MQP-BIO-JBD-0002

Structural and Functional Dissection of Kekkon5: A Putative Novel Component of the Cellular Adhesion Machinery

A Major Qualifying Project Report

Submitted to the Faculty of

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Bachelor of Science

In

Biology and Biotechnology

By

Christina L. Ernst

April 24, 2008

APPROVED:

Joseph B. Duffy, Ph.D.
Biology and Biotechnology
WPI Project Advisor

ABSTRACT

Cellular adhesion is an essential factor in development, serving both structural and communicatory functions, defects in which have been implicated in a wide range of diseases, including cardiovascular and neurodegenerative disorders. LIGs, a small novel class of proteins containing Leucine-rich repeats (LRRs) and Immunoglobulin-like (Ig) domains, have been found to function in several neuronal activities. Integrins are a family of transmembrane proteins that are critical in cellular adhesion, both through physical binding and downstream signaling. Kekkon5 (Kek5), a *Drosophila* LIG protein, has been implicated in integrin and other signaling pathways. Loss of function studies with Kek5 confirm its interaction with integrins, while deletion studies have demonstrated the necessity of both extracellular and intracellular domains for full activity and function, with special regard to the extracellular LRRs and intracellular motifs. Future studies aim to further characterize essential structural elements and functional mechanisms of Kek5.

ACKNOWLEDGEMENTS

First, I would like to thank Duff for being an exemplary advisor, keeping us all in line and motivating us to do the best science we can, all the while maintaining an intimate, enjoyable research environment. Also, thank you for all the career advice and lifelong lessons; I'm sure they will someday prove to be very rewarding. Thank you to veteran and former lab mates Harita Haridas, Prachi Gupta, and Cassie Leduc, who welcomed me when I first joined the Duffy lab and donated their invaluable time to helping me learn the ropes before beginning my MQP. Thanks as well to my newer lab mates Aditi Guatam and Edie Plada for their support and enlightening discussions over the past year. I would also like to thank Reeta Prusty-Rao for recommending I check out the Duffy lab; without your support and guidance, I wouldn't have been able to make my senior year such a success. Thanks to Liz Ryder and Sam Politz for the use of their microscopes, as well as to Liz for being a supportive and enjoyable academic advisor over the past four years. Finally, a huge thank you to the faculty and staff of the Biology and Biotechnology Department for their support throughout my undergraduate career at WPI.

TABLE OF CONTENTS

INTRODUCTION

With the advent of genome sequencing, we have gained invaluable insight into the stunningly complex world of proteins. The structure of proteins is often modular in nature, containing any number and combination of distinct motifs and domains. Even among one family, such as one containing leucine-rich repeats (LRRs), these can range from proteins as simple as Lrrc3, which contains only three LRRs, to immensely complex proteins, such as Pkd1, which contains multiple LRRs, PKD repeats, a REJ domain, a C-type lectin domain, an LH2 domain, and multiple transmembrane domains (Dolan et al., 2007). This proteomic complexity, especially in higher eukaryotes, is achieved in part by the phenomenon known as alternative mRNA splicing, where different combinations of exons from the same genomic transcript are spliced together to create functionally unique proteins. In extreme examples, a single *Drosophila* mRNA transcript may allow for more than 38,000 different proteins, in comparison to its approximately 20,000 total genes.

LIG Proteins: Structure and Function

LIGs represent one class of molecules that contains a distinct modular organization, consisting of two of the most common sequence elements found in both vertebrate and invertebrate proteins: leucine-rich repeats (LRR) and immunoglobulin-like (Ig-like) domains. LRRs, each of which is 20-29 amino acids in length and contains the sequence LxxLxLxxN/CxL, where x may be any amino acid. These repeats are connected by loops, which together form an overall horseshoe-like conformation, and are involved in mediating protein-protein interactions. A significant subset of LRR-containing proteins has been implicated in neuronal outgrowth and synapse differentiation, while others include the Toll-like, NOD-like,

and variable lymphocyte receptors of the immune system (Ko and Kim, 2007; Pålsson-McDermott and O'Neill, 2007).

Ig domains are traditionally associated with antibodies, T-cell receptors and other immune response proteins, but have been found in numerous other protein families, such as cell surface receptors, matrix proteins, and intracellular regulatory proteins, indicating a wide range of activities (Bork et al., 1994). Although LRRs and Ig domains are two of the most abundant sequence elements (approximately 350 and 1100 in humans, and 124 and 228 in flies, respectively), their combination in a single protein is relatively rare; only 35 molecules with an LRR and Ig combination have been identified in humans, and only nine in *Drosophila* (Evans, 2006).

The LIG family is not yet well characterized, but recent evidence points to roles in signal transduction and involvement in neuronal development. For example, Kekkon1 (Kek1) has been demonstrated to interact with the Epidermal Growth Factor Receptor (EGFR), while LINGO-1 has been found to inhibit neuronal outgrowth by interacting with the p75NTR-NgR coreceptor complex (MacLaren et al., 2004; Chen et al., 2006).

Six of the nine LIG proteins found in *Drosophila*, including Kek1, are members of the Kekkon (Kek) family, which contain seven LRRs and single Ig-like domain. Kek1, the founding member of the Kek family and the most well characterized, interacts with the EGFR and is potentially involved in neuronal pathfinding (MacLaren et al., 2004). The remainder of the Kek family has yet to be fully investigated. Although no vertebrate orthologs of the Kek family exist, it is interesting to note that the vertebrate protein family known as the AMIGOs, also contains seven LRRs and one Ig domain, an arrangement of LRRs and Ig domains found only in these two families (Chen et al., 2006).

BMPs & Integrins

Bone Morphogenetic Proteins (BMPs) are part of a family of secreted signaling proteins that are essential for proper development and patterning of both vertebrates and invertebrates. It was first discovered as a regulating factor in bone development in vertebrates (Little and Mullins, 2006). In *Drosophila*, it is now known to regulate diverse developmental events ranging from early dorsal-ventral patterning to crossvein formation in pupal wings. BMPs are synthesized as inactive "pro-proteins," which then undergo molecular modification, resulting in an active form that is able to dimerize and subsequently bind BMP receptors (Little and Mullins, 2006). BMPs often have the capacity to act as a morphogen, which is a secreted molecule that acts in a gradient-dependent fashion. There is extensive extracellular regulation of BMP activity and there are several molecules that inhibit the BMP-receptor interaction, including Sog (Short gastrulation), Tsg (Twisted gastrulation), and Tolloid (Tld), which associate with the BMP ligand and regulate its diffusion and ability to associate with the receptor complex. For example, this method of molecular migration causes an elevated level of BMP in dorsal region of the embryo, resulting in the differentiation of the dorsal epidermis (Little and Mullins, 2006).

Integrins are a family of proteins involved in cellular adhesion, a critical biological function. Integrins function in the attachment of cells to the ECM (extracellular matrix), intracellular signaling, cellular migration, and synthesis of the ECM itself. The functional dimer consists of an α and β subunit, five and two of which have been discovered in *Drosophila*, respectively (Bökel and Brown, 2002). In some cases, the α subunits may be partially functionally redundant, but in cases such as the wing, expression is complimentary, where the α_{PS1} subunit is expressed on the dorsal surface, while the α_{PS2} subunit is expressed on the ventral surface (Brower et al., 1995). At the cellular level, integrins are involved in focal adhesion

complexes that can facilitate cellular adhesion, in which integrin tails will "recruit" proteins that bind actin filaments to create a connection to the ECM (Takada et al., 2007).

Drosophila Wing Development

The *Drosophila* wing has proven to be a useful tool in studying the molecular mechanisms of growth, patterning, and morphogenesis. In particular, the *Drosophila* wing has contributed significantly to our understanding of both the BMP and integrin pathways. The development of the adult wing is a well-defined process, beginning as a flat bilayer of cells known as an imaginal disc (Figure 1). The central pouch of cells, or the disc proper, eventually forms the wing itself, while cells in the periphery contribute to the thorax (Fristrom and Fristrom, 1993).

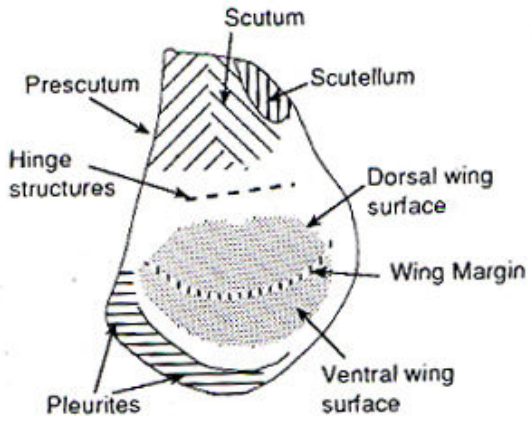


Figure 1: Wing imaginal disc (Fristrom and Fristrom, 1993)

Beginning around two hours after pupariation (AP), cells of the disc proper expand and begin to fold along the wing margin, eventually bending back on itself so that the dorsal and ventral surfaces are nearly touching (Figure 2). By four hours AP, the prepupal wing disc has completely stretched out and folded upon itself (Fristrom and Fristrom, 1993).

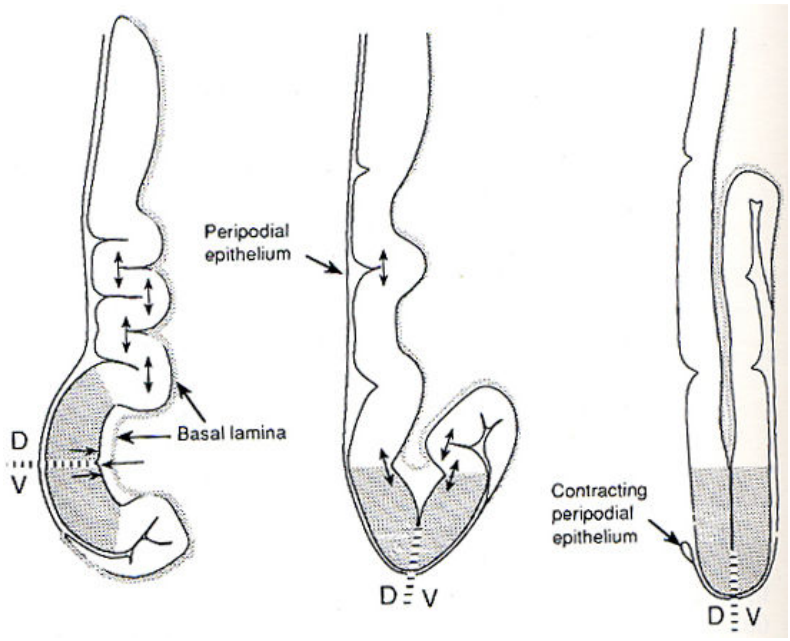


Figure 2: Development of wing imaginal disc to prepupal wing (Fristrom and Fristrom, 1993)

Further prepupal and pupal wing metamorphosis occurs in two series, with four primary steps each: apposition, adhesion, expansion, and separation, summarized in Figure 3. During prepupal development, the dorsal and ventral wing surfaces come into contact during the first four hours AP. The surfaces flatten out and expand during hours 4 to 7, followed by the formation of preveins. Between hours 8 to 11, basal junctions are then proposed to form between the bilayer. Finally at around 11 hours AP, the bilayer separates again, followed by the final mitoses, and eventually, pupation. The second apposition of the dorsal and ventral surfaces occurs from hours 20 to 35, followed by another period of adhesion in intervein regions around 40 hours AP. Between hours 45 to 50, the wing undergoes expansion again, increasing its surface area by two- to three-fold. Finally, at around 60 hours AP, the wing undergoes another, more subtle separation of the bilayers, forming the final conformation of the basal junctions. After eclosion, around 96 hours AP, the wing must unfold and the intervein cells degrade, resulting in the directly apposed layers of the adult wing (Fristrom et al., 1993).

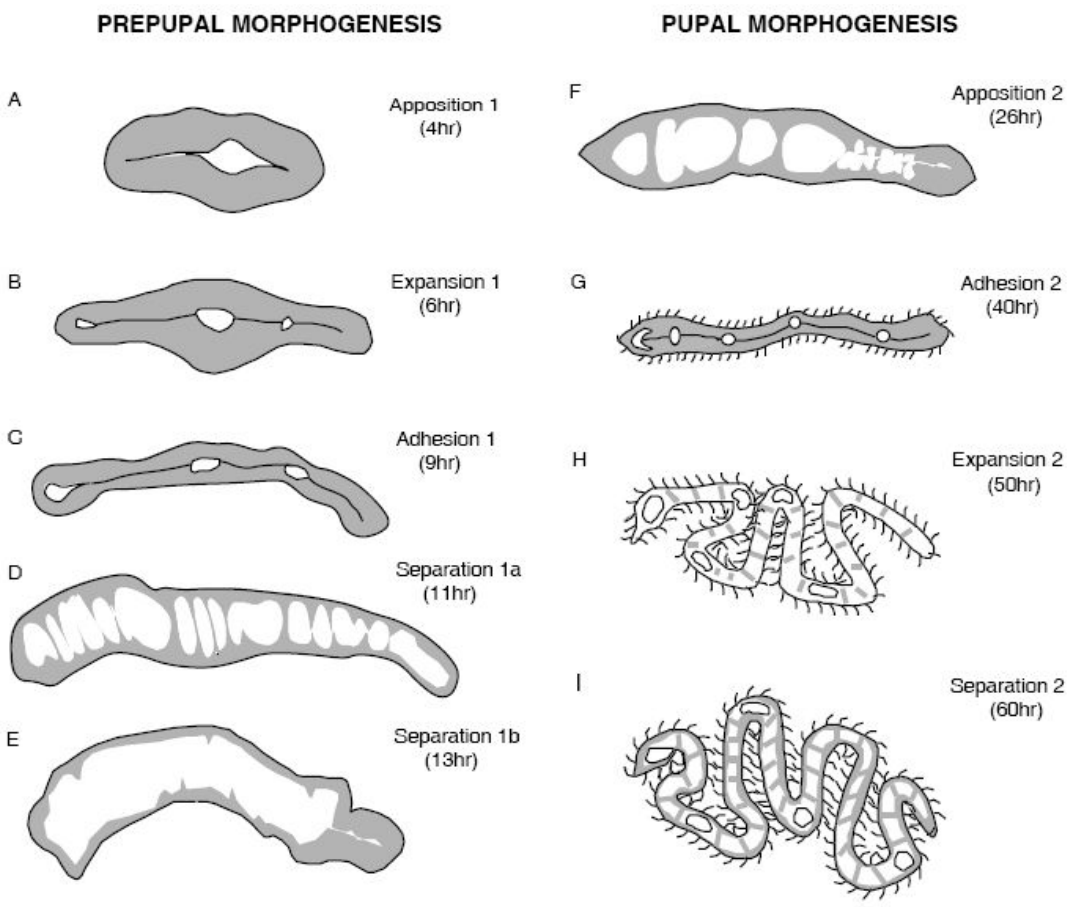


Figure 3: Prepupal and pupal wing development (Fristrom et al., 1993)

Both BMP-like ligands and integrins are highly involved in wing development. Decapentaplegic (Dpp), a *Drosophila* BMP expressed along the anterior-posterior boundary of the imaginal disc, is essential for growth and patterning of the wing, and has significant effects on vein specification (Affolter and Basler, 2007). Integrins, on the other hand, are involved in adhesion of the wing's bilayer. This is evidenced by the classic blister phenotype displayed by hypomorphic variants of the various integrin subunits, along with a less common, more subtle wrinkling phenotype, implying that proper adhesion of the bilayer requires fully functional integrins (Brower et al., 1995). Subunit α_{PS1}, encoded by *multiple edematous wings* (*mew*), is expressed on the dorsal wing surface, while α_{PS2}, encoded by *inflated* (*if*), is expressed on the

ventral wing surface. Integrin dimers on both wing surfaces share the same β_{PS} subunit, encoded by *myospheroid* (*mys*) (Brower et al., 1995).

Kek5: Prior Work

Work on Kek5 was begun previously in the Duffy lab by Timothy Evans. He determined that Kek5 is functionally distinct from Kek1, as it does not affect the EGFR pathway, implicating that it has its own unique role in *Drosophila* development. Null mutants for Kek5 demonstrated its functional significance *in vivo*, as they exhibited reduced viability and wing abnormalities, including altered vein patterning and blisters – phenotypic hallmarks of the BMP and integrin pathways (Evans, 2006).

Through loss- and gain-of-function studies, it was further proposed that Kek5 inhibits BMP signaling, likely through the alteration of extracellular ligand distribution. This was initially suggested by the similarity in wing phenotypes, as Kek5 misexpression exhibits phenotypes similar to those observed in BMP mutants - ranging from small, poorly patterned discs and wings to crossvein defects in the adult wing. Kek5 was also observed to interact with the integrin signaling pathway. Again, this was suggested by similarities in wing phenotypes, with both integrin and Kek5 mutations displaying wing blisters. Further support was provided by genetic interactions between *kek5* and the α_{PS2} integrin subunit *inflated* (*if*), as well as the examination of the phenocritical periods for *kek5* and integrin expression in wing development, which appear to coincide (Evans, 2006).

To better understand the function of Kek5 in the processes previously listed, a series of Kek5 misexpression variants were generated. These included single deletions of the LRRs, the Ig domain, and the cytoplasmic domain, as well as domain swaps with the closely related Kek

family member, Kek4. The analyses of these variants support the following: the LRRs (and not the Ig domain) are important for the function of Kek5 in BMP and integrin signaling, and the presence of an intracellular PDZ domain binding site is essential for Kek5 localization in the cellular membrane. PDZ domains are 80-90 amino acid motifs that regulate protein localization and complex assembly. Kek5 contains a Type I PDZ domain binding site, which is present in proteins implicated in neurological function, including neurotransmitter receptors and voltage-gated ion channels (Hung and Sheng, 2001). Although the PDZ domain binding site of Kek5 is required for localization, in the absence of the remainder of the cytoplasmic domain, it does not restore wild type Kek5 misexpression activity, supporting the presence of additional functional elements within the intracellular domain.

MATERIALS AND METHODS

Generation of DNA Constructs

The intracellular (IC) domain of Kek5 contains six conserved motifs, including the PDZ domain binding site. Flies misexpressing one of two mutant forms of the Kek5 IC domain were previously generated by Tim Evans: pUASTaKek5$^{\Delta IC}$ (Kek5$^{\Delta IC}$), which entirely lacks the IC domain, and Kek5$^{\Delta IC+P}$, which lacks the IC domain except for a partial motif 6, the PDZ domain binding site. I created three novel Kek5 variants to further assess the contribution of the Kek5 intracellular domain: Kek5$^{\Delta 123}$, Kek5$^{\Delta 45}$, and Kek5$^{\Delta IC+PC}$. Kek5$^{\Delta 123}$ lacks IC motifs 1, 2 and 3, Kek5$^{\Delta 45}$ lacks motifs 4 and 5, and Kek $^{\Delta IC+PC}$ is similar to the previous Kek5$^{\Delta IC+P}$ construct, however, the PDZ domain binding site has been extended on the N-terminal end to contain an additional six amino acids. Sequences for the three new constructs may be found in Appendix A.

All constructs were generated from *pUASTaKek5*, a full-length *kek5* construct in a Gateway cloning-adapted *pUAST* vector. All PCR reactions were run in a 50 μL volume with ~1-10 ng of DNA template and ~20 ng of each primer for 35 cycles with a 60°C annealing temperature and a 3:00 minute extension time. Unless specified otherwise, 25 μL of Qiagen's HotStar Taq Master Mix was used for each PCR reaction. The primers used, along with their positions on the construct can be found in Appendix B.

Kek5$^{\Delta 123}$ (1892 bp) was created using stitching PCR. Two separate fragments (A and B) were created using primers 415 and W19, and W20 and W22, respectively. These constructs were run on a 0.8% agarose gel and band-purified using Qiagen's Qiaquick Gel Extraction Kit. The two fragments then underwent a stitching PCR with primers 415 and W22, using 1 μL each of fragments A and B.

Kek5$^{\Delta 45}$ (2238 bp) was created using an extension method. The first fragment was created with primers 415 and W21. This construct was PCR purified using Qiagen's Qiaquick PCR Purification Kit, and was then extended with primer W22 to include the attB2 cloning site.

Kek5$^{\Delta IC+PC}$ (1314 bp) was generated in a similar method as Kek5$^{\Delta 45}$, using primers 415 and W23 to create the first fragment. The fragment was PCR purified with the Qiaquick PCR Purification Kit and was again extended using primer W22. In this case, however, due to problems with Taq fidelity, a reaction with Vent polymerase was used as follows: 0.5 μL Vent, 3 μL 2.5mM dNTPs, and 5 μL 10x buffer per reaction.

After checking all three final fragments for appropriate length on a 0.8% agarose gel, the fragments were band purified with the Qiaquick Gel Extraction Kit. The fragments were then cloned into a *pDONR* vector via the BP Gateway cloning reaction: 7 μL of DNA insert, 1 μL of 150 ng/μL *pDONR*, and 2 μL of BP clonase mix were incubated at 25°C overnight. 5 μL of each construct were then transformed into DH5α cells and plated on LB agar+kanamycin plates. Colonies were grown up in a small liquid culture (LB+kan) and miniprepped using Qiagen's Qiaprep Spin Miniprep Kit. Putative positive clones were analyzed by restriction digestion and positives were then cloned into a *pUASTaGFP* plasmid using the LR Gateway cloning reaction. Similar to the BP reaction, 7 μL of entry DNA, 150 ng of vector (*pUASTaGFP*), and 2 μL of LR clonase mix were incubated at 25°C overnight. These were again transformed in to DH5α cells and plated on agar+ampicillin plates. After miniprepping and verifying positive clones via restriction digestion, a large 400 mL LB+amp liquid culture was grown to maxiprep with Qiagen's Plasmid Maxi Kit. Positive clones of each construct were then sequenced at Yale's DNA Analysis Facility and verified by the lab by constructing contigs of the sequence reads in Sequencher.

Note: Graduate student Harita Haridas assisted in the cloning of *kek5*$^{\Delta 123}$ and *kek5*$^{\Delta 45}$ fragments, specifically with the LR reactions and transformations.

Generation & Mapping of Transgenics

The maxiprepped DNA for all three constructs was sent to Genetic Services, Inc. (Cambridge, MA) to be injected into w^{1118} embryos. The resulting larvae were then returned to us for transgenic screening. Surviving flies (generation G_0) were single-pair mated to either 3-4 w^{1118} females or 2-3 w^{1118} males, depending on the sex of the putative transgenic. Because the transgene confers pigmentation to the eye, any progeny from this mating containing eye color should be transgenic. Such transgenic flies (generation F_1) were collected and separated. For putative transgenics derived from a single injected G_0, 2-3 single-pair matings were set up (preferably with males), again with w^{1118} females or males to amplify the number of transgenic flies on hand (generation F_2). Putative transgenics derived from distinct G_0s were considered independent insertions.

Matings of transgenic males (*P[UAS-Kek5***]*/+; *Kek5 variant) with females of the genotypes w^{-}*; Sp/CyO* (chromosome II) and w^{-}*; Ly/TM3, Sb* (chromosome III) were set up for mapping purposes. Heterozygous progeny from these matings (preferably of the genotypes *P[UAS-Kek5***]/Cyo* or *P[UAS-Kek5***]/TM3, Sb*) were then outcrossed back to w^{1118} to determine the segregation pattern of the transgene versus the marker; e.g., if the *CyO* phenotype consistently segregates away from the transgene, the transgene must be on chromosome II.

Once mapped, stocks of representative transgenic lines were created by first mating a single *P[UAS-Kek5***]*/+ transgenic male to females of its respective balancing stock (w^{-}*; Sp/Cyo* or w^{-}*; Ly/TM3, Sb*) to ensure that the stock's founding members were isogenic. Female and male

progeny of the genotype *P[UAS-Kek5*]/Cyo* or *P[UAS-Kek5*]/TM3, Sb* from this cross were mated to each other to create the final stock. For insertions on the X chromosome (I), females harboring the transgene were mated by *FM6, y w B* males. Female progeny of the genotype *P[UAS-Kek5*]/FM6, y w B* were again mated to *FM6, y w B* males to create the final stock. Note: Identification, mapping, and establishment of transgenic lines were completed jointly with graduate student Harita Haridas.

Gain of Function Studies with the GAL4 System

The main tool used in the misexpression (gain of function) studies is the GAL4 system. This is a system in which a gene of interest can be expressed in a desired pattern using the mated combination of a "driver" (GAL4) and a "responder" (UAS) construct, as shown in Figure 4:

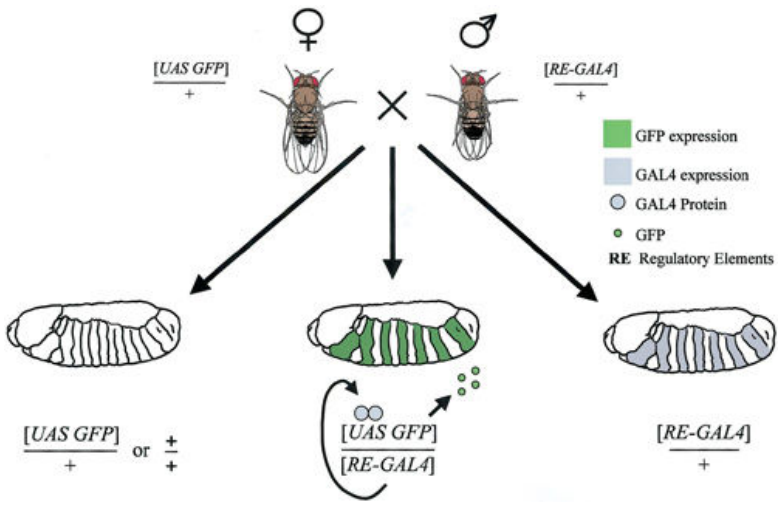


Figure 4: Summary of the GAL4 misexpression system (taken from Duffy, 2002)

The GAL4 gene is linked to a regulatory element, for example, one which expresses in a portion of the wing. As GAL4 is produced in this pattern, it will bind to the upstream activating sequence (UAS), which, if present, should initiate the transcription of the linked gene of interest

(Brand and Phelps, 1998). This is an ideal tool to use in misexpression studies, as it can regulate the expression of a gene both spatially and temporally.

In most cases, GAL4 drivers *A9* and *Ptc*, which express in the dorsal region of the wing and along the anterior-posterior boundary, respectively, were used for these studies. In all cases, 4-5 females with a GAL4 driver were mated to 3-4 males with a UAS-linked gene of interest.

Adult Wing Dissection and Mounting

Adult wings were dissected in 100% ethanol and mounted on a slide in 20% Canada Balsam, 80% methyl salicylate. The wings were then examined at 5X on a Zeiss Axiophot, and images were taken with a Pixelfly camera.

GFP Localization & Expression

Third instar larvae from the *A9.GAL4* x *P[UAS-Kek5*]* variant matings were examined under a fluorescent dissection microscope to determine each construct's level of GFP expression, and were rated on a four-point scale.

Wing imaginal discs were dissected from third instar larvae in 1X PBS, fixed in 3.7% formaldehyde in PBS for 15 minutes, rinsed 5-6 times in PBS, and then mounted on a slide in 70% glycerol in PBS. The discs were then examined at 20X on a Zeiss Imager.Z1 with ApoTome for GFP localization and expression levels. Images were captured using a Zeiss Axiocam and were processed with Zeiss's Axiovision software.

RESULTS

Previous data from the lab demonstrated that deletion of the intracellular domain, excluding the C-terminal PDZ domain binding site, reduces Kek5 activity (Evans, 2006). This, in concert with the blocks of sequence conservation in the intracellular domain among Kek5 orthologs, supports the hypothesis that additional functional elements exist within the intracellular domain of Kek5. Therefore, to better understand the function of Kek5 in adhesion and signal transduction during wing development, a more detailed characterization of the role of the IC domain was undertaken.

Multiple Sequence Elements are Critical for Kek5 Function

Previously with Kek5 misexpression, a wing curling phenotype was noted (Evans, 2006). Using a different GAL4 driver (*A9*) for Kek5 misexpression, I noted a strong wing curling phenotype. To better understand the mechanisms and sequence requirements behind this phenotype, I used the GAL4 system and expressed all of the Kek5 variants with *A9*: Kek5^{wt}, a variant lacking the LRRs (Kek5$^{\Delta LRR}$), a variant lacking the Ig domain (Kek5$^{\Delta Ig}$), a variant lacking the entire IC domain (Kek5$^{\Delta IC}$), and a variant lacking the IC domain except for a six amino acid PDZ domain binding site (Kek5$^{\Delta IC+P}$), as well as double insertions of each construct. These crosses were kept at 28°C to maximize expression and phenotypic effects. Kek5^{wt} exhibited significant upward curling of the wings, along with occasional wing blistering (noted as Moderate, M; Figure 5, B and C). The double insertion of Kek5^{wt} showed even stronger curling, to a similar extent as the *Cy* mutation in the *CyO* balancer line (Strong, S). The remainder of the variants affected wing morphology with varying degrees, ranging from No Effect, NE, to Moderate (Table 1).

A9.GAL4 x	Wing
[UAS.Kek5.GFP]^{16II}	M
[UAS.Kek5.GFP]^{16,52II}	S
[UAS.Kek5$^{\Delta L}$.GFP]$^{43\text{-}1}$	NE
[UAS.Kek5$^{\Delta L}$.GFP]$^{43\text{-}1,16\ R1}$	NE
[UAS.Kek5$^{\Delta I}$.GFP]^{27}	W-M
[UAS.Kek5$^{\Delta I}$.GFP]$^{27,13\text{-}2\ R1}$	M
[UAS.Kek5$^{\Delta IC}$.GFP]^{31A}	NE
[UAS.Kek5$^{\Delta IC}$.GFP]^{31A,55B}	NE
[UAS.Kek5$^{\Delta IC+P}$.GFP]^{33}	NE-W
[UAS.Kek5$^{\Delta IC+P}$.GFP]^{8A,24C}	NE-W

Table 1: Effects of *A9*-driven misexpression of previously generated Kek5 variants on wing morphology

A9-driven misexpression of a variant lacking the LRR motifs (Kek5$^{\Delta LRR}$) resulted in wild type wings (NE; Figure 5, D). In contrast, variant Kek5$^{\Delta Ig}$ retained activity; the curling phenotype was slightly reduced, but not entirely eliminated (W-M; Figure 5, E). This implies that the LRR region is essential to Kek5 function in this gain of function assay, and that the Ig domain is less critical for this activity.

A9-driven misexpression of a variant lacking the IC domain (Kek5$^{\Delta IC}$) produced wild type wings, consistent with an absence of activity (NE; Figure 5, F). This is likely due to low expression levels and a lack of membrane localization for the Kek5 $^{\Delta IC}$ variant, as previously noted (Evans, 2006). Additionally, it was previously shown that adding back the PDZ domain binding site restores expression and localization, but does not confer activity. Consistent with these results, *A9*-driven misexpression of Kek5$^{\Delta IC+P}$ also demonstrated minimal activity, yielding nearly 100% wild type wings (NE-W; Figure 5, G). Thus, these results indicate that additional elements within the IC domain (excluding the PDZ domain binding site) must be critical for Kek5 function.

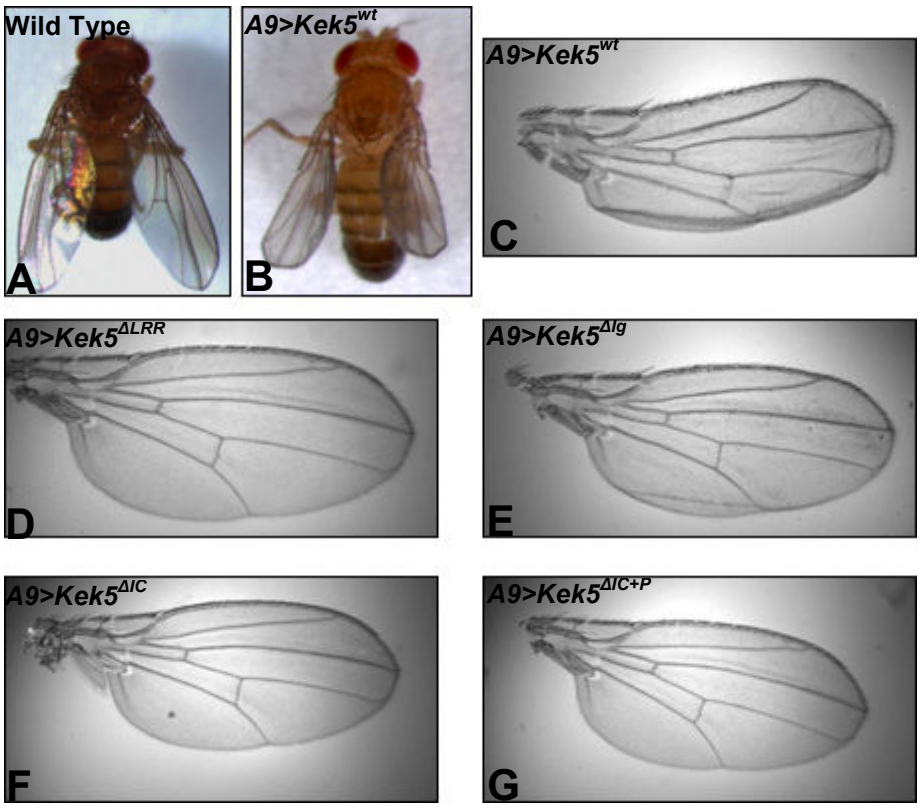


Figure 5: Photo micrographs of *A9*-driven misexpression of Kek5 variants: whole flies and mounted wings
(A) Wild Type, OreR. (B, C) Kek5^{wt}. (D) Kek5$^{\Delta LRR}$. (E) Kek5$^{\Delta Ig}$. (F) Kek5$^{\Delta IC}$. (G) Kek5$^{\Delta IC+P}$.

Distinct Activities Exist within the Intracellular Domain

To further dissect the sequence requirements of the IC domain in Kek5 function, we generated three additional variants. The IC domain was divided approximately in half, yielding two intracellular deletions, an N-terminal deletion and a C-terminal deletion: Kek5$^{\Delta 123}$ and Kek5$^{\Delta 45}$, respectively. In addition, because Kek5$^{\Delta IC+P}$ lacks another six relatively conserved amino acids N-terminal to the PDZ domain binding site, variant Kek5$^{\Delta IC+PC}$ was generated to ensure that an incomplete PDZ domain binding site was not the reason the lack of activity of variant Kek5$^{\Delta IC+P}$ (Figure 6).

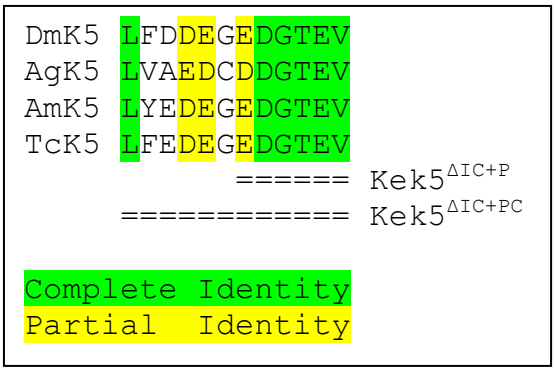


Figure 6: Conservation of the PDZ domain binding site

The schematics of these new variants are found in Figure 7, including any polymorphisms:

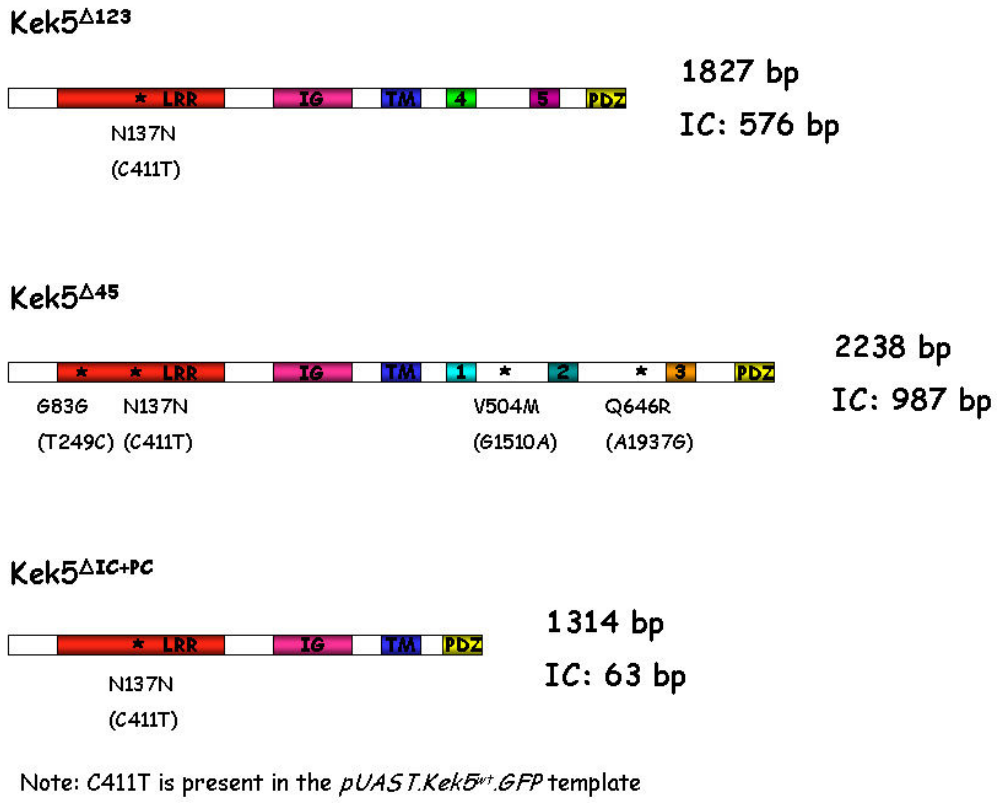


Figure 7: New Kek5 variants: Kek5$^{\Delta 123}$, Kek5$^{\Delta 45}$, and Kek5$^{\Delta IC+PC}$

The constructs were generated via stitching PCR and subsequent cloning into the *pUASTaGFP* vector (Materials and Methods). They were then sent to Genetic Services, Inc. for injection. Injected G_0 larvae were then returned to the lab and adult G_0 flies were single-pair mated with w^{1118} females or males. F_1 progeny with eye color were then selected and again

single-pair mated with w^{1118}, followed by the establishment of stable lines (mapping and stock creation, as detailed in Materials and Methods) At least 33 independent lines were generated for Kek5$^{\Delta 123}$, 29 lines for Kek5$^{\Delta 45}$, and 10 lines for Kek5$^{\Delta IC+PC}$.

The activity of each variant was then tested by crossing with *A9.GAL4* and *Ptc.GAL4* and observing its effect on wing development (Table 2). The two drivers produce distinct misexpression phenotypes, allowing us to better understand the activities of the different variants. While *A9.GAL4*-driven misexpression of Kek5^{wt} tends to produce an upward curling phenotype with relatively little effect on viability, *Ptc.GAL4*-driven misexpression tends to produce anterior crossvein (ACV) defects and more commonly results in reduced viability.

A. CE1 (Kek5$^{\Delta 123}$)		**A9.GAL4**		**Ptc.GAL4**
	Chrom.	**Wing**	**Pen.**	**ACV Pen.**
2F-1M	II	M-S	H	100.0%
3F-1M	II	NE-W	H	60.0%
4F-1M	II	S	H	100.0%
4F-5M	II	S	H	LETHAL
7F-2M	II	M	H	100.0%
10F-2M	II	W	N/A	73.0%
11F-1Mb	X	S	H	LETHAL
12F-1M	III	VS	H	LETHAL
16F-3M	III	VS	H	LETHAL
18F-1Fa	II	W	N/A	74.0%
23M-1Ma	II	VS	H	LETHAL
23M-1Mb	III	LETHAL		LETHAL
24M-1M	II	M	H	93.0%
32M-1M	III	S	H	LETHAL
34M-1Fa	III	S	H	LETHAL

B. CE2 ($Kek5^{\Delta 123}$)

	Chrom.	A9.GAL4 Wing	A9.GAL4 Pen.	Ptc.GAL4 ACV Pen.
4F-1M	X	LETHAL		LETHAL
7F-2M	II	W	H	LETHAL
8F-4Ma	X	S	H	LETHAL
8F-4Mb	II	VS	H	LETHAL
13F-1F	X	S	H	LETHAL
23M-1M	III	M	H	LETHAL
24M-2M	III	W	N/A	ND
24M-3Mc	II	ND	ND	ND
25M-1M	III	S	H	LETHAL
26M-2M	II	VS	H	LETHAL
29M-1Ma	II	ND	ND	ND
29M-1Mb	III	M	H	LETHAL
30M-3Ma	III	S-VS	H	ND
30M-3Mb	III	VS	H	LETHAL
34M-1M	III	S-VS	H	LETHAL
35M-1F	X	S	H	LETHAL
36M-1M	II	S	H	LETHAL
40M-1M	II	VS	H	LETHAL
43F-1F	X	NE	N/A	LETHAL
44F-2M	X	S	H	LETHAL

C. CE3 ($Kek5^{\Delta 45}$)

	Chrom.	A9.GAL4 Wing	A9.GAL4 Pen.	Ptc.GAL4 ACV Pen.
5M-1F	II	W	H	20.8%
6M-1F	III	NE-W	N/A	1.8%
11M-1M	II	W	H	15.3%
14M-1Ma	III	NE-W	N/A	7.5%
16M-1F	III	W-M	H	4.4%
20M-1F	III	M	H	23.5%
25F-1M	III	NE	N/A	WT?
25F-2F	X	ND	ND	ND
30F-2F	II	M	H	15.9%
38F-1M	II	NE-W	N/A	0.0%
40F-1M	III	S	H	100.0%
41F-1F	III	NE-W	N/A	1.7%
42M-1M	II	ND	ND	ND
44M-1M	II	W-M	H	17.3%
49M-1F	III	ND	ND	25.0%
49M-1Fa	II	NE-W	H	1.2%
49M-1Fb	III	W	N/A	ND
53F-1M	II	M	H	15.0%

D. CE4 (Kek5$^{\Delta 45}$)		A9.GAL4		Ptc.GAL4
	Chrom.	Wing	Pen.	ACV Pen.
2F-1F	III	W	N/A	4.0%
2F-3M	III	S	H	100.0%
6F-1Fa	III	W	H	33.0%
6F-2F	II	NE-W	N/A	16.9%
10F-1F	X	W	H	25.4%
15F-1M	III	M	H	0.0%
15F-2M	III	ND	ND	ND
22M-2F	III	W-M	H	13.0%
23M-1M	II	NE	N/A	9.3%
30M-2M	II	W	H	25.0%
30M-2M	III	NE	N/A	0.0%
40M-1Ma	III	ND	ND	ND
40M-1Mb	II	W-M	N/A	ND

E. CE5 (Kek5$^{\Delta IC+PC}$)		A9.GAL4		Ptc.GAL4
	Chrom.	Wing	Pen.	ACV Pen.
1F-1M	III	VS	H	LETHAL
8F-1M	III	NE	N/A	0.0%
8F-2F	III	W	H	0.0%
13M-1M	III	NE-W	N/A	26.7%
15M-1M	II	W	H	0.0%
19M-1M	II	W	H	2.6%

F. CE6 (Kek5$^{\Delta IC+PC}$)		A9.GAL4		Ptc.GAL4
	Chrom.	Wing	Pen.	ACV Pen.
1F-1F	II	W	H	1.3%
6F-1F	III	W	H	3.3%
8F-1F	X	W	H	5.6%
8F-2Fa	III	W	H	0.0%

Table 2: Mapping and functional analysis of new Kek5 variants. (A, B) Kek5$^{\Delta 123}$. (C, D) Kek5$^{\Delta 45}$. (E, F) Kek5$^{\Delta IC+PC}$. (experiments not done noted as ND)

Misexpression of Kek5$^{\Delta 123}$ (lines CE1 and CE2) appeared to have a distinct activity relative to Kek5^{wt}. In many cases, when crossed with *A9.GAL4*, this variant caused a novel phenotype – downward wing curling – and/or severe blistering, as well as reduced viability (Figure 8, C). Most lines were lethal when crossed to *Ptc.GAL4*, and those that were viable showed ACV defects with high penetrance.

Misexpression of Kek5$^{\Delta 45}$ (lines CE3 and CE4) showed slightly weaker activity than Kek^{wt}. With *A9.GAL4*, the most common phenotype was mild upward curling, along with occasional blisters (Figure 8, D). When crossed with *Ptc.GAL4*, this construct was also weaker than Kek5$^{\Delta 123}$, exhibiting better viability and lower ACV defect penetrance.

Misexpression of Kek5$^{\Delta IC+PC}$ (lines CE5 and CE6) yielded similar results as variant Kek5$^{\Delta IC+P}$, with mild or wild type wing phenotypes when driven by *A9.GAL4* (slightly stunted wings, Figure 8, E) and low ACV defect penetrance with *Ptc.GAL4*.

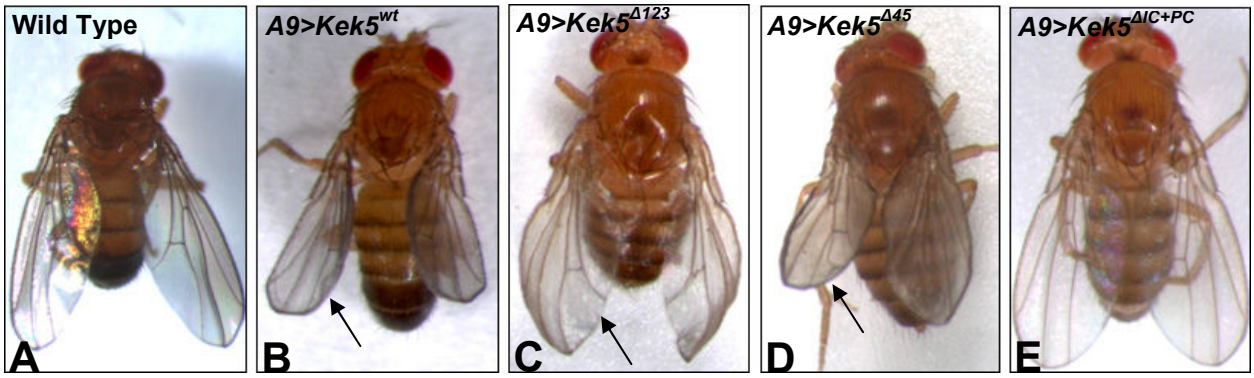


Figure 8: Photo micrographs of *A9*-driven misexpression of new Kek5 variants: whole flies. (A) Wild type, OreR. (B) Kek5^{wt}. (C) Kek5$^{\Delta 123}$. (D) Kek5$^{\Delta 45}$. (E) Kek5$^{\Delta IC+PC}$.

Because of the distinct phenotypes produced by misexpression of Kek5$^{\Delta 123}$ and Kek5$^{\Delta 45}$, larval wing imaginal discs were dissected and examined for possible differences in expression and localization (Figure 9). However, this does not appear to account for the different phenotypes, as both variants are expressed to similar degrees and are membrane localized.

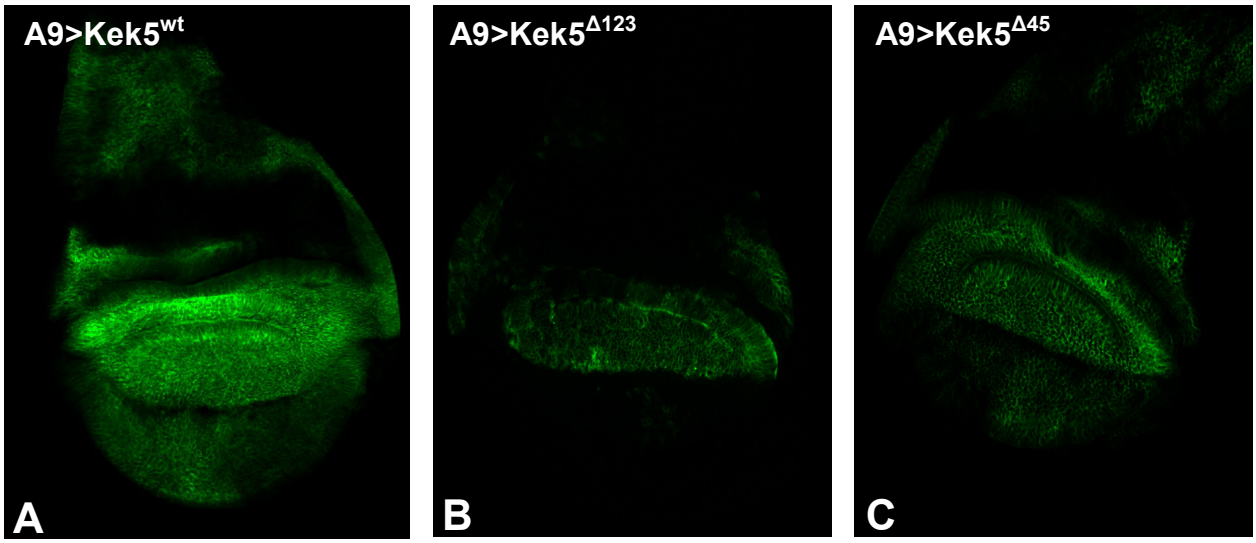


Figure 9: Photo micrographs of *A9*-driven misexpression in larval wing imaginal discs with GFP fluorescence. (A) Kek5^{wt}. (B) Kek5$^{\Delta 123}$. (C) Kek5$^{\Delta 45}$.

Developmental Time-Sensitive Period of Kek5

Given that the GAL4 misexpression system is temperature sensitive, the next goal was to determine when Kek5 misexpression is essential to generate the wing phenotypes observed. Similar phenocritical assays were previously performed with the *Ap.GAL4* driver (expresses in the dorsal region of the wing) and the Kek5^{wt} variant, which produces small, severely blistered wings. It was determined that the critical misexpression period for Kek5^{wt} is approximately four days before eclosion, which coincides with the first round of apposition and adhesion during wing development (Evans, 2006). To determine if the curling phenotype showed a similar temporal profile, this experiment was conducted with the *A9.GAL4* driver, with similar results (Figure 10). In this assay, a scale of 0-4 was used to represent the level of curliness of the wings. The apparent critical expression period with the *A9* driver occurs around three days before eclosion, approximately coinciding with the second series of apposition and adhesion. It was also noted that both periods of critical Kek5^{wt} misexpression correlate with integrin expression and activity in wing development.

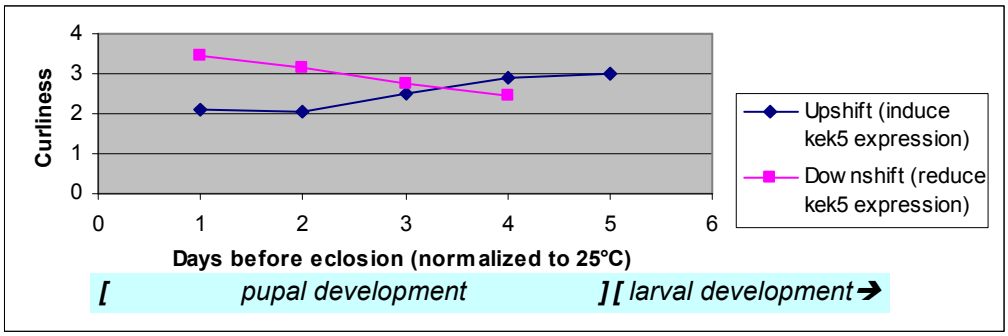


Figure 10: *A9*-driven Kek5^{wt} phenocritical assay

Based on the observation that Kek5^{wt} misexpression may coincide with integrin activity, a recombinant line containing both the *A9.GAL4* driver and the Kek5^{wt} variant was created to screen for potential interactions with Kek5. The *A9.GAL4>P[UAS-Kek5*]* recombinant line was crossed with several hypomorphic alleles of integrins and focal adhesion complex components to test for any dose-dependent interactions with Kek5 misexpression. However, this preliminary screen of approximately 10 integrin-related genes did not result in any significant interactions.

DISCUSSION

LIGs are a relatively new class of proteins containing the unique combination of leucine-rich repeats and Ig-like domains. This small family has not yet been well characterized, but several members have been implicated in neuronal activities, including axon regeneration and synapse pathfinding, as well as in other developmental processes, such as Kek1 in EGFR signaling (Chen et al., 2006; MacLaren et al., 2004). Of note is the fact that most LIG proteins do not contain an intracellular catalytic domain, and thus are likely modulators of receptors and signaling pathways. This study begins to dissect the IC domain of Kek5, a member of the *Drosophila* Kekkon family, and has revealed the importance of multiple distinct elements in the IC domain, as well as confirmed the necessity of specific extracellular elements.

Gain of function studies, including those previously done by Tim Evans, have established that while the intracellular PDZ domain binding site restores membrane localization of Kek5, it does not confer activity, pointing to the necessity of additional elements within the IC domain (2006). This study also confirmed the necessity of the LRR motifs, and not the Ig domain, for Kek5 function. This combination of sequence requirements is consistent with other models, including that of endoglin, a transmembrane modulator of the transforming growth factor-β signaling (TGF-β) system, as well as BMP proteins. Loss of function studies have shown that deletion of either the extracellular or intracellular domains eliminate the interaction between endoglin and TGF-β receptors I and II (Guerrero-Esteo et al., 2002). Endoglin also contains a PDZ domain binding site, shown to be critical for proper function, providing further precedence for this model of receptor-modulator interactions (Koleva et al., 2006).

Further intracellular deletion studies with Kek5 suggest that there are multiple sequence requirements in the IC domain for full restoration of Kek5 activity. Deleting motifs 1, 2, and 3

from the wild type misexpression variant (Kek5$^{\Delta 123}$) produces a novel, even stronger phenotypic effect on wing development than misexpression of Kek5^{wt}, leading to hypothesis that an element(s) within that region may confer autoinhibitory activity. Variant Kek5$^{\Delta 45}$ showed similar but reduced activity, as compared to Kek5^{wt} misexpression, suggesting that an element(s) within that region confers one or more additional distinct activities. Misexpression of Kek5$^{\Delta IC+PC}$ proved to be a useful control, confirming that the addition of an extended PDZ domain binding site does not restore Kek5^{wt} misexpression activity (Figure 11).

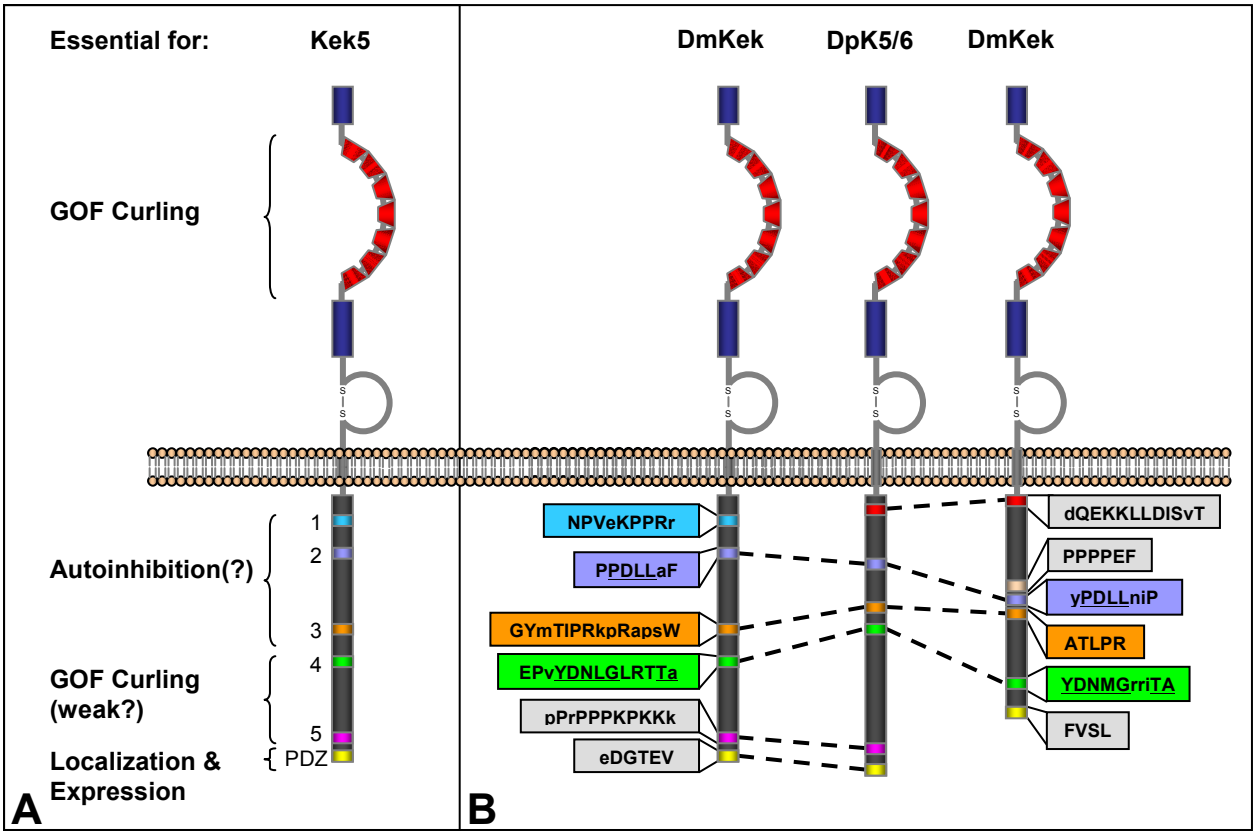


Figure 11: Kek5 models. (A) Kek5 and its essential structural elements. (B) Kek5 IC motif comparison with Kek6 and *Daphnia* Kek5/6.

Based on conservation data with Kek1 and Kek6, as well as *Daphnia pulex* Kek5/6, further variants may prove informative (Figure 11, B). It seems plausible that motifs 2, 3, and 4 may be a single functional unit, as they are consistently found clustered in Kek5, Kek6, and

Daphnia Kek5/6. In contrast, motifs 1 and 5 appear to be functionally independent, as motif 1 is found in Kek1 and Kek5, but not Kek6, and motif 5 is found only in Kek5 and *Daphnia* Kek5/6 (Evans, 2006). Misexpression of various permutations of these motifs should provide further insight into the sequence requirements of the intracellular domain of Kek5.

In addition to the gain of function studies, the temperature sensitive period for Kek5 activity in wing development was also profiled. It was revealed that Kek5 activity may coincide with one or both periods of integrin activity, suggesting that Kek5 may interact with integrins. Ultimately, insight into Kek5 function will require the identification of potential molecular partners through genetic screens for modifiers of Kek5 misexpression phenotypes and biochemical assays, such as co-immunoprecipitations, with Kek5 and its variants. These studies of Kek5 and other Kek family members not only prove informative, but may one day be applied to studies of vertebrate LIG proteins. With the characterization of this small, unique family of proteins, promise for the understanding and treatment of neurodegenerative and other developmental disorders may follow.

REFERENCES

Affolter, M. and Basler, K. (2007). The Decapentaplegic morphogen gradient: from pattern formation to growth regulation. *Nat. Rev. Genet. 8*, 663-674.

Bökel, C. and Brown, N. (2002). Integrins in Development: Moving on, Responding to, and Sticking to the Extracellular Matrix. *Dev. Cell 3*, 311-321.

Bork, P., Holm L. and Sander, C. (1994). The Immunoglobulin Fold: Structural Classification, Sequence Patterns and Common Core. *J. Mol. Biol. 242*, 309-320.

Brand, A. and Phelps, C. (1998). Ectopic Gene Expression in *Drosophila* Using GAL4 System. *Methods 14*, 367-379.

Brower, D., Bunch, T., Mukai, L., Adamson, T., Wherli, M., Lam, S., Friedlander, E., Roote C. and Zusman, S. (1995). Nonequivalent requirements for PS1 and PS2 integrin at cell attachments in *Drosophila*: genetic analysis of the αPS1 integrin subunit. *Development 121*, 1311-1320.

Chen, Y., Aulia, S., Li, L. and Tang, B.L. (2006). AMIGO and friends: An emerging family of brain-enriched, neuronal growth modulating, type I transmembrane proteins with leucine-rich repeats (LRR) and cell adhesion molecule motifs. *Brain Res. Rev. 51*, 265-274.

Dolan, J., Walshe, K., Alsbury, S., Hokamp, K., O'Keeffe, S., Okafuji, T., Miller, S., Tear, G. and Mitchell, K. (2007). The extracellular Leucine-Rich Repeat superfamily; a comparative survey and analysis of evolutionary relationships and expression patterns. *BMC Genomics 8*, 320.

Duffy, J. (2002). GAL4 System in *Drosophila*: A Fly Geneticist's Swiss Army Knife. *genesis 34*, 1-15.

Evans, T. (2006). Characterization of Kekkon5, a *Drosophila* LIG Protein that Modulates BMP and Integrin Function. PhD thesis. Indiana University.

Fristrom, D. and Fristrom, J. (1993). The metamorphic development of the adult epidermis. In The Development of *Drosophila melanogaster* (ed. M. Bate and A.M. Arias), pp. 843-897. New York: Cold Spring Harbor Laboratory Press.

Fristrom, D., Wilcox, M. and Fristrom, J. (1993). The distribution of PS integrins, laminin A and F-actin during key stages in *Drosophila* wing development. *Development 117*, 509-523.

Hung, A. and Sheng, M. (2002). PDZ Domains: Structural Modules for Protein Complex Assembly. *J. Biol. Chem. 277*, 5699-5702.

Guerrero-Esteo, M., Sánchez-Elsner, T., Letamendia, A. and Bernabéu, C. (2002). Extracellular and Cytoplasmic Domains of Endoglin Interact with Transforming Growth Factor-β Receptors I and II. *J. Biol. Chem. 277*, 29197-29209.

Ko, J. and Kim, E. (2007). Leucine-Rich Repeat Proteins of Synapses. *J. Neurosci. Res. 85*, 2824-2832.

Koleva, R., Conley, B., Romero, D., Riley, K., Marto, J., Lux, A. and Vary, C. (2006). Endoglin Structure and Function. *J. Biol. Chem. 281*, 25110-25123.

Little, S. and Mullins, M. (2006). Extracellular Modulation of BMP Activity in Patterning the Dorsoventral Axis. *Birth Defects Res., Part C 78*, 224-242.

MacLaren, C., Evans, T., Alvarado, D. and Duffy, J. (2004). Comparative analysis of the Kekkon molecules, related members of the LIG superfamily. *Dev. Genes. Evol. 214*, 360-366.

Pålsson-McDermott E. and O'Neill L. (2007) Building an immune system from nine domains. *Biochem. Soc. Trans. 35*, 1437-44.

Takada, Y., Ye, X. and Simon, S. (2007). The integrins. *Genome Biol. 8*, 215.1-215.9

APPENDIX A: KEK5 MISEXPRESSION CONSTRUCT SEQUENCES

Kek5^{wt}

ATGATCCTTCTGCTGCTGGGTGTGCTAGTGGTTCTAATGGCCCTACCGCCGCCCACCGCAGGCA
CCACCGATTGGATGCAGAGCTGCGGTACATGCCACTGTCAGTGGAATTCGGGCAAGAAGAGCGC
CGACTGCAAGAACAAGGCGCTAACCAAAATTCCGCAGGACATGAGCAACGAGATGCAGGTGCTG
GACTTTGCCCACAATCAAATACCCGAGCTGCGGCGCGAAGAGTTCCTACTGGCCGGTCTGCCCA
ATGTGCACAAGATCTTTTTGCGCAACTGCACCATCCAGGAGGTGCATCGCGAGGCCTTCAAGGG
TCTGCATATCCTAATCGAGCTGGACCTGTCGGGCAATCGGATACGGGAACTGCATCCGGGCACT
TTCGCCGGCCTGGAGAAGCTGCGCAACATGATCATCAACAACAACGAGATCGAGGTGCTGCCCA
ACCATCTGTTCGTCAACCTGAGCTTCCTGTCGCGCATCGAGTTCCGGAACAATCGATTGCGCCA
GGTGCAGCTGCACGTCTTCGCTGGCACAATGGCGCTGAGCGCCATTTCGCTGGAACAGAACCGC
CTCTCACATCTGCACAAGGAGACATTCAAGGATCTGCAGAAGCTGATGCATCTATCGCTGCAGG
GTAACGCATGGAACTGCAGCTGCGAGCTGCAGGACTTTCGCGACTTTGCGATCAGCAAACGGCT
CTACACACCGCCCACCGATTGCCAGGAGCCGCCACAGCTGCGCGGCAAGCTGTGGAGCGAGGTG
CCATCGGAGAACTTCGCCTGCCGGCCGCGCATTTTGGGTTCCGTGCGCTCCTTCATCGAGGCCA
ATCACGACAATATCTCGCTACCCTGCCGCATTGTCGGCAGTCCGCGTCCCAATGTCACCTGGGT
GTACAACAAGCGGCCATTGCAGCAGTACGACCCGCGTGTGCGTGTCCTCACCTCCGTGGAACAG
ATGCCGGAGCAGCCCTCCCAGGTGCTCACCTCGGAGCTGCGCATCGTGGGCGTACGGGCCTCCG
ACAAGGGTGCCTACACCTGTGTGGCGGATAACCGGGGCGGACGGGCGGAGGCCGAGTTCCAGCT
GCTCGTGAGCGGTGACTATGCCGGCGCGGTATCCGCCTCCGATGGCATGGGCATGGGCGCCATT
GGGGCACCAACCATTGATCCGCAAACGAACATGTTTCTCATCATCTGTCTAATCATTACGACGC
TGCTGCTCCTGCTGCTCGTGGCGGTGCTGACGCTCTTCTGGTACTGCCGTCGCATCAAGACCTA
TCAAAAGGACACCACCATGATGAGCGGCGACGGGCTGATCTCTTCCAAGATGGACAAGACGCAC
AACGGCTCCATGCTCGAGGGTTCCGTCATCATGGAGATGCAGAAGAGCCTGCTCAACGAGGTCA
ATCCAGTCGAGAAGCCGCCACGGCGCACGGACATCGAGAGCGTGGATGGTGGCGATGACGTGCT
CGAGATCAAGAAGACGCTGCTCGACGACACCGTCTATGTGGCCAATCACTCGCGCGACGAAGAA
GCCGTCTCAGTGGCCATGTCGGATACGACGACCACGCCCCGATCTCGACACACCTACGTGGATG
ATGCGTATGCCAATAGCTTGCCACCGGATCTGCTGGCCTTTCCCGCTCGCGTGCCGCCCACCTC
GCCCTCGATGCAATCGTCGCAGTCGAACATACCCGACCAGGTGATCTACGGCATCCGTTCGCCA
CCGTCGCTAACCAGTCCGGTCTACACGCATATGACGCCGCACGGCATCTACGGCACCAAGACGA
TGACGGCTCCGCATAACGGCTTTATGACGCTGCAGCATCCCAAGTCGCGCAACCTGGCGCTCAT
TGCCACCACCAACAGCAGTCGCCAGCACCAGCACCACCATCAGCTGCAGCAGCAGCAGCAGCAC
CACCACCACCACCAGCAGCAACAACAACAGCAGCAGCAGCAGCAACATCCGCTGGCCACCACAT
CGCCCTTCCTGCCCGCACCCGTCGTCTATTCGCCGGCCACGGGTGTGGTCATGAAACAGGGATA
TATGACCATTCCGCGCAAGCCGCGCGCTCCCAGCTGGGCGCCCAGTACTTCCGGTGCCGCTGGC
CACGGATCCATTCAGCTAAGTGAATTCCAGAGCCCCACATCGCCGAATCCCAGCGAGACTGGCA
CCGCCACCACCGCGGAACTGCAGGCGGAGCCAGTGTACGACAACTTGGGATTGCGAACCACTGC
CGGCGGCAACTCCACCCTCAATCTGACCAAGATCGCCGGCTCACAGGGGGGCGCTGGTCAGCAG
TACTCGATGCGGGACCGACCACTTCCGGCCACGCCCAGCCTGACATCGGTGTCCTCGGCGACCA
ATGCCAGTAAGATTTACGAGCCCATACACGAGCTGATTCAGCAGCAACAGCAGTTGCAACAACA
ACAACAGCAGCAGCAGCAGCGACTGGGCTCCATGGACACGGAACCCCTGTACGGAGTTCGGCAA
CAGGGGATCACGATACTGCCCGGCTCGAGCATTAGCGGTGCCGGACTGGGCCACGCCGCCTACC
TTTCACCCGGCTCGGGTGCCGCCGTATCGCCAAGCCACGCCAGCAGCAGCGGTGACTCTCCGAA

GGCCGCCAAGATCCCACCACGCCCACCACCGAAGCCCAAGAAGAAGATGTCCGTGACGACGACG
CGCAGCGGCCAGGGCAGCACCAGCCAGCTCTTCGACGACGAGGGCGAGGATGGCACCGAGGTC

Kek5$^{\Delta 123}$

ATGATCCTTCTGCTGCTGGGTGTGCTAGTGGTTCTAATGGCCCTACCGCCGCCCACCGCAGGCA
CCACCGATTGGATGCAGAGCTGCGGTACATGCCACTGTCAGTGGAATTCGGGCAAGAAGAGCGC
CGACTGCAAGAACAAGGCGCTAACCAAAATTCCGCAGGACATGAGCAACGAGATGCAGGTGCTG
GACTTTGCCCACAATCAAATACCCGAGCTGCGGCGCGAAGAGTTCCTACTGGCCGGTCTGCCCA
ATGTGCACAAGATCTTTTTGCGCAACTGCACCATCCAGGAGGTGCATCGCGAGGCCTTCAAGGG
TCTGCATATCCTAATCGAGCTGGACCTGTCGGGCAATCGGATACGGGAACTGCATCCGGGCACT
TTCGCCGGCCTGGAGAAGCTGCGCAATGTGATCATCAACAACAACGAGATCGAGGTGCTGCCCA
ACCATCTGTTCGTCAACCTGAGCTTCCTGTCGCGCATCGAGTTCCGGAACAATCGATTGCGCCA
GGTGCAGCTGCACGTCTTCGCTGGCACAATGGCGCTGAGCGCCATTTCGCTGGAACAGAACCGC
CTCTCACATCTGCACAAGGAGACATTCAAGGATCTGCAGAAGCTGATGCATCTATCGCTGCAGG
GTAACGCATGGAACTGCAGCTGCGAGCTGCAGGACTTTCGCGACTTTGCGATCAGCAAACGGCT
CTACACACCGCCCACCGATTGCCAGGAGCCGCCACAGCTGCGCGGCAAGCTGTGGAGCGAGGTG
CCATCGGAGAACTTCGCCTGCCGGCCGCGCATTTTGGGTTCCGTGCGCTCCTTCATCGAGGCCA
ATCACGACAATATCTCGCTACCCTGCCGCATTGTCGGCAGTCCGCGTCCCAATGTCACCTGGGT
GTACAACAAGCGGCCATTGCAGCAGTACGACCCGCGTGTGCGTGTCCTCACCTCCGTGGAACAG
ATGCCGGAGCAGCCCTCCCAGGTGCTCACCTCGGAGCTGCGCATCGTGGGCGTACGGGCCTCCG
ACAAGGGTGCCTACACCTGTGTGGCGGATAACCGGGGCGGACGGGCGGAGGCCGAGTTCCAGCT
GCTCGTGAGCGGTGACTATGCCGGCGCGGTATCCGCCTCCGATGGCATGGGCATGGGCGCCATT
GGGGCACCAACCATTGATCCGCAAACGAACATGTTTCTCATCATCTGTCTAATCATTACGACGC
TGCTGCTCCTGCTGCTCGTGGCGGTGCTGACGCTCTTCTGGTACTGCCGACGCATCAAGACCGA
GCCAGTGTACGACAACTTGGGATTGCGAACCACTGCCGGCGGCAACTCCACCCTCAATCTGACC
AAGATCGCCGGCTCACAGGGGGGCGCTGGTCAGCAGTACTCGATGCGGGACCGACCACTTCCGG
CCACGCCCAGCCTGACATCGGTGTCCTCGGCGACCAATGCCAGTAAGATTTACGAGCCCATACA
CGAGCTGATTCAGCAGCAACAGCAGTTGCAACAACAACAACAGCAGCAGCAGCAGCGACTGGGC
TCCATGGACACGGAACCCCTGTACGGAGTTCGGCAACAGGGGATCACGATACTGCCCGGCTCGA
GCATTAGCGGTGCCGGACTGGGCCACGCCGCCTACCTTTCACCCGGCTCGGGTGCCGCCGTATC
GCCAAGCCACGCCAGCAGCAGCGGTGACTCTCCGAAGGCCGCCAAGATCCCACCACGCCCACCA
CCGAAGCCCAAGAAGAAGATGTCCGTGACGACGACGCGCAGCGGCCAGGGCAGCACCAGCCAGC
TCTTCGACGACGAGGGCGAGGATGGCACCGAGGTC

Kek5$^{\Delta 45}$

ATGATCCTTCTGCTGCTGGGTGTGCTAGTGGTTCTAATGGCCCTACCGCCGCCCACCGCAGGCA
CCACCGATTGGATGCAGAGCTGCGGTACATGCCACTGTCAGTGGAATTCGGGCAAGAAGAGCGC
CGACTGCAAGAACAAGGCGCTAACCAAAATTCCGCAGGACATGAGCAACGAGATGCAGGTGCTG
GACTTTGCCCACAATCAAATACCCGAGCTGCGGCGCGAAGAGTTCCTACTGGCCGGTCTGCCCA
ATGTGCACAAGATCTTTTTGCGCAACTGCACCATCCAGGAGGTGCATCGCGAGGCCTTCAAGGG
TCTGCATATCCTAATCGAGCTGGACCTGTCGGGCAATCGGATACGGGAACTGCATCCGGGCACT
TTCGCCGGCCTGGAGAAGCTGCGCAATGTGATCATCAACAACAACGAGATCGAGGTGCTGCCCA
ACCATCTGTTCGTCAACCTGAGCTTCCTGTCGCGCATCGAGTTCCGGAACAATCGATTGCGCCA
GGTGCAGCTGCACGTCTTCGCTGGCACAATGGCGCTGAGCGCCATTTCGCTGGAACAGAACCGC
CTCTCACATCTGCACAAGGAGACATTCAAGGATCTGCAGAAGCTGATGCATCTATCGCTGCAGG

GTAACGCATGGAACTGCAGCTGCGAGCTGCAGGACTTTCGCGACTTTGCGATCAGCAAACGGCT
CTACACACCGCCCACCGATTGCCAGGAGCCGCCACAGCTGCGCGGCAAGCTGTGGAGCGAGGTG
CCATCGGAGAACTTCGCCTGCCGGCCGCGCATTTTGGGTTCCGTGCGCTCCTTCATCGAGGCCA
ATCACGACAATATCTCGCTACCCTGCCGCATTGTCGGCAGTCCGCGTCCCAATGTCACCTGGGT
GTACAACAAGCGGCCATTGCAGCAGTACGACCCGCGTGTGCGTGTCCTCACCTCCGTGGAACAG
ATGCCGGAGCAGCCCTCCCAGGTGCTCACCTCGGAGCTGCGCATCGTGGGCGTACGGGCCTCCG
ACAAGGGTGCCTACACCTGTGTGGCGGATAACCGGGGCGGACGGGCGGAGGCCGAGTTCCAGCT
GCTCGTGAGCGGTGACTATGCCGGCGCGGTATCCGCCTCCGATGGCATGGGCATGGGCGCCATT
GGGGCACCAACCATTGATCCGCAAACGAACATGTTTCTCATCATCTGTCTAATCATTACGACGC
TGCTGCTCCTGCTGCTCGTGGCGGTGCTGACGCTCTTCTGGTACTGCCGTCGCATCAAGACCTA
TCAAAAGGACACCACCATGATGAGCGGCGACGGGCTGATCTCTTCCAAGATGGACAAGACGCAC
AACGGCTCCATGCTCGAGGGTTCCGTCATCATGGAGATGCAGAAGAGCCTGCTCAACGAGGTCA
ATCCAGTCGAGAAGCCGCCACGGCGCACGGACATCGAGAGCGTGGATGGTGGCGATGACGTGCT
CGAGATCAAGAAGACGCTGCTCGACGACACCGTCTATATGGCCAATCACTCGCGCGACGAAGAA
GCCGTCTCAGTGGCCATGTCGGATACGACGACCACGCCCCGATCTCGACACACCTACGTGGATG
ATGCGTATGCCAATAGCTTGCCACCGGATCTGCTGGCCTTTCCCGCTCGCGTGCCGCCCACCTC
GCCCTCGATGCAATCGTCGCAGTCGAACATACCCGACCAGGTGATCTACGGCATCCGTTCGCCA
CCGTCGCTAACCAGTCCGGTCTACACGCATATGACGCCGCACGGCATCTACGGCACCAAGACGA
TGACGGCTCCGCATAACGGCTTTATGACGCTGCAGCATCCCAAGTCGCGCAACCTGGCGCTCAT
TGCCACCACCAACAGCAGTCGCCAGCACCAGCACCACCATCAGCTGCAGCAGCAGCAGCAGCAC
CACCACCACCACCAGCGGCAACAACAACAGCAGCAGCAGCAGCAACATCCGCTGGCCACCACAT
CGCCCTTCCTGCCCGCACCCGTCGTCTATTCGCCGGCCACGGGTGTGGTCATGAAACAGGGATA
TATGACCATTCCGCGCAAGCCGCGCGCTCCCAGCTGGGCGCCCAGTACTTCCGGTGCCGCTGGC
CACGGATCCATTCAGCTAAGTGAATTCCAGAGCCCCACATCGCCGAATCCCAGCGAGACTGGCA
CAGCCACCACCGCGGAACTGCAGGCACTCTTCGACGACGAGGGCGAGGATGGCACCGAGGTC

Kek5$^{\Delta IC+PC}$

ATGATCCTTCTGCTGCTGGGTGTGCTAGTGGTTCTAATGGCCCTACCGCCGCCCACCGCAGGCA
CCACCGATTGGATGCAGAGCTGCGGTACATGCCACTGTCAGTGGAATTCGGGCAAGAAGAGCGC
CGACTGCAAGAACAAGGCGCTAACCAAAATTCCGCAGGACATGAGCAACGAGATGCAGGTGCTG
GACTTTGCCCACAATCAAATACCCGAGCTGCGGCGCGAAGAGTTCCTACTGGCCGGTCTGCCCA
ATGTGCACAAGATCTTTTTGCGCAACTGCACCATCCAGGAGGTGCATCGCGAGGCCTTCAAGGG
TCTGCATATCCTAATCGAGCTGGACCTGTCGGGCAATCGGATACGGGAACTGCATCCGGGCACT
TTCGCCGGCCTGGAGAAGCTGCGCAATGTGATCATCAACAACAACGAGATCGAGGTGCTGCCCA
ACCATCTGTTCGTCAACCTGAGCTTCCTGTCGCGCATCGAGTTCCGGAACAATCGATTGCGCCA
GGTGCAGCTGCACGTCTTCGCTGGCACAATGGCGCTGAGCGCCATTTCGCTGGAACAGAACCGC
CTCTCACATCTGCACAAGGAGACATTCAAGGATCTGCAGAAGCTGATGCATCTATCGCTGCAGG
GTAACGCATGGAACTGCAGCTGCGAGCTGCAGGACTTTCGCGACTTTGCGATCAGCAAACGGCT
CTACACACCGCCCACCGATTGCCAGGAGCCGCCACAGCTGCGCGGCAAGCTGTGGAGCGAGGTG
CCATCGGAGAACTTCGCCTGCCGGCCGCGCATTTTGGGTTCCGTGCGCTCCTTCATCGAGGCCA
ATCACGACAATATCTCGCTACCCTGCCGCATTGTCGGCAGTCCGCGTCCCAATGTCACCTGGGT
GTACAACAAGCGGCCATTGCAGCAGTACGACCCGCGTGTGCGTGTCCTCACCTCCGTGGAACAG
ATGCCGGAGCAGCCCTCCCAGGTGCTCACCTCGGAGCTGCGCATCGTGGGCGTACGGGCCTCCG
ACAAGGGTGCCTACACCTGTGTGGCGGATAACCGGGGCGGACGGGCGGAGGCCGAGTTCCAGCT
GCTCGTGAGCGGTGACTATGCCGGCGCGGTATCCGCCTCCGATGGCATGGGCATGGGCGCCATT
GGGGCACCAACCATTGATCCGCAAACGAACATGTTTCTCATCATCTGTCTAATCATTACGACGC

TGCTGCTCCTGCTGCTCGTGGCGGTGCTGACGCTCTTCTGGTACTGCCGTCGCATCAAGACACT
CTTCGATGACGAGGGCGAGGATGGCACCGAGGTC

APPENDIX B: PRIMER SEQUENCES AND LOCATIONS

415
ACAAGTTTGTACAAAAAAGCAGGCTCCAGGAAAATGATCCTTCTGCTGCTGGGTGT

W19
CCCAAGTTGTCGTACACTGGCTCGGTCTTGATGCGTCGGCAGTACC

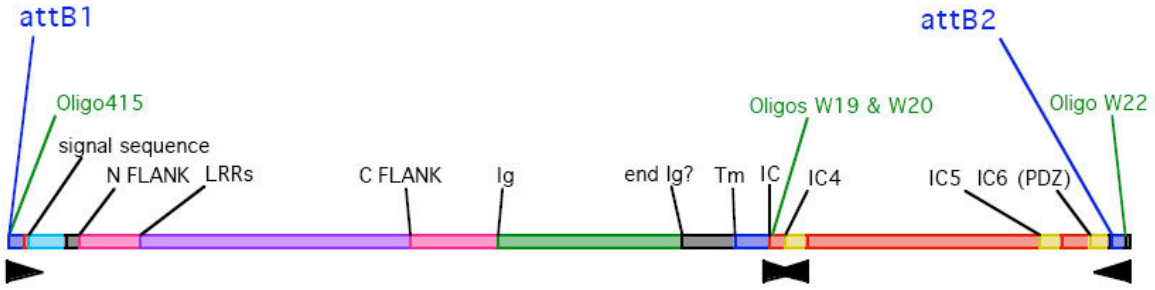
W20
GGTACTGCCGACGCATCAAGACCGAGCCAGTGTACGACAACTTGGG

W21
GCCATCCTCGCCCTCGTCGTCGAAGAGTGCCTGCAGTTCCGCGGTGGTGGCTGTGCC

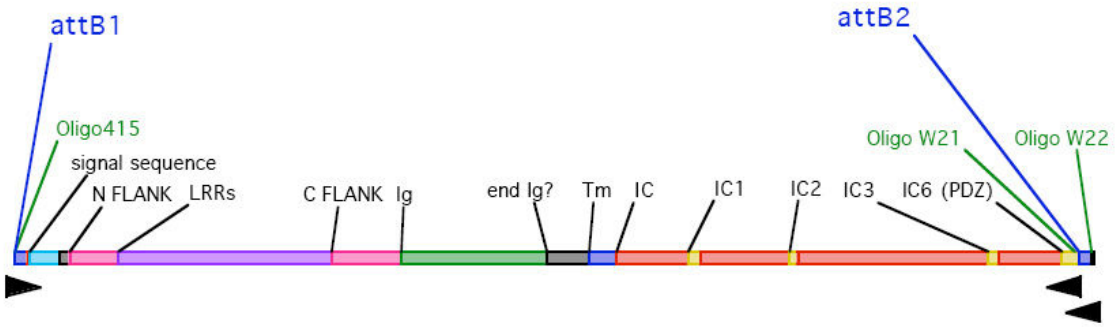
W22
GGGGACCACTTTGTACAAGAAAGCTGGGTCGACCTCGGTGCCATCCTCGCCCTCGTC

W23
GCCATCCTCGCCCTCGTCATCGAAGAGTGTCTTGATGCGACGGCAGTACCAGAAGAG

Kek5$^{\Delta 123}$



Kek5$^{\Delta 45}$



Kek5$^{\Delta IC+PC}$