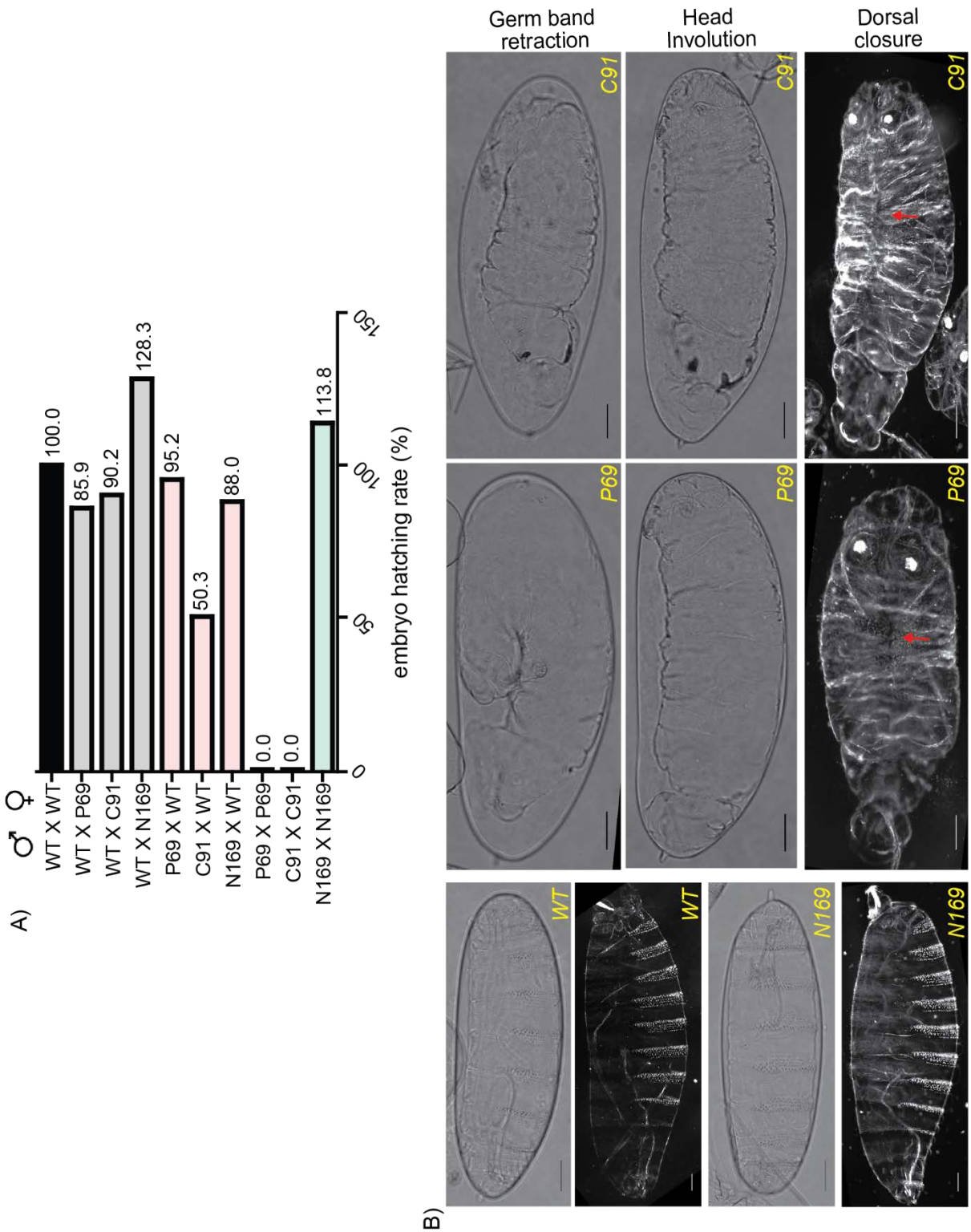


Figure 3.5 *CG5830* deletion causes embryonic lethality and segment polarity defects.

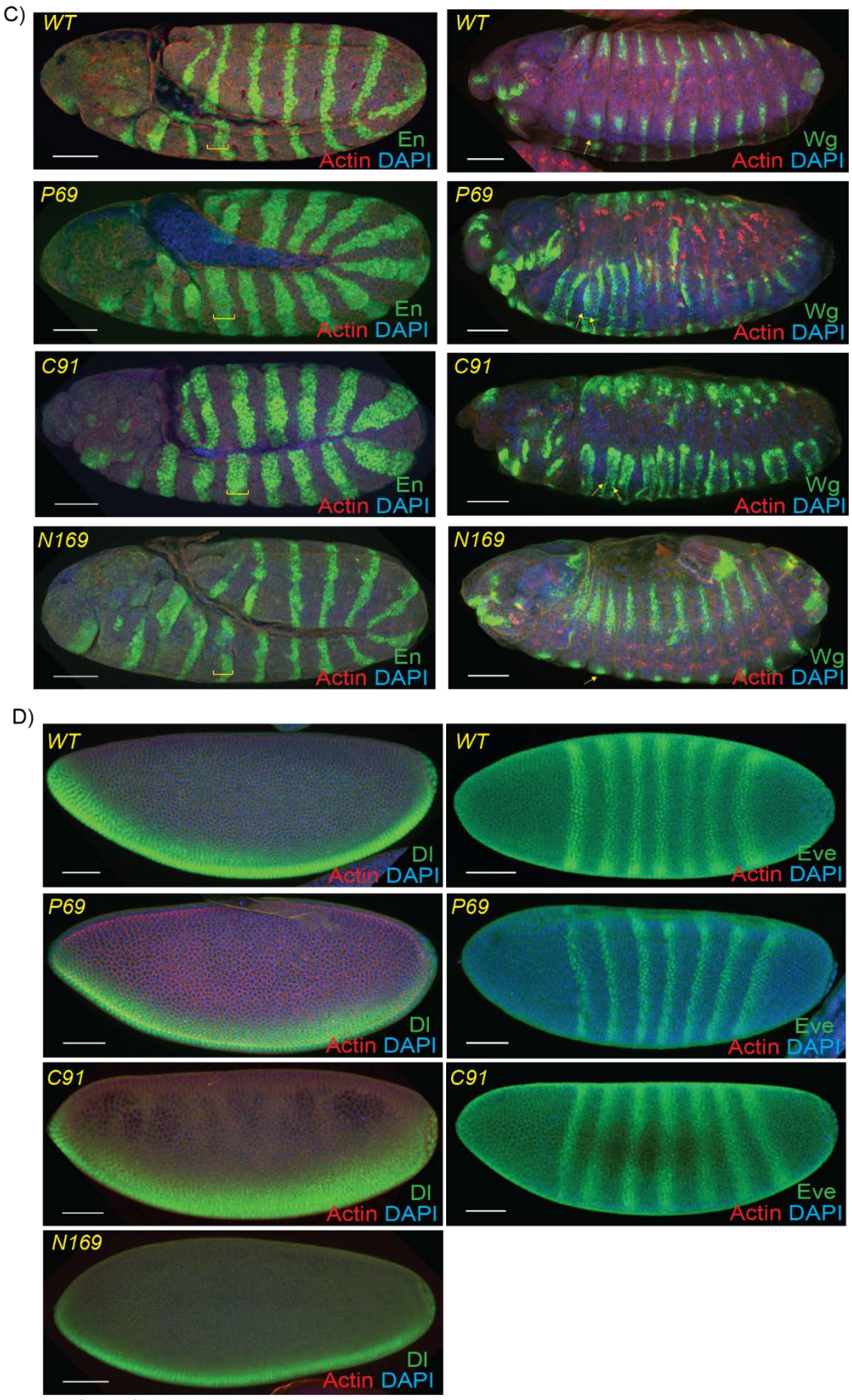


Figure 3.5 Continued.

Figure 3.5 *CG5830* deletion causes embryonic lethality and segment polarity defects.

A) Embryo hatch rate (number of hatched embryos over total number of embryos counted) for each group of crosses was normalized by the embryo hatch rate of wild type flies and was plotted as percentages.
B) Cuticle images showing representative examples of mutant phenotypes and wild type. There is no denticles present in *P69* and *C91* compared to wild type and *N169*. (Scale bar: 50um)
C) Top panel: Embryos stained for En (green). *P69* and *C91* mutants show an enlarged *en* stripe in each parasegment compared to wild type and *N169*, indicated by yellow brackets. Bottom panel: Embryos stained for Wg (green). *P69* and *C91* mutants show an ectopic *wg* stripe in each segment compared to wild type and *N169*, indicated by yellow arrows. Red: actin, Blue: DAPI.
D) Left panel: Embryos stained for Dorsal (green). There is no change in D-V axis in any of the mutants. Right panel: Embryos stained for Eve (green There is no change in early A-P axis patterning in any of the mutants. Red: actin, Blue: DAPI.

Together these data suggest that *CG5830* has a role in establishing the A-P axis of each segment as a segment polarity gene, acting as an antagonist of *wg* signaling during embryogenesis. Moreover, lack of any defect in the deletion mutant $CG5830^{N169}$ implies that this function of *CG5830* likely requires its conserved HAD phosphatase motif. Therefore, in further characterization of CG5830 protein, I focused on embryogenesis of *Drosophila melanogaster.*

3.2.3 CG5830 is a ubiquitous membrane protein and forms higher order assemblies on the membrane after gastrulation

To monitor the expression and subcellular distribution of CG5830 protein during *Drosophila* embryogenesis, I tagged *CG5830* gene C-terminally with either 3XHA or EGFP tags at its endogenous genomic locus using *CRISPR*/Cas9 gene-editing system (Figure 3.6A). Tagging did not alter protein function since there was no observable developmental defect and expression of tagged CG5830 protein in embryos was confirmed by western blotting. Protein extracts of 0-20-hour old embryos were prepared from the following four samples: $CG5830^{C91}$, wild type, *CG5830-3XHA*, *CG5830-EGFP* and, and similar amount of total protein was loaded on a SDS-PAGE gel. From the immunoblots of these samples with anti-CG5830, anti-HA and anti-GFP, I detected 37

kD, 41 kD and 67 kD bands corresponding to the predicted molecular weight of CG5830 in each version, respectively (Figure 3.6A).

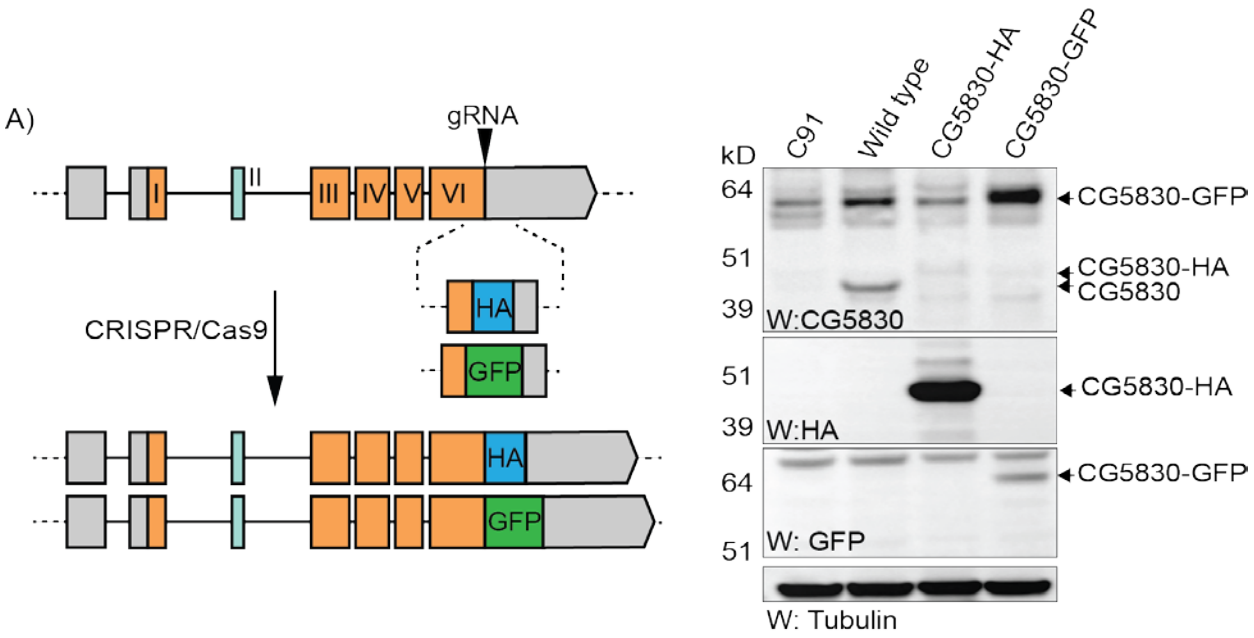


Figure 3.6 CG5830 is a ubiquitous membrane protein and form higher order assembly in the membrane after gastrulation.

A) Left panel: Scheme of the generation of tagged flies by CRISPR/Cas9. Right panel: protein expression of CG5830 in 0-20-hour old mutant CG5830^{C91}, wild type (WT), and tagged embryos (CG5830-HA and CG5830-GFP). Arrowheads indicate specific bands for CG5830 in each blot. Tubulin is detected as a loading control. W: Western, kD: kilodaltons.

B) CG5830 is ubiquitously and constitutively expressed on the membrane during embryogenesis. CG5830-HA embryos of different ages were collected and stained with anti-HA antibody (green) and DAPI (blue). St: Stage, WT: wild type. Scale bar: 50um.

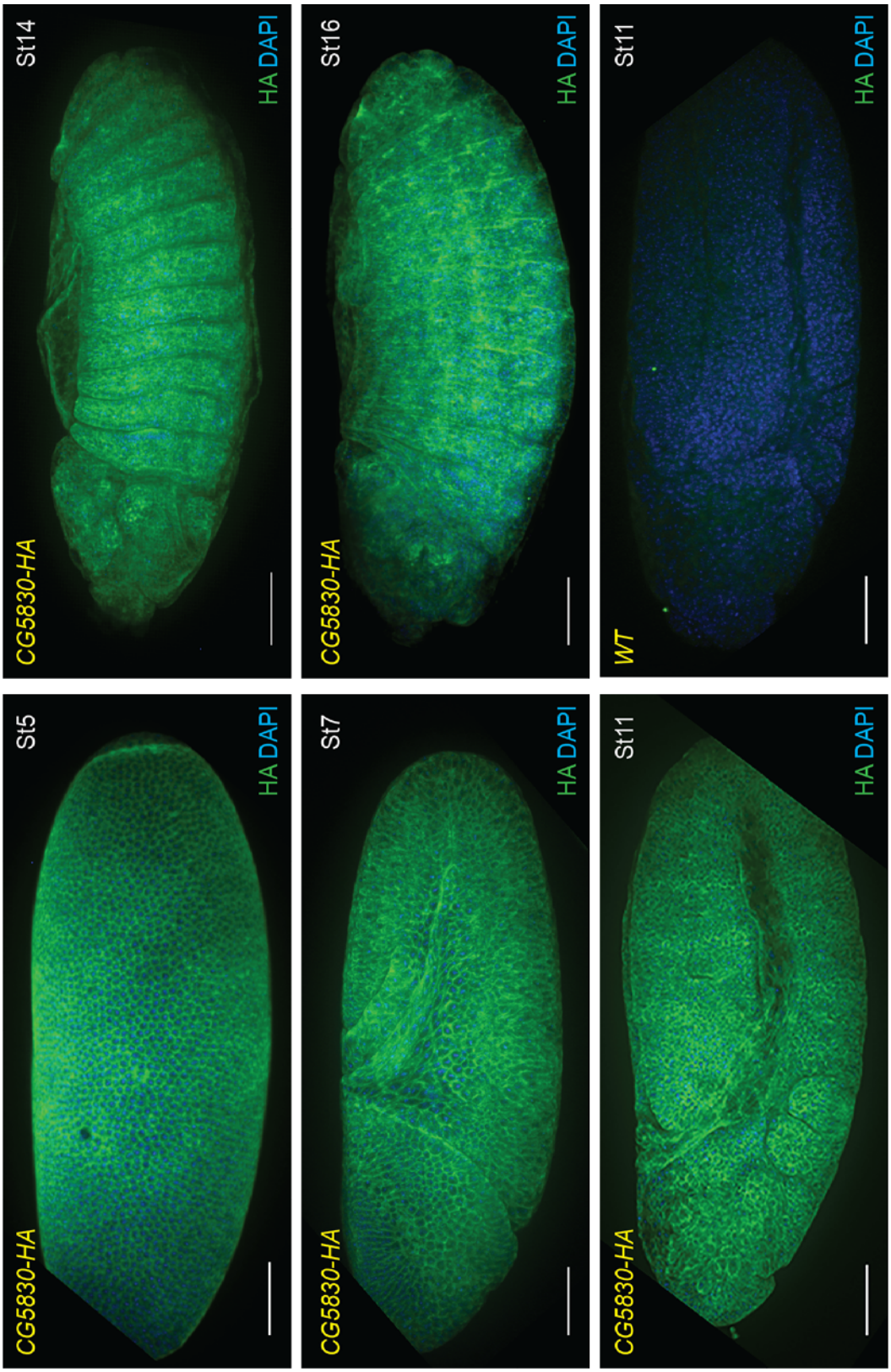
C) CG5830 has a diffused form on the membrane until 3hour, in 3-7h it forms puncta and this puncta appearance persists in 7-10 and 10-13h. CG5830-HA embryos of different ages were collected and paraffinized. Paraffin sections (10um) were stained with anti-HA antibody (green), anti-discs large (dlg, red) (membrane marker), and DAPI (blue). St: Stage, WT: wild type. Scale bar: 10um.

D) CG5830 has a diffused form on the membrane until gastrulation. Snapshots of CG5830-EGFP live imaging at indicated time points during 0-20h period. The green spots indicated with yellow arrowheads in time 0 is auto-fluorescence due to yolk.

E) CG5830 forms aggregates/foci on the membrane at gastrulation as shown by CG5830-EGFP::Histone2A-RFP embryo live imaging. Left panel: whole embryo at 0 and 105 min. Right panel: Zoomed version of timepoints: 0, 15, 30, 45, 60, 75, 90, and 105min. Yellow arrowheads show puncta emerging. AS: Amnioserosa, which is autoflourescent.

F) CG5830 aggregates dynamically assemble and disassemble on the membrane. Focus is on an abdominal segment of a 13h embryo. Embryo movement is registered to a point by image-J and movement of aggregates followed over time. Different color arrowheads show different puncta assembling-disassembling over time.

G) CG5830 forms an SDS and heat resistant oligomer after gastrulation. Protein expression in embryos was followed at time points 0-30min, 30min-1, 1-2, 2-3, 3-4, 4-5, 5-7, 7-10 and 10-13-hour by western blotting with anti-GFP, anti-CG5830, and anti-HA antibodies and in both soup and pellet fractions. Arrowheads indicate CG5830 specific bands.



B)

Figure 3.6 Continued.

C)

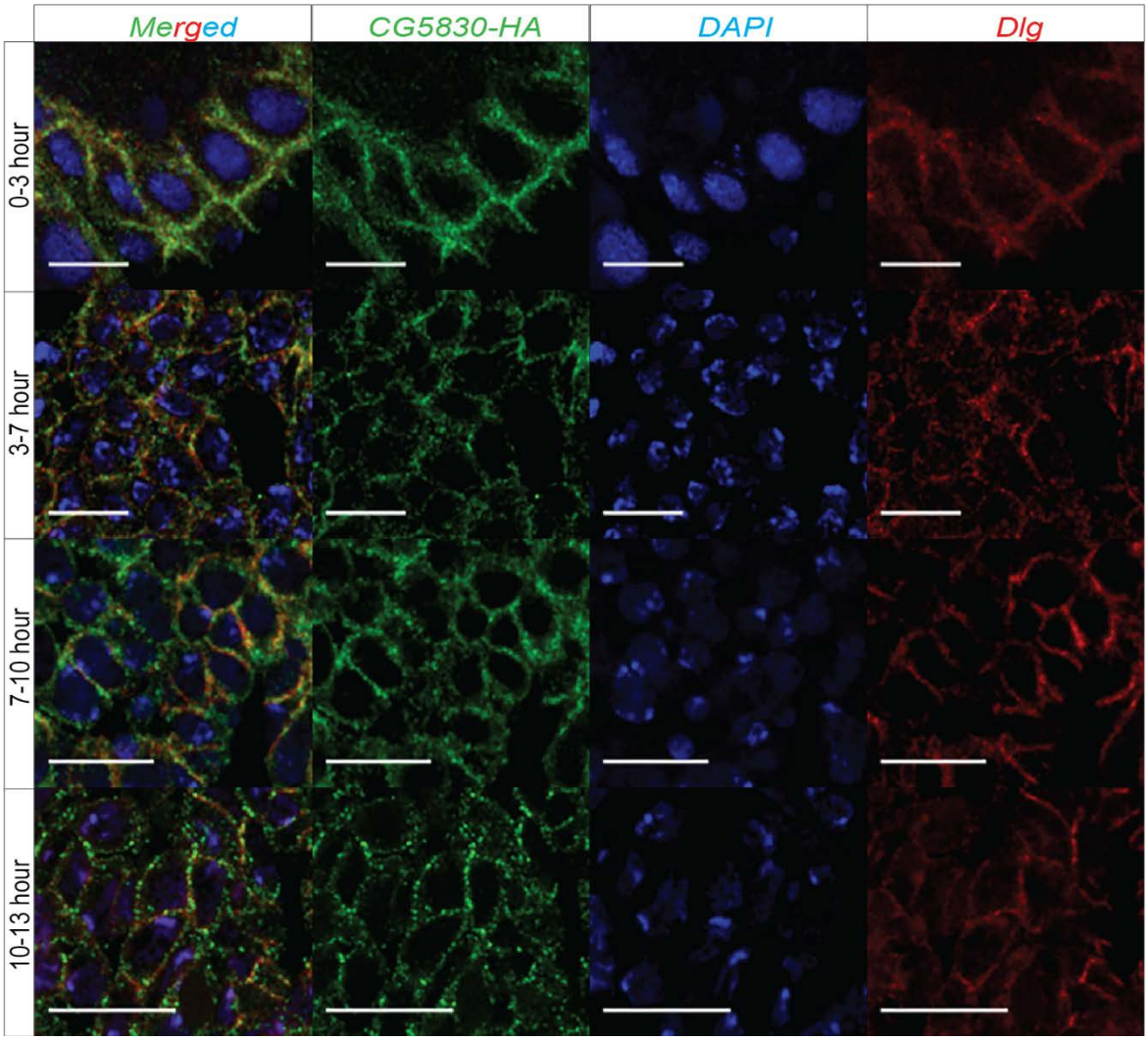


Figure 3.6 Continued.

D)*CG5830-EGFP*

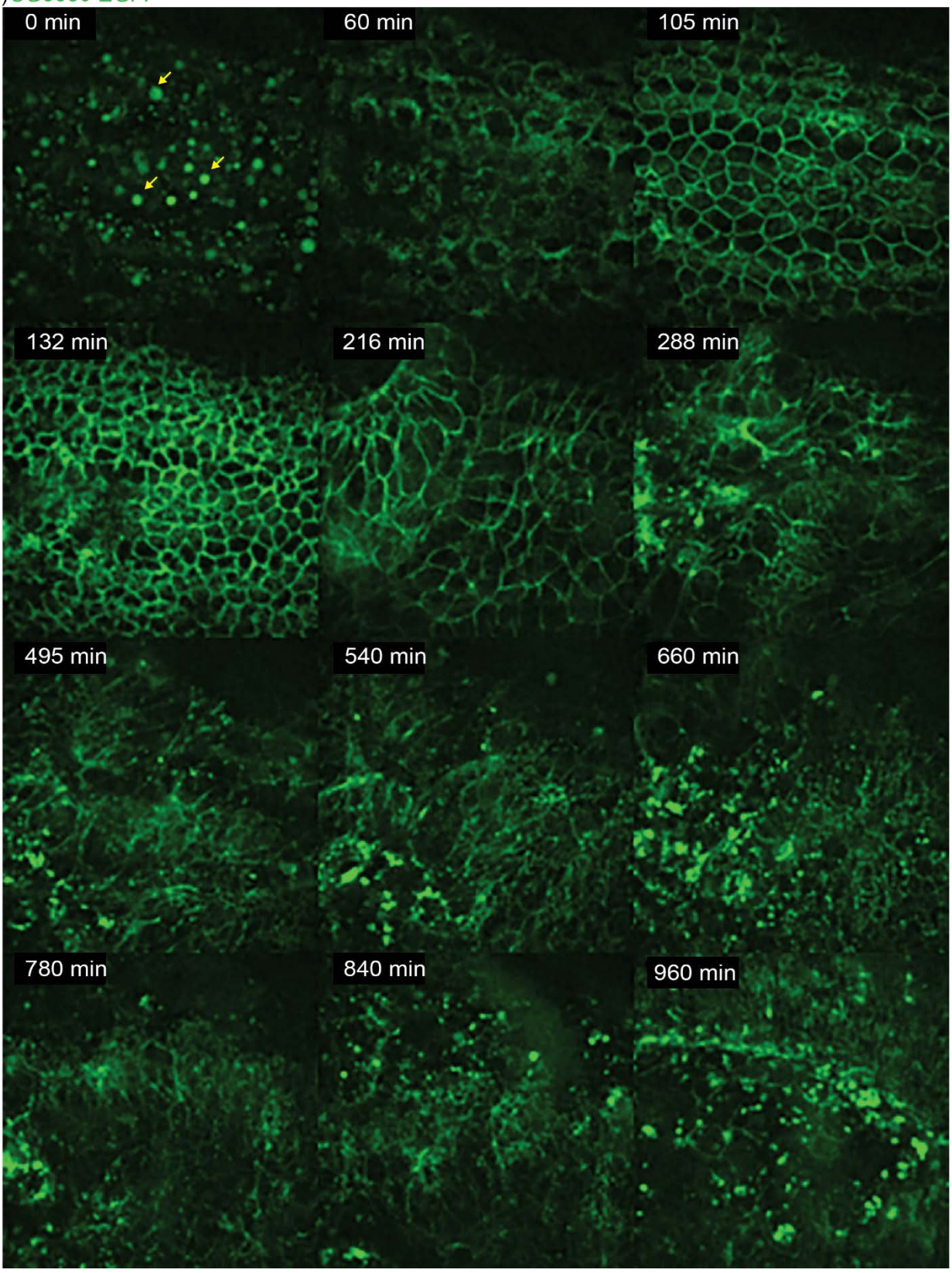
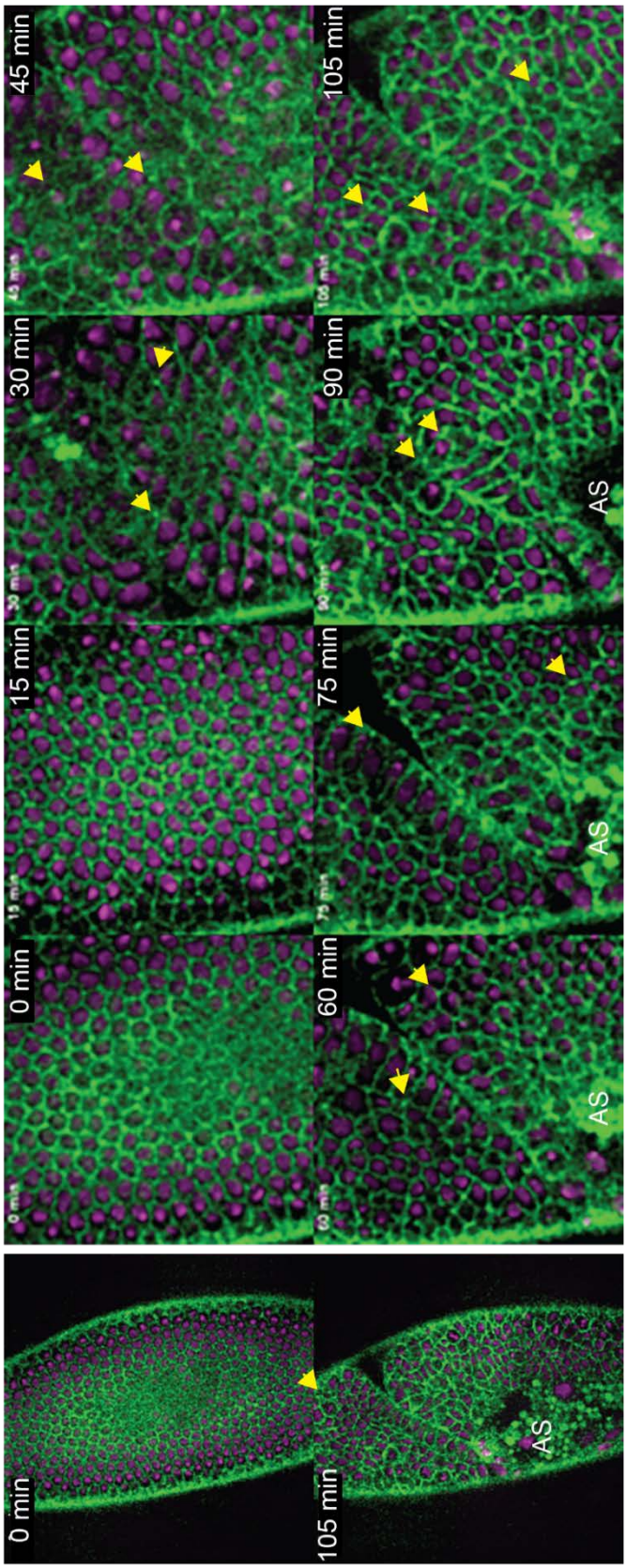


Figure 3.6 Continued.



E) *Histone2A-RFP; CG5830-EGFP*

Figure 3.6 Continued.

F) *CG5830-EGFP*

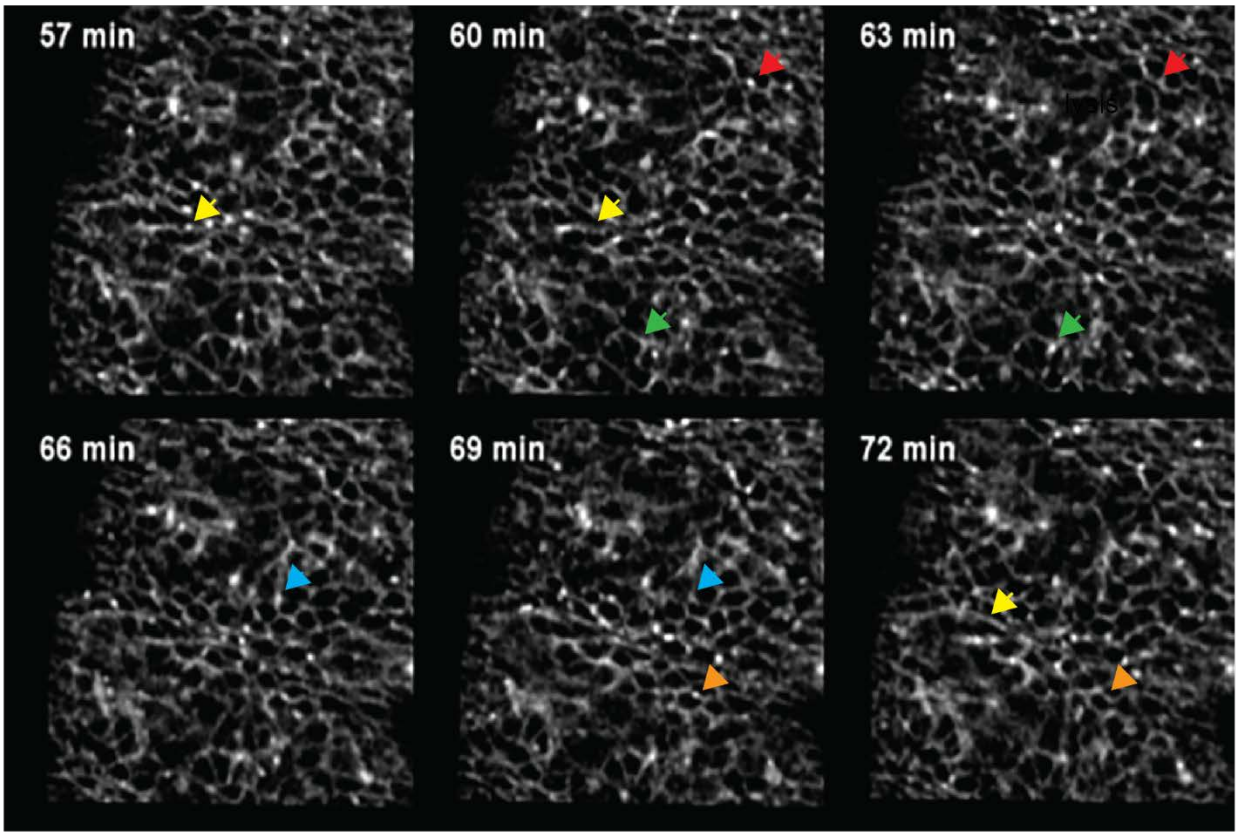


Figure 3.6 Continued.

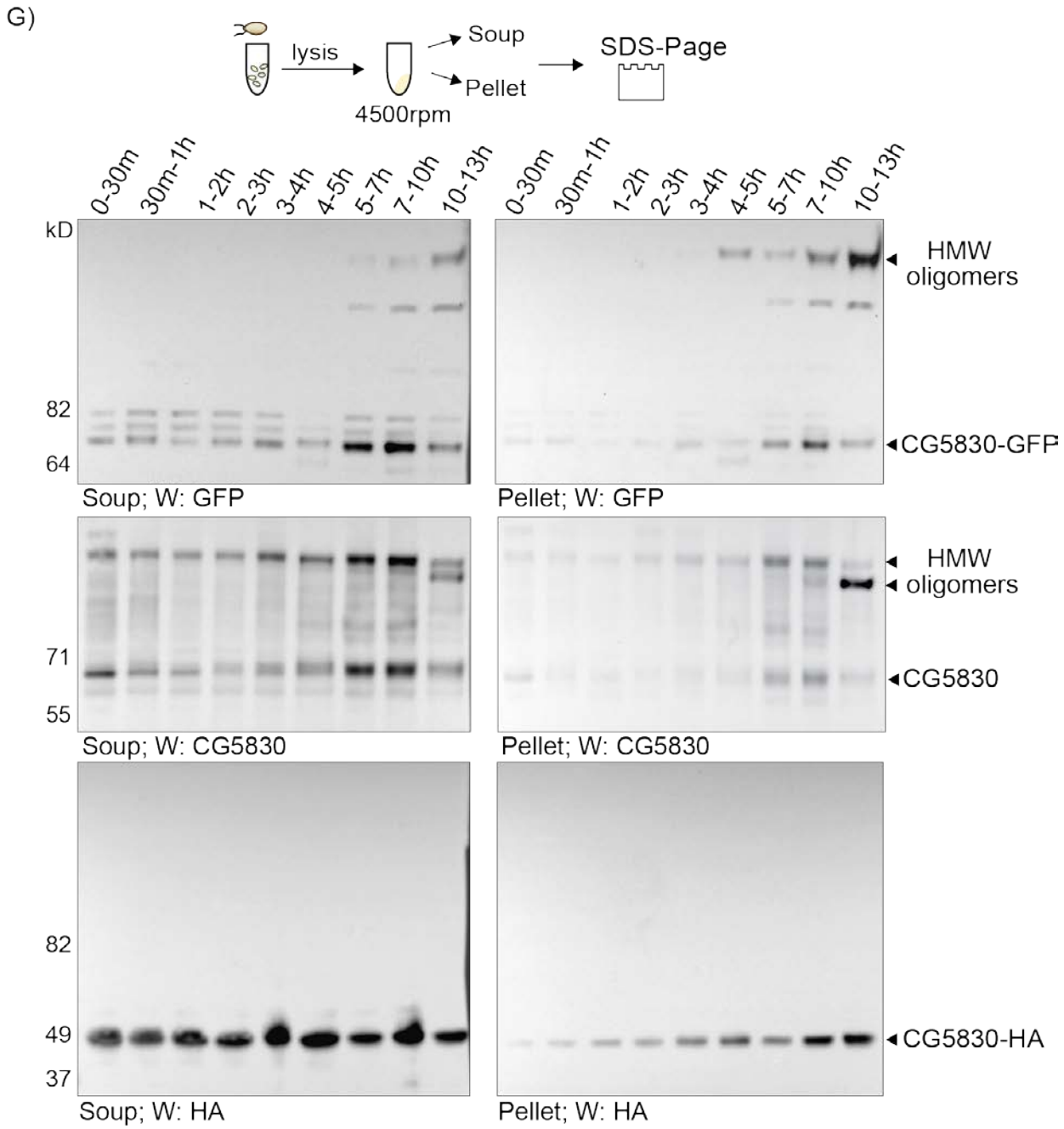


Figure 3.6 Continued.

Immunostaining of *CG5830-HA* embryos of different ages with anti-HA antibody showed that it is ubiquitously localized to membrane throughout embryonic development (Figure 3.6B). With a closer look at *CG5830-HA* embryos at these different stages with high resolution microscopy, we see that CG5830 has a diffused morphology on the membrane in 0-3-hour old embryos and it

has visible puncta/foci on the membrane after this point in 3-7, 7-10 and 10-13-hour old embryos (Figure 3.6C). To understand exactly when this punctate appearance starts and how the puncta behave on the membrane, I did live imaging with *CG5830-EGFP* and *CG5830-EGFP:Histone 2A-RFP* embryos (Figure 3.6D and E). I found that punctate appearance begins during gastrulation (3hour 10min), which coincides with the initiation of *wg* signaling. As the embryo continues to develop, these puncta persist. More detailed analysis of the images revealed that some of these puncta are dynamic such that they form and disappear along the membrane within minutes (3-6 mins) (Figure 3.6F).

In the previous experimental screen for candidate prion-like proteins, it was shown that exogenous CG5830 forms visible puncta in the cell. Puncta appearance is a common feature of prion-like proteins (Section 1.1) and it implies a conformational change. However, many proteins can form several different forms of assemblies without any conformational change, such as multimerization. To understand if the puncta of CG5830 protein results from multimerization of the protein or if there is a conformational change, I followed CG5830 protein in embryos over time with western blotting. To this end, I extracted total protein from 0-30min, 30min-1, 1-2, 2-3, 3-4, 4-5, 5-7, 7-10 and 10-13-hour old, which corresponds to major developmental events. I loaded both light (supernatant) and heavy (pellet) fractions of these samples since oligomeric species of many prion-like proteins is mostly recovered in heavy fractions. With different antibodies, including anti-GFP, anti-HA and anti-CG5830, CG5830 protein is detected with increasing expression over time in all stages of development (Figure 3.6G). Expression at the very early stages of embryonic development indicates translation of maternally deposited mRNA, which is consistent with the maternal effect phenotype observed with the mutant flies. Starting from 3-4-hour (gastrulation), a higher molecular weight CG5830 protein appears in addition to its monomeric form, detected only by anti-GFP and anti-CG5830 antibodies. This time point coincides exactly with the time point that CG5830 becomes punctate in live imaging. Interestingly,

anti-HA antibodies recognize only a band that corresponds to monomeric size of the protein. The anti-HA antibody also reveals that, although there is significant amount of monomeric protein in the initial stages of embryonic development (0-3h), they don't form the higher order assembly. This implies that aggregates are not a simple correlate of the amount of protein, but there are some other events that control the conversion of the protein from monomer to oligomer. Moreover, differential recognition of unique forms of CG5830 with different antibodies suggests there has been a conformational shift between these forms. In addition, the high molecular weight CG5830 protein is resistant to both high concentration of (2%) SDS in sample buffer and elevated temperatures during boiling, suggesting that the structure of the assembly once formed is quite stable.

Taken together, immunoblot analysis and imaging suggest that CG5830 exists in two different states during embryonic development: low molecular weight monomers and high-molecular weight oligomers which start to appear after gastrulation at a time when the role of CG5830 in segment polarity is required. Considering both the stability of CG5830 oligomers and dynamic nature of CG5830 puncta, these data also implies that aggregation and disaggregation is highly regulated.

3.2.4 Identification of CG5830 protein complex

Previously, I observed that CG5830 dephosphorylated CTD tail of RNAPII in isolation and that it acts as an antagonist against *wg* pathway. However, it is unclear: what signaling pathway(s) does CG5830 regulate? and what proteins does it dephosphorylate? To answer these questions and to understand the molecular function of CG5830, I used a proteomics-based approach to perform a comprehensive search for CG5830 interacting proteins. To this end, WT and PD CG5830 were expressed stably in an embryonic S2 cell line, as C-terminal HA/Flag-tagged proteins. The CG5830 protein complex was immunopurified using anti-Flag antibodies from both cell lines and the components of CG5830 protein complex were identified using Multidimensional

Protein Identification technology (MudPIT) (Figure 3.7). Two types of statistical analyses were performed to define significantly enriched proteins in the WT CG5830 Flag-affinity purifications (APs) compared to Flag negative controls and Flag-APs with PD CG5830. Proteins were considered significantly enriched if their STN (signal-to-noise) values were greater than 0.9 with a p-value lower than 0.05; and/or if their log2 Fold Change values were greater than 2 with a p-value lower than 0.05. Proteins also had to be detected in at least 2 of the three replicate analyses. Proteins passing these thresholds are listed in Appendix 1 (Table 7.1) and a String protein-protein interaction network for these proteins is given in Figure 3.7B. The String analysis demonstrates both known and predicted interactions of proteins, which are depicted by distinct colors of the connecting lines between proteins. This analysis of CG5830 interactome shows that RpII215 (RNA polymerase II 215kD subunit) is a putative CG5830 interactor. However, this interaction between CG5830 and RpIII215 could not be verified by co-immunoprecipitation experiment (Figure 3.7C).

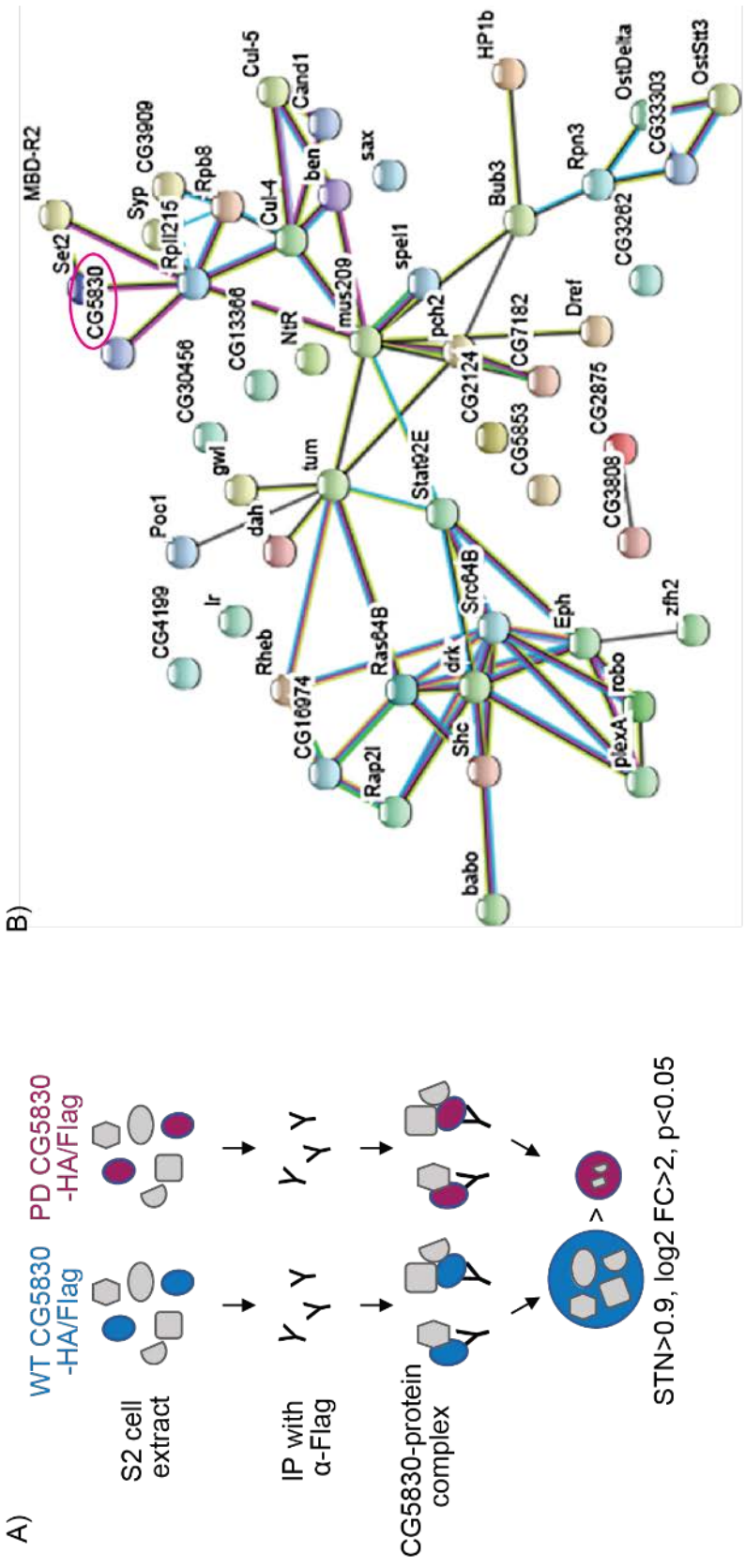


Figure 3.7 S2 cell proteomics identified CG5830 interacting proteins.

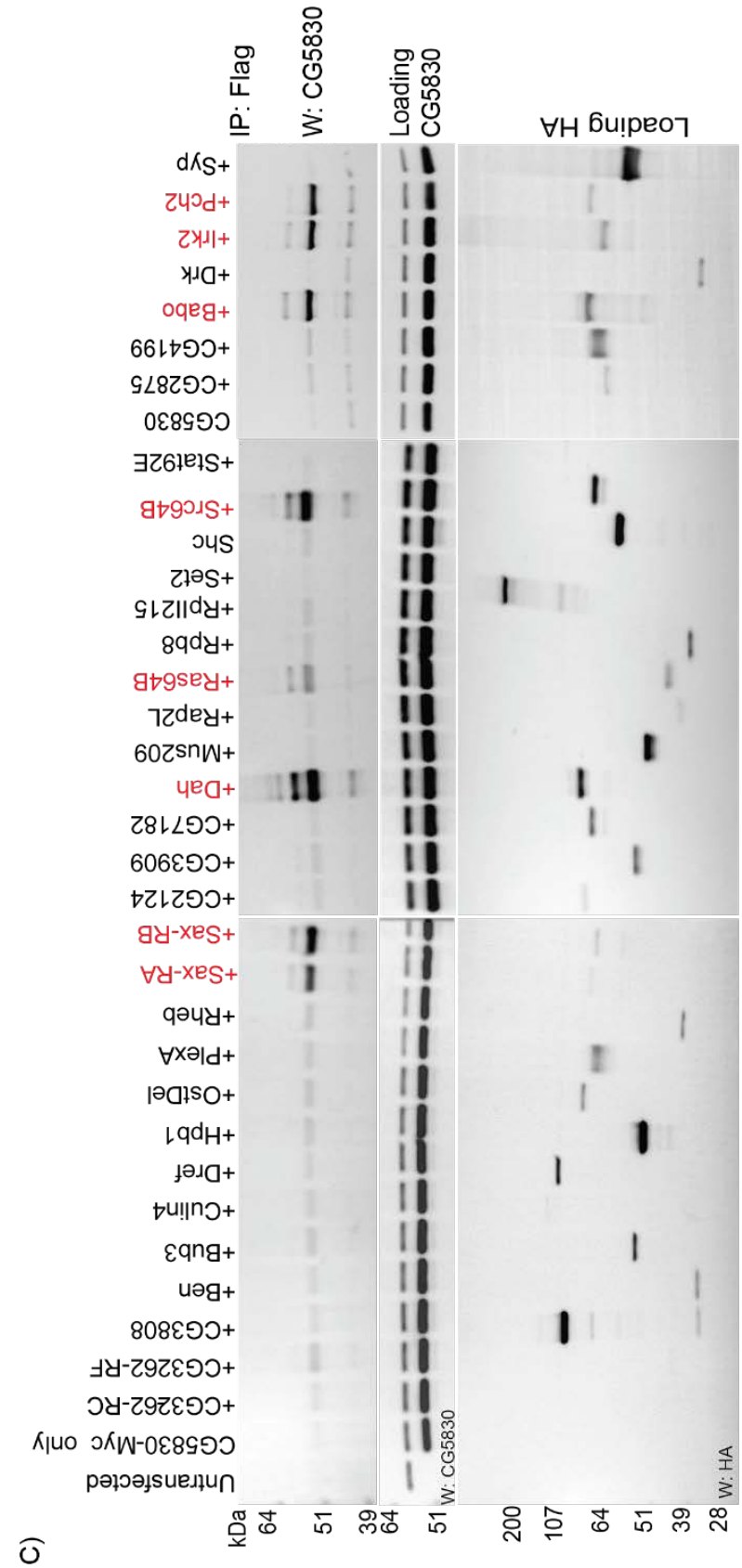


Figure 3.7 Continued.

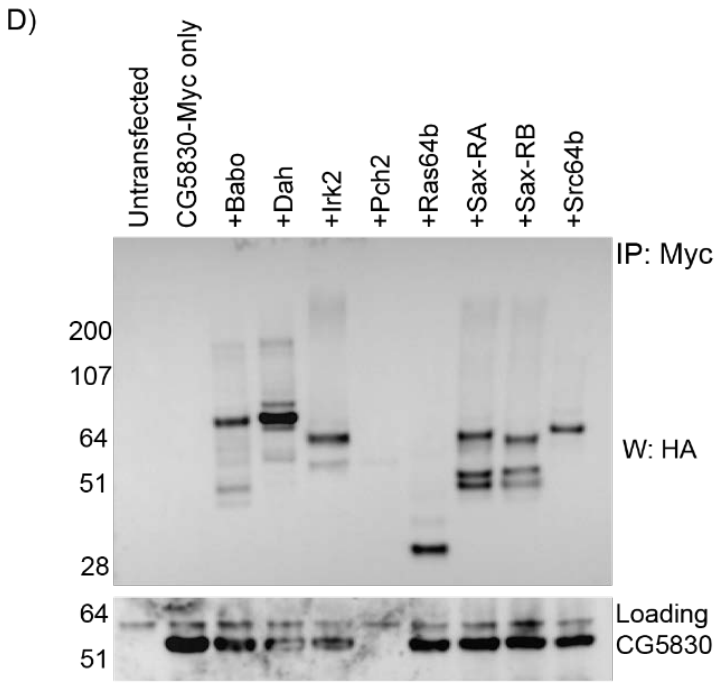


Figure 3.7 S2 cell proteomics identified CG5830 interacting proteins.

A) Schematic of the proteomics analyses from S2 cells. WT and PD CG5830 interacting proteins from S2 cells were immunoprecipitated with Flag antibody, and proteins significantly enriched in WT immunoprecipitate were further analyzed.

B) Network analysis results for significantly enriched proteins together with CG5830 (circled in pink) visualized by STRING. Different colors of the lines represent the types of evidence for association: green line, neighborhood evidence; red line, fusion evidence; purple line, experimental evidence; light blue line, database evidence; black line, co-expression evidence; blue line, co-occurrence evidence; and yellow line, text-mining evidence.

C) Pairwise interaction study of 31 interacting proteins (IP for interactors, blot for CG5830) revealed Saxophone (Sax), Discontinuous actin hexagon (Dah), Ras64b, Src64B, Baboon (Babo), Inwardly rectifying potassium channel (Irk2), and Pachytene checkpoint protein 2 (Pch2). The low level of interaction between Ras64B and CG5830 is due to lower expression of Ras64B. Lower panels show the loading controls.

D) Reciprocal co-immunoprecipitation for CG5830 interacting proteins (IP for CG5830, blot for interactors. The lack of interaction between Pch2 and CG5830 is due to lack of expression of CG5830 in that sample. Lower panel show the loading controls.

To validate the interactions which were found significant in MudPIT analysis, I did pairwise interaction studies using Co-immunoprecipitation (Co-IP) assays in S2 cells (Figure 3.7B and C). 7 out of the 31 proteins tested (Table 7.1), namely Baboon (Babo), Discontinuous actin hexagon (Dah), Inwardly rectifying potassium channel 2 (Irk2), Pachytene checkpoint protein 2 (Pch2), Ras oncogene at 64B (Ras64B), Saxophone (Sax) and Src oncogene at 64B (Src64B), formed a complex with WT CG5830. Reciprocal co-immunoprecipitation experiment also validated these interactions. Except Pch2, all these interactors are either located on the membrane (Babo, Sax, and Irk2) or associates with membrane proteins (Ras64B, Src64B and Dah) based on uniprot and FlyBase.

When we look at the pathways where these interacting partners function, there are four main groups: TGF-β signaling (Babo, Sax and Irk2), EGF/FGF signaling (Ras64B and Src64B) and cell division (Dah and Pch2), which are detailed in section 1.7. Moreover, except Pch2, all of these interacting partners have been previously reported to be phosphorylated (Lusk et al., 2017; Parker et al., 2003; Wouda, Bansraj, de Jong, Noordermeer, & Fradkin, 2008; C. X. Zhang et al., 2000) suggesting they may be substrates of CG5830.

3.2.5 A system to detect CG5830 enzymatic activity

To determine whether any of these interacting proteins are putative substrates of CG5830, I first designed an experiment which would allow me to detect both phosphorylation and the changes in the phosphorylation status of the interactors. To this end, I took advantage of Phos-tag™ agarose beads, which specifically binds and enriches phosphoproteins and phosphorylated peptides at a physiological pH of 7.5. I first analyzed the phosphorylation status of the interacting proteins in S2 cells. For this, C-terminal HA-Flag tagged interactors were expressed in S2 cells and all phosphorylated proteins in the total lysates were immunoprecipitated using Phos-tag™ beads and probed for bead-bound HA-tagged substrates. Among the seven CG5830-interacting proteins, Dah, Irk2 and Src64b were found to be readily phosphorylated in S2 cells (Figure 3.8A).

To determine whether this assay system can be used to detect dephosphorylation of these three targets, I treated the cell extract with nonspecific phosphatases, calf intestinal phosphatase and lambda phosphatase, followed by Phos-tag™ immunoprecipitation. Prior treatment with phosphatases significantly decreased the binding of target protein to Phos-tag™ beads compared to untreated controls (Figure 3.8B).

Since we know that CG5830 is expressed early in embryogenesis and changes its state during gastrulation, for next set of experiments I picked proteins based on two criteria: reported protein expression during early embryonic development (0-4h) and reported phosphorylation during this developmental time. Except Irk2, which is expressed in the larva, both Src and Dah fit into both criteria (Casas-Vila et al., 2017). I specifically focused on Dah because: it is highly phosphorylated during cellularization and gets dephosphorylated during gastrulation (C. X. Zhang et al., 2000). To test the effect of CG5830 on Dah phosphorylation, I overexpressed HA-Flag tagged Dah in the presence of either WT or PD CG5830 and pulled down phosphorylated proteins in the total lysates using Phos-tag™ beads and detected the phosphorylation status of Dah with anti-HA antibody. After quantification and normalization of the phosphospecific bands to the total amount of Dah in the lysate, I observed that while WT CG5830 decreased the level of phosphorylated Dah, PD CG5830 increased its level suggesting a dominant negative effect of enzymatically dead protein on the endogenous CG5830 (Figure 3.8C). The dominant negative effect of PD version also shows that Dah is specifically dephosphorylated by CG5830.

In the following section, using the same experimental setup with Dah as a substrate, I investigated the effect of conformational change of CG5830 on its enzymatic activity.

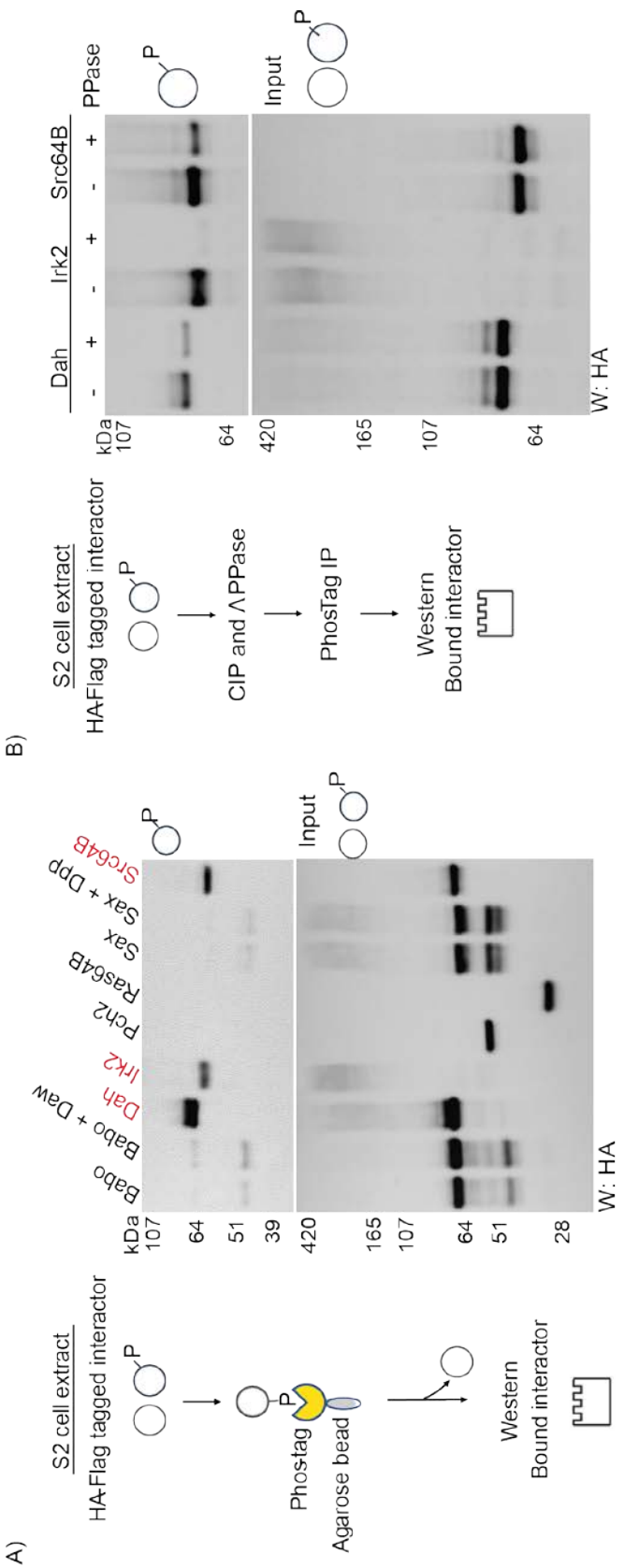


Figure 3.8 A system to detect CG5830 phosphatase activity.

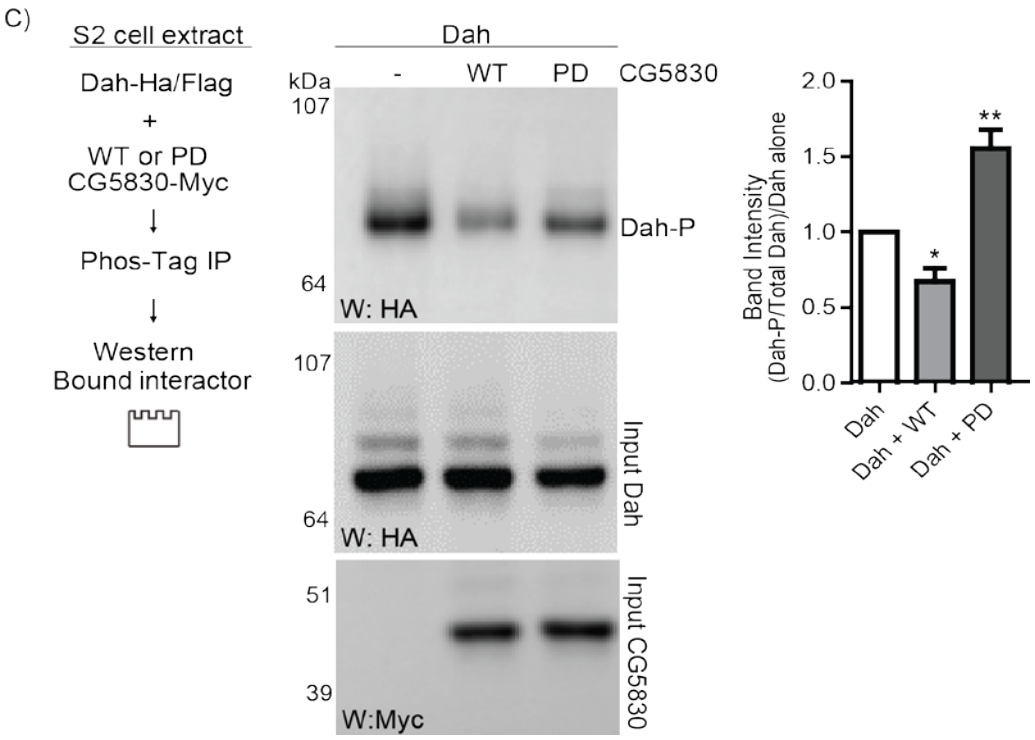


Figure 3.8 A system to detect CG5830 phosphatase activity.

A) Phosphorylation status of interactors in S2 cells. Left panel: Schematic of the experiment. Right top: HA-Flag tagged proteins are expressed in S2 cells and phosphorylated forms are pulled down by phosphospecific agarose beads, and blotted with anti-HA antibody. Daw (Dawdle) and Dpp (Decapentaplegic) ligands are used to induce phosphorylation of Babo and Sax receptors, respectively. Dah, Irk2 and Src64B showed phosphorylation. Right bottom panel show the loading controls.

B) Detection of dephosphorylation by nonspecific phosphatases. Left panel: Schematic of the experiment. Right top: Phosphorylated proteins are pulled down from lysates (Dah, Irk2, Src64B) treated or not treated with phosphatases and compared with each other. Right bottom panel show the loading controls.

D) Detection of phosphatase activity of CG5830 on Dah. Left top panel: Schematic of the experiment. Right top: Dah phosphorylation status was detected in the lysates co-transfected with either wild type (WT) or phosphatase dead (PD) CG5830. Intensity of phosphorylated bands were measured and normalized by the band intensity of total Dah in each sample. Normalized band intensities for each group from 4 independent experiments were plotted as mean ± SEM, and statistical significance was determined by one-way ANOVA for multiple groups, $p<0.01$.

Chapter IV: Characterization of prion-like properties of *CG5830* and their effect on its function

4.1 Objectives and Rationale

Prion-like proteins are known to exist in at least two stable conformational states in the cell: monomers and oligomers. Because of the self-assembling dominant nature of the oligomers, the conformational state change of these proteins acts like a switch. This switch does not only affect the structure of the protein, but also it can alter the functional state of it from one function to the other either causing a loss of function or gain of a new function. In section 1.8, I gave details of such structural-functional protein switches and how they can regulate normal processes, such as adaptation to new environmental conditions through regulation of transcriptional or translational changes (yeast Swi1, Cyc8, Sup35, Ger, Mot3), immune response (MAVS), and memory formation through regulation of translation (Orb2, CPEB). This suggests that prion-like protein switches can be applied to many different normal physiological processes. However, because of the lack of *in vivo* evidence, especially in multicellular eukaryotes, it remains unclear how commonly this mechanism is utilized in normal physiology of organisms.

In the previous chapter I demonstrated that CG5830, a strong candidate for a prion-like protein, is a ubiquitous membrane phosphatase required for segment polarity during embryogenesis. I identified its interactors and potential substrates, and I found that it specifically dephosphorylates Dah. I also found that CG5830 is diffuse on the membrane until gastrulation and after that point it forms puncta. Most importantly, these two forms are conformationally distinct from each other, as demonstrated by different antibodies, and a high molecular weight state can be detected at the time point that corresponds to protein activity in gastrulation. This suggests that CG5830 may utilize a prion-like switch to alter its function during embryogenesis and regulate segment polarity. Therefore, in this part of my thesis, to understand whether the prion-like properties of CG5830 contribute to its function, I set the following objectives:

a) Determine the prion domain of CG5830, and its effect on enzymatic activity of CG5830

b) Characterize structural and functional differences of monomeric and polymeric CG5830

4.2 Results

4.2.1 N-terminal PrD of CG5830 is important for its oligomerization and membrane localization

Prion-like proteins often contain a modular disordered domain called prion-like domain or PrD, which is necessary and sufficient for self-propagation of the protein but can be dispensable for its biochemical function (Alberti et al., 2009). CG5830 has two disordered domains at N- and C-terminal ends, which are predicted to be unfolded by the *FoldIndex* program (Prilusky et al., 2005) as shown in Figure 4.1A. The middle domain corresponds to the conserved FCP homology domain. To determine the relationship between various domains of CG5830 and its biophysical properties, I assigned these regions as N (1-78), M (79-244) and C (245-329) domains and expressed these domains of CG5830 alone or in combination, namely N, NM, M, MC, C, as C-terminal-tagged proteins and addressed three questions; i) do they interact with full-length protein?; ii) does this interaction enhance full-length protein oligomerization?; iii) do they form stable foci in the cell?

To answer the first question, I did a co-immunoprecipitation experiment. Full-length *CG5830-Myc* tagged was transfected alone and in combination with each *HA-Flag* tagged construct to S2 cells. First, full-length protein was immunoprecipitated with anti-Myc beads and blotted for the truncated versions of CG5830 with anti-HA antibody. I observed that except C domain, N and M domains of CG5830 interact with full-length protein Figure 4.1B.

Second, I sought to determine SDS-resistant oligomer formation of CG5830 full-length protein in the presence of the truncated versions. Whole cell lysates were prepared from co-transfected and full-length alone transfected S2 cells and run on SDS-agarose gel. Compared to regular SDS-

PAGE gels, these agarose gels have bigger pores which can accommodate higher molecular weight polymers. Moreover, in this assay, prion-like proteins due to their resistance to SDS would migrate slowly and appear as smears on western blot. By immunoblotting with anti-CG5830 antibody, I observed increased oligomer formation compared to full-length protein alone only in the presence of N domain (Figure 4.1C).

To understand what part of CG5830 forms stable foci in the cell, I expressed GFP tagged full length and truncated versions of CG5830 in S2 cells. Fluorescence microscopy analysis of these domains alone showed that N-GFP and NM-GFP formed several, and M-GFP and MC-GFP formed one large fluorescent puncta in cells, while C-GFP showed a diffused profile (Figure 4.1D). FRAP analysis on cytoplasmic puncta of all constructs, except C-GFP, showed significantly less recovery after photobleaching compared to glutamine repeats, indicating a more stable structure. Through this analysis, I also observed that membrane localization of CG5830 happened only in the presence of N domain. I concluded that N domain of CG5830 is important for its oligomerization and required for its membrane localization.

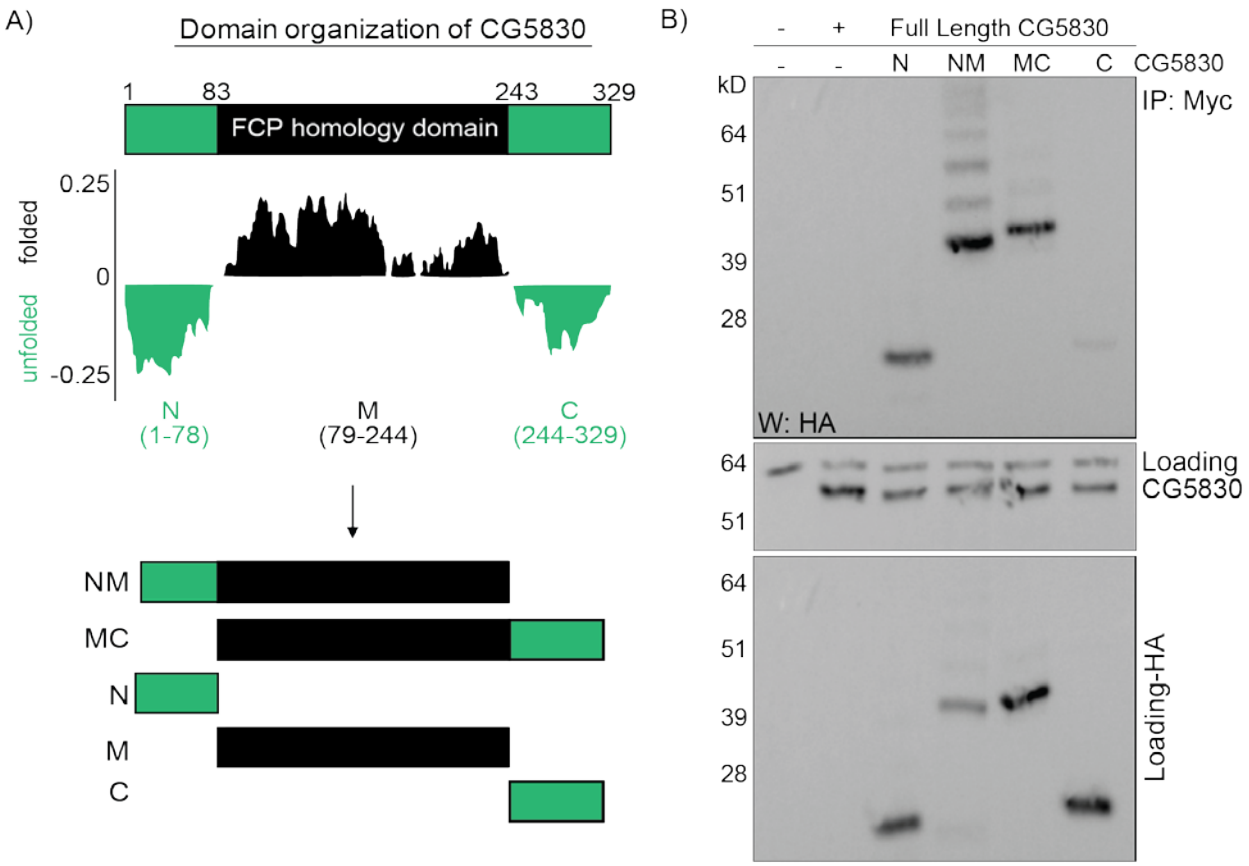


Figure 4.1 N-terminal PrD of CG5830 is important for its oligomerization and membrane localization.

A) Domain organization of CG5830. The N- (0-78) and C- (245-329) terminal ends of CG5830 are predicted to be unfolded using the Foldindex program (Prilusky et al., 2005). Middle region corresponding to the conserved FCP homology domain is folded.

B) Interaction of Full length CG5830 with truncated versions of itself. Top panel: N and M domains are interaction domains of CG5830. Bottom panel show the loading controls.

C) Oligomerization of CG5830 is enhanced by its N-terminal PrD. CG5830 oligomers detected by anti-CG5830 with SDS-Agarose gel electrophoresis (1.5%) in total lysates of full-length alone and full-length co-transfected with truncated versions. Loading control blots shown in (B) are loading controls of this blot, too.

D) Confocal microscopy images (top) and FRAP analysis (bottom) of different versions of CG5830 shows that it forms stable puncta in the cell. N terminal PrD is also required for membrane localization. Blue: DAPI. Scale bar: 5um

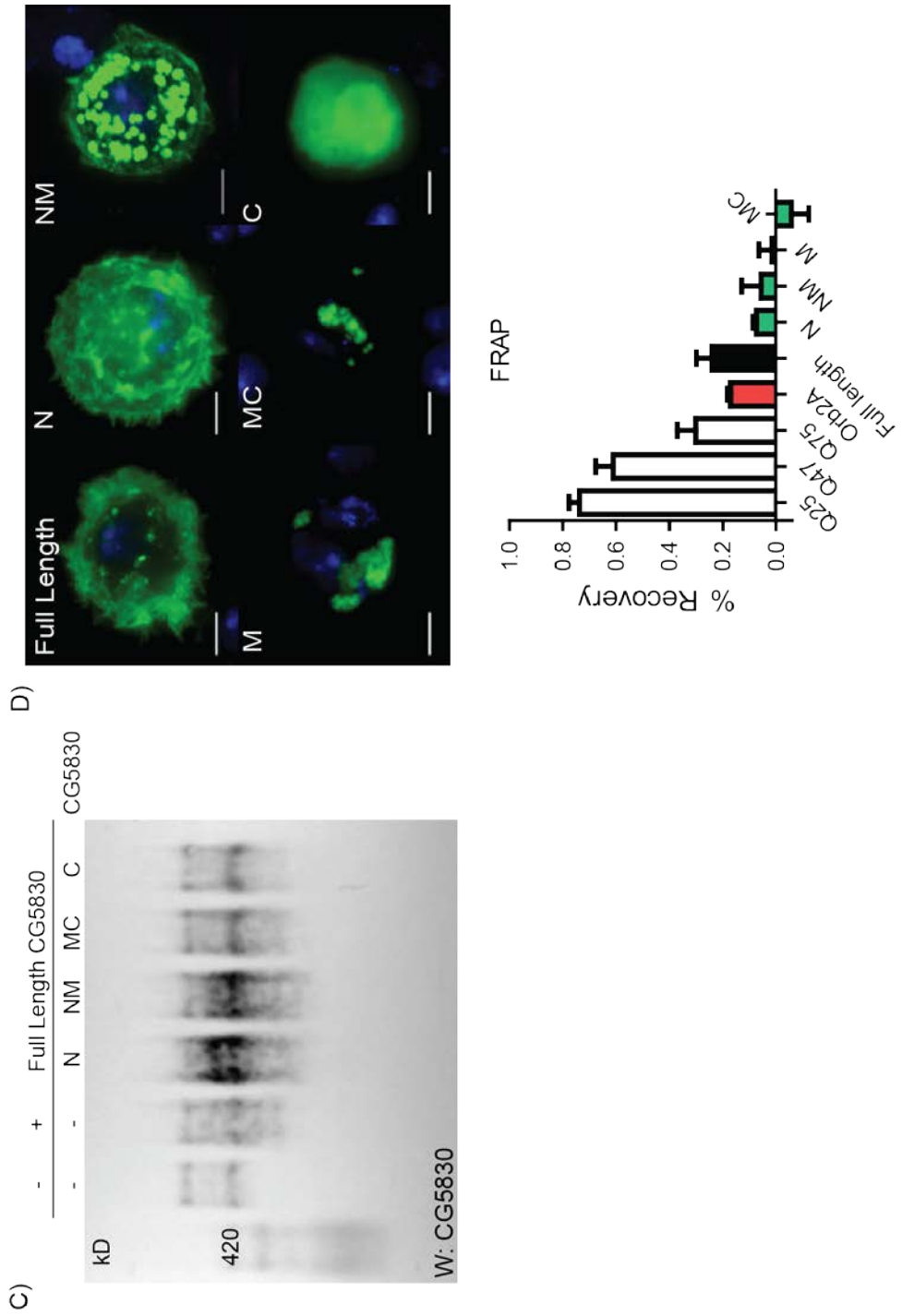


Figure 4.1 Continued.

4.2.2 CG5830 phosphatase activity is modulated by its N-terminal PrD

To characterize the effect of oligomerization on the phosphatase enzymatic function of CG5830, I utilized the same experimental assays of enzymatic activity of WT- and PD- CG5830 on its specific substrate Dah. To this end I immunoprecipitated Dah using Phos-tag™ beads in the presence of different truncated versions of CG5830. I detected increased dephosphorylation whenever N domain is present, namely in full length, N and NM domain co-expressed cells (Figure 4.2A).

The N-domain can enhance phosphatase activity in three ways: i) increased substrate recognition through N domain; ii) increased substrate availability through membrane localization by N domain; iii) increased enzyme activity through oligomerization by N domain. To distinguish between these possibilities, I first evaluated which domain of CG5830 recognizes Dah by immunoprecipitating Dah in the presence of different regions of CG5830 and scored for interaction. I observed that M domain is required for Dah to be recognized by CG5830 (Figure 4.2B). To test for the second putative mechanism, I analyzed possible membrane targeting sites in N terminal domain and found two cysteines at aa positions 46 and 47 that can be potentially palmitoylated. Palmitoylation is known to induce membrane association (Veit & Schmidt, 1998). I mutated both cysteine residues to alanine by site-directed mutagenesis. Additionally, I targeted MC domain to cell membrane by attaching a nonspecific membrane targeting site, CAAX motif of *Drosophila* Ras2, to the N terminal of MC domain of CG5830 (Choy et al., 1999; Kakihara, Shinmyozu, Kato, Wada, & Hayashi, 2008). Fluorescence microscopy of these GFP tagged mutant constructs showed that C46C47 to A46A47 mutation did not change the localization (Figure 4.2C) or activity of CG5830 (Figure 4.2D). Interestingly, MC domain that by itself does not localize to the membrane, upon addition of CAAX motif resulted in membrane localization of MC domain (Figure 4.2C). However, membrane localization did not rescue the reduced phosphatase activity of MC (Figure 4.2D), suggesting membrane localization alone is not sufficient to increase

enzymatic activity. This leaves the third possibility that oligomerization in the membrane through N terminal domain increases enzymatic activity of CG5830. Since my attempts to eliminate oligomerization without changing subcellular localization of CG5830 failed, with this experimental assay above I could not reach a clear conclusion that the effect on enzymatic activity only results from polymerization of CG5830. Hence, I decided to test the activity of separated CG5830 monomers and oligomers *in vitro*.

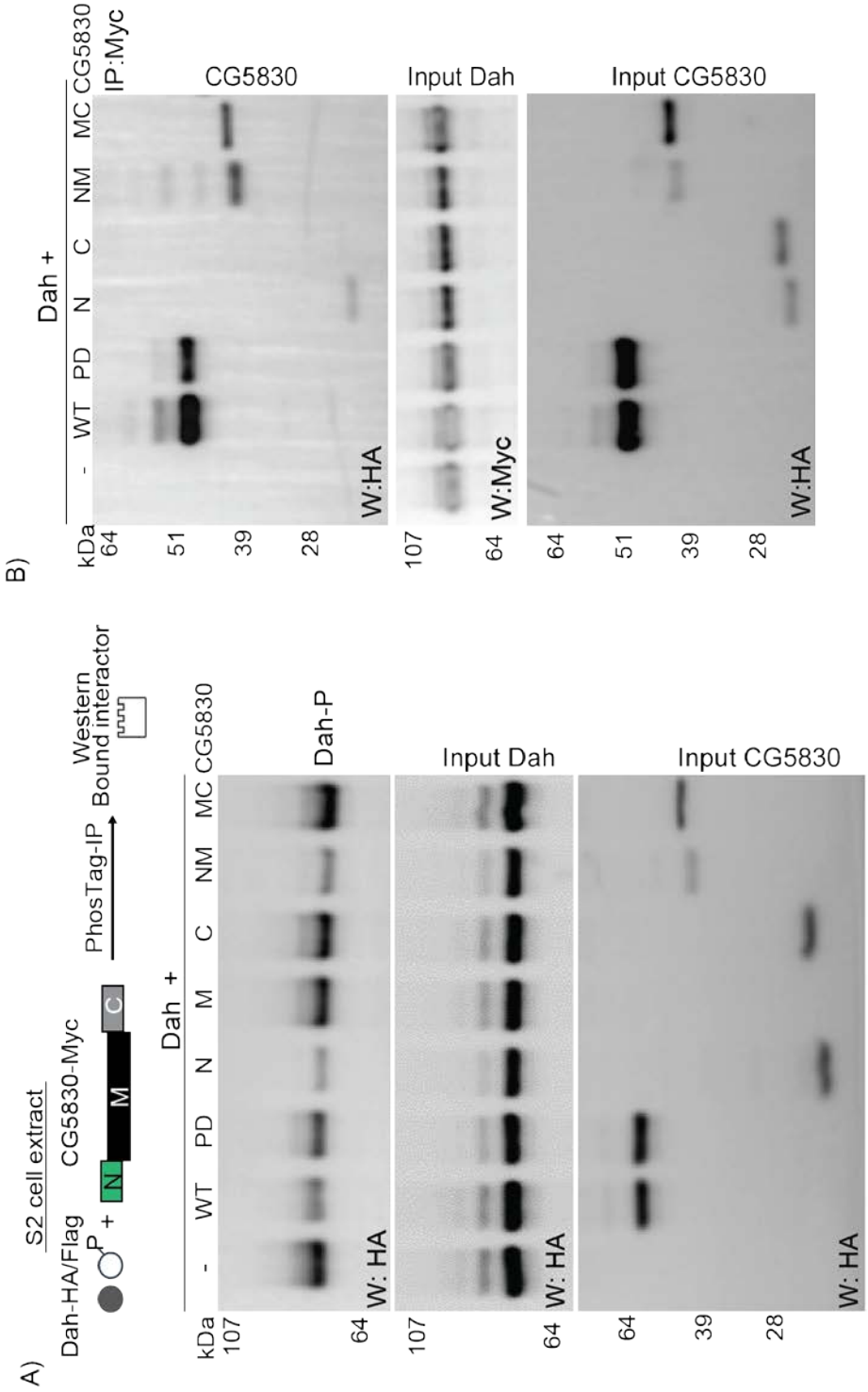


Figure 4.2 CG5830 phosphatase activity is modulated by its N-terminal PrD

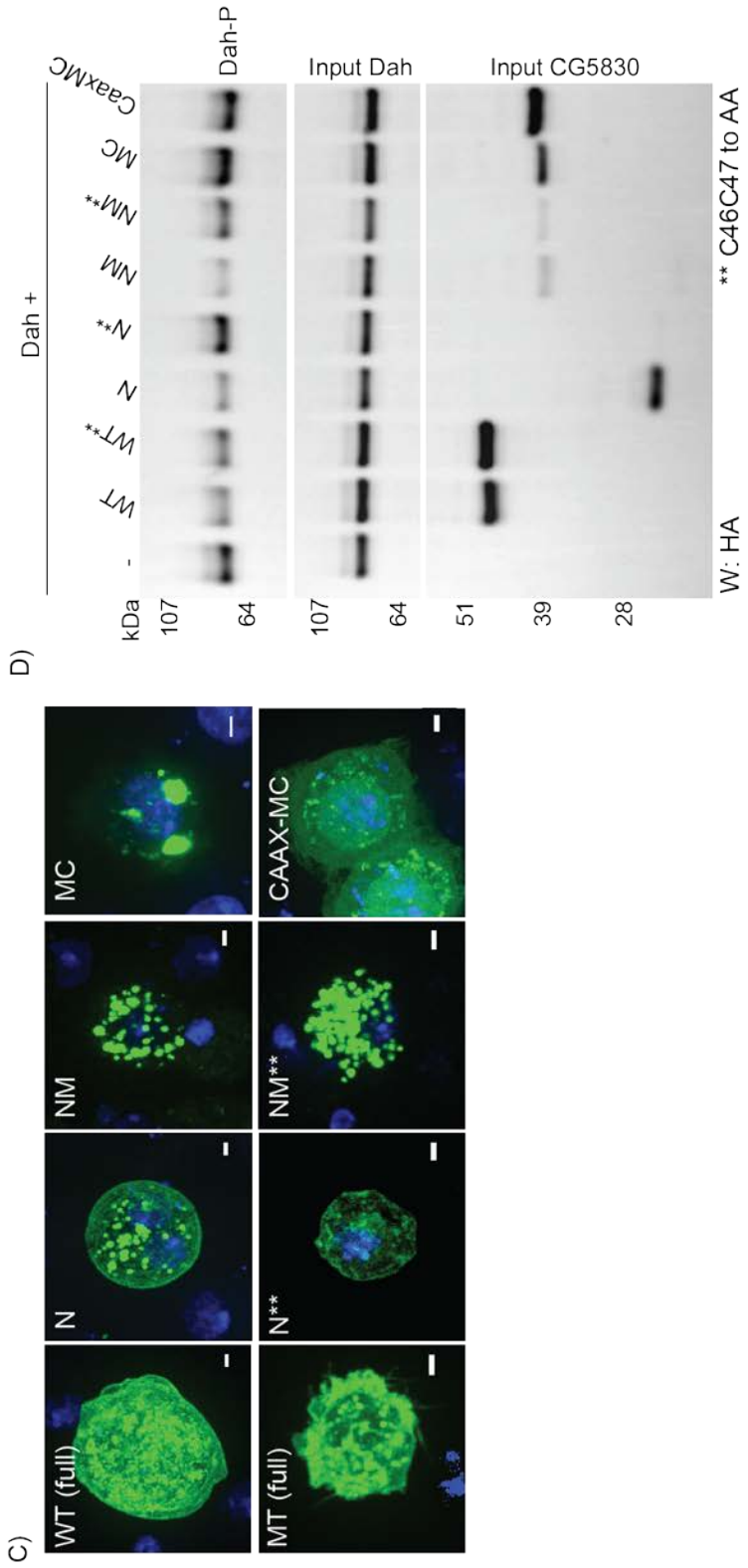


Figure 4.2 Continued.

Figure 4.2 CG5830 phosphatase activity is modulated by its N-terminal PrD

A) Effect of truncated versions of CG5830 on Dah phosphorylation was measured by Phosho-Tag assay. In presence of N terminal PrD, amount of phosphorylated Dah decreases compared to others. M domain construct does not express and serves as a negative control. Bottom panels show loading controls.
B) CG5830 recognizes its substrate through its catalytic domain, M. Co-immunoprecipitation of Dah with different versions of CG5830 shows interaction in the presence of M domain.
C) Fluorescence microscopy of membrane targeting chimeric and mutant constructs. There is no alteration in subcellular localization of CG5830, N and NM domains with C46C47 to A46A47 (indicated as **) point mutations. Addition of CAAX to N terminal site of MC domain localizes it to membrane. Top panel: wild type controls of each version. Blue: DAPI. Scale bar: 5um.
D) Effect of mutant and chimeric constructs of CG5830 on Dah phosphorylation. Membrane localization of MC domain did not change the activity. High phosphorylation seen in N** sample is because of low expression level of N** compared to N. Bottom panels show loading controls.

4.2.3 Endogenous CG5830 oligomers are amyloid-like and shows phosphatase activity

Previous experimental *in vitro* assays using exogenous CG5830 showed that it has amyloid-like structure and high aggregation propensity. Additionally, it forms visible and stable puncta in cultured cells and the oligomeric species is heat and detergent resistant (Figure 1.1). It is known that cellular milieu is a critical factor for protein states/aggregation. To analyze if the high-molecular weight state of CG5830 detected in embryos has similar characteristics to recombinant CG5830, we isolated the CG5830 protein from 0-20-hour old *CG5830-HA* embryos using different chromatographic approaches. In addition to CG5830-HA protein, we purified protein from *CG5830*^{C91} flies to use as a negative control in activity and structural assays. This mutant line expresses a truncated version of CG5830 which contains only a 11 KDa N-terminal PrD region while missing the phosphatase domain.

Whole-embryo lysates from both full-length and N-terminal samples were separated by using an anion exchange chromatography followed by a cation exchange chromatography. To obtain both the monomeric and oligomeric CG5830 species separated from each other, the fractions that contain the protein were combined and fractionated using a superose 6-size exclusion column (Figure 4.3A). Selected fractions were run on SDS-agarose gel electrophoresis and blotted with anti-HA (for CG5830-HA) or anti-CG5830 (for CG5830^{C91}) antibodies. Both CG5830-

HA and CG5830^{C91} protein were recovered in fractions consistent with two distinct physical states: monomeric and SDS-resistant oligomeric species (Figure 4.3B).

I wanted to explore how polymerization affects CG5830 phosphatase activity. I entertained three possibilities; i) monomer is inactive and it only becomes active upon oligomerization; ii) oligomerization increases the activity of the monomer by changing avidity; or iii) both monomer and oligomers are active, however they have altered substrate specificity. To distinguish between these possibilities, I used CG5830-HA monomeric and oligomeric conformations separated during the gel filtration and tested their phosphatase activity *in vitro* to understand if there is any difference in enzymatic activity between CG5830 monomers and oligomers. For this, I used a biochemical substrate specific for acid phosphatases called DiFMUP (6,8-Difluoro-4-Methylumbelliferyl Phosphate). I tested the activity of various oligomeric and monomeric of CG5830-HA. To remove any contaminant effect, I normalized the activity of CG5830-HA fractions with the activity of corresponding fractions in CG5830^{C91}. Results indicate that only oligomeric CG5830 has phosphatase activity. This supports the first possibility, indicating that oligomerization of CG5830 leads to a gain of enzymatic function (Figure 4.3C).

Next, I asked what is the nature of the oligomeric CG5830 species? To address this question, fractions containing monomeric or oligomeric CG5830-HA and CG5830^{C91} proteins were used in several biophysical assays. Oligomeric CG5830 showed higher binding to Thioflavin T (Figure 4.3D), which is a benzothiazole dye that exhibits increased fluorescence upon binding to amyloid structures (Khurana et al., 2005). Moreover, only oligomeric forms of both proteins showed reactivity with both an anti-oligomeric antibody A11 (Kayed et al., 2003) and an anti-fibrillar antibody OC, which has been tested to recognize amyloid fibers for different amyloid-like proteins involved in neurodegeneration (Kayed et al., 2007) (Figure 4.3E). Another feature of prion-like proteins is high resistance to proteinase-K treatment. Monomeric and oligomeric fractions were treated with increasing concentrations of proteinase-K. While monomeric CG5830-HA and

CG5830^{C91} proteins were digested at low concentrations of proteinase-K, oligomeric form was still present at high proteinase-K concentration (Figure 4.3F).

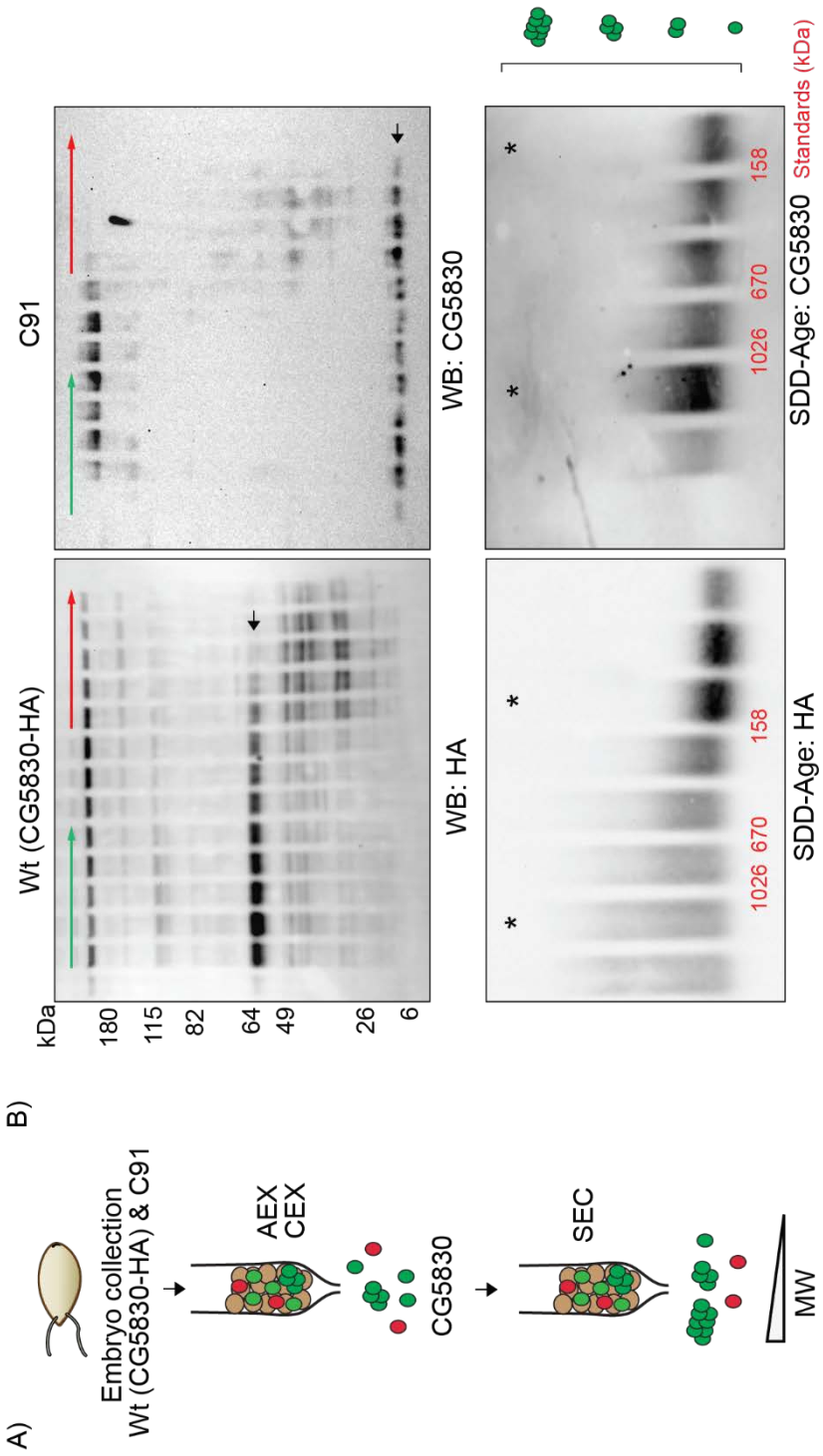


Figure 4.3 Endogenous CG5830 oligomers have phosphatase activity.

Figure 4.3 Continued.

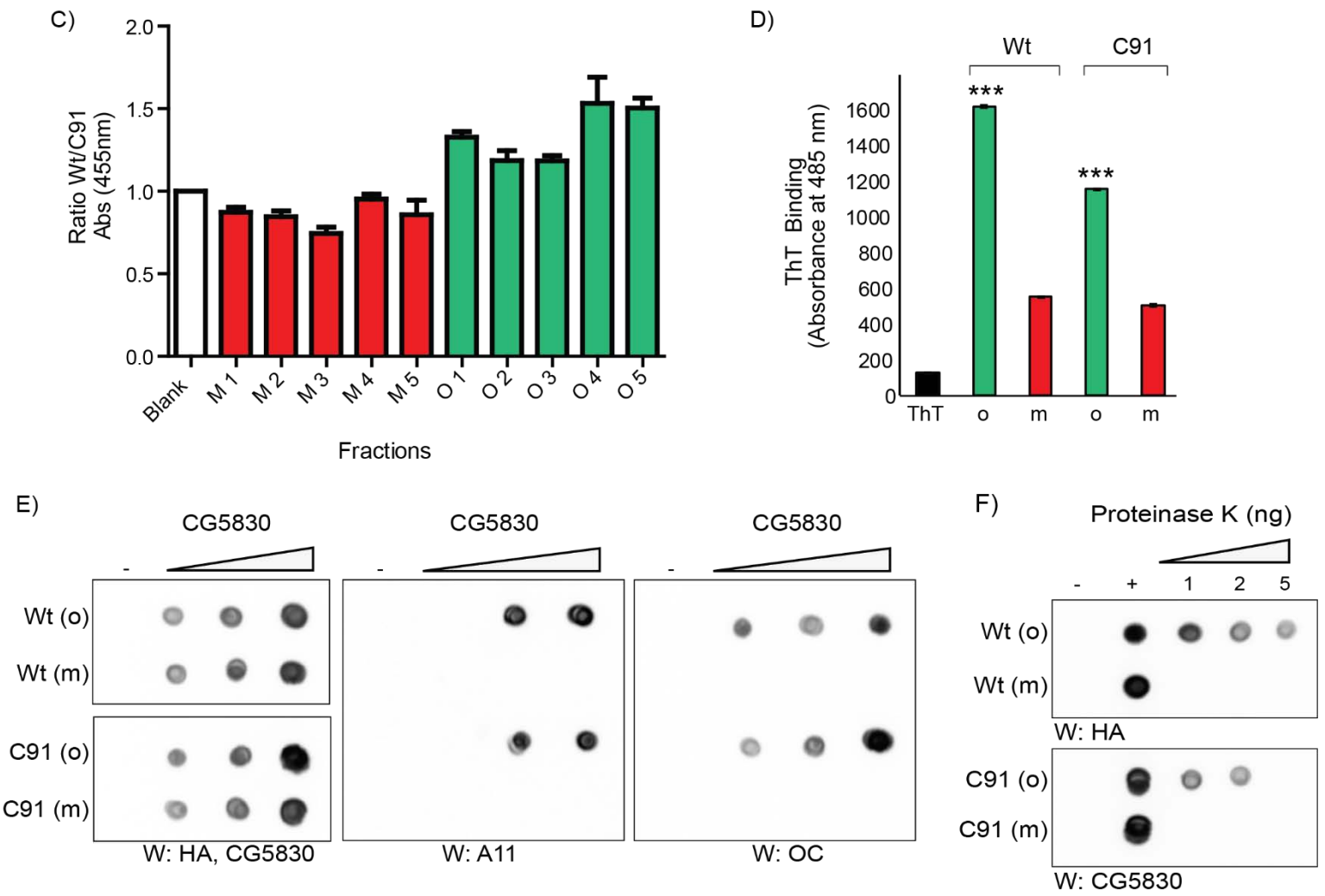


Figure 4.3 Endogenous CG5830 oligomers have phosphatase activity.

A) Isolation or CG5830 monomer and oligomers using a multistep chromatography approach. Scheme of the experiment. AEX: Anion exchange chromatography. CEX: Cation exchange chromatography. SEC: Size exclusion chromatography. MW: Molecular weight. Wt stands for CG5830-HA and C91 stands for CG5830^{C91}.
B) Western blot of various fractions obtained after the SEC by SDS-PAGE (top panel) and SDS-Agarose (bottom). Arrows in top panel indicate specific bands for CG5830-HA and CG5830^{C91} proteins. The relative size of the protein standards is indicated at the bottom (red). Fractions used in the following in vitro assays are indicated with stars.
C) The phosphatase activity of monomeric (m) and oligomeric (o) fractions of CG5830-HA protein. Phosphatase activity was measured in the presence of 100uM DifMUP and normalized to activity of corresponding fractions in CG5830^{C91}.
D) Thioflavin T (ThT) binding of monomeric (m) and oligomeric (o) CG5830. Fractions containing polymeric CG5830 show enhanced ThT binding compared to monomeric fractions. Data is expressed as mean ± SEM. The statistical significance was measured by unpaired two-tailed t-test, $p < 0.001$.
E) Proteinase K digestion and dot blot analysis of different amounts of CG5830 monomer (m) and oligomeric (o) fractions. The monomer is readily digested upon enzyme treatment while the oligomer shows protease resistance.
F) Dot-blot analysis of CG5830 monomer (m) and oligomer (o) fractions using anti-oligomeric A11 and anti-fibrillar antibody OC antibodies. Both A11 and OC reactivity was observed only in CG5830 oligomer fractions, but not in monomer fractions.

Since CG5830 is not purified to homogeneity, it is still possible that these measurements were influenced by putative contaminants. To rule out this possibility, I performed an HA-pull down experiment to further purify CG5830-HA protein after AEX-CEX chromatography (Figure 4.4A, Right panel). I used similarly purified CG5830^{C91} as a negative control, since the bound/eluted samples would contain the same putative contaminants while only the CG5830-HA will be specifically attached to the HA-beads (Figure 4.4A, Left panel). I found only the CG5830-HA protein is resistant to heat, detergents, and Proteinase-K (Figure 4.4B and D). It specifically binds to Thioflavin T and shows reactivity to both anti-A11 and OC antibodies (Figure 4.4C and E)

Altogether, this set of data suggests (in the absence of atomic-level 3D structural analysis of endogenous CG5830 protein) that in addition to phosphatase activity, the high molecular weight oligomeric CG5830 has SDS and proteinase-K resistant amyloid-like structure.

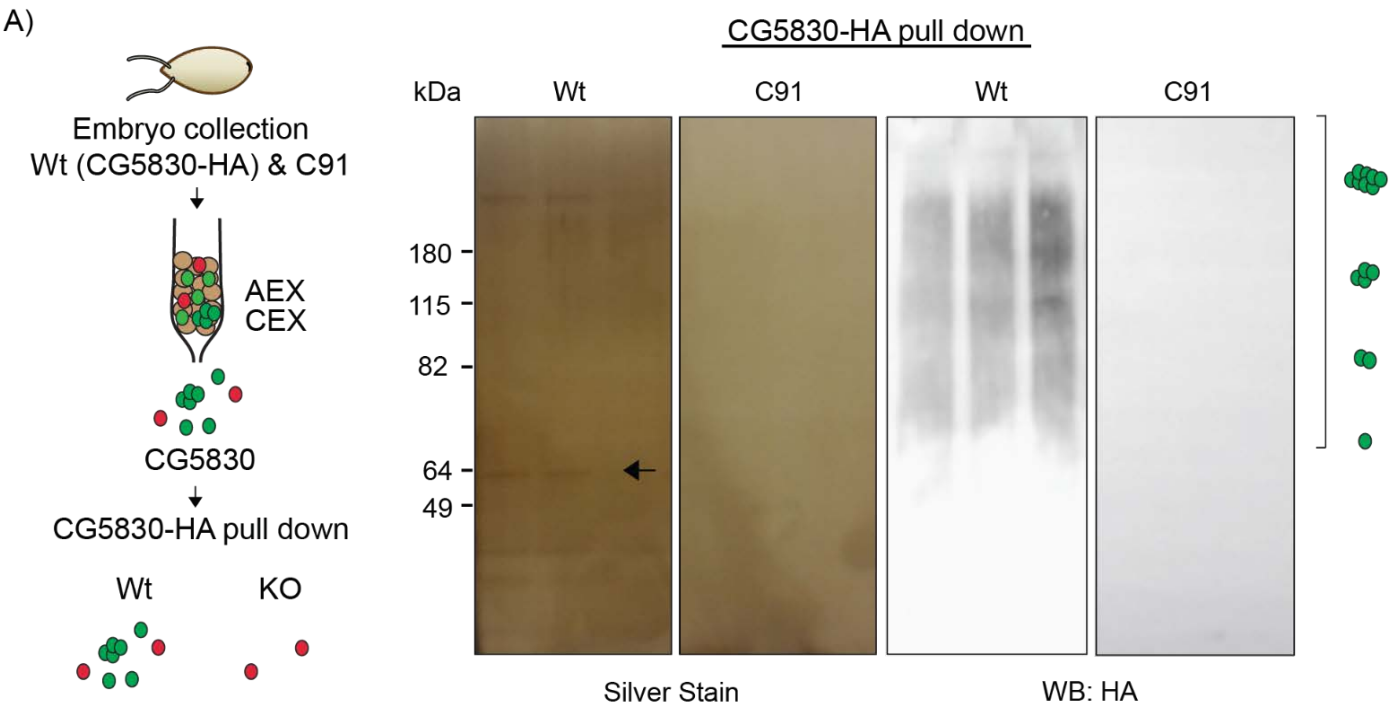


Figure 4.4 Endogenous CG5830 has amyloid-like structural features.

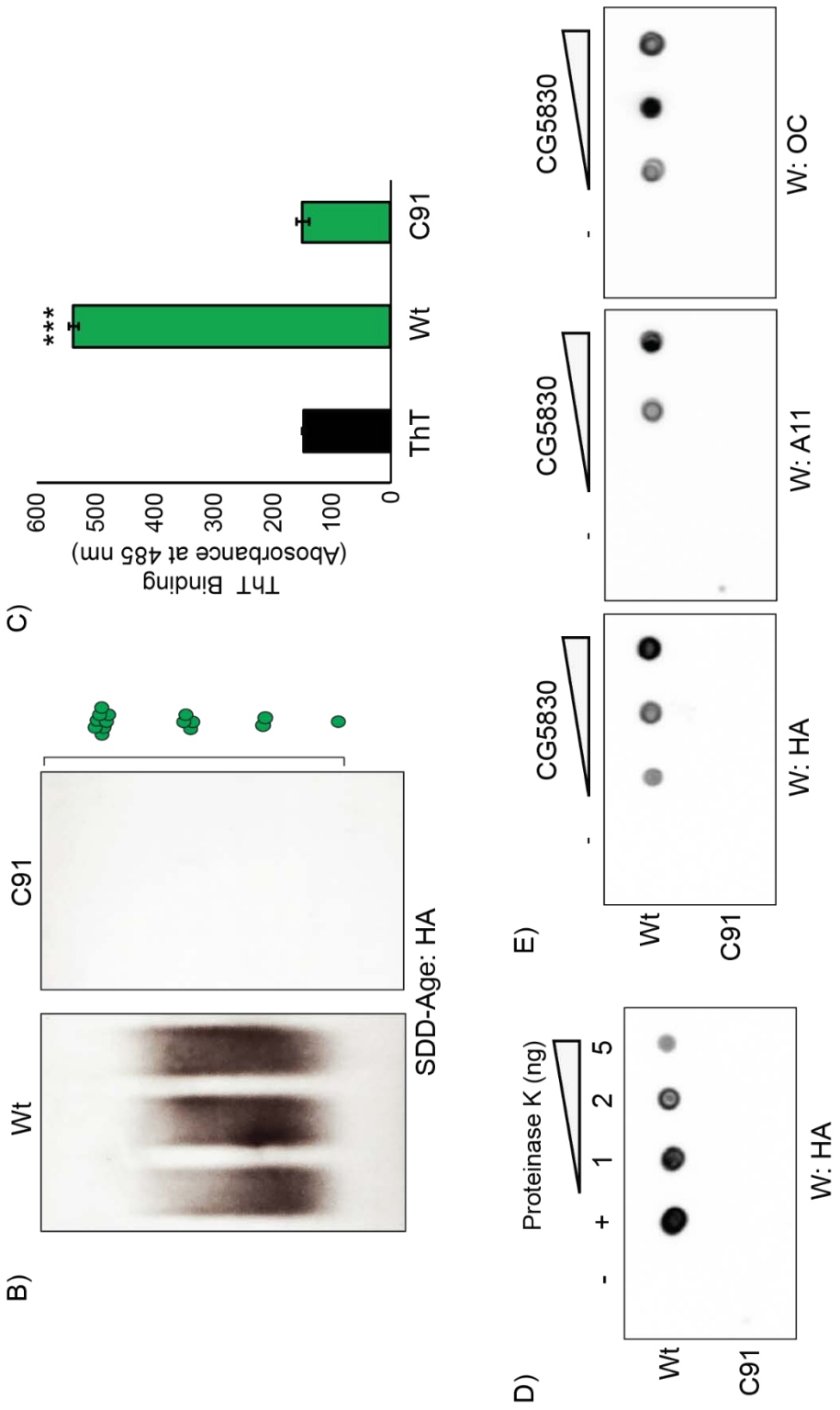


Figure 4.4 Continued.

Figure 4.4 Endogenous CG5830 has amyloid-like structural features.

A) Isolation or CG5830 monomer and polymers using a multistep chromatography followed by immunoprecipitation with anti-HA agarose beads. Left panel: Scheme of the experiment. AEX: Anion exchange chromatography. CEX: Cation exchange chromatography. Wt stands for CG5830-HA and C91 stands for $CG5830^{C91}$. Right: Silver staining and western blot of different elutions.

B) Endogenous CG5830 polymers are SDS resistant shown by western blot of different elutions following SDS-Agarose gel electrophoresis.

C) Thioflavin T (ThT) binding of CG5830. One eluate containing CG5830-HA shows enhanced ThT binding compared to $CG5830^{C91}$. Data is expressed as mean ± SEM. The statistical significance was measured by unpaired two-tailed t-test, $p < 0.001$.

D) Proteinase K digestion and dot blot analysis of different amounts of CG5830-HA and $CG5830^{C91}$ elutions. CG5830-HA protein shows protease resistance.

E) Dot-blot analysis of CG5830-HA and $CG5830^{C91}$ elutions using anti-oligomeric A11 and anti-fibrillar antibody OC antibodies. Both A11 and OC reactivity was observed only in CG5830-HA elution.

Chapter V: Discussion and Future directions

Prion-like protein switches exist in many different organisms including bacteria, yeast, *Drosophila*, *Aplysia*, and human (Hou et al., 2011a; Khan et al., 2015; Pallarès & Ventura, 2017; Shorter & Lindquist, 2005; Si et al., 2010; Reed B Wickner et al., 2007). Such switches have been shown to be involved in many different processes such as stress response, memory formation, and immune response working as transcriptional regulators, translational regulators and signaling molecules (Si et al. 2010; Hou et al. 2011; Holmes et al. 2013). However, most of these proteins have been studied in yeast at unicellular level. Because of the lack of *in vivo* evidence in multicellular eukaryotes, it is still elusive what other processes these sort of protein switches can be involved in at systems level. Here, by endogenously following CG5830 in *Drosophila* embryo, by characterizing CG5830 function at both molecular and organismal levels, and by testing activity of different states of CG5830, I find that CG5830 changes its conformational state from monomers to amyloid like oligomers in the membrane during gastrulation. I conclude that it is a membrane phosphatase required for embryonic patterning and that conformational state change of CG5830 leads to gain of enzymatic function. Taken together, this study shows an example of how a prion-like protein switch may be utilized during development of an organism to control signaling pathways.

5.1 CG5830 protein switch: From inactive monomers to active oligomers

Previous *in silico* and experimental screen using exogenously expressed proteins for prion-like proteins uncovered CG5830 as a candidate. To understand the behavior of CG5830 protein *in vivo*, I followed endogenous CG5830 during embryogenesis when its function is required the most.

CG5830 is a ubiquitous membrane protein expressed throughout embryonic development and has no discernable expression pattern both in space and time. However, it changes its physical state on the membrane from a diffused profile to a punctate form during embryogenesis. Live

imaging of GFP-tagged CG5830 embryos showed that it is diffusely localized along the membrane during very early stages (embryonic stage 1-5, until the end of cellularization) of embryonic development. During gastrulation (stage 6-9), it changes its expression pattern on the membrane to a punctate form and once formed the punctate pattern continues to persist. This does not result from the aggregation tendency of GFP tag since immunostaining of HA tagged CG5830 embryos similarly showed diffused appearance between 0-3h and punctate form during the rest of the embryonic development. Moreover, immunoblot analysis of CG5830 protein in these embryos showed that CG5830 exists in two different physical states: low molecular weight monomers and high-molecular weight oligomers, which starts appearing at the time of gastrulation, a time course similar to the appearance of puncta in the embryo. The differential antibody reactivity of the monomeric and oligomeric state also suggest that they are conformationally distinct. Since *CG5830* is an essential gene and the tagged *CG5830* animals had no observable defect, this suggests that tagging did not alter protein conformation or function. Further characterization of the endogenous purified CG5830 showed that high molecular weight oligomeric CG5830 has an SDS and proteinase-K resistant amyloid-like structure.

One common feature of prion-like proteins is that they often contain a modular prion-domain, which is sufficient for oligomerization. *Ex vivo* experiments using truncated proteins showed that oligomerization and membrane localization of CG5830 is dependent on its N terminal PrD. This is further supported by the observation that purified $CG5830^{C91}$ protein, which corresponds to 11 KDa N-terminal PrD region, can form high molecular weight oligomers in the embryo that are SDS and proteinase-K resistant.

What are the functional consequences of the state change of CG5830? We assessed both the general phosphatase activity *in vitro* on a generic substrate as well as specific activity on a substrate *ex vivo*. Phosphatase activity assay of various oligomeric and monomeric CG5830 fractions showed that only oligomeric CG5830 has phosphatase activity. Similarly, in absence of

N terminal PrD, CG5830 loses its enzymatic activity on its specific substrate Dah. The middle domain of CG5830 interacts with Dah, suggesting the lack of enzymatic activity in the absence of the N-terminal PrD domain, is not due to a failure to recognize or bind the substrate. This is also not due to altered membrane localization in absence of N-terminal PrD, since membrane targeting of the rest of the CG5830 protein with a nonspecific membrane localization motif did not rescue enzymatic activity.

Taken together, these results suggest that CG5830 protein switches its conformation from a monomer to an enzymatically active amyloid-like oligomer during gastrulation in embryonic development. This conformational switch is mediated by its N-terminal PrD. However, the *in vitro* and *ex vivo* experiments above only show us the activity difference between two states of CG5830 and do not tell us if CG5830 requires this enzymatic activity *in vivo*. At this point, I conclude that phosphatase activity of CG5830 is essential for its function during embryonic development, since the deletion mutant $CG5830^{N169}$ containing both N-terminal PrD and conserved HAD phosphatase motif is completely viable. In mutants expressing only the PrD there is embryonic lethality. Viability of $CG5830^{N169}$ mutants, however, does not completely address whether it is only the requirement for enzymatic activity or another functional outcome through the phosphatase motif. Future experiments to discern between these two possibilities would be to genetically introduce point mutations to the phosphatase motif of CG5830, which has been shown to effectively eliminate phosphatase activity in this study, and follow the effect both on the development of the organism and on the phosphorylation levels of its potential substrates. Also, it is unclear at this stage how this conformational switch is triggered *in vivo*: Is it stochastic, is it regulated by a post-translational modification of CG5830, or is it a phase separation induced by interaction with the substrates? Tools that would modify only the oligomerization of CG5830, possible post-translational modification sites, or possible interaction sites with the substrates are necessary to address these questions.

5.2 CG5830 as a regulator of segment polarity in Drosophila melanogaster

What is the molecular function of CG5830 during gastrulation? I addressed this question in two ways: creating mutant flies missing *CG5830* and looking at its interacting partners via a proteomics approach.

First, I observed that mutant embryos expressing only N-terminal PrD of CG5830 die before going into differentiation, based on the lack of denticles. They show severe segment polarity phenotypes, with diffused *en* expression in each parasegment at stage 10-11, and an ectopic stripe of *wg* at the anterior side of each segment later at stage 13-14 embryos. In fact, this extra *wg* stripe requires both wg and hh (en) signaling activity, indicating that these signaling pathways are still intact. Moreover, the enhanced expression of *en* is not caused by an earlier defect, since these embryos have normal patterning of *eve* and *dl* in initial stages of embryonic development. These embryos also have a normal number of segments, indicating both para-segmental and segmental registers are not altered. A similar phenotype has been seen previously in embryos lacking both maternal and zygotic contributions from any of the three genes that antagonize wg signaling: *zw3/GSK* (Siegfried et al., 1992), *D-axin* (Hamada et al., 1999), and *D-Apc2* (McCartney et al., 1999), and in embryos lacking a zygotic copy of *nkd*. Altogether, this tells us that CG5830 normally antagonizes wg signaling at the anterior end of each segment. In absence of CG5830, this inhibition is released, leading to elevated wg signaling and in association excess hh activity and *en* expression. Most interestingly, this phenotype correlates with the conformational and functional switch of CG5830 during gastrulation, which is the time when wg signaling starts during embryogenesis. However, it is still unclear exactly how CG5830 is involved in this pathway. Considering membrane localization of CG5830, I speculate that its antagonistic effect on wg signaling happens either at the level of ligand-receptor interaction, at the level of receptor-co-receptor interaction, or at the level of destruction complex, since the inhibitor of this complex is initially membrane associated. Considering the phosphatase activity of CG5830, this

antagonistic effect is likely to exert itself at the level of receptor-co-receptor or destruction complex, since phosphorylation of the wg receptor, frizzled, has not been reported in *Drosophila*. At the receptor-coreceptor level, the CG5830 inhibitory effect can be on co-receptor *arrow*, a protein which is known to be activated by phosphorylation at its cytoplasmic tail by membrane associated GSK3 and CK1 (Davidson et al., 2005; X. Zeng et al., 2005). At the level of destruction complex, CG5830 inhibitory effects could be on dsh or axin, since the activity of both of these proteins is regulated by phosphorylation (Gao & Chen, 2010; Tacchelly-Benites et al., 2018). A potential way to explore specific effects of CG5830 on wg pathway would be to drive expression of CG5830-WT or CG5830-PD version at the anterior end of each segment with specific gal4 drivers and test its enzymatic activity on above mentioned three components of wg pathway. This will help to discern the importance of its phosphatase function within this pathway.

Second, proteomics analysis followed by pairwise interaction assays for CG5830 interacting partners led to the identification of 7 proteins, namely Babo, Dah, Irk, Pch, Ras64, Sax and Src64B. Surprisingly, I did not find any specific components of wg pathway in this analysis. But I found that interacting partners of CG5830 are involved in major developmental pathways regulating embryonic patterning: TGF-β signaling (Babo, Sax and Irk2), EGF/FGF signaling (Ras64B and Src64B) and cell division (Dah and Pch2). Considering the complex interplay of signaling pathways with each other during embryonic patterning, I argue that CG5830 indirectly affects wg pathway through its involvement in one of these pathways. Moreover, with the exception of Pch2, all these interactors are either located on the membrane (Babo, Sax, and Irk2) or associates with membrane proteins (Ras64B, Src64B and Dah) similar to CG5830. They have also been previously reported to be phosphorylated (Lusk et al., 2017; Parker et al., 2003; Wouda et al., 2008; C. X. Zhang et al., 2000). Based on phosphatase assays, I found CG5830 dephosphorylates one of these proteins, Dah. Interestingly, in the embryo dephosphorylation of Dah occurs after cellularization, which coincides with oligomerization and activation of CG5830

phosphatase activity (C. X. Zhang et al., 2000). It is possible that for all these ligands and their respective pathways, the effect of CG5830 and its phosphatase activity can be inhibitory and/or activating. All of these ligands are activated via phosphorylation as detailed in section 1.7, hence removal of potentiating phosphorylations by CG5830 could lead to inhibition of the pathway. Dephosphorylation could also lead to recycling of these proteins for another round of signaling, indirectly exerting an activating effect. Interestingly, it has been shown that TGF-β signaling, EGF/FGF signaling and cell division have crosstalk with the wg signaling pathway. TGF-β signaling has been shown to repress wg pathway by inducing expression of its antagonist nkd (Yang, Meng, Ma, Xie, & Fang, 2013). The wg pathway inhibits EGF signaling and leads to differentiation of ventral epidermis (FranÇois Payre et al., 1999; Urban, 2004) and works together with FGF pathway to induce cell fate determination (Katoh & Katoh, 2006). This double-sided crosstalk also suggests potential for a dual role for CG5830 phosphatase activity in these pathways.

5.3 Model for CG5830 mediated embryonic patterning

CG5830 is a ubiquitous membrane protein changing its conformational state to an enzymatically active amyloid-like oligomer during gastrulation (stage 6-9). It is also constitutively expressed throughout the life-cycle of *Drosophila*. However, its interacting partners are expressed in a very dynamic pattern (Figure 5.1). Considering this pattern difference between CG5830 and its targets, and its aggregation and disaggregation on the membrane, I hypothesize that CG5830 forms functional foci on the membrane, and these foci change their conformation dynamically from monomers to amyloid-like oligomers making functional complexes with different targets over time. This hypothesis is supported in part by the fact that CG5830 oligomers dynamically form and disappear on the cell membrane over time in the live imaging, while CG5830 monomers exist throughout embryonic development as seen by immunoblotting. To test this model, interaction studies of CG5830 with different targets and dependency of these interactions on CG5830

structure over time in the embryo are required. In addition to this, the phosphorylation status of these targets should be assayed for in the absence and presence of CG5830.

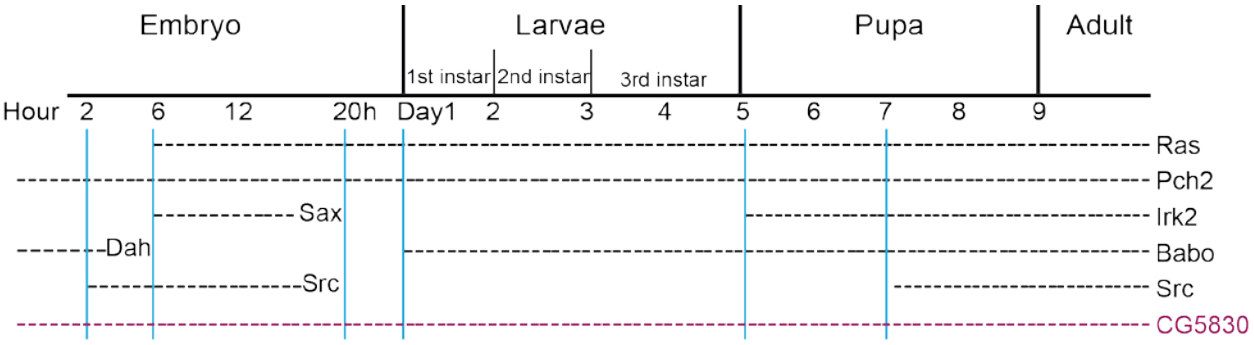


Figure 5.1 Protein expression pattern of CG5830 and its interacting partners throughout life cycle of *Drosophila*.

It is unclear why an enzyme, CG5830, with different targets would form a seemingly stable structure like amyloid. Even though amyloids are known to be one of the most stable protein structures, in recent years, with the advancement of techniques in structural biology, it has been found that the same polypeptide can fold into multiple, distinct amyloid conformations in the cell and that this structural diversity can lead to distinct heritable states (Toyama & Weissman, 2011). From recent atomic level structure analysis of various amyloid-like proteins, it is emerging that in addition to canonical cross-β-sheet, there are also α-helical amyloid-like states with charged residues in the core that confers flexibility (Stroobants et al., 2017). However, the physiological relevance of such structural flexibility remains unclear. I speculate that such flexibility and heterogeneity allow an enzyme like CG5830 to adopt stable yet dynamic conformational states as well as alter the function. Development, which needs to accommodate the changing environment, might utilize such molecular stability and flexibility to tune the time course of development. Moreover, this structural feature may also lead to functional diversity, such that a protein with this structural feature can form distinct functional units, with even opposing functions, in the same cell and can maintain this functional diversity with the help of its stability. Also, one can argue that this dual nature of an amyloid protein would be particularly beneficial for a protein

with multiple functions or with multiple partners, like phosphatases. As stated earlier, the amount of phosphatases encoded by the genome is relatively small compared to kinases considering that almost one third of all proteins in the cell are regulated by phosphorylation (Caenepeel et al., 2004; Morrison et al., 2000; Snyder et al., 2000). For years, this led to an assumption that phosphatases are promiscuous and redundant enzymes. Even though some genetic studies showed distinct phenotypes for specific phosphatases (Mustelin, 2007), this has not answered the question of how they achieve substrate specificity, particularly, for phosphatases, like CG5830, with no recognizable regulatory element in their structure. Amyloid folds of CG5830 can also be part of an explanation for the mechanism of substrate specificity of such enzymes that belongs to the SCP family of phosphatases.

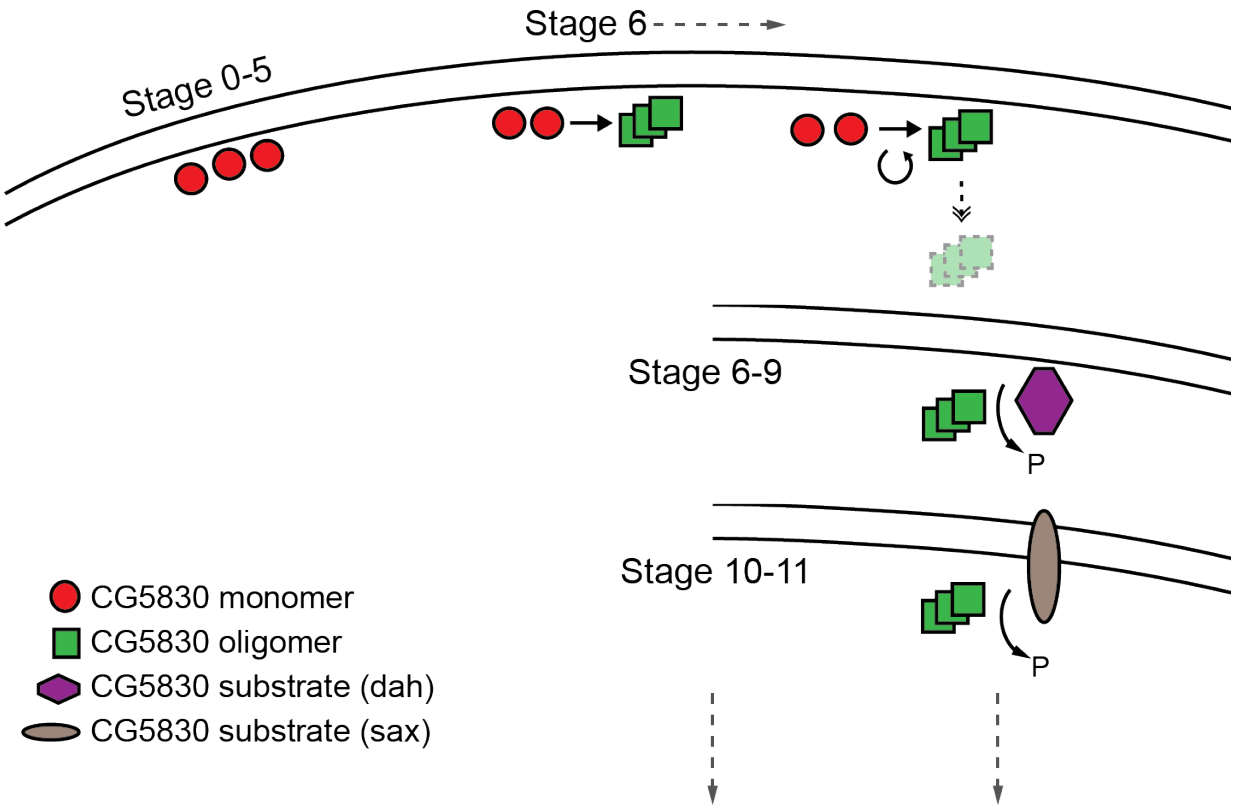


Figure 5.2 A model for CG5830 protein switch during embryogenesis.
CG5830 monomers switch their conformation to an enzymatically active amyloid like oligomer during gastrulation (stage 6). Dynamic creation of these oligomers leads to formation of distinct functional foci on the membrane at different stages of embryogenesis with different substrate specificities.

5.4 Prion-like conformational functional switches as regulators of animal development

Prion-like proteins have been implicated in different physiological processes including adaptation to changing environmental conditions (Halfmann et al., 2012), immune response (Hou et al., 2011b), and memory formation (Majumdar et al., 2012). In all of these processes, they act as transcriptional regulators, translational regulators or signaling components at molecular level, leading to a global change in cellular response. Most of these proteins are well studied yeast prions, often determinants of heritable phenotypes. Even though yeast provides us with the ability to follow functional outputs of prion-like protein switches both quantitatively (molecular) and qualitatively (phenotype), because of its unicellular nature, it does not give us information at a system level. The examples of such protein switches in multicellular eukaryotes are both restricted and, except Orb2, studied *in vitro*. In this study, I found that protein switch of CG5830 *in vivo* leads to a gain of enzymatic function and controls embryonic patterning. This is the first example of a prion-like protein switch with a critical role in animal development. Most interestingly, CG5830 has several potential functions at the intersection of multiple signaling pathways. It is known that development is regulated a complex interplay of different signaling pathways, which are inherently dynamic but create stable information to carry development further forward (Perrimon et al., 2012). Considering the stable nature of prion-like protein switches, this could be a plausible mechanism for the organism to stably maintain the altered activity states of proteins in a particular pathway. This could impact the identity of a group of cells which would develop into a particular organ even through fluxes in the availability of ligands involved in signaling. Given that these proteins can assume multiple conformational states, they could produce multiple activity states, therefore generating variation in activity. Therefore, my characterization of the CG5830 protein switch in this study illustrates how protein switches can be important regulators for animal physiology and lends support to the idea that there can be other prion-like conformational functional switches regulating other important developmental events.

Appendix-List of proteins enrichment comparison

List of proteins significantly enriched in WT-CG5830 Flag-APs compared to PD-CG5830 and/or Flag-Negative controls.

Gene	Locus Tag	Description
OstStt3	**CG7748**	oligosaccharyl transferase 3 [Drosophila melanogaster]
Irk1	**CG44159**	inwardly rectifying potassium channel, isoform B [Drosophila melanogaster]
CG16974	**CG16974**	CG16974, isoform B [Drosophila melanogaster]
pch2	**CG31453**	pch2 [Drosophila melanogaster]
babo	**CG8224**	baboon, isoform B [Drosophila melanogaster]
robo1	**CG13521**	roundabout, isoform A [Drosophila melanogaster]
Src64B	**CG7524**	Src oncogene at 64B, isoform I [Drosophila melanogaster]
Shc	**CG3715**	SHC-adaptor protein [Drosophila melanogaster]
Cul4	**CG8711**	Cullin-4, isoform B [Drosophila melanogaster]
sax	**CG1891**	saxophone, isoform A [Drosophila melanogaster]
Rap2l	**CG3204**	Ras-associated protein 2-like [Drosophila melanogaster]
OstDelta	**CG6370**	oligosaccharide transferase delta subunit [Drosophila melanogaster]
Dref	**CG5838**	DNA replication-related element factor, isoform B [Drosophila melanogaster]
Bub3	**CG7581**	Bub3 [Drosophila melanogaster]
spel1	**CG4215**	spellchecker1, isoform D [Drosophila melanogaster]
dah	**CG6157**	discontinuous actin hexagon, isoform A [Drosophila melanogaster]
Rpn3	**CG42641**	regulatory particle non-ATPase 3, isoform B [Drosophila melanogaster]
ben	**CG18319**	bendless, isoform E [Drosophila melanogaster]
CG3909	**CG3909**	CG3909 [Drosophila melanogaster]
Rheb	**CG1081**	rheb, isoform B [Drosophila melanogaster]
CG7182	**CG7182**	CG7182 [Drosophila melanogaster]
CG3262	**CG3262**	CG3262, isoform E [Drosophila melanogaster]
PCNA/mus209	**CG9193**	mutagen-sensitive 209, isoform B [Drosophila melanogaster]
CG3808	**CG3808**	CG3808 [Drosophila melanogaster]
CG2124	**CG2124**	CG2124 [Drosophila melanogaster]
Eph	**CG1511**	Eph receptor tyrosine kinase, isoform E [Drosophila melanogaster]
gwl	**CG7719**	greatwall, isoform A [Drosophila melanogaster]
RpII215	**CG1554**	RNA polymerase II 215kD subunit [Drosophila melanogaster]
Ras64B	**CG1167**	Ras oncogene at 64B [Drosophila melanogaster]
Cand1	CG5366	Cullin-associated and neddylation-dissociated 1, isoform B [Drosophila melanogaster]
Rpb8	**CG11246**	Rpb8 [Drosophila melanogaster]
CG33303	**CG33303**	CG33303, isoform B [Drosophila melanogaster]
PlexA	CG11081	plexin A, isoform A [Drosophila melanogaster]
NtR	CG6698	NtR, isoform A [Drosophila melanogaster]
Stat92E	CG4257	Signal-transducer and activator of transcription protein at 92E, isoform B [Drosophila melanogaster]

HP1b	CG7041	heterochromatin protein 1b, isoform C [Drosophila melanogaster]
CG5853	**CG5853**	CG5853, isoform B [Drosophila melanogaster]
Syp	**CG17838**	syncrip, isoform I [Drosophila melanogaster]
MBD-R2	**CG10042**	MBD-R2, isoform A [Drosophila melanogaster]
zfh2	**CG1449**	Zn finger homeodomain 2, isoform A [Drosophila melanogaster]
drk	**CG6033**	downstream of receptor kinase, isoform F [Drosophila melanogaster]
CG30456	CG30456	CG30456, isoform B [Drosophila melanogaster]
CG2875	CG2875	CG2875, isoform A [Drosophila melanogaster]
tum	CG13345	tumbleweed [Drosophila melanogaster]
Poc1	CG10191	proteome of centrioles 1, isoform A [Drosophila melanogaster]
Set2	CG1716	Set2, isoform A [Drosophila melanogaster]
CG4199	CG4199	CG4199, isoform A [Drosophila melanogaster]
CG13366	CG13366	CG13366, isoform B [Drosophila melanogaster]
Cul5	CG1401	Cullin-5 [Drosophila melanogaster]

STN>0.9 with PLGEM p<0.05; log2 Fold Change>2 with QPROT p<0.05. Proteins with validated physical interactions are indicated in red.

References

Abrams, J. M., White, K., Fessler, L. I., & Steller, H. (1993). Programmed cell death during Drosophila embryogenesis. *Development*, *117*(1), 29–43. Retrieved from http://dev.biologists.org/content/117/1/29.short

Adkins, N. L., Hagerman, T. A., & Georgel, P. (2006). GAGA protein: a multi-faceted transcription factor. *Biochemistry and Cell Biology = Biochimie et Biologie Cellulaire*, *84*(4), 559–567. https://doi.org/10.1139/o06-062

Alberti, S., Halfmann, R., King, O., Kapila, A., & Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. *Cell*, *137*(1), 146–58. https://doi.org/10.1016/j.cell.2009.02.044

Alberti, S., Halfmann, R., & Lindquist, S. (2010). Biochemical, Cell Biological, and Genetic Assays to Analyze Amyloid and Prion Aggregation in Yeast. *Methods in Enzymology*, *470*, 709–734. https://doi.org/10.1016/S0076-6879(10)70030-6

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Drosophila and the Molecular Genetics of Pattern Formation: Genesis of the Body Plan. In *Molecular biology of the cell* (4th editio). New York: Garland Science. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK26906/

Almo, S. C., Bonanno, J. B., Sauder, A. J. M., Spencer, A., Ae, E., Dilorenzo, T. P., … Burley, S. K. (2007). Structural genomics of protein phosphatases. https://doi.org/10.1007/s10969-007-9036-1

Anfinsen, C. B. (1973). Principles that govern the folding of protein chains. *Science (New York, N.Y.)*, *181*(4096), 223–30. https://doi.org/10.1126/SCIENCE.181.4096.223

Ballas, N., & Mandel, G. (2005). The many faces of REST oversee epigenetic programming of neuronal genes. *Current Opinion in Neurobiology*, *15*(5), 500–6. https://doi.org/10.1016/j.conb.2005.08.015

Batlle, C., Iglesias, V., Navarro, S., & Ventura, S. (2017). Prion-like proteins and their computational identification in proteomes. *Expert Review of Proteomics*, *14*(4), 335–350. https://doi.org/10.1080/14789450.2017.1304214

Bejsovec, A. (2018). Wingless signaling: A genetic journey from morphogenesis to metastasis. *Genetics*, *208*(4), 1311–1336. https://doi.org/10.1534/genetics.117.300157

Bejsovec, A., & Martinez Arias, A. (1991). Roles of wingless in patterning the larval epidermis of Drosophila. *Development (Cambridge, England)*, *113*(2), 471–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1782860

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300. https://doi.org/10.2307/2346101

Benoit, B., He, C. H., Zhang, F., Votruba, S. M., Tadros, W., Westwood, J. T., … Theurkauf, W. E. (2009). An essential role for the RNA-binding protein Smaug during the Drosophila maternal-to-zygotic transition. *Development*, *136*(6), 923–932. https://doi.org/10.1242/dev.031815

Bier, E., Harrison, M. M., O'connor-Giles, K. M., & Wildonger, J. (2018). Advances in engineering the fly genome with the CRISPR-Cas system. *Genetics*, *208*(1), 1–18. https://doi.org/10.1534/genetics.117.1113

Brown, J. C. S., & Lindquist, S. (2009). A heritable switch in carbon source utilization driven by an unusual yeast prion. *Genes & Development*, *23*(19), 2320–32. https://doi.org/10.1101/gad.1839109

Caenepeel, S., Charydczak, G., Sudarsanam, S., Hunter, T., & Manning, G. (2004). The mouse kinome: discovery and comparative genomics of all mouse protein kinases. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(32), 11707–12.

https://doi.org/10.1073/pnas.0306880101
Cai, X., & Chen, Z. J. (2014). Prion-like polymerization as a signaling mechanism. *Trends in Immunology*, *35*(12), 622–630. https://doi.org/10.1016/j.it.2014.10.003
Campos-Ortega, J. A., & Hartenstein, V. (1997). *The Embryonic Development of Drosophila melanogaster*. *Springer-Verlag Berlin*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-22489-2
Carpenter, K., Bell, R. B., Yunus, J., Amon, A., & Berchowitz, L. E. (2018). Phosphorylation-Mediated Clearance of Amyloid-like Assemblies in Meiosis. *Developmental Cell*, *45*(3), 392–405.e6. https://doi.org/10.1016/j.devcel.2018.04.001
Casas-Vila, N., Bluhm, A., Sayols, S., Dinges, N., Dejung, M., Altenhein, T., … Butter, F. (2017). The developmental proteome of Drosophila melanogaster. *Genome Research*, *27*(7), 1273–1285. https://doi.org/10.1101/gr.213694.116
Celniker, S. E., Dillon, L. A. L., Gerstein, M. B., Gunsalus, K. C., Henikoff, S., Karpen, G. H., … Waterston, R. H. (2009, June 18). Unlocking the secrets of the genome. *Nature*. https://doi.org/10.1038/459927a
Ceresa, B. P., & Vanlandingham, P. A. (2008). Molecular Mechanisms that Regulate Epidermal Growth Factor Receptor Inactivation. *Clinical Medicine. Oncology*, *2*, 47–61. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21892266
Choi, H., Fermin, D., & Nesvizhskii, A. I. (2008). Significance Analysis of Spectral Count Data in Label-free Shotgun Proteomics. *Molecular & Cellular Proteomics*, *7*(12), 2373–2385. https://doi.org/10.1074/mcp.M800203-MCP200
Choy, E., Chiu, V. K., Silletti, J., Feoktistov, M., Morimoto, T., Michaelson, D., … Philips, M. R. (1999). Endomembrane trafficking of ras: The CAAX motif targets proteins to the ER and Golgi. *Cell*, *98*(1), 69–80. https://doi.org/10.1016/S0092-8674(00)80607-8
Chuang, E., Hori, A. M., Hesketh, C. D., & Shorter, J. (2018). Amyloid assembly and disassembly. *Journal of Cell Science*, *131*(8), jcs189928. https://doi.org/10.1242/jcs.189928
Colby, D. W., Giles, K., Legname, G., Wille, H., Baskakov, I. V., DeArmond, S. J., & Prusiner, S. B. (2009). Design and construction of diverse mammalian prion strains. *Proceedings of the National Academy of Sciences*, *106*(48), 20417–20422. https://doi.org/10.1073/pnas.0910350106
Cox, B. S. (1965). Ψ, A cytoplasmic suppressor of super-suppressor in yeast. *Heredity*, *20*(4), 505–521. https://doi.org/10.1038/hdy.1965.65
Csiszar, A., Vogelsang, E., Beug, H., & Leptin, M. (2010). A Novel Conserved Phosphotyrosine Motif in the Drosophila Fibroblast Growth Factor Signaling Adaptor Dof with a Redundant Role in Signal Transmission. *Molecular and Cellular Biology*, *30*(8), 2017–2027. https://doi.org/10.1128/MCB.01436-09
Dahal, G. R., Rawson, J., Gassaway, B., Kwok, B., Tong, Y., Ptacek, L. J., & Bates, E. (2012). An inwardly rectifying K+ channel is required for patterning. *Development*, *139*(19), 3653–3664. https://doi.org/10.1242/dev.078592
Daniels, D. L., & Weis, W. I. (2005). β-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nature Structural and Molecular Biology*, *12*(4), 364–371. https://doi.org/10.1038/nsmb912
Davidson, G., Wu, W., Shen, J., Bilic, J., Fenger, U., Stannek, P., … Niehrs, C. (2005). Casein kinase 1γ couples Wnt receptor activation to cytoplasmic signal transduction. *Nature*, *438*(7069), 867–872. https://doi.org/10.1038/nature04170
De Robertis, E. M. (2008, January 25). Evo-Devo: Variations on Ancestral Themes. *Cell*. https://doi.org/10.1016/j.cell.2008.01.003
DiNardo, S., Heemskerk, J., Dougan, S., & O'Farrell, P. H. (1994). The making of a maggot: patterning the Drosophila embryonic epidermis. *Current Opinion in Genetics and Development*, *4*(4), 529–534. https://doi.org/10.1016/0959-437X(94)90068-E
DiNardo, S., & O'Farrell, P. H. (1987). Establishment and refinement of segmental pattern in the

Drosophila embryo: spatial control of engrailed expression by pair-rule genes. *Genes & Development*, *1*(10), 1212–1225. https://doi.org/10.1101/gad.1.10.1212

Driever, W., & Nüsslein-Volhard, C. (1988). The bicoid protein determines position in the Drosophila embryo in a concentration-dependent manner. *Cell*, *54*(1), 95–104. https://doi.org/10.1016/0092-8674(88)90183-3

Drillien, R., Aigle, M., & Lacroute, F. (1973). Yeast mutants pleiotropically impaired in the regulation of the two glutamate dehydrogenases. *Biochemical and Biophysical Research Communications*, *53*(2), 367–372. https://doi.org/10.1016/0006-291X(73)90671-2

Drozdova, P., Rogoza, T., Radchenko, E., Lipaeva, P., & Mironova, L. (2014). Transcriptional response to the [ISP+] prion of Saccharomyces cerevisiae differs from that induced by the deletion of its structural gene, SFP1. *FEMS Yeast Research*, *14*(8), 1160–1170. https://doi.org/10.1111/1567-1364.12211

Du, Z., Park, K. W., Yu, H., Fan, Q., & Li, L. (2008). Newly identified prion linked to the chromatin-remodeling factor Swi1 in Saccharomyces cerevisiae. *Nature Genetics*, *40*(4), 460–465. https://doi.org/10.1038/ng.112

Edgar, B. A., & Schubiger, G. (1986). Parameters controlling transcriptional activation during early drosophila development. *Cell*, *44*(6), 871–877. https://doi.org/10.1016/0092-8674(86)90009-7

Eisenberg, D., & Jucker, M. (2012). The Amyloid State of Proteins in Human Diseases. *Cell*, *148*(6), 1188–1203. https://doi.org/10.1016/J.CELL.2012.02.022

Eisenberg, D. S., & Sawaya, M. R. (2017). Structural Studies of Amyloid Proteins at the Molecular Level. https://doi.org/10.1146/annurev-biochem

Flores, G. V, Duan, H., Yan, H., Nagaraj, R., Fu, W., Zou, Y., … Banerjee, U. (2000). Combinatorial signaling in the specification of unique cell fates. *Cell*, *103*(1), 75–85. https://doi.org/10.1016/S0092-8674(00)00106-9

Ford, A. F., & Shorter, J. (2015, October 8). Fleeting Amyloid-like Forms of Rim4 Ensure Meiotic Fidelity. *Cell*. NIH Public Access. https://doi.org/10.1016/j.cell.2015.09.049

Fowler, D. M., Koulov, A. V, Alory-Jost, C., Marks, M. S., Balch, W. E., & Kelly, J. W. (2005). Functional Amyloid Formation within Mammalian Tissue. *PLoS Biology*, *4*(1), e6. https://doi.org/10.1371/journal.pbio.0040006

Franklin, B. S., Bossaller, L., De Nardo, D., Ratter, J. M., Stutz, A., Engels, G., … Latz, E. (2014). The adaptor ASC has extracellular and “prionoid” activities that propagate inflammation. *Nature Immunology*. https://doi.org/10.1038/ni.2913

Gao, C., & Chen, Y. G. (2010, May). Dishevelled: The hub of Wnt signaling. *Cellular Signalling*. https://doi.org/10.1016/j.cellsig.2009.11.021

Gilbert, S. C. (2000). *Developmental Biology* (6th editio). Sunderland, MA: Sinauer Associates. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK9983/

Gohla, A., Birkenfeld, J., & Bokoch, G. M. (2005). Chronophin, a novel HAD-type serine protein phosphatase, regulates cofilin-dependent actin dynamics. *Nature Cell Biology*, *7*(1), 21–29. https://doi.org/10.1038/ncb1201

Graba, Y., Aragnol, D., Rothbächer, U., & Pradel, J. (2013). Wnt/Wingless Signaling in Drosophila. Landes Bioscience. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK6231/

Hajheidari, M., Koncz, C., & Eick, D. (2013). Emerging roles for RNA polymerase II CTD in Arabidopsis. *Trends in Plant Science*, *18*(11), 633–43. https://doi.org/10.1016/j.tplants.2013.07.001

Hakimi, M.-A., Bochar, D. A., Chenoweth, J., Lane, W. S., Mandel, G., & Shiekhattar, R. (2002). A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(11), 7420–5. https://doi.org/10.1073/pnas.112008599

Halfmann, R., Jarosz, D. F., Jones, S. K., Chang, A., Lancaster, A. K., & Lindquist, S. (2012).

Prions are a common mechanism for phenotypic inheritance in wild yeasts. *Nature*, *482*(7385), 363–368. https://doi.org/10.1038/nature10875

Hall, E. T., Pradhan-Sundd, T., Samnani, F., & Verheyen, E. M. (2017). The protein phosphatase 4 complex promotes the Notch pathway andwinglesstranscription. *Biology Open*, *6*(8), 1165–1173. https://doi.org/10.1242/bio.025221

Hamada, F., Tomoyasu, Y., Takatsu, Y., Nakamura, M., Nagai, S. I., Suzuki, A., … Akiyama, T. (1999). Negative regulation of wingless signaling by D-Axin, a Drosophila homolog of Axin. *Science*, *283*(5408), 1739–1742. https://doi.org/10.1126/science.283.5408.1739

Harbi, D., & Harrison, P. M. (2014). Classifying prion and prion-like phenomena. *Prion*. https://doi.org/10.4161/pri.27960

Harrison, P. M., & Gerstein, M. (2003). A method to assess compositional bias in biological sequences and its application to prion-like glutamine/asparagine-rich domains in eukaryotic proteomes. *Genome Biology*, *4*(6), R40. https://doi.org/10.1186/gb-2003-4-6-r40

Hartenstein, V. (1993). Atlas of Drosophila Development. *Atlas of Drosophila Development*, 1–57. Retrieved from http://www.worldcat.org/title/atlas-of-drosophila-development/oclc/29631848

Hiramatsu, H., Lu, M., Matsuo, K., Gekko, K., Goto, Y., & Kitagawa, T. (2010). Differences in the molecular structure of β2-microglobulin between two morphologically different amyloid fibrils. *Biochemistry*, *49*(4), 742–751. https://doi.org/10.1021/bi901536j

Holmes, D. L., Lancaster, A. K., Lindquist, S., & Halfmann, R. (2013a). Heritable remodeling of yeast multicellularity by an environmentally responsive prion. *Cell*, *153*(1), 153–65. https://doi.org/10.1016/j.cell.2013.02.026

Holmes, D. L., Lancaster, A. K., Lindquist, S., & Halfmann, R. (2013b). Heritable remodeling of yeast multicellularity by an environmentally responsive prion. *Cell*, *153*(1), 153–65. https://doi.org/10.1016/j.cell.2013.02.026

Horwich, A. L., & Weissman, J. S. (1997, May 16). Deadly conformations - Protein misfolding in prion disease. *Cell*. Cell Press. https://doi.org/10.1016/S0092-8674(00)80232-9

Hou, F., Sun, L., Zheng, H., Skaug, B., Jiang, Q.-X., & Chen, Z. J. (2011a). MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell*, *146*(3), 448–61. https://doi.org/10.1016/j.cell.2011.06.041

Hou, F., Sun, L., Zheng, H., Skaug, B., Jiang, Q.-X., & Chen, Z. J. J. (2011b). MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell*, *146*(3), 448–61. https://doi.org/10.1016/j.cell.2011.06.041

Hughes, M. P., Sawaya, M. R., Boyer, D. R., Goldschmidt, L., Rodriguez, J. A., Cascio, D., … Eisenberg, D. S. (2018). Atomic structures of low-complexity protein segments reveal kinked b sheets that assemble networks. *Science*, *359*(6376), 698–701. https://doi.org/10.1126/science.aan6398

Ishimaru, D., Andrade, L. R., Teixeira, L. S. P., Quesado, P. A., Maiolino, L. M., Lopez, P. M., … Silva, J. L. (2003). Fibrillar aggregates of the tumor suppressor p53 core domain. *Biochemistry*, *42*(30), 9022–9027. https://doi.org/10.1021/bi034218k

Itoh, S., & ten Dijke, P. (2007). Negative regulation of TGF-β receptor/Smad signal transduction. *Current Opinion in Cell Biology*. https://doi.org/10.1016/j.ceb.2007.02.015

Jacobs, J. L., & Coyne, C. B. (2013). Mechanisms of MAVS Regulation at the Mitochondrial Membrane. *Journal of Molecular Biology*, *425*, 5009–5019. https://doi.org/10.1016/j.jmb.2013.10.007

Jaramillo-Tatis, S., Bamm, V. V, Vassall, K. A., & Harauz, G. (2014). Over-expression in E. coli and purification of functional full-length murine small C-terminal domain phosphatase (SCP1, or Golli-interacting protein) in the oligodendrocyte lineage) Golli-interacting protein Intrinsically-disordered protein Myelin basic p. https://doi.org/10.1016/j.pep.2014.05.013

Johnston, D. S., & Nüsslein-Volhard, C. (1992, January 24). The origin of pattern and polarity in the Drosophila embryo. *Cell*. https://doi.org/10.1016/0092-8674(92)90466-P

Jürgens, G., Wieschaus, E., Nüsslein-Volhard, C., & Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle in Drosophila melanogaster - II. Zygotic loci on the third chromosome. *Wilhelm Roux's Archives of Developmental Biology*, *193*(5), 283–295. https://doi.org/10.1007/BF00848157

Kakihara, K., Shinmyozu, K., Kato, K., Wada, H., & Hayashi, S. (2008). Conversion of plasma membrane topology during epithelial tube connection requires Arf-like 3 small GTPase in Drosophila. *Mechanisms of Development*, *125*(3–4), 325–336. https://doi.org/10.1016/j.mod.2007.10.012

Katoh, M., & Katoh, M. (2006). Cross-talk of WNT and FGF signaling pathways at GSK3beta to regulate beta-catenin and SNAIL signaling cascades. *Cancer Biology & Therapy*, *5*, 9, 1059–1064. https://doi.org/3151 [pii]

Kayed, R., Head, E., Sarsoza, F., Saing, T., Cotman, C. W., Necula, M., … Glabe, C. G. (2007). Fibril specific, conformation dependent antibodies recognize a generic epitope common to amyloid fibrils and fibrillar oligomers that is absent in prefibrillar oligomers. *Molecular Neurodegeneration*, *2*(1), 18. https://doi.org/10.1186/1750-1326-2-18

Kayed, R., Head, E., Thompson, J. L., McIntire, T. M., Milton, S. C., Cotman, C. W., & Glabel, C. G. (2003). Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, *300*(5618), 486–489. https://doi.org/10.1126/science.1079469

Kelly, J. W. (1998). The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Current Opinion in Structural Biology*, *8*(1), 101–106. https://doi.org/10.1016/S0959-440X(98)80016-X

Khan, M. R., Li, L., Pé Rez-Sá Nchez, C., Slaughter, B. D., Unruh, J. R., Si, K., … Florens, L. (2015). Amyloidogenic Oligomerization Transforms Drosophila Orb2 from a Translation Repressor to an Activator. *Cell*, *163*, 1468–1483. https://doi.org/10.1016/j.cell.2015.11.020

Khurana, R., Coleman, C., Ionescu-Zanetti, C., Carter, S. A., Krishna, V., Grover, R. K., … Singh, S. (2005). Mechanism of thioflavin T binding to amyloid fibrils. *Journal of Structural Biology*, *151*(3), 229–238. https://doi.org/10.1016/J.JSB.2005.06.006

Kim, S. Y., Jung, K. Il, Kim, S. H., & Jeon, S.-H. (2008). Dpp represses eagle expression at short-range, but can repress its expression at a long-range via EGFR signal repression. *Molecules and Cells*, *26*(6), 576–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18779662

Knockaert, M., Sapkota, G., Alarcón, C., Massagué, J., & Brivanlou, A. H. (2006). Unique players in the BMP pathway: small C-terminal domain phosphatases dephosphorylate Smad1 to attenuate BMP signaling. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(32), 11940–5. https://doi.org/10.1073/pnas.0605133103

Kölling, N. (2015). Quantitative genetics of gene expression during fruit fly development. Retrieved from https://www.ebi.ac.uk/sites/ebi.ac.uk/files/shared/documents/phdtheses/Kolling.pdf

Lancaster, A. K., Nutter-Upham, A., Lindquist, S., & King, O. D. (2014). PLAAC: A web and command-line application to identify proteins with prion-like amino acid composition. *Bioinformatics*, *30*(17), 2501–2502. https://doi.org/10.1093/bioinformatics/btu310

Lee, K., Florens, L., Swanson, S., Washburn, M., & Workman, J. (2005). The deubiquitylation activity of Ubp8 is dependent upon Sgf11 and its association with the SAGA complex. *Molecular and Cellular Biology*, *25*(3), 1173–1182. https://doi.org/10.1128/MCB.25.3.1173-1182.2005

Leptin, M. (1999, June 15). Gastrulation in Drosophila: The logic and the cellular mechanisms. *EMBO Journal*. Wiley-Blackwell. https://doi.org/10.1093/emboj/18.12.3187

Levanon, D., Goldstein, R. E., Bernstein, Y., Tang, H., Goldenberg, D., Stifani, S., … Groner, Y. (1998). Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(20), 11590–11595. https://doi.org/10.1073/pnas.95.20.11590

LeVine, H. (1999). Quantification of β-sheet amyloid fibril structures with thioflavin T. *Methods in*

Enzymology, *309*, 274–284. https://doi.org/10.1016/S0076-6879(99)09020-5
Li, C., Scott, D. A., Hatch, E., Tian, X., & Mansour, S. L. (2007). Dusp6 (Mkp3) is a negative feedback regulator of FGF-stimulated ERK signaling during mouse development. *Development*, *134*(1), 167–176. https://doi.org/10.1242/dev.02701
Li, L., Sanchez, C. P., Slaughter, B. D., Zhao, Y., Khan, M. R., Unruh, J. R., … Si, K. (2016). A Putative Biochemical Engram of Long-Term Memory. *Current Biology : CB*, *26*(23), 3143–3156. https://doi.org/10.1016/j.cub.2016.09.054
Lin, X., Buff, E. M., Perrimon, N., & Michelson, A. M. (1999). Heparan sulfate proteoglycans are essential for FGF receptor signaling during Drosophila embryonic development. *Development (Cambridge, England)*, *126*(17), 3715–3723. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10433902
Lusk, J., Lam, V., & Tolwinski, N. (2017). Epidermal Growth Factor Pathway Signaling in Drosophila Embryogenesis: Tools for Understanding Cancer. *Cancers*, *9*(2), 16. https://doi.org/10.3390/cancers9020016
Majumdar, A., Cesario, W. C., White-Grindley, E., Jiang, H., Ren, F., Khan, M. R., … Si, K. (2012). Critical role of amyloid-like oligomers of Drosophila Orb2 in the persistence of memory. *Cell*, *148*(3), 515–29. https://doi.org/10.1016/j.cell.2012.01.004
Malartre, M. (2016, May 2). Regulatory mechanisms of EGFR signalling during Drosophila eye development. *Cellular and Molecular Life Sciences*. https://doi.org/10.1007/s00018-016-2153-x
Mallo, M., & Alonso, C. R. (2013). The regulation of Hox gene expression during animal development. *Development*, *140*(19), 3951–3963. https://doi.org/10.1242/dev.068346
Martinez-Arias, A., & Lawrence, P. A. (1985, February 21). Parasegments and compartments in the Drosophila embryo. *Nature*. Nature Publishing Group. https://doi.org/10.1038/313639a0
McCartney, B. M., Dierick, H. A., Kirkpatrick, C., Moline, M. M., Baas, A., Peifer, M., & Bejsovec, A. (1999). Drosophila APC2 is a cytoskeletally-associated protein that regulates wingless signaling in the embryonic epidermis. *Journal of Cell Biology*, *146*(6), 1303–1318. https://doi.org/10.1083/jcb.146.6.1303
McEwen, D. G. (2005). Puckered, a Drosophila MAPK phosphatase, ensures cell viability by antagonizing JNK-induced apoptosis. *Development*, *132*(17), 3935–3946. https://doi.org/10.1242/dev.01949
Michelitsch, M. D., & Weissman, J. S. (2000). A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(22), 11910–5. https://doi.org/10.1073/pnas.97.22.11910
Morgan, T. H. (1909). What are “Factors” in Mendelian explanations? *Journal of Heredity*, *os-5*(1), 365–368. https://doi.org/10.1093/jhered/os-5.1.365
Morgan, T. H. (1910). Sex-limited inheritance in Drosophila. *Science*, *32*(812), 120–122. https://doi.org/10.1126/science.32.812.120
Morrison, D. K., Murakami, M. S., & Cleghon, V. (2000). Analysis Protein Kinases and Phosphatases in the Drosophila Genome Identification and Classification of Drosophila Protein Kinases and Phosphatases. *The Journal of Cell Biology*, *150*(2), 57–62. Retrieved from http://www.jcb.org
Muha, V., & Müller, H. A. J. (2013, March 14). Functions and mechanisms of fibroblast growth factor (FGF) signalling in Drosophila melanogaster. *International Journal of Molecular Sciences*. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/ijms14035920
Mustelin, T. (2007). Introduction to Protein Phosphatase Families A Brief Introduction to the Protein Phosphatase Families. In G. Moorhead (Ed.), *Protein Phosphatase Protocols. Methods in Molecular Biology, vol 365.* (pp. 9–22). Totowa, NJ: Springer. https://doi.org/https://doi.org/10.1385/1-59745-267-X:9

Narasimha, M., & Brown, N. H. (2006). Confocal Microscopy of Drosophila Embryos. *Cell Biology: A Laboratory Handbook*, 77–86.

Naruse, Y., Aoki, T., Kojima, T., & Mori, N. (1999). Neural restrictive silencer factor recruits mSin3 and histone deacetylase complex to repress neuron-specific target genes. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(24), 13691–6. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=24126&tool=pmcentrez&rendertype=abstract

Nesti, E., Corson, G. M., McCleskey, M., Oyer, J. A., & Mandel, G. (2014). C-terminal domain small phosphatase 1 and MAP kinase reciprocally control REST stability and neuronal differentiation. *Proceedings of the National Academy of Sciences*, *111*(37), E3929-3936. https://doi.org/10.1073/pnas.1414770111

Nüsslein-volhard, C., & Wieschaus, E. (1980). Mutations affecting segment number and polarity in drosophila. *Nature*, *287*(5785), 795–801. https://doi.org/10.1038/287795a0

Nüsslein-Volhard, C., Wieschaus, E., & Kluding, H. (1984). Mutations affecting the pattern of the larval cuticle inDrosophila melanogaster. *Wilhelm Roux's Archives of Developmental Biology*, *193*(5), 267–282. https://doi.org/10.1007/BF00848156

O'Connor, M. B. (2005). Shaping BMP morphogen gradients in the Drosophila embryo and pupal wing. https://doi.org/10.1242/dev.02214

Ostrowski, S., Dierick, H. A., & Bejsovec, A. (2002). Genetic control of cuticle formation during embryonic development of Drosophila melanogaster. *Genetics*, *161*(1), 171–182. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12019232

Pallarès, I., & Ventura, S. (2017). The Transcription Terminator Rho: A First Bacterial Prion. *Trends in Microbiology*. https://doi.org/10.1016/j.tim.2017.03.008

Pandey, U. B., & Nichols, C. D. (2011). Human disease models in Drosophila melanogaster and the role of the fly in therapeutic drug discovery. *Pharmacological Reviews*, *63*(2), 411–36. https://doi.org/10.1124/pr.110.003293

Parker, L., Stathakis, D. G., & Arora, K. (2003). Regulation of BMP and Activin Signaling in Drosophila. *Progress in Molecular and Subcellular Biology*, *34*, 73–101. https://doi.org/10.1007/978-3-642-18670-7_4

Patel, B. K., Gavin-Smyth, J., & Liebman, S. W. (2009). The yeast global transcriptional co-repressor protein Cyc8 can propagate as a prion. *Nature Cell Biology*, *11*(3), 344–349. https://doi.org/10.1038/ncb1843

Pavelka, N., Fournier, M. L., Swanson, S. K., Pelizzola, M., Ricciardi-Castagnoli, P., Florens, L., & Washburn, M. P. (2008). Statistical Similarities between Transcriptomics and Quantitative Shotgun Proteomics Data. *Molecular & Cellular Proteomics*, *7*(4), 631–644. https://doi.org/10.1074/mcp.M700240-MCP200

Payre, F. (2004, January 1). Genetic control of epidermis differentiation in Drosophila. *International Journal of Developmental Biology*. UPV/EHU Press. https://doi.org/10.1387/ijdb.15272387

Payre, F., Vincent, A., & Carreno, S. (1999). ovo/svb Integrates Wingless and DER pathways to control epidermis differentiation. *Nature*, *400*(6741), 271–275. https://doi.org/10.1038/22330

Perrimon, N., Pitsouli, C., & Shilo, B. Z. (2012). Signaling mechanisms controlling cell fate and embryonic patterning. *Cold Spring Harbor Perspectives in Biology*, *4*(8), a005975. https://doi.org/10.1101/cshperspect.a005975

Petit, V., Nussbaumer, U., Dossenbach, C., & Affolter, M. (2004). Downstream-of-FGFR is a fibroblast growth factor-specific scaffolding protein and recruits Corkscrew upon receptor activation. *Molecular and Cellular Biology*, *24*(9), 3769–3781. https://doi.org/10.1128/MCB.24.9.3769–3781.2004

Poulson, D. F. (1937). Chromosomal Deficiencies and the Embryonic Development of Drosophila Melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, *23*(3), 133–137. https://doi.org/10.1073/pnas.23.3.133

Preibisch, S., Saalfeld, S., & Tomancak, P. (2009). Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics*, *25*(11), 1463–1465. https://doi.org/10.1093/bioinformatics/btp184

Prilusky, J., Felder, C. E., Zeev-Ben-Mordehai, T., Rydberg, E. H., Man, O., Beckmann, J. S., … Sussman, J. L. (2005). FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics (Oxford, England)*, *21*(16), 3435–8. https://doi.org/10.1093/bioinformatics/bti537

Prusiner, S. (1982). Novel proteinaceous infectious particles cause scrapie. *Science*, *216*(4542), 136–144. https://doi.org/10.1126/science.6801762

Prusiner, S. (1991). Molecular biology of prion diseases. *Science*, *252*(5012), 1515–1522. https://doi.org/10.1126/science.1675487

Prusiner, S. B. (2013). Biology and Genetics of Prion Causing Neurodegeneration. *Annual Review of Genetics*, *47*, 601–23. https://doi.org/10.1146/annurev-genet-110711-155524

Ramakrishnan, A.-B., & Cadigan, K. M. (2017). Wnt target genes and where to find them. *F1000Research*, *6*, 746. https://doi.org/10.12688/f1000research.11034.1

Raveendra, B. L., Siemer, A. B., Puthanveettil, S. V, Hendrickson, W. A., Kandel, E. R., & McDermott, A. E. (2013). Characterization of prion-like conformational changes of the neuronal isoform of Aplysia CPEB. *Nature Structural and Molecular Biology*, *20*(4), 495–501. https://doi.org/10.1038/nsmb.2503

Rebay, I. (2002). Keeping the receptor tyrosine kinase signaling pathway in check: Lessons from Drosophila. *Developmental Biology*. https://doi.org/10.1006/dbio.2002.0806

Rebay, I., Silver, S. J., & Tootle, T. L. (2005, March 1). New vision from Eyes absent: Transcription factors as enzymes. *Trends in Genetics*. Elsevier. https://doi.org/10.1016/j.tig.2005.01.005

Richter, J. D. (2007, June 1). CPEB: a life in translation. *Trends in Biochemical Sciences*. Elsevier. https://doi.org/10.1016/j.tibs.2007.04.004

Rogoza, T., Goginashvili, A., Rodionova, S., Ivanov, M., Viktorovskaya, O., Rubel, A., … Mironova, L. (2010). Non-Mendelian determinant [ISP+] in yeast is a nuclear-residing prion form of the global transcriptional regulator Sfp1. *Proceedings of the National Academy of Sciences*, *107*(23), 10573–10577. https://doi.org/10.1073/pnas.1005949107

Ross, E. D., Minton, A., & Wickner, R. B. (2005). Prion domains: sequences, structures and interactions. *Nature Cell Biology*, *7*(11), 1039–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16385730

Rothwell, W. F., Zhang, C. X., Zelano, C., Hsieh, T., & Sullivan, W. (1999). The Drosophila centrosomal protein Nuf is required for recruiting Dah, a membrane associated protein, to furrows in the early embryo. *Journal of Cell Science*, *112*, 2885–2893.

Roy, S., Ernst, J., Kharchenko, P. V, Kheradpour, P., Negre, N., Eaton, M. L., … Lowdon, R. F. (2010). Identification of functional elements and regulatory circuits by Drosophila modENCODE. *Science*, *330*(6012), 1787–1797. https://doi.org/10.1126/science.1198374

Sabate, R., Rousseau, F., Schymkowitz, J., & Ventura, S. (2015). What Makes a Protein Sequence a Prion? *PLoS Computational Biology*, *11*(1), e1004013. https://doi.org/10.1371/journal.pcbi.1004013

Sacco, F., Perfetto, L., Castagnoli, L., & Cesareni, G. (2012). The human phosphatase interactome: An intricate family portrait. *FEBS Letters*, *586*(17), 2732–2739. https://doi.org/10.1016/j.febslet.2012.05.008

Sanson, B. (2001, December 1). Generating patterns from fields of cells: Examples from Drosophila segmentation. *EMBO Reports*. EMBO Press. https://doi.org/10.1093/embo-reports/kve255

Sapkota, G., Knockaert, M., Alarcón, C., Montalvo, E., Brivanlou, A. H., & Massagué, J. (2006).

Dephosphorylation of the Linker Regions of Smad1 and Smad2/3 by Small C-terminal Domain Phosphatases Has Distinct Outcomes for Bone Morphogenetic Protein and Transforming Growth Factor-$_{NL}$ Pathways *. https://doi.org/10.1074/jbc.M610172200

Schoenherr, C. J., Paquette, A. J., & Anderson, D. J. (1996). Identification of potential target genes for the neuron-restrictive silencer factor. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(18), 9881–6. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=38523&tool=pmcentrez&rendertype=abstract

Schüpbach, T., & Wieschaus, E. (1986). Maternal-effect mutations altering the anterior-posterior pattern of the Drosophila embryo. *Roux's Archives of Developmental Biology*, *195*(5), 302–317. https://doi.org/10.1007/BF00376063

Seifried, A., Schultz, J., & Gohla, A. (2013). Human HAD phosphatases: structure, mechanism, and roles in health and disease. *FEBS Journal*, *280*(2), 549–571. https://doi.org/10.1111/j.1742-4658.2012.08633.x

Serizawa, H., Conaway, R. C., & Conaway, J. W. (1992). A carboxyl-terminal-domain kinase associated with RNA polymerase II transcription factor delta from rat liver. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(16), 7476–80. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=49733&tool=pmcentrez&rendertype=abstract

Shenolikar, S. (2007). Analysis of Protein Phosphatases: Toolbox for Unraveling Cell Signaling Networks. In G. Moorhead (Ed.), *Protein Phosphatase Protocols. Methods in Molecular Biology, vol 365.* (pp. 1–8). Totowa, NJ: Springer. https://doi.org/10.1385/1-59745-267-X:1

Shorter, J., & Lindquist, S. (2005). Prions as adaptive conduits of memory and inheritance. *Nature Reviews. Genetics*, *6*(6), 435–50. https://doi.org/10.1038/nrg1616

Si, K. (2015). Prions: What Are They Good For? *Annu. Rev. Cell Dev. Biol*, *31*, 149–69. https://doi.org/10.1146/annurev-cellbio-100913-013409

Si, K., Choi, Y.-B., White-Grindley, E., Majumdar, A., & Kandel, E. R. (2010). Aplysia CPEB can form prion-like multimers in sensory neurons that contribute to long-term facilitation. *Cell*, *140*(3), 421–35. https://doi.org/10.1016/j.cell.2010.01.008

Si, K., Giustetto, M., Etkin, A., Hsu, R., Janisiewicz, A. M., Miniaci, M. C., … Kandel, E. R. (2003). A Neuronal Isoform of CPEB Regulates Local Protein Synthesis and Stabilizes Synapse-Specific Long-Term Facilitation in Aplysia. *Cell*, *115*(7), 893–904. https://doi.org/10.1016/S0092-8674(03)01021-3

Siegfried, E., Chou, T.-B., Perrimon, N., Akam, M., Baker, N. E., Bejsovec, A., … Woodgett, J. R. (1992). wingless signaling acts through zeste-white 3, the drosophila homolog of glycogen synthase kinase-3, to regulate engrailed and establish cell fate. *Cell*, *71*(7), 1167–1179. https://doi.org/10.1016/S0092-8674(05)80065-0

Simcox, A., & Roberts, I. (1989). Imaginal discs can be recovered from cultured embryos mutant for the segment-polarity genes engrailed, naked and patched but not from wingless. *Development*, *107*(4), 715–722. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23862219

Snyder, M., Zhu, H., Klemic, J. F., Chang, S., Bertone, P., Casamayor, A., … Reed, M. A. (2000). Analysis of yeast protein kinases using protein chips. *Nature Genetics*, *26*(3), 283–289. https://doi.org/10.1038/81576

Soto, C. (2012). Transmissible proteins: expanding the prion heresy. *Cell*, *149*(5), 968–77. https://doi.org/10.1016/j.cell.2012.05.007

Staaf, J., Lindgren, D., Vallon-Christersson, J., Isaksson, A., Göransson, H., Juliusson, G., … Ringnér, M. (2008). Segmentation-based detection of allelic imbalance and loss-of-heterozygosity in cancer cells using whole genome SNP arrays. *Genome Biology*, *9*(9), R136. https://doi.org/10.1186/gb-2008-9-9-r136

Städeli, R., & Basler, K. (2005). Dissecting nuclear Wingless signalling: Recruitment of the transcriptional co-activator Pygopus by a chain of adaptor proteins. *Mechanisms of Development*, *122*(11), 1171–1182. https://doi.org/10.1016/j.mod.2005.07.004

Stamos, J. L., & Weis, W. I. (2013). The β-catenin destruction complex. *Cold Spring Harbor Perspectives in Biology*, *5*(1), a007898. https://doi.org/10.1101/cshperspect.a007898

Stroobants, K., Kumita, J. R., Harris, N. J., Chirgadze, D. Y., Dobson, C. M., Booth, P. J., & Vendruscolo, M. (2017). Amyloid-like Fibrils from an α-Helical Transmembrane Protein. *Biochemistry*, *56*(25), 3225–3233. https://doi.org/10.1021/acs.biochem.7b00157

Swarup, S., Pradhan-Sundd, T., & Verheyen, E. M. (2015). Genome-wide identification of phospho-regulators of Wnt signaling in Drosophila. *Development*, *142*(8), 1502–1515. https://doi.org/10.1242/dev.116715

Swarup, S., & Verheyen, E. M. (2012). Wnt/wingless signaling in drosophila. *Cold Spring Harbor Perspectives in Biology*, *4*(6), 1–15. https://doi.org/10.1101/cshperspect.a007930

Tacchelly-Benites, O., Wang, Z., Yang, E., Benchabane, H., Tian, A., Randall, M. P., & Ahmed, Y. (2018). Axin phosphorylation in both Wnt-off and Wnt-on states requires the tumor suppressor APC. *PLoS Genetics*, *14*(2), e1007178. https://doi.org/10.1371/journal.pgen.1007178

Tariq, M., Wegrzyn, R., Anwar, S., Bukau, B., & Paro, R. (2013). Drosophila GAGA factor polyglutamine domains exhibit prion-like behavior. *BMC Genomics*, *14*(1), 374. https://doi.org/10.1186/1471-2164-14-374

Tayeb-Fligelman, E., Tabachnikov, O., Moshe, A., Goldshmidt-Tran, O., Sawaya, M. R., Coquelle, N., … Landau, M. (2017). The cytotoxic Staphylococcus aureus PSMα3 reveals a cross-α amyloid-like fibril. *Science*, *355*(6327), 831–833. https://doi.org/10.1126/science.aaf4901

Toyama, B. H., & Weissman, J. S. (2011). Amyloid Structure: Conformational Diversity and Consequences. *Annual Review of Biochemistry*, *80*(1), 557–585. https://doi.org/10.1146/annurev-biochem-090908-120656

True, H. L., Berlin, I., & Lindquist, S. L. (2004). Epigenetic regulation of translation reveals hidden genetic variation to produce complex traits. *Nature*, *431*(7005), 184–7. https://doi.org/10.1038/nature02885

Tsuda, L., Kaido, M., Lim, Y.-M., Kato, K., Aigaki, T., & Hayashi, S. (2006). An NRSF/REST-like repressor downstream of Ebi/SMRTER/Su(H) regulates eye development in Drosophila. *The EMBO Journal*, *25*(13), 3191–202. https://doi.org/10.1038/sj.emboj.7601179

Turner, F. R., & Mahowald, A. P. (1976). Scanning electron microscopy of Drosophila embryogenesis: 1. The structure of the egg envelopes and the formation of the cellular blastoderm. *Developmental Biology*, *50*(1), 95–108. https://doi.org/10.1016/0012-1606(76)90070-1

Urban, S. (2004). EGF receptor signalling protects smooth-cuticle cells from apoptosis during Drosophila ventral epidermis development. *Development*, *131*(8), 1835–1845. https://doi.org/10.1242/dev.01058

Veit, M., & Schmidt, M. F. (1998). Membrane targeting via protein palmitoylation. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 88, pp. 227–39). New Jersey: Humana Press. https://doi.org/10.1385/0-89603-487-9:227

Verheyen, E. M., & Gottardi, C. J. (2010). Regulation of Wnt/beta-catenin signaling by protein kinases. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *239*(1), 34–44. https://doi.org/10.1002/dvdy.22019

Visvanathan, J., Lee, S., Lee, B., Lee, J. W., & Lee, S. K. (2007). The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. *Genes and Development*, *21*(7), 744–749. https://doi.org/10.1101/gad.1519107

Wani, S., Sugita, A., Ohkuma, Y., & Hirose, Y. (2016). Human SCP4 is a chromatin-associated CTD phosphatase and exhibits the dynamic translocation during erythroid differentiation. *Journal of Biochemistry*, *160*(2), 111–120. https://doi.org/10.1093/jb/mvw018

Weigmann, K., Klapper, R., Strasser, T., Rickert, C., Technau, G., Jäckle, H., … Klämbt, C. (2003). FlyMove - A new way to look at development of Drosophila. *Trends in Genetics*, *19*(6), 310–311. https://doi.org/10.1016/S0168-9525(03)00050-7

Wharton, K. A., Yedvobnick, B., Finnerty, V. G., Tsakonas, S. A., & Artavanis-Tsakonas, S. (1985). opa : A Novel Family of Transcribed Repeats Shared by the Notch Locus and Other Developmentally Regulated Loci in D. melanogaster. *Cell*, *40*(1), 55–62. https://doi.org/10.1016/0092-8674(85)90308-3

White-Grindley, E., Li, L., Mohammad Khan, R., Ren, F., Saraf, A., Florens, L., & Si, K. (2014). Contribution of Orb2A Stability in Regulated Amyloid-Like Oligomerization of Drosophila Orb2. *PLoS Biology*, *12*(2), e1001786. https://doi.org/10.1371/journal.pbio.1001786

Wickner, R. (1994). [URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae. *Science*, *264*(5158), 566–569. https://doi.org/10.1126/science.7909170

Wickner, R. B., Edskes, H. K., Shewmaker, F., & Nakayashiki, T. (2007). Prions of fungi: inherited structures and biological roles. *Nature Reviews. Microbiology*, *5*(8), 611–8. https://doi.org/10.1038/nrmicro1708

Wickner, R. B., Edskes, H. K., Shewmaker, F. P., Kryndushkin, D., Nemecek, J., Mcglinchey, R., & Bateman, D. (2010). The relationship of prions and translation. *Wiley Interdisciplinary Reviews: RNA*. NIH Public Access. https://doi.org/10.1002/wrna.8

Wickner, R. B., Taylor, K. L., Edskes, H. K., & Maddelein, M. L. (2000). Prions: Portable prion domains. *Current Biology : CB*, *10*(9), R335-7. https://doi.org/10.1016/S0960-9822(00)00460-7

Wieschaus, E., & Nüsslein-Volhard, C. (2016). The Heidelberg Screen for Pattern Mutants of Drosophila : A Personal Account. *Annual Review of Cell and Developmental Biology*, *32*(1), 1–46. https://doi.org/10.1146/annurev-cellbio-113015-023138

Wodarz, A. (2008). Extraction and Immunoblotting of Proteins From Embryos. In *Drosophila Methods and Protocols* (pp. 335–345). https://doi.org/10.1007/978-1-59745-583-1_21

Wong, K. H., & Struhl, K. (2011). The Cyc8-Tup1 complex inhibits transcription primarily by masking the activation domain of the recruiting protein. *Genes and Development*, *25*(23), 2525–2539. https://doi.org/10.1101/gad.179275.111

Wouda, R. R., Bansraj, M. R. K. S., de Jong, A. W. M., Noordermeer, J. N., & Fradkin, L. G. (2008). Src family kinases are required for WNT5 signaling through the Derailed/RYK receptor in the Drosophila embryonic central nervous system. *Development (Cambridge, England)*, *135*(13), 2277–2287. https://doi.org/10.1242/dev.017319

Wrighton, K. H., Willis, D., Long, J., Liu, F., Lin, X., Feng, X.-H., & Debakey, M. E. (2006). Small C-terminal Domain Phosphatases Dephosphorylate the Regulatory Linker Regions of Smad2 and Smad3 to Enhance Transforming Growth Factor- B Signaling. *Journal of Biological Chemistry*, *281*(50), 38365–38375. https://doi.org/10.1074/jbc.M607246200

Wu, H., & Fuxreiter, M. (2016). The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell*. https://doi.org/10.1016/j.cell.2016.05.004

Wu, Y., Mark Evers, B., & Zhou, B. P. (2009). Small C-terminal domain phosphatase enhances snail activity through dephosphorylation. *Journal of Biological Chemistry*, *284*(1), 640–648. https://doi.org/10.1074/jbc.M806916200

Xia, F., Li, J., Hickey, G. W., Tsurumi, A., Larson, K., Guo, D., … Hafen, E. (2008). Raf activation is regulated by tyrosine 510 phosphorylation in Drosophila. *PLoS Biology*, *6*(5), 1115–1129. https://doi.org/10.1371/journal.pbio.0060128

Xia, L., Jia, S., Huang, S., Wang, H., Zhu, Y., Mu, Y., … Chen, D. (2010). The Fused/Smurf complex controls the fate of Drosophila germline stem cells by generating a gradient BMP response. *Cell*, *143*(6), 978–990. https://doi.org/10.1016/j.cell.2010.11.022

Xu, H., He, X., Zheng, H., Huang, L. J., Hou, F., Yu, Z., … Jiang, Q.-X. (2014). Structural basis for the prion-like MAVS filaments in antiviral innate immunity. *ELife*, *3*, e01489.

https://doi.org/10.7554/eLife.01489
Yang, L., Meng, F., Ma, D., Xie, W., & Fang, M. (2013). Bridging Decapentaplegic and Wingless signaling in Drosophila wings through repression of naked cuticle by Brinker. *Development*, *140*(2), 413–422. https://doi.org/10.1242/dev.082578
Yeo, M., Lee, S.-K., Lee, B., Ruiz, E. C., Pfaff, S. L., & Gill, G. N. (2005). Small CTD phosphatases function in silencing neuronal gene expression. *Science (New York, N.Y.)*, *307*(5709), 596–600. https://doi.org/10.1126/science.1100801
Yeo, M., Lin, P. S., Dahmus, M. E., & Gill, G. N. (2003). A novel RNA polymerase II C-terminal domain phosphatase that preferentially dephosphorylates serine 5. *The Journal of Biological Chemistry*, *278*(28), 26078–85. https://doi.org/10.1074/jbc.M301791200
Yuan, A. H., & Hochschild, A. (2017). A bacterial global regulator forms a prion. *Science*, *355*(6321). https://doi.org/10.1126/science.aai7776
Zambrano, R., Conchillo-Sole, O., Iglesias, V., Illa, R., Rousseau, F., Schymkowitz, J., … Ventura, S. (2015). PrionW: A server to identify proteins containing glutamine/asparagine rich prion-like domains and their amyloid cores. *Nucleic Acids Research*, *43*(W1), W331–W337. https://doi.org/10.1093/nar/gkv490
Zeng, W., Wharton, K. A., Mack, J. A., Wang, K., Gadbaw, M., Suyama, K., … Scott, M. P. (2000). naked cuticle encodes an inducible antagonist of Wnt signalling. *Nature*, *403*(6771), 789–795. https://doi.org/10.1038/35001615
Zeng, X., Tamai, K., Doble, B., Li, S., Huang, H., Habas, R., … He, X. (2005). A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation. *Nature*, *438*(7069), 873–877. https://doi.org/10.1038/nature04185
Zhang, C. X., Rothwell, W. F., Sullivan, W., & Hsieh, T. S. (2000). Discontinuous actin hexagon, a protein essential for cortical furrow formation in Drosophila, is membrane associated and hyperphosphorylated. *Molecular Biology of the Cell*, *11*(March), 1011–1022.
Zhang, M., Cho, E. J., Burstein, G., Siegel, D., & Zhang, Y. (2011). Selective inactivation of a human neuronal silencing phosphatase by a small molecule inhibitor. *ACS Chemical Biology*, *6*(5), 511–9. https://doi.org/10.1021/cb100357t
Zhao, L., Wang, L., Chi, C., Lan, W., & Su, Y. The emerging roles of phosphatases in Hedgehog pathway, 15 Cell Communication and Signaling 35 (2017). BioMed Central. https://doi.org/10.1186/s12964-017-0191-0