GFAT1/ZEPPELIN IS AN ESSENTIAL HETEROCHROMATIC GENE INVOLVED IN CUTICLE FORMATION IN *D. MELANOGASTER*.

by

Catherine Jackson
B.Sc, Simon Fraser University, 2005

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
Molecular Biology and Biochemistry Department

SIMON FRASER UNIVERSITY

2007

APPROVAL

Name: **Catherine Jackson**

Degree: **Master of Science**

Title of Thesis: ***GFAT1/zeppelin* is an essential heterochromatic gene involved in cuticle formation in *D. melanogaster.***

Examining Committee:

Chair: Dr N. Forde, Assistant Professor

Dr. B. M. Honda, Senior Supervisor
Professor,
Department of Molecular Biology and Biochemistry

Dr. N. Harden, Supervisor
Professor,
Department of Molecular Biology and Biochemistry

Dr. D. A. R. Sinclair, Supervisor
Lecturer,
Department of Molecular Biology and Biochemistry

Dr. D Baillie, Supervisor
Professor,
Department of Molecular Biology and Biochemistry

Dr. Andrew Beckenbach,
Internal Examiner
Professor, Department of Biological Sciences

Date Defended/Approved: August 2, 2007

ABSTRACT

Many essential genes exist in heterochomatin despite the high composition of repeat or 'junk' DNA in these regions. The gene *zeppelin* (*zep*) was genetically mapped to a distal segment of 3R heterochromatin by the Bejsovec lab. As well as being essential, it is characterised by an abnormal expansion of the embryo cuticle, a 'blimp' phenotype. *Gfat1* is part of the hexosamine pathway that uses glucose to form chitin, a major component of the exoskeleton and therefore a candidate for causing the 'blimp' phenotype. Using genetic and molecular approaches, I have determined that the *zep* gene coincides with *Gfat1* and further alleles have also been identified and characterized. *Inter se* complementation has revealed an interesting semi-viable *Splayed (Spl)* phenotype, characterised by weak legs and melanotic deposits at leg joints. Reproduction of the *Spl* phenotype via RNAi has further confirmed the correspondence of the *zeppelin* locus with *Gfat1*.

Keywords: Gfat1, zeppelin, splayed, cuticle, heterochromatin, drosophila ;

DEDICATION

For my daughter Emma, whose easygoing nature and bright disposition

I love to be around.

ACKNOWLEDGEMENTS

Many thanks to Dr Barry Honda and my committee members Dr Don Sinclair, Dr Dave Baillie and Dr Nick Harden who have all provided some great ideas and questions on the project as well as being supportive and encouraging.

Thanks also to members of the Honda lab who have been a great team of people to work with and especially thanks to people that have contributed to the zeppelin/Gfat1 project - Sarah, Nav, Anshu, Celeste, Kelsey, Kevin, Rob.

Thanks also to Kathleen and Monika who have been a great source of advice on topics genetic and molecular throughout the project.

I would like to acknowledge the generous interest in this work expressed by Dr Amy Bejsovec who took the time to make and photograph cuticle preparation slides of mutant zep allele embryos.

Finally I would like to thank my friends and family for their support and encouragement throughout my graduate and undergraduate work, especially my daughter Emma, sister Julie, and Mom and Dad.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

Flybase	A database of *Drosophila* genes and genomes
Fru6P	Fructose 6-Phosphate
FruN6P	Fructosimine 6-phosphate
GlcN6P	Glucosamine 6-Phosphate
GlmS	Glucosamine-6-phosphate synthase
Haplosufficient	One copy of the gene is sufficient for normal function
UDP-NacGlc	Uridine 5' diphosphate-N-acetylglucosamine
SV	Semi-viable (> 50% of expected class size viable)
SL	Semi-lethal (> 50% of expected class size lethal)

CHAPTER 1: INTRODUCTION

Constitutive heterochromatin

Heterochromatin may be distinguished from euchromatin by its compact state at interphase, as first noted by Heitz in 1928. Characterised by darkly stained chromosome regions, this condensed arrangement appears to be maintained throughout the cell cycle. Constitutive heterochromatin is found in a pericentric location in nearly all higher eukaryotes, forming 30% of the human genome and 30% of the genome of *Drosophila melanogaster* (Corradini, Rossi, Giordano, et al. 2007). It is distinct from facultative heterochromatin which refers to the condensed state of whole chromosomes e.g. the mechanism of X inactivation in mammals (Eberl, Duyf, Hilliker. 1993).

Constitutive heterochromatin has historically been characterised as a region unlikely to contain any actively expressed single copy genes (Yasuhara, Wakimoto. 2006). Indeed because of the repetitive nature of the DNA in these regions and lack of expressed genes mapped to heterochromatin by early geneticists, it has often been referred to as a region of 'junk' DNA. However, in recent years the idea that this is a genomic region devoid of any significant protein coding sequences has been challenged and the few genes that have been mapped to this region viewed as more than an anomaly. In *D. melanogaster* several hundred (230-256) protein-coding genes are predicted

based on genomic sequencing, of which 137 are supported by the existence of full-length cDNA clones (Smith, Shu, Mungall, et al. 2007). Moreover, approximately 32 of these genes have been characterised as essential and mapped to mitotic heterochromatin cytologically. However very few of these have been characterised further in terms of function (Corradini, Rossi, Giordano, et al. 2007).

Heterochromatin structure

The condensed nature of heterochromatin appears to rely on the presence of middle and highly repetitive DNA sequences such as transposable elements and satellite sequences (Grewal, Jia. 2007). At a biochemistry level heterochomatin structure is distinguished from that of euchromatin by the methylation of histone H3 at lysine 9 (H3K9me) and hypoacetylation. Euchromatin is characterised by acetylation of histone H4 and methylation of H3 at lysine 4. H3K9me leads to the recruitment of histone modifying proteins such as HP1. Contrary to the euchromatic environment, these epigenetic mechanisms of compaction have been found to be essential for heterochromatic gene transcription, and are necessary for correct chromosome segregation and DNA repair (Grewal, Jia. 2007).

Position effect variegation (PEV)

Another unusual feature associated with heterochromatin is the phenomenon of position effect variegation (PEV). This was first observed by H.J.

Muller in 1930 and refers to the ability of heterochromatin to silence genes that have been placed within or in close proximity to it by chromosome rearrangements. The PEV silencing effect of heterochromatin has also contributed to the perception of this region as being inert in terms of gene expression (Eberl, Duyf, Hilliker. 1993). PEV operates in a clonal manner leading to a variegated phenotype, e.g. when the euchromatic gene *white* is relocated to the heterochromatin environment, a patched red and white eye phenotype results due to the clonal expression causing red eye color, and silencing causing white.

PEV offers an opportunity to study the mechanism involved in silencing of genes associated with heterochromatin which may allow a better understanding of the complexities of active expression of genes normally residing in this repressive environment. It is an interesting paradox that approximately 32 essential actively functioning genes reside in heterochromatin, while at the same time euchromatic genes moved to this environment suffer from silencing by PEV. Furthermore, it appears that heterochromatic genes actually need to be positioned in close proximity to a large block of heterochromatin in order to maintain normal function (Eberl, Duyf, Hilliker. 1993).

Genetic screens for modifiers of PEV have highlighted two highly conserved genes coding for chromatin modifying proteins (Sinclair, Lloyd, Grigliatti. 1989); (Yasuhara, Wakimoto. 2006); *Suppressor of variegation 3-9 (Su(var)3-9)* codes for a methyltransferase which methylates Lys9 of histone H3

(H3K9me) and *Supressor of variegation 2-5 (Su(var)2-5)* codes for the chromodomain protein heterochromatin protein 1 (HP1).

Model of heterochromatin formation

One suggestion put forward for the mechanism of heterochromatin formation and maintenance is HP1 binds H3K9me chromatin and causes compression through protein-protein interaction (Yasuhara, Wakimoto. 2006; Grewal, Jia. 2007). This appears to be a self-reinforcing process which continues along the chromatin strand. With the continued recruitment of Su(var) methyltransferases and HP1, the repressive effect of condensed chromatin is exerted over neighbouring sequences. It has been proposed that this may be a sequence independent method of coordinated control of loci that would otherwise not be able to recruit these silencing proteins independently (Grewal, Jia. 2007).

Many other factors have also been implicated in the formation of the condensed state of heterochromatin; it appears that the RNA interference (RNAi) machinery may play a key role (Lippman, Martienssen. 2004). Small interfering RNAs (siRNA) generated from non-coding repeat sequence DNA in heterochromatin regions are thought to be involved in the nucleation of heterochromatin formation. Their transcription and binding to local sequences leads to the subsequent recruitment of methyltransferases and HP1 proteins (Grewal, Jia. 2007). This model could be one explanation for the connection between the condensed constitutive heterochromatin state with the repeat sequences found so abundantly in the pericentric region.

Transcription of heterochromatic genes

As well as involvement in the formation and maintenance of condensed chromatin, evidence suggests that HP1 is also critical for the transcription of genes resident in heterochromatin. As reviewed by Grewal and Jia (2007), two essential heterochromatic genes, *light(lt)* and *rolled(rl)* require HP1 in addition to their location in proximity to large blocks of heterochromatin for expression (Lu, Emtage, Duyf, et al. 2000). In HP1 mutants *lt* and *rl* transcription was reduced 2.5 fold. However the mechanism of HP1 in this capacity is unknown. It is possible that proteins such as HP1 are necessary for the recruitment of specialised transcription factors adapted to operate in this environment and/or the nature of the heterochromatin structure itself allows transcription.

A model put forward by Yasuhara et al. (2005) supports the second proposal based on evidence gained from a comparative evolutionary study of the *lt* gene among *Drosophila* species. Comparison of *lt* and its gene cluster in *D. melanogaster* with euchromatic orthologs in more ancient *Drosophila* species revealed that there are no specialized adaptations within the promoter regions that are unique to genes relocated to heterochromatin (Yasuhara, Wakimoto. 2006). They suggest that instead, accumulation of transposable elements has lead to a separation of enhancer regions from the promoter site. For gene expression a mechanism has evolved where the remote enhancers are brought into close proximity to the promoter by the compression of the interspersed repeated sequences, facilitated by heterochromatin Su(var) proteins such as HP1.

Gene mapping in *D. melanogaster* heterochromatin

The endeavour to understand the significance of heterochromatin and how it influences gene expression has mostly focussed on the model organism *D. melanogaster*. Several factors related to the repetitive nature of the genomic sequence inherent to constitutive heterochromatin and its pericentric location have meant that genetic and molecular methods available for euchromatic gene mapping are not amenable for work in heterochromatin and until recently very little gene mapping of these regions has been possible.

Close proximity to the centromere has meant genetic mapping through meiotic recombination is not possible. Instead early geneticists developed a method using genetic complementation with known heterochromatic deficiencies created by irradiation to map essential loci. Amongst the first genes mapped to hetrochromatin in this way were *light (lt)* and *rolled (rl)* on the second chromosome; (Hilliker. 1976) and reviewed in (Fitzpatrick, Sinclair, Schulze, et al. 2005).

Ultimately, a genetic locus can be related to a unique cytological location through hybridisation of a labelled sequence probe with chromosome DNA that has been stained with various chemicals to produce a chromatin banding pattern. The cytological location contributes to a cytogenetic map that is useful for positioning a locus with respect to other genes and deficiencies in the region as well as in reference to repeat sequences such as satellite sequences and transposable elements and also to relate it to the physical sequence map.

Euchromatin band staining uses polytene chromosomes retrieved from salivary glands. These highly replicated chromosomes are joined at the centromeres and are therefore not as useful to distinguish bands in heterochromatic regions. Heterochromatin staining uses smaller mitotic chromosomes which gives a more detailed resolution of the pericentric region. Bands produced in this way are labelled from h1-h60.

The repetitive nature of heterochromatic regions has hindered the construction of a comprehensive contiguous genomic sequence and has also made molecular work challenging. Genetic loci mapped to deficiencies cannot be aligned with genomic sequence until a high quality sequence database is available. Once a locus has been identified with a predicted gene sequence using molecular methods, analysis with respect to neighbouring genes and sequences is not possible until the overall scaffold order has been resolved. The first release of the 120Mb of *D. melanogaster* genomic sequence omitted most of the 60Mb of heterochromatin sequence (Adams, Celniker, Holt, et al. 2000). Release 3 gave 21 Mb of heterochromatin sequence in short unlinked scaffolds, that had many gaps and low quality sequence because of the difficulty in aligning multiple repeat sequences (Hoskins, Carlson, Kennedy, et al. 2007).

The recently released Release 5 *D.melanogaster* genomic sequence has resolved many of these gaps and alignment problems. Hoskins et al. (2007) have described the generation of 15Mb of finished or improved sequence out of the possible 20Mb of heterochromatin amenable to mapping. Whole genome shotgun (WGS3) scaffolds from release 3 were assembled into larger scaffolds

with improved quality of sequence and gaps were filled with a new set of scaffolds. The use of five bacterial artificial chromosome (BAC) libraries has allowed more scaffolds to be linked into BAC contigs covering 13.4 Mb of pericentric sequence.

24 Mb of the release 5 genomic sequence has been annotated by Smith et al. (2007) rendering at least 230 to 254 predicted protein coding sequences which are supported by their conservation in other Drosophila lineages and divergent species. It was also found that 77% of heterochromatin sequence is composed of transposable elements or fragments of transposable element sequence and other repeat sequences interspersed between genes and within intron sequence.

Initial characterisation of a number of essential genes residing in the densely packed heterochromatic environment has generated interest in the pursuit of understanding any common mechanism(s) that may be necessary for their transcription, and to understand the biological significance of heterochromatin itself. These questions have given momentum to the development of a set of complementary molecular, cytological and genetic tools to characterise remaining heterochromatic genes at a functional level (Fitzpatrick, Sinclair, Schulze, et al. 2005). Also, further improvement in the quality of the genomic sequence available will facilitate ongoing gene characterisation by providing a physical map anchor point for new loci and an entry point for reverse genetic methods such as RNAi. As more heterochromatic genes are mapped and characterised, identification of any common features in structure, function

and mode of operation amongst them may highlight specific adaptations that genes have needed to accumulate to counteract the repressive heterochromatin environment.

Heterochromatin of the third chromosome

Work towards a complete genetic and molecular profile of heterochromatin on the third chromosome has identified approximately 20 genes (Fitzpatrick, Sinclair, Schulze, et al. 2005). Mapping of this region began with the identification of 12 complementation groups identified with the use of a set of autosome detachment deficiencies produced by Marchant and Holm (1988). The ongoing identification of new genetic loci and generation of deficiencies through the use of EMS, X-ray mutagenesis and P-element tranposition screens has further refined gene map positions (Sinclair, Schulze, Silva, et al. 2000), and extended the region covered by deficiencies, as well as providing overlapping deficiencies.

This comprehensive set of deficiencies on the third chromosome has provided a base for genetic studies from which the characterisation of newly discovered genes can be done using complementary methods such as molecular studies, genomic sequencing, cytological placement, RNAi and transgenic rescue (Schulze, Sinclair, Fitzpatrick, et al. 2005). Currently at least 16 of the loci on 3L and 3R are known to be essential and many have been characterised at the functional level (Schulze, Sinclair, Fitzpatrick, et al. 2005).

zeppelin

zeppelin (*zep*) was isolated coincidentally by the Bejsovek lab, in an F3 lethal genetic screen for EMS mutations that would reverse the *wg* phenotype. During the screen, cuticle preparations were found that did not alter the *wg/wnt* pathway but did have an altered embryo phenotype (Ostrowski, Dierick, Bejsovec. 2002).

The Bejsovec lab found that the recessive lethal mutation was characterised by initial hyperactivity and later failure to hatch. When mutant embryos were mechanically dechorionated and devitellinized, the cuticle was found to stretch to a much greater extent than the wild-type cuticle (Ostrowski, Dierick, Bejsovec. 2002). This new mutation, with a 'blimp' phenotype was named *zeppelin* (allele *LP13*). The expanded cuticle phenotype suggested that a component of the exoskeleton may be affected in *zep* mutants.

I chose to focus my research on *zep* because it is a particularly interesting gene in terms of investigation into how essential genes operate in heterochromatin. According to earlier screens by Lindsley et al. (1972) and Tasaka and Suzuki (1973), *zep* is located in an area of distal 3R heterochromatin that appears to be quite gene poor. It was therefore unusual to find that despite the repressive heterochromatin environment, a gene such as *zep* appeared to be not only an essential single copy gene, but may also be involved in cuticle formation, a function that would require specific spatial and temporal control. The fortuitous generation of a phenotype associated with lethality (blimp),

allowed some speculation on *zep* function and gave some indication of putative candidate genes located in the distal 3R heterochromatin region.

zep^{LP13} mapped to 3R deficiencies.

The Bejsovek lab found *zep* failed to complement Df(3R)4-75 which is characterised as an inversion which has disrupted loci *81Fa* and *81Fb* in h58 on 3R (Marchant, Holm. 1988). A number of 3R deficiencies were used to characterise the *zep* locus further. The most extensive lesion used for mapping heterochromatin on the third chomosome is *Df(3L+3R)6B-29* which uncovers heterochromatin sequence on both sides of the centromere (Koryakov, Zhimulev, Dimitri. 2002). Some smaller deficiencies used for screening putative *zep* alleles were made in the Honda lab by X-ray mutagenesis and may only delete one or two essential genes to result in homozygous lethality e.g. *Df(3R)9B-919* is a small deficiency on 3R (Fitzpatrick. 2005).

zep^{LP13} was also found to be lethal in combination with 3R deficiencies *Df(3R)XM3*, Df(3R)A321, *Df(3R)7B-90* and *Df(3R)9B-919* (figure 3.1) (Fitzpatrick. 2005). It complemented all deficiencies on 3L, and *Df(3R)10-65* on 3R (see Table 3.1).

zep alleles from the Leptin lab

Two putative *zep* alleles were acquired from the Leptin lab (Fitzpatrick. 2005). Leptin *L-1* (*l400-1*) and Leptin *L-8* (*l400-8*) were generated in an EMS screen for alleles of *abstrakt (abs)* which is located in band 82A1 (Irion, U. and

Leptin, M. 1999). The complementation map for these two alleles was identical to that found with *zep*^{LP13}. Lethality was found in combination with 3R deficiencies *Df(3R)XM3*, *Df(3R)A321*, *Df(3R)7B-90* and *Df(3R)9B-919* (Fitzpatrick. 2005). Non-complementation with *zep*^{LP13} showed *l400-1* and *l400-8* were allelic with *zep* (Table 3.1).

zep P-element alleles

P-element lesions are generated when a transposable element is induced to jump with the availability of transposase. This event eliminates the intervening sequence, hence it is often dubbed a 'male recombination' event.

In the transposition events for *zep* alleles *8740-20*, 8740-22 and *EP-167R*, the distal end of the element remained in its original position while the proximal end has reinserted itself in a position closer to the centromere. Two *zep* alleles resulted from the mobilisation of P-element KG08740: *8740-20* and *8740-22*. The P-element originated distal to the heterochromatin-euchromatin border, 29bp within the proximal end of scaffold AE003607, at cytological band 81F (Fitzpatrick. 2005).

A third P-element deletion *EP167-R*, was generated from the mobilisation of P-element *EP(3)3632*, (Fitzpatrick. 2005). This had the same complementation map as the above 8740 P-element deletions. *EP(3)3632* originates 16kb from the end of AE003607.

These three lesions failed to complement deficiencies *Df(3L+3R)6B-29*, *Df(3R)XM3*, Df(3R)A321, Df(3R)4-75, *Df(3R)7B-90* and *Df(3R)9B-919* and the EMS generated alleles of zep^{LP13}, *l400-1* and *l400-8*. They also failed to complement *Df(3R)4-75* but are viable with *Df(3R)10-65* which is known to delete part of h58 (Marchant, Holm. 1988; Fitzpatrick. 2005).

Candidate genes on 3R

As mentioned above, very few essential loci have been revealed on 3R by genetic complementation methods (Marchant, Holm. 1988). *zep* was originally mapped based on its lethality in combination with *Df(3R)4-75* which is located in the heterochromatin cytological band h58, close to the heterochromatin-euchromatin border. Predicted candidate genes in the genetically mapped area include *lethal (l(3)81Fb)*, CG12581, and CG12449. All are localized to polytene band 81F according to Flybase (Graack, Cinque, Kress. 2001) (http://flybase.bio.indiana.edu/).

Of the two loci mapped to 3R by Marchant and Holm (1988), the gene designated as *l(3)81Fb* is a candidate for *zep* based on its known lethality with the same deficiency.

The original work by Marchant and Holm (1988), identified 10 EMS mutants that failed to complement *Df(3R)4-75* but complemented *Df(3R)10-65.* The 10 mutants together defined the *l(3)81Fb* locus. These alleles were further divided into two complementary groups. According to Flybase (Gelbart, Crosby, Matthews, et al. 1997), the *Df(3R)4-75* deficiency is located in cytological band

81F, and the *l(3)81Fb* locus is in the same cytological band based on this non-complementation. However *l(3)81Fb* has not been localised to the genome sequence and a phenotype is not described. Unfortunately no alleles for this locus are currently available as the original mutants have been lost.

CG12581 located distal to the heterochromatin-euchromatin border in 81F6 was another candidate gene based on its position in relation to the 4-75 3R deficiency. Flybase annotation of this gene is limited. There are no expression data or alleles listed. Molecular function is unknown and there are two predicted transcripts.

Glutamine:fructose-6-phosphate aminotransferase 1 (*Gfat1)* is identified as computed gene numbers *CG12449* and *CG40197* on Flybase (Graack, Cinque, Kress. 2001). It is proximal to CG12581 and was therefore considered a stronger possibility for correspondence to *zep* because of its close alignment with 4-75 and 10-65 genetic complementation data. Also in its favour was the reported biological function for GFAT1 which included involvement in embryonic cuticle formation as discussed below.

Glutamine:fructose-6-phosphate aminotransferase 1 (*Gfat1)*

Correspondence of the *zep* locus isolated by genetics (Ostrowski, Dierick, Bejsovec. 2002) with the physical map gene sequence *Gfat1* was suspected based on its location at cytological band 81F, *zep* failure to complement *Df(3R)4-75,* and the known function of GFAT1 in the synthesis of chitin, the main component of the embryo and adult cuticle (Fitzpatrick. 2005).

GFAT1 is the rate limiting enzyme in the hexosamine pathway and functions to convert glucose to glucosamine-6-phosphate (Glc6P). The pathway is common in all species because of its dual role as a monitor of cell glucose levels and its formation of UDP-*N*-acetyl-D-glucosamine (UDP-NacGlc) for glycosylation of proteins. In insects this pathway contributes to a third function, the production of UDP-NacGlc for polymerisation and chitin formation, here it is called the chitin synthesis pathway. Chitin is the major component of the fruitfly cuticle and tracheae. In fruit flies, the highly glycosylated salivary gland secretion glue proteins (Sgs) also requires NacGlc.

Human *Gfat1* and diabetes

Human *Gfat1* has been implicated in connection with diabetes through its function in the hexosamine sugar monitoring and regulation pathway. It has also been implicated in Alzheimer's disease when hyperphosphorylation and aggregation of tau takes place as a result of insufficient UDP-NacGlc availablity for tau glycosylation. The production of UDP-NacGlc requires glucose, UDP and glutamine and therefore levels of the end product of the hexosamine pathway, UDP-NacGlc reflect the level of nutrients and energy available in the cell. The quickly responding level of UDP-NacGlc act as a sugar level sensor that passes the signal forward by its availability for glycosylation of regulatory proteins.

The enzyme OGT (UDP–*N*-acetylglucosaminyl transferase) is present in all cells but especially found in the pancreas and brain where it catalyzes the addition of UDP-NacGlc to regulatory proteins. Approximately 2-5% of the

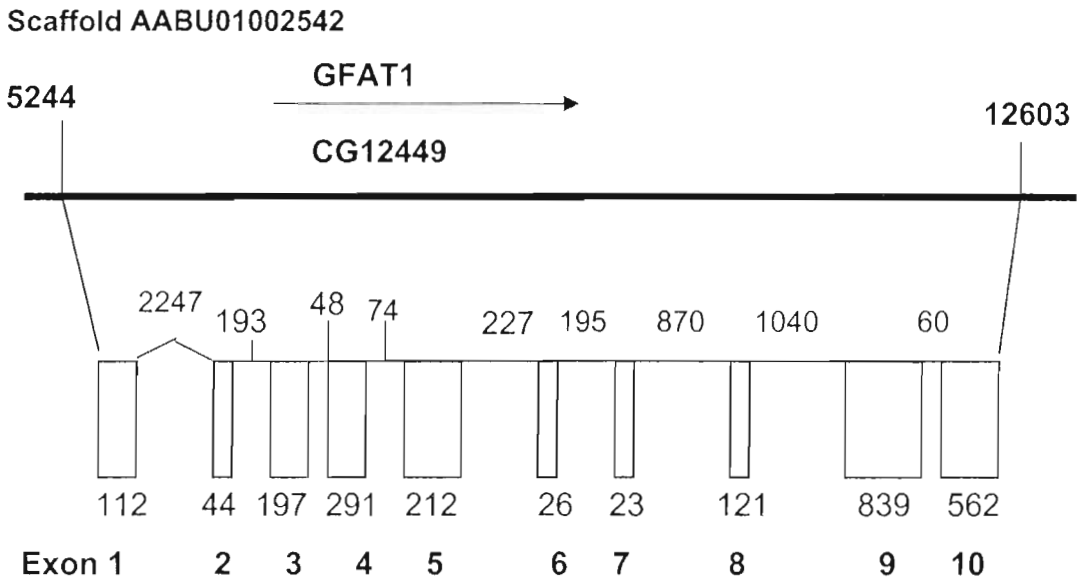
glucose entering the cell is diverted to the hexosamine signalling pathway for the purpose of sugar level monitoring (Marshall. 2006; Hart, Housley, Slawson. 2007).

Diabetes and Alzheimers disease are both associated with abnormal O-GlcNacylation of proteins (Hart, Housley, Slawson. 2007). O-GlcNacylation prevents phosphorylation by kinases which alters the activation of essential regulatory proteins. When glucose levels in the cell are abnormally high, there is an increase in O-GlcNacylated regulatory proteins and diabetes associated pathologies such as decreased glucose transporter expression, insulin signalling and glycogen manufacture result. Thus hyperglycemia can be a self-perpetuating state. In the brain when glucose levels are lower than normal, decreased O-GlcNacylation occurs and tau is hyperphosphorylated which is associated with Alzheimer's disease (Hart, Housley, Slawson. 2007).

Sequence analysis of *Gfat1*

The *Gfat1* genomic sequence is ~7360 nucleotides long and is located on 3R scaffold AABU01002542 which is 48.5 kb long. There are a total of ten exons (Figure 1.1) and six alternate splice transcripts are predicted according to Gadfly Release 3 version 1 from 2003 (Flybase). Intron sequences vary from 47bp to ~2kb and the translation start codon is at genomic position 7641, in exon 2.

Figure 1.1 *Gfat1* Organisation



A total of eleven *Gfat1* EST clones have been described that match the *Gfat1* genomic sequence on Flybase(http://flybase.bio.indiana.edu/) , however only one cDNA is listed, RE72989, which corresponds to the predicted CG12449-RA splice variant, using all ten exons. Two of the ESTs were found to be full length cDNAs available from DGRC, GH03520 and LP07309. The GH03520 clone contained intron sequence when sequenced. RE72989 and LP07309 cDNA clones both contained errors and will be discussed below in reference to the rescue experiment. Two alternate polyadenylated *Gfat1* transcripts were located by Graack et al. in 2001, they were 2406 nt and 2413 nt in length. The longest ORF is 2082 nucleotides. There is a TATAAA box, found 40bp 5' of the transcription start site at genomic position 5251.

Gadfly Release 3 database annotation shows that there are two other computed genes in close proximity to *Gfat1*. CG40338, is located at position 12615, following exon 10 of *Gfat1* on scaffold AABU01001252, and in opposite orientation to *Gfat1*. CG40338 has two introns and is predicted to be 759 nucleotides long. However, six-frame translation shows that the longest ORF is only 210 nucleotides long therefore the predicted protein may not be a functional protein. CG40198 at position 36228 on scaffold AABU001252 has one intron and is predicted to be 727 nucleotides long in reverse orientation to *Gfat1*. The six-frame translation for this sequence shows that the longest ORF is 555 nucleotides long. Neither CG40338 nor CG40198 have been characterised further and there is no predicted functional information on Flybase.

Since heterochromatic scaffolds depicted on Flybase are unordered, it is not possible to determine the orientation of scaffold AABU01002542 with respect to the centromere or the heterochromatin-euchromatin boundary or to other predicted genes located on nearby scaffolds.

Heterochromatin cytological bands on 3L are numbered h47-h52 moving from the euchromatin border moving proximally towards the centromere. Those on 3R are numbered h53-h58 starting at the centromere moving distally toward the euchromatin border. The polytene chromosome band 81F encompasses all of the 3R heterochromatin bands as well as the first part of euchromatin on 3R.

Currently Flybase places the *Gfat1* locus in the heterochromatin cytological bands h54-h55. However Flybase also shows the euchromatic

cytological band placement as 81F, based on in situ hybridization experiments with a biotin labelled probe to polytene chromosomes by Graack et al. in 2001.

The two different methods of staining, using polytene or mitiotic chromosomes, make it difficult to define the exact cytological locations of genetic loci in this region. *Gfat1* is either localised in hetrochromatin bands h54-h55 according to Flybase or in h58 according to non-complementation with *Df(3R)10-65* mentioned above. The placement within the euchromatin cytological band 81F does not provide any greater resolution, as it encompasses the whole h53-h58 region.

The *Gfat1* paralog *Gfat2* (CG1345) resides in 3R euchromatin at cytological band 98C4 on scaffold AE003765. *Gfat2* cDNA is 79% homologous with *Gfat1*, but has only a single intron compared to the nine *Gfat1* introns. From an evolutionary perspective it is likely that *Gfat2* and *Gfat1* have resulted from either a gene duplication event or reverse transcription and insertion of *Gfat1* sequence in euchromatin to form *Gfat2*. *Gfat2* has a single predicted transcript, 2552bp long.

The *Drosophila melanogaster* GFAT1 (Dmel/GFAT1) protein sequence is 694 amino acids long, *Drosophila melanogaster* GFAT2 (Dmel/GFAT2) is 683 amino acids. There is 81% homology between the two sequences which increases to 88% when functional substitutions are considered. There are an additional 23 amino acids in GFAT1 absent in GFAT2 that correspond to exons 4 and 5. They form an extended variation of the hinge region between the two

main functional domains. There is 62% homology between human GFAT1 and Dmel/GFAT1, increasing to 77% when similar functional groups are considered. The level of homology suggests that they are likely to both be in the hexosamine pathway, however the biological process they contribute to could be different depending on when they are expressed and how they are regulated. This issue is further addressed below and in the discussion.

Since the GFAT1 ortholog Glucosamine 6-phosphate synthase (*GlmS*) in *E.coli* has been extensively studied, the domains and main functional residues of the protein are well characterised. Alignment of the polypeptides between *Glms*, *Gfat1* and *Gfat2* in *D. melanogaster*, and human *Gfat1* allows a greater understanding of *Gfat1* in *D. melanogaster* and identifies regions of importance by comparison of conserved residues (Figure 1.2).

Figure 1.2 ***D. melanogaster*** **GFAT1 amino acid alignment with** ***D. melanogaster*** **GFAT2, human GFAT1 and** ***E. coli GlmS*****.**

D.mel/Gfat1 and D.mel/Gfat2 are 81% homologous.
Residue conservation: *= fully conserved, := strong group conserved, .= weak group conserved. Gaps are filled with dashes.

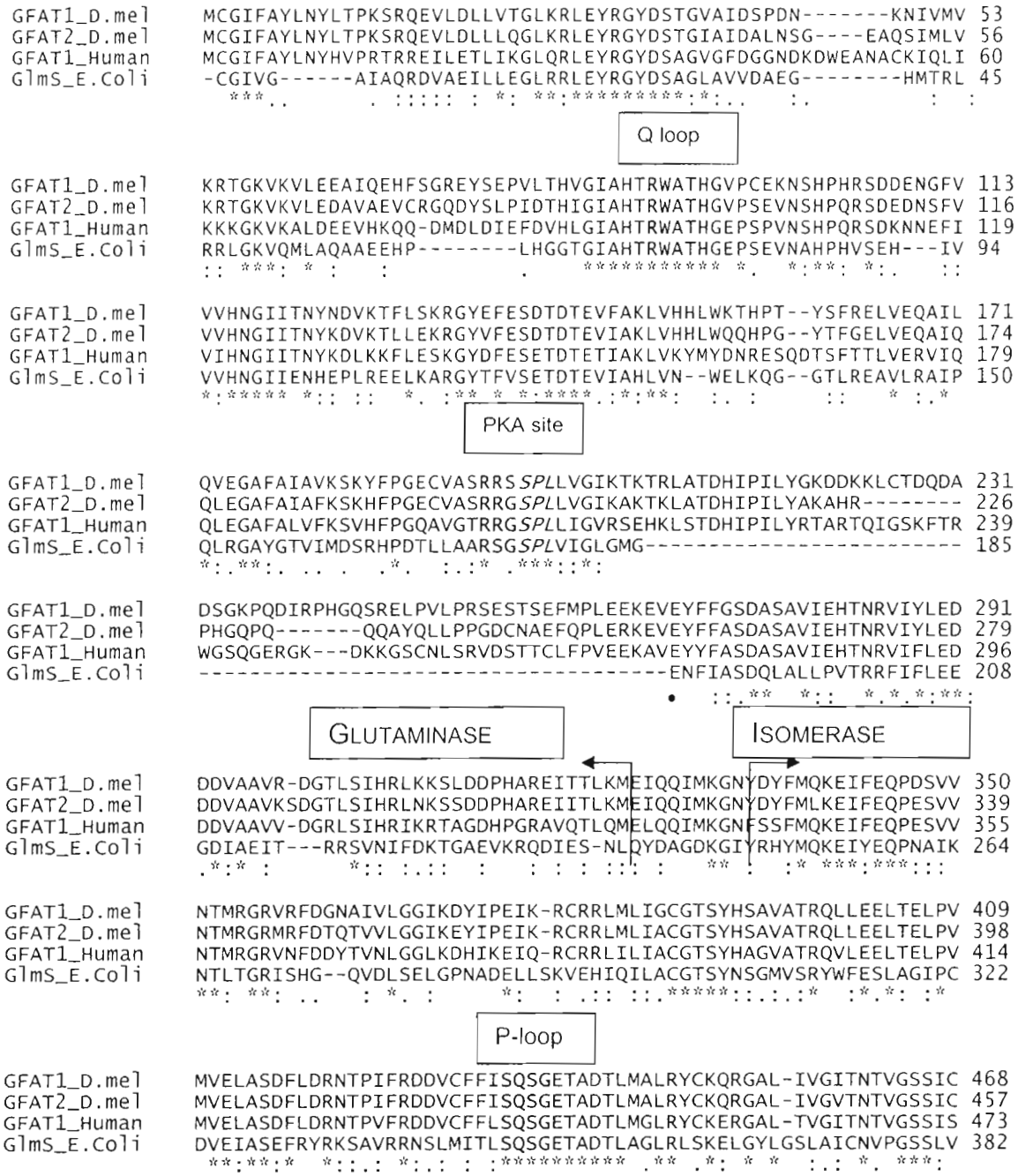


Figure 1.2 (continued)

There are two structurally and functionally distinct domains (Teplyakov et al., 2001). The first 307 amino acids in Dmel/GFAT1 forms the glutamine binding domain which catalyses the hydrolysis of glutamine to glutamate and ammonia. This includes the Q-loop at residues 88-95 that functions to shield the active site once glutamine is bound (Figure 1.2).

The C-terminal isomerase domain at residues 317-637 uses the ammonia released by the glutaminase domain to add to fructose 6-phosphate (Fru6P) and performs an isomerisation of fructosimine 6-phosphate (FruN6P) to glucosamine 6-phosphate (GlcN6P). All GFAT proteins must associate as homodimers in order to function as the catalytic residue for opening the Fru6P sugar ring is on the opposing dimer. His*590 is the only residue on this subunit that has an

identified functional role. The core of the dimer is the pair of isomerase domains, the glutaminase domains are on each side.

The GlmS crystal structure shows that both substrates, glutamine and Fruc6P, bind to the same subunit. The P-loop (residues 534-539), forms H-bonds with the phosphate group of the sugar in an open conformation and stabilizes it for the isomerisation reaction. In order to shield the ammonia produced by the glutaminase domain from protonation by contact with the solvent, it travels through a channel between the two active sites. Both domains contribute hydrophobic residues to the channel, and the central part is formed by the residues of the C-loop at 687-693.

Gfat1 expression

As well as the documented function of GFAT1 in the chitin synthesis pathway, the GFAT1 expression pattern makes it a good candidate sequence to correspond with zep as it coincides with that expected of a protein involved in cuticle formation that might contribute to the *zep* phenotype. Expression begins at stage 16 in the developing tracheae and extends to the whole cuticle by stage 17. It is also expressed in late third instar larva salivary glands that are producing the highly glycosylated sgs proteins for the attachment of forming pupa cases (Graack, Cinque, Kress. 2001). Expression patterns for Gfat2 are currently not available.

Gfat1 regulation

The *Gfat1* enzyme is likely to contribute to several cell processes making the regulation of this enzyme an intriguing area of study. The issue is further complicated by the location of the *Gfat1* gene within heterochromatin. The gene must be explicitly expressed in order to contribute to the proper function of Sgs proteins in corpus cells of the salivary glands as well as more generally expressed for cuticle formation in epithelial cells. Since the Sgs genes have such a strict regime of expression, they have been widely used as a model to study how systemic hormone release is translated and refined for the co-ordinated expression of proteins at the tissue level (Renault, King-Jones, Lehmann. 2001).

Temporal and spatial regulation of the Sgs genes is under the control of 20-hydroxyecdysone (20E), therefore it would be expected that enzymes associated with their glycosylation might also have the same control mechanism. However according to Graack et al. (2001) and Kato (2002) there are no Ecdysone response motifs in the upstream sequence of *Gfat1*. Broad-complex (BR-C) is a zinc finger transcription factor that has four isoforms, Z1-Z4. BR-C is directly activated by 20E and is also required for activation of the early genes, late genes and Sgs genes in metamorphosis (Fletcher, Thummel. 1995). Fork head (Fkh) has also been found to be required for Sgs gene expression. It is a member of the HNF-3/Fkh family of transcription factors associated with the specific expression of ecotodermal parts of the gut and salivary glands.

Further analysis of the *Gfat1* promoter region has revealed motifs that closely correspond (85-89%) to the BR-C isoform Z1 response element and have 90-95% correspondence with isoforms Z2 and Z4 (Kato, Dasgupta, Smartt, et al. 2002). However the Fkh consensus sequence (TnnGTnT) is contained within the BR-C sequence (underlined); Z1 A<u>T</u>TT<u>GT</u>T<u>T</u>ATTA. BR-C has been found to be required for *sgs* gene activation and repression. However absence of BR-C at *sgs* gene regulatory sites was confirmed by the lack of polytene chromosome in situ binding by a BR-C protein complex probe in the *sgs* gene regulatory region as well as by mobility shift DNA assays. This suggests that BR-C could be indirectly regulating the *sgs* genes, possibly through Fkh (Renault, King-Jones, Lehmann. 2001). Analysis of the promoter region in *Gfat1* shows that the Fkh consensus sequence is present at genomic scaffold position 5042, ~200bp upstream from the transcription site. Therefore it is possible that Fkh may also regulate *Gfat1* in the corpus cells of salivary glands to coordinate *Gfat1* expression with sgs proteins in order for glycosylation to take place.

The regulation of *Gfat1* in chitin synthesis is still unclear. Graack et al. (2001) have suggested that *Gfat1* could be constitutively regulated in its role in cuticle production and therefore the PKA site, located at residues 194-199 may or may not be necessary for *Gfat1* activity in chitin formation.

To understand the role of PKA and the negative feedback loop established by UDP-NacGlc, Graack et al. (2001) have investigated *Gfat1* regulation by assay of G6P in cell-free extracts from yeast cells expressing the *Gfat1* cDNA. They found that cAMP increased the basal activity 1.7 fold by activation of the

endogenous yeast PKA present. The addition of UDP-NacGlc repressed activity independently of cAMP addition and therefore shows that negative regulation by UDP-NacGlc can overcome the PKA activation. The level of repression by UDP was almost 100% compared to GFAT in other systems such as human GFAT1 where it is 51%.

Chitin synthase performs the final step in the chitin synthesis pathway. It is a transmembrane protein that forms the chitin polymer from UDP-NacGlc monomers, and deposits it to the external cell matrix. (Fletcher, Thummel. 1995) have reported that chitin synthase is directly regulated by ecdysone. Therefore Fkh or a similar mediator of the 20E signal may also initiate *Gfat1* transcription in epithelial cells for chitin synthesis. Once chitin synthase is activated the basal concentration of the regulatory UDP-NacGlc molecule decreases in the cell, which would cause GFAT1 activity to increase. Since O-glycosylation of regulatory proteins often takes place on the same serine as phosphorylation by kinase (Hart, Housley, Slawson. 2007), a similar mechanism may exist in *Gfat1* for chitin synthesis, a decrease in UDP-NacGlc may directly free serine 196 for activation by PKA, thereby increasing GFAT1 activity as supported by the cell free assay results above.

How *Gfat1* and *Gfat2* are regulated for the above sgs and chitin synthesis functions and also function to provide UDP-NacGlc for O-glycosylation and sugar level monitoring is unknown. However it seems likely that a basal level of expression of *Gfat1* or *Gfat2* or both must exist for these 'housekeeping' purposes and therefore the promoter region must accommodate this function

also. In the mouse, *Gfat1* and *Gfat2* have been reported as expressed in separate tissues. In both *Gfat1* and *Gfat2* upstream regulatory sequences there are three Sp1 sites and three Ap-1 sites. The absence of a TATAAA box and the presence of the GC-rich promoter sequence in the mouse suggests that they are both housekeeping genes. Both Sp1 and AP-1 proteins have been confirmed at their respective binding sites by electrophoretic mobility shift assays (EMSA). A silencer element was also confirmed in *Gfat2* which represses expression by greater than 50% (Yamazaki, Mizui, Oki, et al. 2000).

The promoter region for *Gfat1* in *D. melanogaster* includes three putative Sp1 regulatory protein binding sites with sequence GGGCGG known as a GC box (this is the consensus sequence in mammals, however it is not known whether this is the same sequence for the Drosophila Sp1 (D-Sp1) transcription factor). Two putative Activating Protein (AP-1) binding sites were also identified ~400bp upstream (Graack, Cinque, Kress. 2001).

Thus it seems possible that basal expression of *Gfat1* in *D. melanogaster* operates in a similar manner to that found in mouse *Gfat1* and *Gfat2* using the Sp1 and AP-1 binding sites, and that *Gfat1* expression for chitin synthesis and glycosylation of sgs proteins uses the TATAAA box and Fkh motifs. Investigation of *Gfat2* promoter regions and expression patterns would be useful in determining if there is a separation of the various functions of GFAT between *Gfat1* and *Gfat2*. In mouse and humans, *Gfat1* expression predominates over *Gfat2* in most tissues, apart from in kidney cells, where there is a higher level of expression of *Gfat2* (Yamazaki, Mizui, Oki, et al. 2000).

Correspondence of *Gfat1* with *zep*

Initial PCR experiments showed that Gfat1 was missing in the deficiency *Df(3L+3R)6B-29* and present in *Df(3R)10-65* which supported the correspondence of *zep* with *Gfat1* (Fitzpatrick. 2005).

To further support the identification of *Gfat1* (CG12449) with *zep* I have used single embryo PCR experiments to characterise the point mutations found in *Gfat1* in *zep* mutants. Additionally I have shown that the *Gfat1* sequence is missing from P-element generated *zep* alleles.

The *Gfat1* sequence was mapped to distal 3R deficiencies in heterochromatin using PCR and the resulting map coincides with the genetic localisation of *zep* mutants in this region by non-complementation. Also, since the deficiency *Df(3R)10-65* was previously localised cytogenetically to h58, combining this with the genetic and molecular data gained from *zep* alleles created by P-element deletions, the cytogenetic placement of *Gfat1*/*zep* is now possible.

Alignment of Gfat cDNA sequences from *E. coli* with *D. melanogaster* highlighted conserved regions and some important functional residues have been identified in Dmel *Gfat1*. This has allowed me to relate the nature of the *Gfat1*/*zep* point mutations to the degree of severity of the embryo and transheterozygote adult phenotypes. The correlation of mutation with phenotype has given some insight into the function of *Gfat1* as a homodimer. It has also further supported the correspondence of *Gfat1* with *zep* by showing that the

nature of the severity of the molecular mutation is reflected in the degree of severity of the phenotype.

RNAi experiments have further supported the correspondence of *zep* with *Gfat1* by showing that *Gfat1* is an essential gene and by the replication of *splayed* (*Spl),* the *zep* transheterozygote adult phenotype.

Results indicate that *zep* corresponds to *Gfat1* and this has allowed some speculation about the regulation and function of *Gfat1* within the chitin synthesis pathway and its contribution to cuticle formation. Based on map localisation, *Gfat1/zep* likely corresponds to the *l(3)81Fb* locus. Based on the similarity of the transheterozygote adult phenotype, the *zep* locus may also be associated with the *Spl* locus identified by Lindsley et al in 1972.

CHAPTER 2: MATERIALS AND METHODS

EMS screen for new *zeppelin* alleles

Ethyl methanesulphonate (EMS) mutagenized flies were obtained from the Zuker lab (Koundakjian, Cowan, Hardy, et al. 2004) and screened for non-complementation with deficiencies that covered all of 3L and 3R. After initial identification of non-complementation with 3R hetrochromatin deficiency *Df(3L+3R)6B-29* narrower deficiencies were selected to progressively define the mutant locus. Deficiencies used for the initial screen for heterochromatic genes of interest were *Df(3R)XM3* and *Df (3L)Delta AK*^{1} *Df(3R)7B-90* and *Df(3R)9B-919,* these are smaller deficiencies specific to 3R. *zep* alleles were ultimately confirmed by their non-complementation with the original *zep*^{LP13} allele. Fly stocks used were from the Berkeley Drosophila Genome Project (BDGP) Center.

Fly stocks

All flies were kept in vials or bottles containing a standard fly food cornmeal-sucrose mix with tegosept or proprionic acid as a mould inhibitor. Crosses were made at room temperature or 25 degrees.

Descriptions of all mutations and deficiencies may be found on the Flybase website ; (http://flybase.bio.indiana.edu/).

Bacterial strains and vectors

The *E. coli* lab stock DH5α was used to propagate plasmids (pBluescript, pTZ57R, pUAST, pSympUAST). Transformation was done by heatshock for 20 seconds at 42 ^{0}. Bacterial cultures were grown in LB medium (Sambrook et al. 1989), with the appropriate concentration of antibiotic (100 ug/mL ampicillin, 50ug/mL kanamycin, 25ug/mL chloramphenicol).

Isolation of plasmid DNA

Plasmid DNA was isolated using standard alkaline lysis procedures (Sambrook, Fritsch, Maniatis. 1989).

Single embryo collection and PCR

This technique was used to collect homozygous DNA from EMS mutant *zep* alleles for sequencing, to carry out deficiency mapping, and for P-element exon testing in P-element *zep* mutants.

All *zep* mutants are homozygous lethal and therefore kept as heterozygotes over a balancer chromosome. The Zuker flies were initially balanced over TM6, a rearranged third chromosome that includes the *humeral (hu)* and *tubby (tb)* markers. They were rebalanced over a TM3 balancer carrying the *Green Fluorescent Protein (GFP)* and *serrate (Ser)* markers. The *GFP* marker could then be used to test the DNA processed from embryos for homozygosity.

Single embryos were collected from plates made with an agar and apple juice solution. Embryos were lysed in 13ul of embryo lysis solution with 2ul of 20 ug/mL stock proteinase K in 0.6ml tubes. Sterile technique between embryos was maintained throughout collection. Tubes were incubated at 37^0 for 25 minutes and transferred to 99^0 for 15 minutes.

Single embryo PCR reactions were carried out using two sets of primers, GFP-F GFP-R for homozygote testing as described above and Grp84-01 Grp84-02 corresponding to a gene on the X-chromosome that tests for DNA template quality. 2ul of embryo DNA were used for each reaction.

PCR reactions for sequencing and exon testing in homozygotes used 1 ul of genomic homozygote DNA or ½ ul of DNA amplified using the genomi-Phi kit (Amersham). 50ul PCR reactions were run with a 30 cycle amplification. 10ul of the resulting PCR product was run on an agarose gel to check product concentration and band size before sending for sequencing.

Sequencing of EMS *zep* alleles

Primers were designed that would provide sequence that extended beyond each *Gfat1* exon intron splice site. The resulting PCR products together covered the complete coding sequence for the *Gfat1* gene. The PCR fragments generated varied from the smallest for exon 1 at 500bp to the longest at ~2kb, which covered exons 3 – 7. Forward and reverse sequencing primers were used

within each PCR product, approximately at 500bp intervals, in both directions, to provide duplicate sequence for analysis.

After purification using the PCR purification kit supplied by Invitrogen, samples were sent with 5mMol concentration primers to Macrogen for sequencing. A BLASTN comparison of *Gfat1* genomic sequence with the PCR sequence result was performed, if a base change was detected the sequence results were compared against the *Gfat1* cDNA sequence based on the RE72989 cDNA sequence (see Appendix B for RE72989 sequence).

Sequencing was repeated twice using a total of three different embryos to confirm each point mutation found.

Plasmid restriction digests

All restriction digest enzymes were from Invitrogen or NEB. The amount of DNA used was between 500-1000ng in a volume of between 25-50ul. Digests were carried out at 37 ^{0} for a minimum of 2 hours.

Agarose gel electrophoresis

Agarose gel concentration varied between 0.7% and 1.5% depending on the size of the fragments visualized. Ethidium bromide was added to a concentration of 500ng/ul. All gels were run in 0.5X TBE (Sambrook, Fritsch, Maniatis. 1989) and visualized using a UVP documentation system.

RNAi construct

A 290 bp fragment was amplified by PCR from position 628-928 of the RE72989 cDNA sequence for insertion into the psympUAST vector. The forward primer was GFATRNAIF – GACTCCTTCCTCGAGCTGT where the underlined T replaced G of the original sequence in order to amplify a modified fragment incorporating a XhoI restriction site. The reverse primer was GFATRNAIB – TCAGAATTCCTTTCCGAACGG where the underlined C replaced G and the underline T replaced G creating an EcoRI site. The fragment generated was unique to *Gfat1* i.e. there was no homology with the predicted *Gfat2* transcript sequence.

The 290bp fragment was inserted into pTZ57R, digested with EcoRI and XhoI, and inserted into the RNAi vector pSympUast. The construct was then prepared using the Qiagen maxiprep kit, sequenced and sent to Bestgene at a final concentration of 5070ng/ul for injection.

The RNAi construct was injected into w^{1118} embryos. Adults with a red eye phenotype were crossed back to w^{1118} flies and balanced over a chromosome two balancer chromosome carrying the *CyO* marker, or the third chromosome balancer TM3 carrying the *Ser* marker.

Ten lines carrying the RNAi transgenic contruct were produced, five on the second and five on the third chromosome. Crosses with various GAL4 transcriptional activators (driver) lines generated expression in a tissue specific manner as discussed below.

Rescue construct

Erroneous sequence was found in both the LP07309 and RE72989 cDNAs received from the Drosophila Genomics Resource Centre (DGRC) (see below for an explanation of errors). To correct the errors the first 2138 bp corresponding to the EcoRI internal cut site were used from RE72989, spliced with the last 315 bp from LP07309. pBluescript was used for cloning the two fragments together as it had the advantage of blue/white selection. Once the correct size and orientation of the *Gfat1* cDNA was verified by restriction mapping and sequencing, it was inserted into the pUAST expression vector and sent to Bestgene for injection. Appendix A is the corrected cDNA sequence. As discussed below injections with this rescue construct failed to produce red eye adult flies. A new construct has been made using a new vector source and is now ready for injection.

CHAPTER 3: RESULTS

Genetic Analysis of *zep* alleles

All *zep* mutants are homozygous lethal, and map to the distal end of 3R heterochromatin. The genetic map depicted in Figure 3.1 shows the *zep* locus with respect to known deficiencies. Complementation results confirm that all new *zep* alleles fit within the complementation map constructed with the original *zep* allele which positioned zep^{LP13} genetically by its non-complementation with Df(3R)4-75 (Ostrowski, Dierick, Bejsovec. 2002). A summary of this data is shown in Table 3.1; the origin and nature of all *zep* alleles is also included. The original genetic work on *zep* using alleles and P-element lesions LP13, *8740-20*, *8740-22*, *EP-167R, l400-1,* and *l400-8* from work in the Honda lab, is shown for completeness (Fitzpatrick. 2005).

There are exceptions to complete inter-allelic lethality, the five *zep* alleles isolated in the Zuker screen produced an *inter se Splayed(Spl)* phenotype with variable penetrance. *XM3* and *9B-919* deficiencies crossed with *Z1914* and *Z352* also produce the occasional *Spl* phenotype.

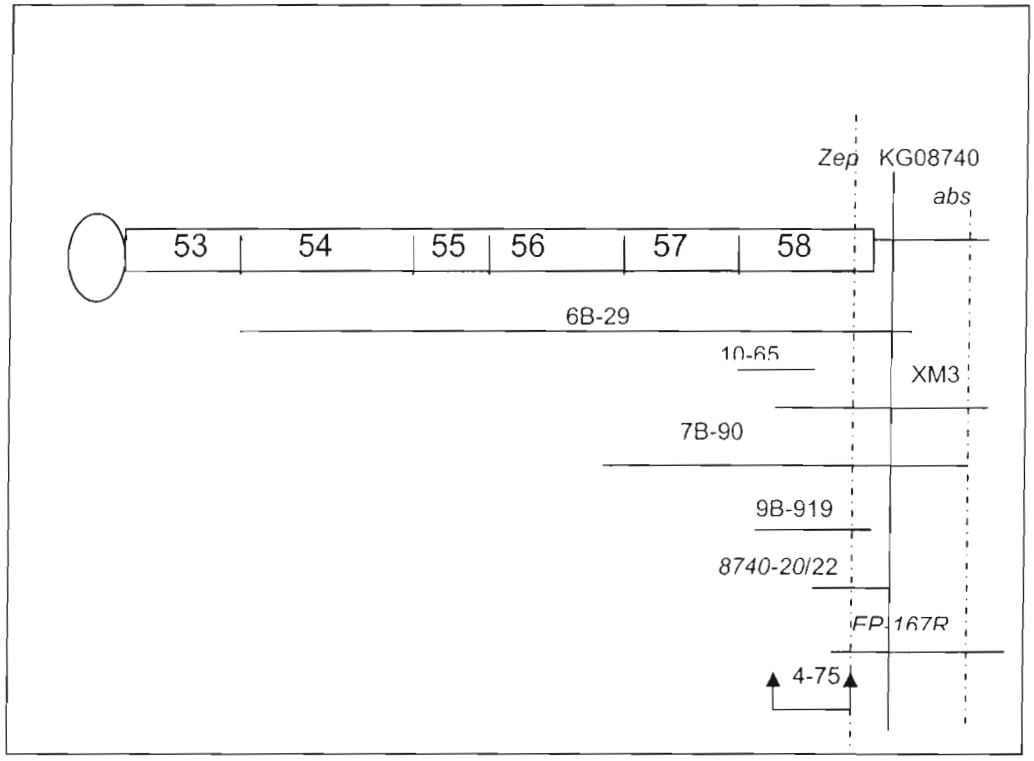
Table 3.1 Complementation of *zep* alleles with known 3R deficiencies.

L=lethal, SL=semi-lethal, V=viable

	6B-29	*7B-90*	*XM3*	*9B-919*	*4-75*	*10-65*	*817*	*Method*	*Origin*
LP13	L	L	L	L	L	V	V	EMS	Ostrowski
Z1904	L	L	L	SL		V		EMS	Zuker
Z1914	L	L	SL	SL		V		EMS	Zuker
Z1014	L	L	L	L		V		EMS	Zuker
Z1608	L	L	L	L		V		EMS	Zuker
Z352	L	L	SL	L		V		EMS	Zuker
I400-1	L		L	L	L	V		EMS	Leptin
I400-8	L		L	L	L	V		EMS	Leptin
EP-167R	L	L	L	L	L	V		P-element	Fitzpatrick
8740-20	L	L		L	L	V		P-element	Fitzpatrick
8740-22	L	L		L	L	V		P-element	Fitzpatrick

Figure 3.1 Deficiency map of 3R heterochromatin with respect to the *zeppelin* locus

The circle represents the centromere and bar represents 3R heterochromatin. Numbers correspond to the cytological bands. Deficiencies are represented by horizontal lines. The positions of the genes *zeppelin (zep)* and *abstrakt (abs)* are shown as vertical lines. KG08740 is the original location for the P-element used to create the *8740-20* and *8740-22* lesions affecting *zep*. Deficiency 4-75 is thought to be an inversion (Marchant and Holm, 1988)



Zuker EMS screen for *zep* alleles

Five new *zep* alleles were isolated through a screen of EMS mutant flies donated by the Zuker lab (Koundakjian, Cowan, Hardy, et al. 2004). Zuker lines were first crossed with the largest heterochromatic deficiency, *Df(3L+3R)6B-29*. This deficiency is well characterised and is known to be missing portions of h49 and h50 as well as all of h52 on 3L, and all of 3R, including most of the last heterochromatic band, h58 (Koryakov, Zhimulev, Dimitri. 2002). The Zuker lines that were lethal with *Df(3L+3R)6B-29* were crossed with progressively narrower deficiencies. Zuker *Z352*, *Z1014*, *Z1608*, *Z1904* and *Z1914* were viable with 3L deficiencies *Df(3L)K-2, Df(3L)FX53, Df(3L)γ28, Df(3L)0-1, Df(3L)1-166* and *Df(3L)Delta AK*^{1}. These Zuker lines were also viable with *Df(3R)10-65* on 3R but lethal or semi-lethal with *Df(3R)XM3*, *Df(3R)7B-90* and *Df(3R)9B-919* (Table 3.1). Final non-complementation with *zep*^{LP13} confirmed that Zuker fly lines *Z352*, *Z1014*, *Z1608*, *Z1904* and *Z1914* were allelic with *zep*. *Df(3R)10-65* is located in 81F according to Flybase and is listed as including the complementation group *l(3)81Fa*, but not *l(3)81Fb* (http://flybase.bio.indiana.edu/).

zep trans-heterozygote interactions

Table 3.2 shows that most *zep inter se* combinations are lethal as would be expected of alleles of a single complementation group. However some crosses involving Zuker EMS alleles produced a percentage of flies with a *Spl* phenotype that was originally described by Lindsley et al. in 1972.

Table 3.2 *Inter se* complementation of *zeppelin* alleles.

	LP13	l400-1	l400-8	EP-167R	8740-20	8740-22	Z1904	Z1914	Z1014	Z1608	Z352
LP13	L	L	L	L	L	L	SL	SL	L	L	L
l400-1		L	L	L	L	L	L	L	L	L	L
l400-8			L	L	L	L	L	L	L	L	L
EP-167R				L	L	L					
8740-20					L	L					
8740-22						L					
Z1904							L	L	SL	SL	SV
Z1914								L	SL	SL	SV
Z1014									L	L	SL
Z1608										L	L
Z352											L

Spl was initially characterised as a dominant phenotype where flies have extended legs and dried black material accumulating at leg joints. It was first seen as a result of haplo-insufficiency produced by a translocation at the euchromatic polytene cytological band 81F (Lindsley et al.,1972). Later work by Tasaka and Suzuki (1973) created two temperature sensitive alleles described with the same phenotype. These alleles have been lost and at the beginning of this study no lesions existed which gave a *Spl* phenotype.

The Zuker *Spl* flies isolated in this study were occasionally able to eclose and survived for up to 24 hours. They were characterised by their inability to walk properly and a black deposit found mostly at the knee joint (melanotic tumor) which seemed to exude from the joint itself. Their legs were outspread and weak, hence the splayed description. In most cases the flies died as they eclosed and were covered with a sticky glue-like substance. All other characteristics appeared normal in these flies (Figure 3.2).

Figure 3.2 Trans-heterozygote *Splayed* phenotype

a. *Z1014/Z352* b. *Z1014/Z352*

c. *Z1914/Z352* d. *Z1914/Z352*

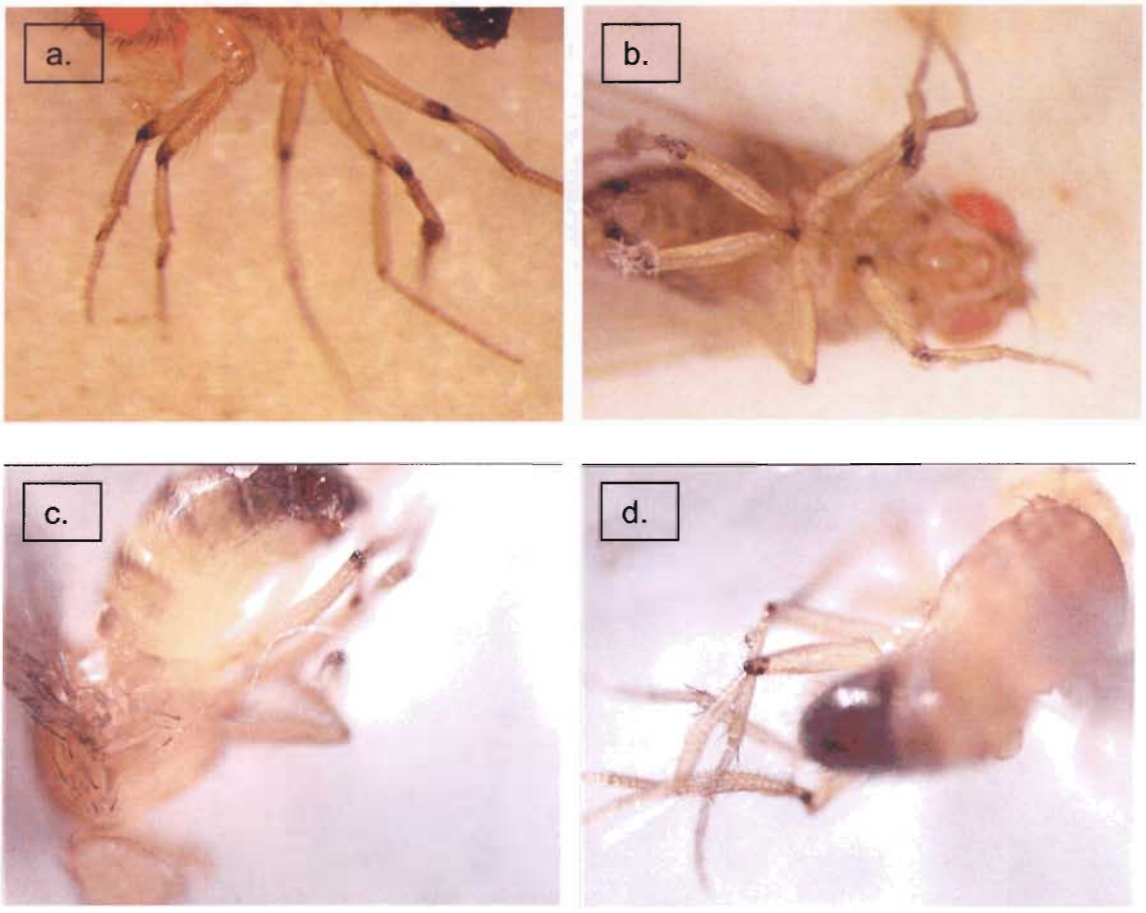


Table 3.3 ***Inter se*** **complementation results between Zuker *zeppelin* alleles.**

Cross	Balancer	*Splayed*	%*Splayed*	Total
Z1914 X *Z352*	140	52	74	192
Z1904 X *Z352*	169	58	69	227
Z1608 X *Z352*	165	0	0	165
Z1014 X *Z352*	149	14	31	163
Z1914 X *Z1904*	166	0	0	166
Z1914 X *Z1608*	201	15	15	216
Z1904 X *Z1608*	127	9	14	136
Z1914 X *Z1014*	198	23	23	221
Z1904 X *Z1014*	382	51	26	433
Z1608 X *Z1014*	165	0	0	165
Z1914 X 9B-919	223	9	8	232
Z1904 X 9B-919	186	6	6	192
Z1914 X LP13	223	29	26	252

Some trans-heterozygote combinations produced more *Spl* flies than others (Table 3.3). This could be because some severe mutant GFAT1 subunits were less complementary when in trans with any other mutant subunit. The greatest number of *Spl* survivors were *Z1904/Z352* and *Z1914/Z352* (Figure 3.2 c-d), at 69% and 74% of the expected wild type class number. *Z1904/Z1014* and *Z1914/Z1014* were 26% and 23% *Spl* of expected class number.

The most severe combination involved any cross that included *Z1608*. *Z1608* with *Z352* and *Z1608* with *Z1014* were completely lethal combinations. When crossed with the weakest alleles *Z1914* and *Z1904*, *Z1608* gave 15% and 14% of expected class number *Spl* respectively. The frequency of the *Spl* phenotype is reduced in *Z1608* reflecting the nature of the mutation; in trans complementation appears to be insufficient for normal Glc6P production and even semi-viability is not possible.

Z352 was another weak *zep* allele and produced *Spl* flies in every combination except with *Z1608*. *Z1014* also showed semi-viability with the other three Zuker *zep* alleles except *Z1608*. *Z1014/Z352* generated 31% of expected class number *Spl* (Figure 3.2 a-b). The *Z1904/Z1914* combination was 100% lethal.

Interestingly, homozygotes for all of these Zuker EMS alleles did not survive and no *Spl* phenotypes were seen in any stocks. This could be because second site mutations exist in these EMS lines or because the subunits are unable to complement as dimers when the identical mutation is present. However some stocks did show a small melanotic tumor 'dot' on the junction of legs and abdomen on flies heterozygous with the TM3 balancer chromosome, this was seen in stocks *Z1914*, *Z1904*, and *Z352*.

Deficiencies *XM3* and 9B-919 also produced a small number of *Spl* flies in combination with Zuker alleles *Z352*, *Z1914* and *Z1904*. *Z352* with *XM3* showed the *Spl* phenotype in 1% of the expected class size. *Z1914* with *XM3* gave 6%

Spl. *Z1914* with 9B-919 gave 8% *Spl* and *Z1904* with 9B-919 gave 6% *Spl*. The original EMS allele, zep^{LP13} also produced 26% of the expected class number *Spl* when in combination with *Z1914*.

Blimp phenotype analysis of *zeppelin* alleles

The original zep^{LP13} allele had shown remarkable cuticle expansion as an embryo compared to wild type, therefore the new zep alleles were also examined for this defective phenotype. Cuticle preparations were prepared by the Bejsovec lab following the protocol outlined in Ostrowski et al. (2002). According to these authors, cuticle preparations for the same genotype can vary due to differences in the mechanical devitellinization process. However when viewing a number of preparations of the same allele, a typical preparation can be selected as representative. The variation in the degree of stretchiness of the cuticle layer between several *zep* alleles is considerable and therefore it is possible to distinguish three distinct groups.

Some embryos are close to wild type in length, but have an expanded cuticle, and some are much longer than wild type and also show greater expansion. Since 8740-20 and 8740-22 have been shown to lack the *Gfat1* coding sequence completely, these mutants represent a null phenotype. The head skeleton and denticle markings are close to wild type in all genotypes. Three blimp phenotypes classes were determined based on the level of stretchiness. Class 1: six *zep* alleles, *l400-8*, *Z1014, Z1608, 167-R, 8740-22* and *8740-20* were extremely elastic and often had constrictions around anterior and

posterior segments. They looked similar to the original *zep*^{LP13} phenotype. Class 3: *Z1904, Z1914* and *Z352* had a much milder stretchy phenotype, with less expansion in length and almost wild type appearance. Class 2: *I400-1* exhibited an intermediate stretchy phenotype. Figures 3.3–3.6 show representative phenotypes for each allele.

Figure 3.3 Class 1 *zeppelin* severe blimp phenotype - *8740-22, LP13, 8740-20, Z1608, I400-8, EP167-R.*

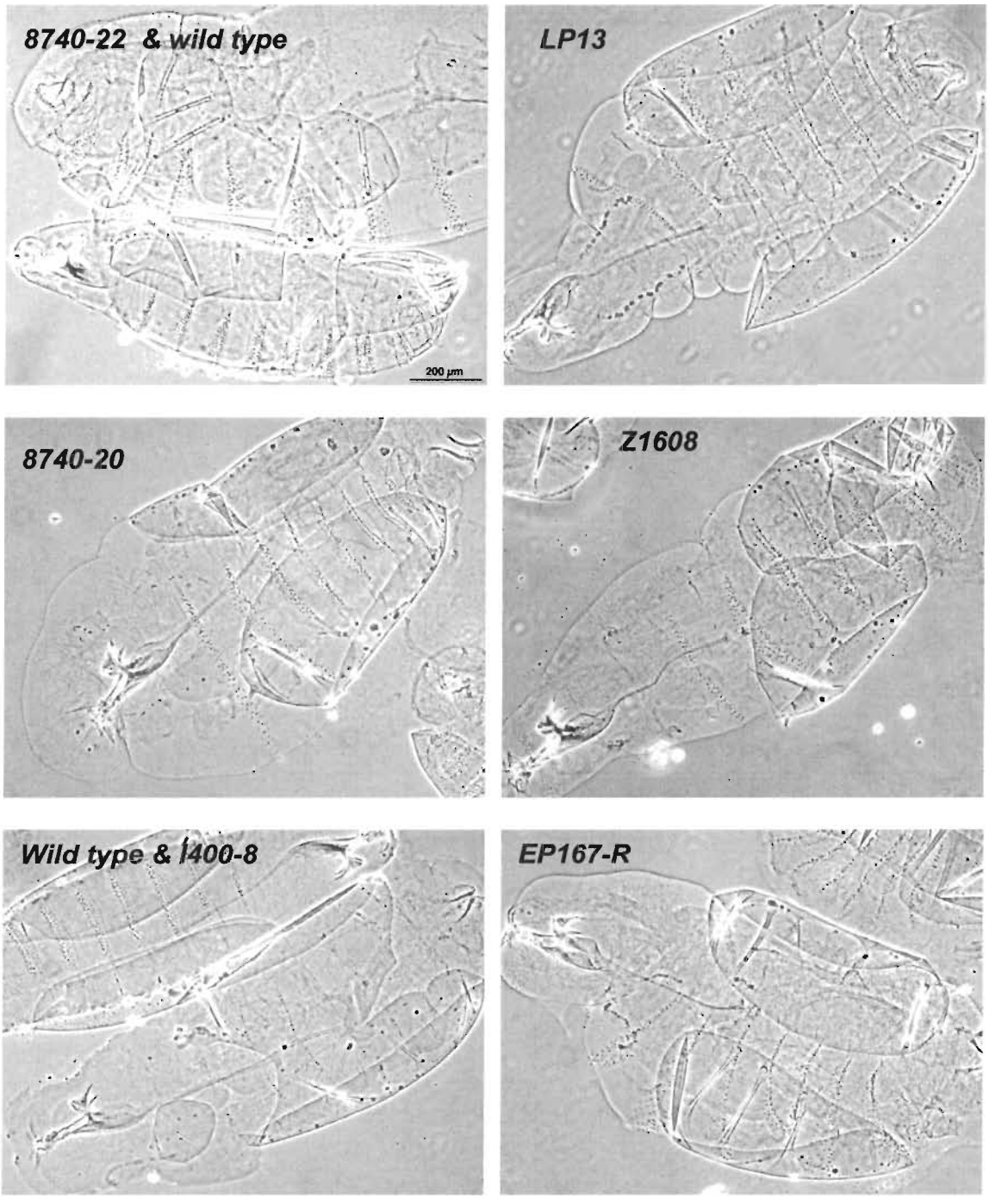


Figure 3.4 Class 1 *zeppelin* severe blimp phenotype - *Z1014, 8740-22*

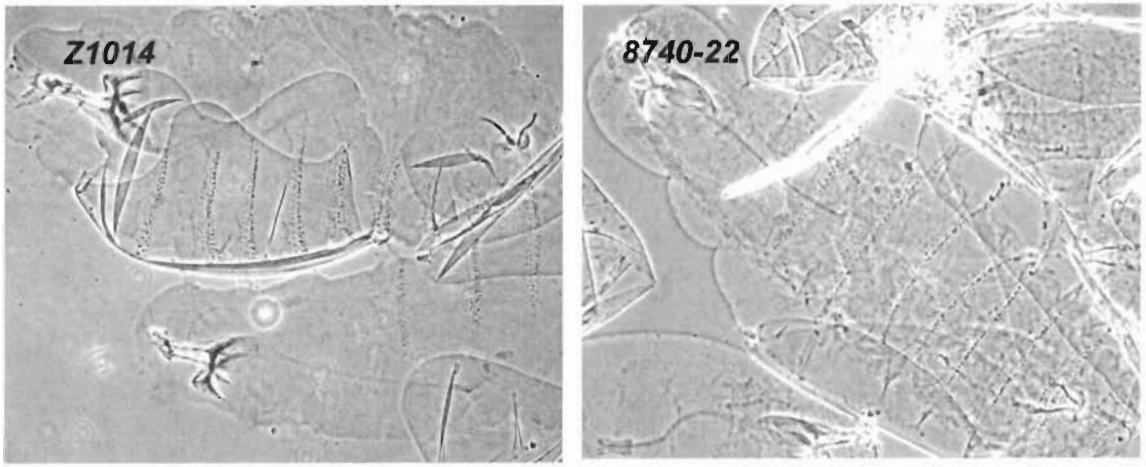


Figure 3.5 Class 2 *zeppelin* embryo showing less expansion - *I400-1.*

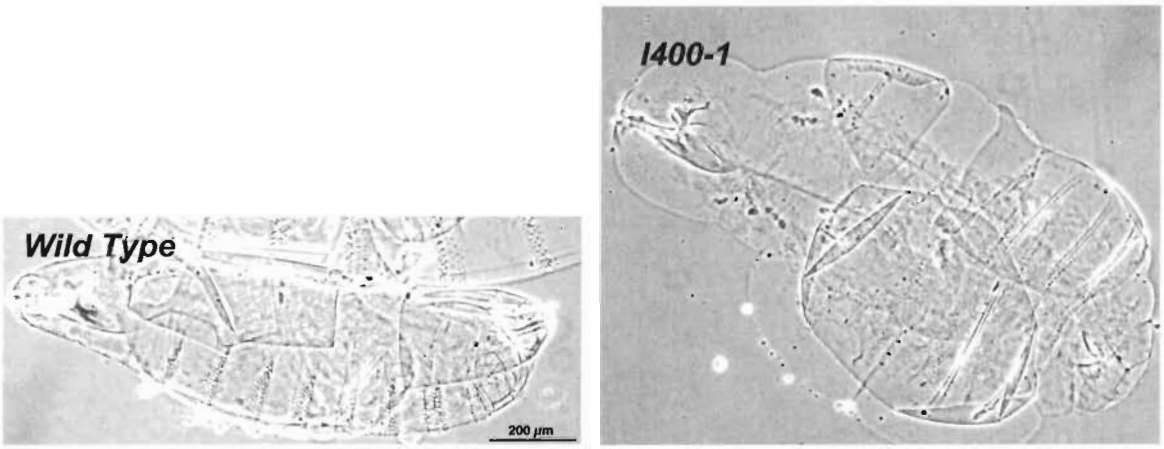
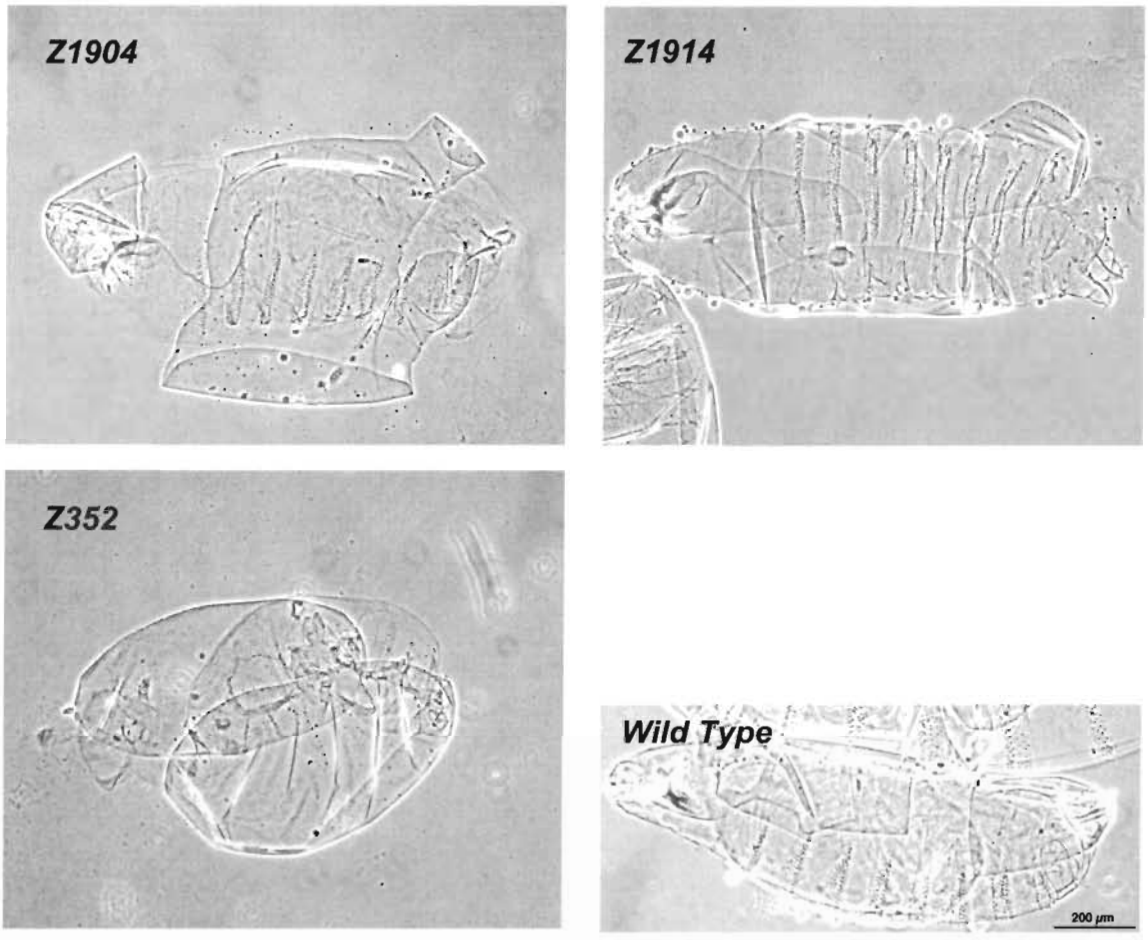


Figure 3.6 Class 3 *zeppelin* embryo showing very little expansion – *Z1904, Z1914, Z352.*



PCR mapping of *Gfat1* exons in P-element deletion *zeppelin* alleles

As a first step towards showing correspondence between *zep* and *Gfat1*, I attempted to establish that *Gfat1* coding sequence was deleted in the P-element transposition *zep* alleles. These deficiencies were produced by a male recombination event of the original P-element when exposed to transposase. The re-insertion of the transposon and consequential loss of the intervening DNA

was therefore expected to result in the partial deletion or complete absence of *Gfat1* in the *zep* alleles *8740-20* and *8740-22*.

PCR experiments used DNA extracted from single embryos as explained in materials and methods. Using *Gfat1* specific primers for exon 1, Gfat1F and Gfat1B (see Appendix A, Table A1 for sequence) and exon 10, Gfat14F and Gfat24R, it was determined that the sequence at both ends of *Gfat1* was absent in p-element deletion *8740-22*. In figure 3.7 the same individual template DNA was used for each successive PCR test e.g. lane 1 uses the same embryo template for exon 1, exon 10, GFP and Grp84 PCR experiments. Since GFP is present only on the balancer, it was used to determine the homozygosity of the embryos. Grp84 shows that template quality for each embryo was good.

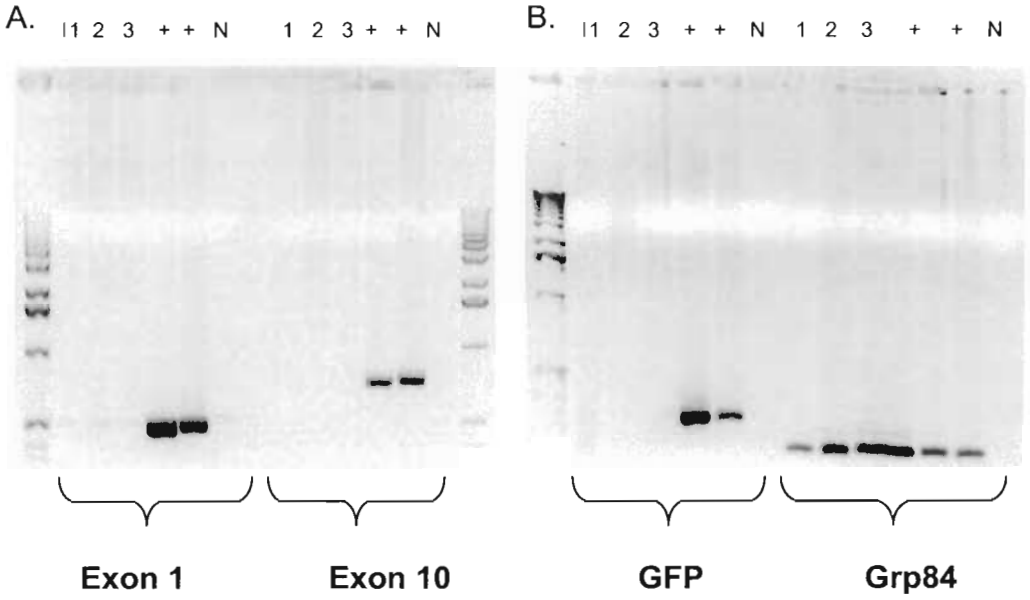
Figure 3.7 Exon specific PCR of *Gfat1* in p-element deletion allele *8740-22*.

Exon 1: Lanes 1, 2, 3, show 3 different homozygote embryos.
Lanes +, +, heterozygote controls.
N Negative 1F1B control.

Exon 10: the same template DNA as lanes 1-5 above.
N Negative 14F 24R control.

GFP: the same template DNA as lanes 1-5 above.
N Negative GFP control.

Grp84: the same template DNA as lanes 1-5 above
N Negative Grp84 control.



Using primers Gfat15 and Gfat6B designed for exons 3-7 it was confirmed that the *Gfat1* sequence was completely absent in *8740-20* and *8740-22* (Figure 3.8).

Figure 3.8 Exons 3-7 are deleted in homozygote embryos *8740-20* and *8740-22*.

156B – exons 3-7, **Grp84** – template control, **GFP** – homozygote control ***8740-20*** lane 1, **8740-22** lane 2, **heterozygous control** +, **no template** N

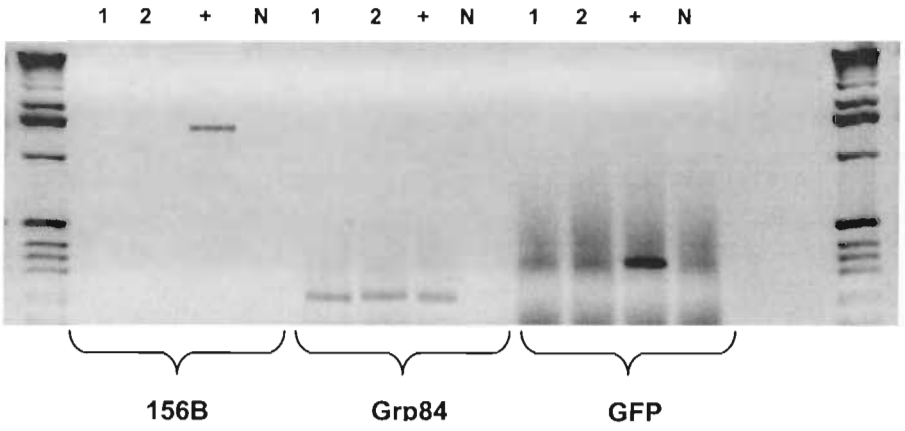


Table 3.4 is a summary of mapped exon deletions using various primer sets to confirm the absence of *Gfat1* in *8740-20* and *8740-22* by PCR. Since these mutations are a result of male recombination that has removed the whole of *Gfat1*, it is expected that the phenotypes represent a complete loss of function of *Gfat1* in these *zep* alleles.

Table 3.4 Absence of *Gfat1* in P-element deletion *zeppelin* alleles detected by PCR.

Genotype	*Primers*	*Exons*
8740-20	1f 1B	1
8740-20	2F 13	3, 4, 5
8740-20	15 6B	3 - 7
8740-20	10F 6B	6, 7
8740-20	14F24R	9, 10
8740-22	1F 1B	1
8740-22	15 6B	3 - 7
8740-22	14F 24R	9, 10

See appendix A for primer sequences

PCR mapping using *Gfat1* specific primers in 3R deficiencies

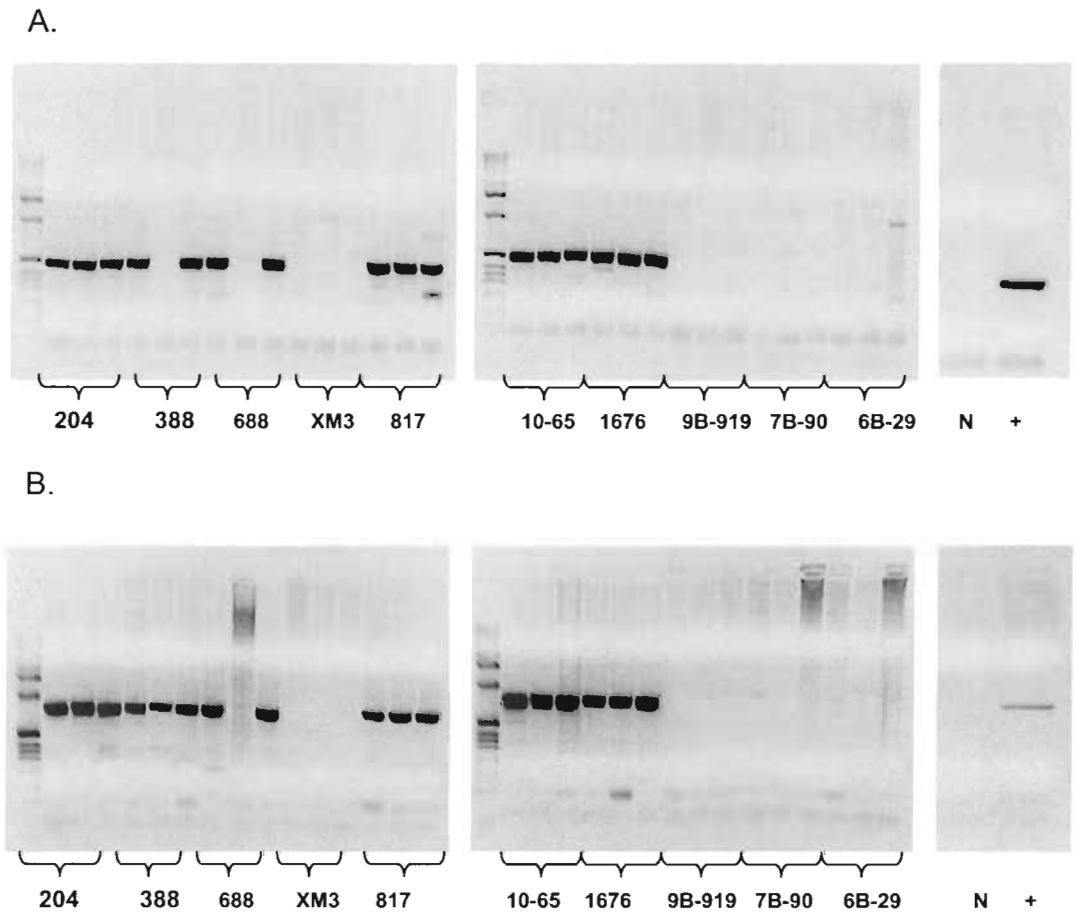
Using *Gfat1* specific primers for exon 1, Gfat1F and Gfat1B and exon 10, Gfat14F and Gfat24R (see Appendix B for oligonucleotide sequence), various deficiencies that were established as non-complementary with *zep* by genetics were tested for the presence or absence of the *Gfat1* sequence (Figure 3.9).

Gfat1 was found absent in 3R deficiencies *XM3*, *9B-919* and *7B-90* and absent in the deficiency *6B-29* covering 3L and 3R. It was present in the X-ray deficiency on 3R 10-65, and present in smaller deficiencies located within the 10-65 lesion: *204, 388, 688, 817* and *1676*.

Figure 3.9 *Gfat1* Deficiency Map by PCR.

shows 3 homozygous embryos for each genotype

A. 5' end *Gfat1* Exon 1 PCR
B 3' end *Gfat1* Exon 10 PCR



Sequencing of EMS *zep* alleles

Eight of the eleven *zep* alleles were isolated by EMS. In order to further establish the correspondence of *Gfat1* with *zep*, the *Gfat1* gene was sequenced in each of the mutants with the expectation that point mutations would be present in this gene in each *zep* allele. Table 3.5 presents a summary of the mutations found and Figure 3.10 highlights these mutations with respect to the GFAT1 functional domains. It is interesting to note that no mutations were found in the glutaminase domain

Table 3.5 Summary of point mutations found in *zeppelin* EMS mutants.

scaffold position	exon	nucleotide change	amino acid position	amino acid change	genotype
11273	9	t > a	348	Tyr > Stop	*LP13*
11511	9	g > a	445	Ala > Thr	*Z1904, Z1914*
11874	9	c > t	535	Glu > Stop	*I400-8*
12092	10	t > a	588	Leu > Met	*Z352*
12297	10	g > a	656	Cys > Tyr	*Z1608*
	Polymorphism sites				
8617	5	a > g		His > Arg	*)*
8660	5	a > c		noncoding	*) Z446, Z1904, Z1914*
8996	5	t > g		noncoding	*) Z1608, Z1014*

The original *zep* mutant *zep*^{LP13}, a Class 1 embryo, had a point mutation of thymine to adenine at scaffold position 11273, this results in a codon change that translates to a STOP codon in place of tyrosine at amino acid position 348. The truncation occurs at the start of the isomerase domain. Since nonsense transcripts can be degraded through nonsense-mediated mRNA decay, the possibility of a truncated, partially functional GFAT1 protein is low (Alberts, Johnson, Lewis, et al. 2002).

The *Z1904* and *Z1914* Class 3 Zuker *zep* alleles were both found to have the same point mutation at scaffold position 11511 in exon 9, where guanine was changed to adenine. This results in an amino acid change of alanine to threonine. This amino acid is conserved in *D. melanogaster Gfat1* and *Gfat2*, Yeast and Human *Gfat2*. It is not unexpected that this amino acid replacement would cause significant disruption of the GFAT1 protein fold due to the replacement of the nonpolar side chain of alanine with the larger and polar hydroxyl group of threonine. Erroneous hydrogen bonding with neighbouring main and side chain groups may be introduced. The change is located at amino acid position 445 that is within 6 residues of the P-loop. The P-loop functions to stabilize the phosphate group of fructose-6-phosphate within the isomerase domain.

Since it is unusual to find the same point mutation in different fly lines, both experimental stocks were replaced by the original lab stock received from the Zuker lab and rebalanced over TM3 with the GFP marker and resequenced. The results remained the same and therefore it could be concluded that the EMS affected chromosome in the male underwent replication in the first stage of spermatogenesis to produce duplicate chromosomes carrying the same point mutation. Two separate fly lines in the original Zuker lab screen, i.e. Class 3 *Z1904* and *Z1914* could then have been isolated from separate fertilization events. However, it is more likely that the stocks Z1904 and Z1914 were mislabelled in the Zuker lab as the stock numbers are very similar. Other Zuker

zep alleles were screened for this change and it was confirmed to be absent in *Z1608*, *Z352* and the control Z446 fly lines.

Class 3 allele *Z352* was found to have a change of thymine to adenine at scaffold position 12092 in exon 10. This results in an amino acid change of a conserved leucine to methionine at amino acid position 588. The leucine is one residue away from the histidine loop in the opposing dimeric subunit. The histidine loop opens the fructose sugar ring and is the only structure on the second subunit that is involved directly in catalyzing the isomerase reaction apart from contributing to a stable quarternary structure necessary for the functional enzyme. It is possible that the change has altered the efficiency of the GFAT1 enzyme such that the rate of formation of UDP-NacGluc is close to normal levels of chitin synthesis for embryo cuticle formation causing a nearly wild type phenotype.

Class 1 mutant *Z1608* has a point mutation of guanine to adenine at scaffold position 12297 exon 10. This results in an amino acid substitution of a conserved cysteine to tyrosine at amino acid position 656. The fact that cysteine is conserved suggests that it may have an important role in maintenance of the tertiary structure through its participation in di-sulfide bridges.

I400-8 is an EMS *zep* allele from the Leptin lab. A cytosine to thymine change was found at scaffold position 11874. This translates to a stop codon at amino acid position 535 and another loss of function *zep* allele.

No sequence errors were found in the last Zuker *zep* allele, *Z1014* or in *I400-1* from the Leptin lab. All exon sequences in *Gfat1* as well as intron sequences at splice sites were sequenced in these flies. It is possible that an extraneous splice site may have been created by a point mutation within intron intervals, or that upstream regulatory sequence has been disrupted.

Three polymorphisms exist in the Zuker alleles within the *Gfat1* gene exon 5 sequence, when matched with flybase sequence. At scaffold position 8617, there is a consistent change of adenine to guanine that translates to a change in the peptide sequence from histidine to arginine. The polymorphisms at 8660 and 8996 reside in intron sequence. To determine that these three changes were consistent throughout the Zuker lab flies, an unrelated Zuker stock Z446, genetically complementary with *zep* was used to sequence this region to establish positive control sequence.

Figure 3.10 *Gfat1* EMS point mutations in relation to GFAT1 functional domains.

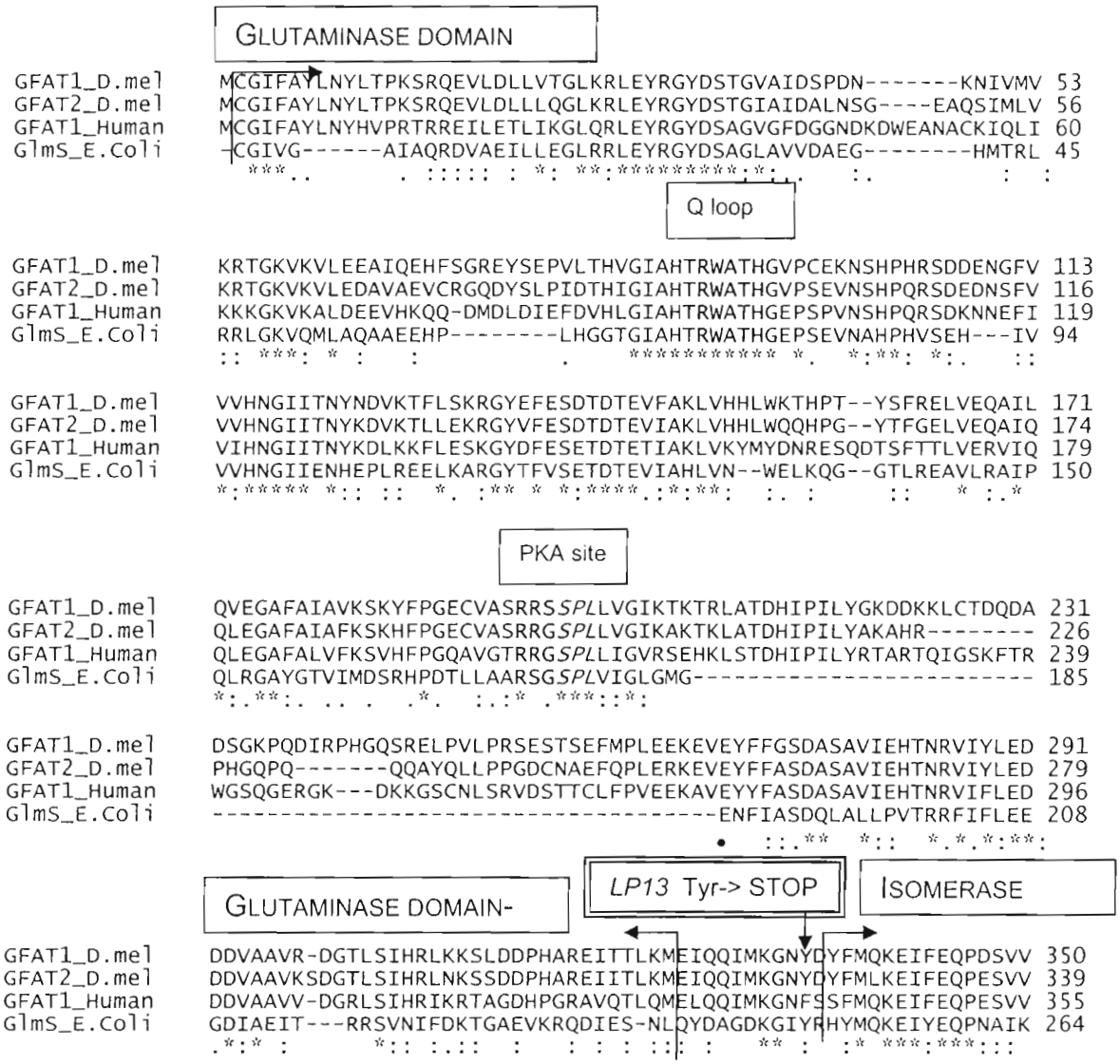
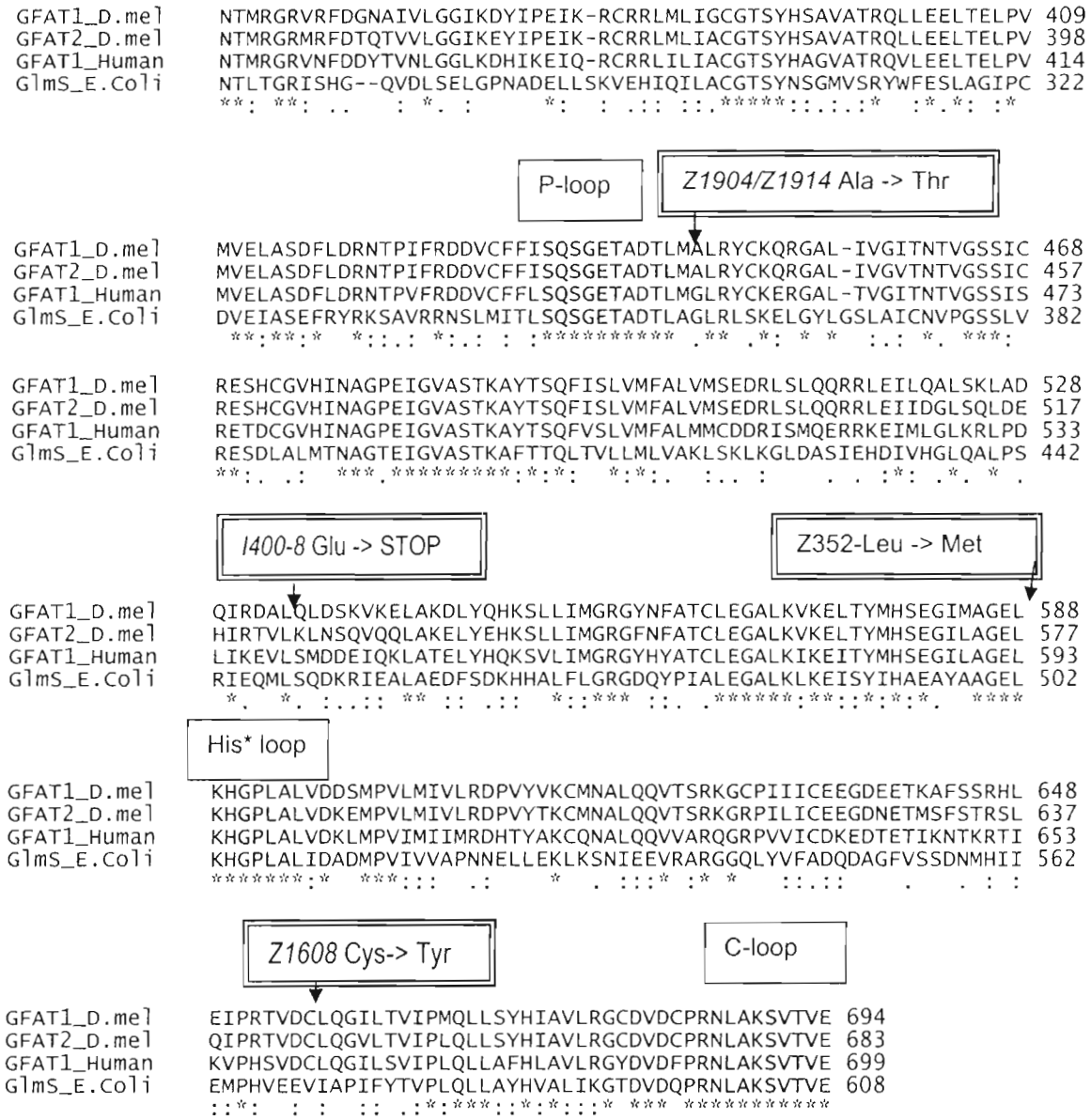


Figure 3.10 continued.



Gfat1 disruption by RNA interference (RNAi)

According to whole-mount *in situ* hybridization by Graack et al. (2001), *Gfat1* expression begins at embryo stage16 in the tracheae and has extended to the whole epidermis by stage 17. It is again expressed in late third instar larvae in the corpus cells of the salivary gland to provide NacGlc for the glycosylation of the Sgs proteins and during pupation to form the hard cuticle layer.

The highly glycosylated Sgs proteins are needed for the adhesion of the pupa case in the last stage of development. This is confirmed by the observation that trans-hetrozygote (*Spl*) cross experiments usually had many pupa cases accumulate at the bottom of the vials rather than stuck to the sides.

To further support the correlation of *zep* with *Gfat1* a 290 bp cDNA fragment was inserted into pSympUast RNA interference (RNAi) expression vector to create knockdown of *Gfat1* expression. Since the genetically isolated *zep* alleles showed *Gfat1* to be an essential gene, lethality was expected with the RNAi transgene expression. However, since the *Spl* non-lethal *Gfat1* expression phenotype had also been observed, it was possible that this phenotype could also be replicated. Transcription of the Gfat1 fragment was induced by crossing the transgenic lines with diverse drivers with varied spatial and temporal expression.

Ten fly lines with a stable insertion of the *Gfat1* RNAi expression construct were generated (Table 3.6) by Bestgene (see Materials and Methods). Five were inserted on the second chromosome and five on the third chromosome.

They were then balanced over TM2 with the *curly wing* (*CyO*) marker, and TM3 with the *serrate (Ser)* marker respectively.

Table 3.6 Transgenic RNAi construct insertion - initial expression data. Chromosome 2 insertions crossed with Gal4-Actin/Tm2 Cyo. Chromosome 3 insertions crossed with Gal4-Tubulin/Tm3 Ser.

Transgenic insertion	*Balancers*	*Wild Type*
Chromosome 2		
2664-1-1M-CH2	Not done	
2664-1-2M-CH2	100	0
2664-1-5M-CH2	55	0
2664-1-9M-CH2	42	0
2664-1-10M-CH2	11	0
Chromosome 3		
2664-1-3M-CH3	17	0
2664-1-4M-CH3	27	0
2664-1-6M-CH3	58	0
2664-1-7M-CH3	47	0
2664-1-8M-CH3	40	0

Initial crosses between nine lines with the ubiquitous gal4 drivers Actin-4414-Gal4/TM2 (http://flybase.bio.indiana.edu/) and tubP-Gal4/TM3 (hereafter referred to as actin and tubulin drivers) gave complete lethality (Table 3.6). Since the RNAi sequence was unique and specific to *Gfat1* these results indicated that this is an essential gene. Complete lethality with the ubiquitous expression of RNA interference does not necessarily link the *Gfat1* gene with the *zep* locus isolated by genetics. However, the possibility that secondary genes were unintentionally disrupted by the insertion of the RNAi construct can be eliminated since all nine crosses were 100% lethal.

To gain further support for the correspondence of *Gfat1* with *zep*, various drivers were selected to drive the RNAi construct that may replicate the trans-

heterozygote *Spl* phenotype and melanotic tumors found with *zep inter se* crosses. The drivers would direct partial suppression of *Gfat1* expression in more defined regions and periods of development.

Table 3.7 Transgenic RNAi construct insertion – Heat shock gal-4 driver

Cross	*Temp (0C)*	*Balancer*	*Wild Type*	*Spl*
2m-ch2/cyo x hs-gal4/cyo	29	205	125	0
2m-ch2/cyo x hs-gal4/cyo	25/29	231	86	0
5m-ch2/cyo x hs-gal4/cyo	18/25/29	57	30	4

In combination with a ubiquitous heat shock driver, small melanotic tumors were generated in 4 out of 91 flies. The flies were first kept at 18°C for 17 days, then moved to 25 °C for another 5 days and only moved to 29 °C, which is the optimum temperature for Gal4 for the last days of development. It was anticipated that the 18 °C period would allow a slower rate of development for embryos in the first stages and the 25 °C period would provide less than adequate GFAT1 protein levels but not complete knockdown. However as Table 3.7 shows, the RNAi transgene with the heat shock driver had inconsistent effects. Even when at the optimum temperature of 29 °C from day 1, wild type flies would still eclose. Four flies from this experiment survived and showed the *Spl* phenotype.

Figure 3.11 Melanotic tumors present in RNAi/gal-4 heat shock driver flies.

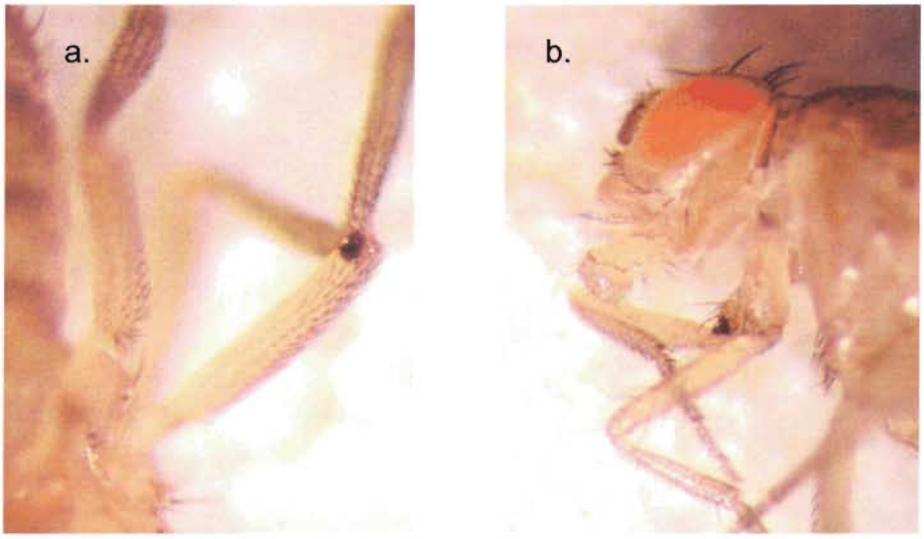


Figure 3.11 shows tumors on two of the heat shock flies, they are small black patches found at the knee and elbow joints which cannot be distinguished from the *Splayed* phenotype found in the Zuker inter se crosses. These flies eclosed 11 days after being moved to 29 degrees. For comparison Figure 3.2 shows the melanotic tumors present in the Zuker *zep* trans-heterozygote flies.

Table 3.8 Transgenic RNAi construct insertion – Various gal-4 drivers. See Appendix A for full genotype description.

Cross	Driver Description	Bal	wt	Comments
1878/CyO x 1M-CH2/CyO	Ubiquitous in third instar imaginal discs	130	0	0 *Spl*
1967/CyO x 5M-CH2/CyO	Embryonic salivary glands, leg & wing imaginal discs	49	76	3 *Spl* alive
3732/CyO x 10M-CH2/CyO	Salivary gland, segmentally repeated pattern at stage 15, third instar: crescent in leg disc	79	101	0 *Spl*
Prd/TM3 x 8M/TM3	Paired protein segmentally repeated pattern in 14 stripes at gastrulation	86	36	0 *Spl*
8142/TM3 x 8M-CH3/TM3	A leg disc ring and portions of antennae, wing and haltere discs	74	42	0 *Spl*

Of the five tissue specific drivers used for *Gfat1* RNAi construct expression, two gave positive results. 1878 has ubiquitous expression in third instar imaginal discs and gave complete lethality in combination with construct 1M-CH2. 1967 is expressed in the salivary gland, in a segmentally repeated pattern at stage 15, and also in a crescent in the leg disc at third instar. The 5M-CH2 construct produced three flies with a *Spl* phenotype out of the class of 79 with this Gal4 driver (Table 3.8). The 1967/5M-CH2 flies had tumors that are more pronounced than those produced in trans-heterozygotes or in flies expressing the RNAi construct with the heat shock driver. Figure 3.12 shows black deposits that extend from joints in one fly and completely envelope the lower leg in another.

Figure 3.12 *Spl* flies result from 1967/5M partial leg disc expression.

The RNAi experiments did not produce a great number of *Spl* phenotype flies. However these rare events show that with reduced rather than complete suppression of *Gfat1* expression at the critical time and location, melanotic tumors indistinguishable from those of *zeppelin* transheterozygote mutants can occur.

Embryo death analysis of RNAi mutants

The RNAi experiments above show that *Gfat1* is essential and that the amount present at the critical point in development in the leg imaginal discs is important. Since *Gfat1* is normally expressed at stages 16 to 17 in the forming cuticle layer (Graack, Cinque, Kress. 2001), it is expected that the RNAi *Gfat1* construct with ubiquitous actin and tubulin drivers would show lethality that would be visible at the embryo stage.

Table 3.9 Analysis of embryonic death RNAi embryos using the ubiquitous driver Actin.

Cross	# of embryos	# failed to hatch	Percentage failed to hatch
Actin/CyO x 2M-CH2/2M-CH2	50	10	20%
Actin/CyO x 2M-CH2/2M-CH2	30	7	23%
Ore-R control	30	4	13%

Initial analysis supports this with the observation that 20% (sample size 50) and 23% (sample size 30) of Actin-Gal4/5M-CyO embryos in two separate plates failed to hatch to first instar compared to 13% (sample size 30) of Ore-R

embryos (Table 3.9). Insufficient data were collected to make a comparison of embryo phenotype compared to zep mutants.

Gfat1 cDNA constructs for rescue of *zeppelin* phenotypes.

Rescue of the blimp and lethal phenotypes by expression of a transgenic *Gfat1* sequence would provide ultimate confirmation of the correspondence of *Gfat1* with the *zep* gene.

GH03520 cDNA was obtained from DGRC and sequenced for insertion into the expression vector pUAST cDNA expression vector. GH03520 was found to have intron sequence present throughout and was therefore abandoned as a prospective rescue transcript. LP07309 was the next cDNA to be examined. It was found to be missing exon 6 producing a transcript that would terminate prematurely just before the protein kinase A (PKA) binding site and lack the isomerase domain.

RE72989 cDNA was then sequenced and found to be correct until exon 10, where approximately 1000 bp of non-*Gfat1* coding sequence was encountered. A BLASTN search showed that this corresponded to a transposable element insertion, TN10.

To create a cDNA sequence that would provide inframe translation to a fully functional protein, the first part of RE72989 and the last part of LP07309 were spliced at an EcoRI site to align with the original RE72989 sequence as specified on the DGRC database and used in the Gfat homologue comparison by

Graack in 2001. The complete RE72989 cDNA was inserted in pBluescript then sequenced in pUAST and sent to Bestgene for injection. The injection experiment did not produce any adults with a stable insertion of the construct, detected by the red eye marker on the pUast construct. The rescue insert with pUast has now been reconstucted using a new pUast vector and is ready for a second injection.

CHAPTER 4: DISCUSSION

The correspondence of the *zep* locus isolated by genetics (Ostrowski, Dierick, Bejsovec. 2002) with the candidate gene *Gfat1* (CG12449) has been established in this work using genetic, molecular and phenotype analyses.

Four new *zeppelin* alleles have been isolated from a screen of EMS mutant lines obtained from the Zuker lab. *Z1914* and *Z1904* were found to be indistinguishable by PCR, genetics and phenotype analysis and can therefore be classified as a single allele.

Together with alleles and P-element deletions that were previously known there are now ten identified *zep* lesions that are all homozygous lethal. Knockdown of the gene by RNA interference has further supported *Gfat1/zep* correspondence by establishing *Gfat1* is an essential gene and by the generation of transgenic 'phenocopies' of the *Spl* phenotype seen with *zep inter se* crosses.

The correspondence of *zep* with *Gfat1* was initially supported by the absence of *Gfat1* coding sequence in embryos homozygous for *zep* P-element deletions and non-complementary 3R deficiencies, shown by PCR. It has also been shown by the sequencing of point mutations in *Gfat1* coding sequence in homozygous single embryo *zep* mutants that have been generated by EMS. The degree of severity of the phenotypes observed in embryos and adult stages has

been related to the base change or absence of the *Gfat1* sequence, and likely effects of these mutations on enzyme activity have been proposed. RNAi *Gfat1* transgene expression has also given insight into the correlation of the *zep* phenotype with *Gfat1* enzyme function.

Splayed (Spl) phenotype

Inter se complementation has revealed an interesting *Spl* phenotype (Table 3.3). *Spl* flies develop to the pharate adult stage and sometimes eclose. However they do not survive for long and are characterised by the presence of melanotic deposits at leg joints. This delay in lethal phase could be explained by some *zep* EMS mutants complementing each other *in trans* in the GFAT1 homodimeric protein. Another explanation could be that *Gfat2* expression is partially compensating for compromised *Gfat1* expression.

Spl was originally described as a dominant phenotype where the flies had awkward motion and outspread legs, with melanized patches at the joints. Lindsley et al. (1972) used an haploinsufficiency experiment which translocated elements of the Y-chromosome to displace segments of autosomes resulting in zygotes that either received one copy or three copies of the affected region. *Spl* resulted from a dominant haplo-insufficiency phenotype encompassing bands 81F-82A. Exact breakpoint positions were not established because of the tendency of heterochromatin regions to aggregate around the centromere in polytene chromosomes.

Another *Spl* mutant, *1(3)9* ^{hts} was isolated by EMS temperature sensitive experiments examining the third chromosome (Tasaka, Suzuki. 1973). It mapped to the same region as that uncovered by the deletion of 81F-82A by Lindsley et al. (1972). In this case the phenotype was characterised as recessive. It had a permissive temperature of 22^{0}C. Of interest is that the *Spl* phenotype mutant is lost at 29^{0}C and death occurs in the mid to late pupae stages, but at 25^{0}C the lethal phase appeared later and flies were able to eclose with the characteristic melanotic tumors and splayed leg appearance.

Marchant and Holm (1988) later identified a corresponding gene, *l(3)81Fb* through an EMS mutation screen of 3R heterochromatin. They noted that it was an essential gene but did not record any survivor phenotype. Given the scarcity of complementation groups in this region (Marchant, Holm. 1988), (Fitzpatrick. 2005), it is suggested that the *zep/Gfat1* locus corresponds to the *l(3)81Fb* locus already described for which none of the original ten alleles now exist.

The *Spl* phenotype amongst *zep* transheterozygotes had variable severity and penetrance depending on the alleles present. Adults are wild type in appearance apart from melanotic tumors found at leg joints, and some are able to walk, though not very well, most die as they eclose suggesting that they lack the strength to push out of the pupa case.

RNAi replicates *Spl* phenotype

The three *Spl* flies resulting from the RNAi transgene driver combination 5M/1967 had very severe tumors but had enough strength to eclose. The rest of

the class were completely wild type which suggests that at the critical time and location for chitin manufacture the *Gfat1* enzyme was mostly not affected by the RNAi construct. This could be explained if the driver expression was timed too early or too late with respect to chitin formation in the imaginal discs, or if the driver was not activated in the right tissue location, i.e. epithelial cells. The replication of the *Spl* phenotype using RNAi suggests that it is the quantity of GlcN6P produced by the enzyme rather than a defective product that causes the *Spl* phenotype.

The heat shock driver in the RNAi experiments did not give the expected result of lethality at 29 ^{0}C. In one of the experiments 4 *Spl* flies resulted from a cross which was started at 18^{0}C then moved to 25^{0}C and finally to 29^{0}C, the rest of the class were wild type. In another cross started at 25^{0}C and moved to 29^{0}C , the wild type class was reduced to 74% of expected levels. The results were inconsistent but the few resulting *Spl* flies and partial lethality suggests that knockdown was occasionally sufficient to reduce GlcN6P synthesis and compromise the chitin synthesis pathway.

Correspondence of phenotypes with the molecular nature of mutations.

The *zep* mutant homozygote embryos were found to stretch to as much as three times the wild type size. Three classes describing the severity of this blimp phenotype have been defined, based on the degree of stretchiness of the dechorionated and devitellinized embryo. The classification of *zep* mutants into the three blimp classes also corresponded with the expected effect of error found

in the sequence of each allele. Class 1 was the most severe and included *zep*^{LP13}, the P-element deletions *8740-20*, *8740-22*, *EP-167R*, *Z1608*, *Z1014* and *l400-8*. Class 2 had a moderate amount of stretchiness and only included one allele, *l400-1*. Class 3 had very little stretchiness and included *Z352* and *Z1904/Z1914*.

Genetic *inter se* experiments have also shown that the degree of penetrance of the *Spl* phenotype varies depending on the allellic combination (Table 3.3). This can be explained by the necessary formation of GFAT1 as a homodimer for activity.

Z1608 was classified as having a severe class 1 blimp embryo phenotype, and only produced *Spl* progeny when in combination with the weakest EMS allele, *Z1914/Z1904* (~15%). *Z1608* had a base change that substitutes tyrosine for cysteine. This residue has been found to be conserved in GFAT1 and GFAT2 in many species apart from *E. coli* GlmS (Kato, Dasgupta, Smartt, et al. 2002). From these results it could be concluded that the cysteine residue contributes significantly to enzyme integrity. The 1914/1904 mutation (ala->thre) appears to partially complement the *Z1608* defective subunit to allow some embryos (~15%) to make enough GlcN6P for sufficient chitin production. Consequently they are strong enough to emerge from the egg case and go on to larva stages.

Conversely, *l400-8*, which was also classified as having a severe blimp phenotype, was completely lethal in combination with all *zep* alleles including the weakest, *Z1914/Z1904* and *Z352*. The EMS treatment produced a stop codon in

the isomerase domain that would eliminate this essential function even if the protein could be folded and not degraded.

One of the least severe alleles, *Z352* classified as blimp class 3, was close to wild type in embryo length and had very little cuticle expansion. *Spl* progeny were ~70% of expected wild type class levels in combination with the second weak allele, *Z1904/Z1914*. *Z352* had a mutation of leu to met at position 588, one residue away from the His loop that functions to open the Fru6P sugar ring. Although this is expected to be an important role, it is the only function performed by the opposite subunit in the homodimer and therefore it could be partially complemented on a molecular level *in trans* with an alternative mutant subunit. This is supported by the mutation acting as homozygous lethal, indicating that the change is significant enough to not be self-complementary.

The above describes how some *inter se* combinations may allow partial complementation and *Spl* flies are sometimes close to expected wild type class levels. It may be concluded from these results that GlcN6P is sometimes present in sufficient quantity for the chitin synthesis pathway to deposit embryo cuticle with enough strength for embryo survival. However, chitin levels are insufficient for the normal transition from pupa to adult and thus melanotic tumors result.

Trans acting subunits occasionally produce dominant Splayed phenotype

Some stocks had flies with a small melanotic tumor or 'dot' on the junction between the abdomen and legs when heterozygous with the TM3 balancer

chromosome (frequency ~1:10). This was seen in stocks of *Z1914, Z1904,* and *Z352* and could be the result of the occasional event where the number of compromised homodimeric subunits has lowered the overall GFAT1 activity levels in this tissue to a critical point. This corresponds with the original Splayed phenotype which also acted in a dominant manner (Lindsley, Sandler, Baker, et al. 1972).

Based on the above, it appears that the amount of functional GFAT1 protein at the pupa-adult metamorphosis stage is not critical until below 50% of wild type levels. For the weaker *zep* alleles that are able to contribute a slightly defective protein, the less than fully functional homodimer cannot make sufficient GlcN6P for the downstream components of the chitin biosynthetic pathway to consistently result in wild type flies. In this case *Z1914/Z1904,* and *Z352* alleles represent a dominant negative mutation.

Gfat1 role in exoskeleton formation – *Spl* phenotype analysis

Gfat1 is the rate limiting enzyme of the chitin biosynthesis pathway (Badet, Vermoote, Haumont, et al. 1987), and will therefore influence the amount of chitin available for cuticle deposition at the various development stages from embryo to adult.

New cuticle is formed in three distinct layers, which are deposited by the epithelial cells preceding ecdysis, on cue with a high titre pulse of 20-hydroxyecdysone. The outer envelope layer, made from waxes and lipids, is secreted first at the tips of epithelial cell microvillae (Moussian, Veerkamp,

Muller, et al. 2007), followed by the epicuticle which contains proteins crosslinked with quinones (sclerotinization), and lastly the procuticle which is a chitin matrix composed of bundles of chitin fibers with sclerotonin and melanin proteins attached forming laminae arranged in a helical manner parallel to the epidermal surface (Merzendorfer. 2006).

When reviewing the functions of chitin in exoskeleton morphogenesis (Moussian, Schwarz, Bartoszewski, et al. 2005), it is evident that the cuticle envelope and epicuticle rely on the integrity of the procuticle for normal morphology. Chitin anchors the cuticle layer to the epidermis and provides lateral tension to the exoskeleton. When the procuticle layer is lost the epicuticle layer above becomes flaccid as is evidenced by the embryonic blimp phenotype in *zep* mutants. The variance in the blimp phenotype amongst the *zep* alleles could therefore be attributed to the amount, or complete lack of chitin available for procuticle layer formation in embryogenesis. The 'retroactive' phenotype where the embryo is extremely active and sometimes reversed in the egg case could be interpreted as evidence that the embryo is normal in all other respects but lacks the strength of the cuticle layer provided by chitin to push its way out.

When chitin production is disrupted in a *chitin synthase (kkv)* mutant, the epicuticle and envelope layers become detached and morphologically rearranged though the layer order is still maintained. Proteins that are normally lodged in the procuticle chitin laminae, such as sclerotinization and melanization proteins are dissociated in the amorphous procuticle layer. It is also apparent from TEM images that the cuticle layer contains more electron dense material (proteins),

which Moussian et al. (2005) suggest may be a compensation mechanism for the lack of cuticle structure.

It is possible that the melanotic tumors at knee and other leg joints shown in the adult stage *Spl* phenotype are a result of this morphological disruption of the cuticle layers by lack of chitin in the procuticle. The melanotic deposits observed could be composed of the amorphous epi- and pro- cuticle layers with abnormally deposited melanin and other proteins that must surface at joints. The strength of the outermost envelope layer could explain why there is not an overall 'porosity' to the leg cuticle in regard to malformed cuticle components but rather the deposits must accumulate at joints where the envelope could be weaker. Since the outermost protective layer, the envelope is still in tact (Moussian, Schwarz, Bartoszewski, et al. 2005), it is possible that it forms a barrier against the exit of epicuticle and procuticle layer components.

The nature of the *zep* mutant alleles has been characterised at a molecular level, and the predicted degree of severity of the mutations on the GFAT1 enzyme function related to the mutant embryo and adult phenotypes. These results show that the *zep* locus corresponds to the gene *Gfat1* (CG12449) for which there are no alleles currently identified on Flybase. The range of phenotypes in embryo and adult stages have helped to elucidate GFAT1 function in the context of chitin synthesis and the consequences of reduced and complete loss of GFAT1 function on cuticle deposition in *Drosophila melanogster*.

Since GFAT1 is the rate-limiting enzyme in the chitin synthesis pathway, if the GFAT1 Fru6P processing is slow because the protein is defective, the whole pathway is affected and less final product is available. The same result may occur if the active site is inhibited by substrate that cannot be released; all enzyme becomes unusable over time. Given the relationship observed between the nature of the mutant alleles at a molecular level, and the degree of mutant phenotype severity, it appears to be essential that sufficient amounts of Gln6P are made available in the correct location and time frame needed for normal cuticle deposition in both embryo and adult.

Zep/Gfat1 cytogenetic placement

Flybase currently describes *Gfat1* as located in heterochromatin cytological band h53, the closest band to the centromere(Gelbart, Crosby, Matthews, et al. 1997). However with the identification of *zep* alleles as *Gfat1* mutants this suggests that this is misplaced. By combining the results of the *zep* complementation map with reference to smaller 3R deficiencies on the distal region of 3R, together with the positioning results provided by the P-element generated *zep* alleles, it is possible to resolve the cytological location of *zep/Gfat1*, (Fitzpatrick. 2005). The deficiency *Df(3R)10-65* provides an anchor point on the proximal side of *Gfat1* and the P-element origins KG0840 and *EP(3)3632* provide anchor points on the distal side.

Genetically, *zep* complements the 3R deficiency *Df(3R)10-65* and does not complement the deficiency Df(3R)4-75. *Df(3R)10-65* is characterised

cytologically as a deficiency in h58 (Koryakov, Zhimulev, Dimitri. 2002) (Dimitri, Corradini, Rossi, et al. 2003), Df(3R)4-75 is characterised as an inversion that has breakpoints in h53 and h58. Non-complementation between these two lesions supports these positions as correct. The deficiencies generated from P-elements originating in the first euchromatin scaffold distal to the heterochromatin/euchromatin boundary (*8740-20*, *8740-22* and *EP-167R*) do not complement *zep* but do complement deficiency *Df(3R)10-65*. Therefore the P-element deletion end points appear to be distal to the distal boundary of this deficiency in h58, as shown in figure 3.1. This places the *zep* locus in the distal boundary region of heterochromatin, band h58 rather than h53.

Future work

It can not be assumed that GFAT1 is the only enzyme producing Gln6P in the chitin synthesis pathway, as the function of the paralog GFAT2 is unknown. However it is clear that GFAT1 at least contributes significantly to the production of UDP-NacGlc required for cuticle production, as *Gfat1* mutant embryos normal function is disrupted and they appear to have the same mutant phenotype seen in mutants of other components of the chitin synthesis pathway (Moussian, Schwarz, Bartoszewski, et al. 2005). Since GFAT proteins also participate in the production of UDP-NacGlc for the glycosylation of proteins, it is possible that GFAT2 operates in more of a housekeeping capacity. Further work could include *Gfat2* RNAi to determine if it is an essential gene and investigation of its

function in various tissues as well as whole mount in situ hybridisation with a Gfat2 cDNA probe to determine expression patterns.

In order to investigate the role of GFAT1 in embryo cuticle formation further it would be useful to prepare slides of the transheterozygote cross embryos to determine if the blimp phenotype is reduced compared to embryos homozygote for single alleles. Embryo preparations of the RNAi transgene with actin or tubulin drivers may also show that the blimp phenotype can be replicated. A stain to compare any changes in the cuticle morphology between homozygous *zep* mutants and homozygous *chitin synthase (kkv)* mutants would provide an interesting comparison as the two enzymes are in the same pathway.

Identification of the lethal phase with the *Gfat1* transgene and actin driver would further support that the transgenic lethality with this combination is exclusively due to knockdown of *Gfat1* expression. Using whole mount *in situ* hybridization Graack et al. (2001) established that *Gfat1* expression begins in stage 16, therefore the transgenic embryos should fail to hatch. The lethal phase analysis described in results above showed that a greater percentage of transgenic embryos failed to hatch compared to Ore-R, however a larger sample size is needed to detect the lethal phase accurately.

Base changes in *Z1014* and *l400-1* have yet to be found. Since *Z1014* has a class 1 blimp phenotype the sequence change could have produced an additional splice site and a truncated GFAT1 protein. This could be further investigated by sequencing or in a Western blot. *l400-1* is in blimp class 2 and

is predicted to have a less severe base change in the coding or upstream promoter regions.

Retrieval of the sequence adjacent to the P-element generated *zep* alleles, *8740-20*, *8740-22* or *EP-167R*, may provide the orientation and placement of the *Gfat1* scaffold AAB01002542. This could be done using an inverse PCR experiment.

Rescue of *zep* mutant lethality and correction of the blimp phenotype in homozygous mutants using a *Gfat1* transgene would provide ultimate evidence that the *zep* locus corresponds to *Gfat1*. Use of various drivers for the rescue, specific to tissue type and stages of development, would contribute to the understanding of the role of *Gfat1* in cuticle, trachea formation and Sgs protein glycosylation.

Based on RNA *in situ* hybridization experiments published on Flybase, CG8708 has an expression pattern that coincides with the *Gfat1* pattern of expression in salivary glands at stages 13-16. It is described as having a molecular function of beta-1,3-galactosyltransferase activity and to be involved in amino acid glycosylation. It may therefore be a good candidate for study in relation to researching the role of *Gfat1* expression in protein glycosylation in the embryonic developing salivary glands. CG8708 is located on 2R between cytological bands 44B5 and 44B7 and has two annotated transcripts. CG8708 cDNA AY051529 corresponds to the predicted CG8708-RA transcript and has

an ORF of 928 nucleotides. One allele is reported but there are no phenotypic data for this gene.

Unique features associated with transcription of genes in the condensed heterochromatin environment are yet to be determined. The *Gfat1* gene offers an opportunity to ascertain the differences between gene activation and regulation in the euchromatin *(Gfat2)* versus heterochromatin *(Gfat1)* environment as both genes have the same biochemical function and may also have the same pattern of expression. The *Gfat1* and *Gfat2* genes could be the result of a chromosome segment duplication event or a reverse transcription and insertion of *Gfat1* sequence as *Gfat2*. *Gfat1* appears to have evolved to develop its own specialised domain of operation as it appears to be an essential gene whose function is not made redundant by *Gfat2*.

Investigation of the upstream promoter region of *Gfat1* and *Gfat2* to determine which transcription factors are binding to activate the Gfat genes in various tissues would be a valuable inquiry. They are predicted to be regulated by Fkh based on the Fkh consensus sequences present in both as well as the TATAAA box sequences. Proteins associated with housekeeping such as SP1 and AP-1 may also be present. In addition, *Gfat2* has 90-95% upstream sequence homology with the E74 recognition sequence, a factor associated with metamorphosis, activated by ecdysone.

A closer investigation of what allows the fkh (or other) transcription factor to operate in the heterochromatin environment to activate *Gfat1*, but at the same

time allows it to operate in the euchromatin environment to activate *Gfat2*, would be useful to compare to understand how genes must adapt to the heterochromatin environment.

It is possible that *Gfat1* resides in a region of less condensed chromatin within the heterochromatin environment or that it employs specialized proteins to act as mediators in the 5' upstream regulatory region. It has been found that a reduction in the dose of *Suppressor of Variegation Su(var)2-5*, (HP1 protein), compromises the activity of the heterochromatin gene *Rpl15 (lethal 2)* which codes for a ribosomal protein (Schulze, McAllister, Sinclair, et al. 2006). This was demonstrated by observing enhanced wing defects produced by having only a single dose each of HP1 with Rpl15 and Notch, compared to the normal dose produced by two functional copies of *Su(var)2-5* and a single dose each of Rpl15 and Notch.

The same type of investigation could be used with Gfat1/zep to ask if the intermediate (dot) phenotype occasionally observed in weaker *zep* alleles could be enhanced with a reduced dose of HP1. The effects of a reduced HP1 dose on a single mutant copy of *Gfat1* could also be compared to the effects on its euchromatin counterpart *Gfat2.* However *Gfat2* mutants would first need to be generated by EMS or P-element mobilisation, and characterised.

APPENDICES

Appendix A: Gfat1 cDNA used for the characterisation of mutant alleles and transgenic Rescue experiment

```
LOCUS       >RE72989.c   2440 BP
DEFINITION  >RE72989.c
ORIGIN
        1 GAGGTAGACT CAGTGTTATT TAAACAGCGT ACCTGTGCCT GTTCGAGACG CCAGCAACAA
       61 CCAGCTTGAC AGCGACAACC CAGTACTGCA GACCAGCGAA TTAGGCAAAT CAAATATTTA
      121 CTAACATCTT TAATTACAAA TATCTATAAA AATGTGTGGA ATATTTGCAT ATCTAAATTA
      181 TCTGACACCC AAGTCCCGCC AGGAGGTGCT GGACCTTCTG GTCACGGGCT TGAAGAGATT
      241 GGAATACCGC GGCTACGACT CTACTGGAGT GGCAATTGAC TCTCCGGATA ATAAAAACAT
      301 CGTGATGGTC AAGCGAACGG GCAAGGTCAA AGTGCTTGAA GAAGCAATTC AAGAGCACTT
      361 CAGCGGAAGA GAATACAGCG AACCCGTCCT GACCCACGTT GGCATTGCTC ACACCCGCTG
      421 GGCCACCCAT GGTGTTCCTT GCGAGAAGAA CTCCCACCCA CACCGTTCGG ACGACGAAAA
      481 TGGTTTCGTC GTAGTTCATA ACGGCATCAT CACCAACTAC AATGATGTAA AGACTTTTCT
      541 TTCGAAGCGA GGTTATGAGT TTGAGTCGGA TACTGATACA GAGGTATTTG CCAAGCTAGT
      601 ACACCACCTG TGGAAAACCC ACCCCACCTA CTCCTTCCGC GAGCTGGTCG AGCAAGCCAT
      661 CCTTCAAGTG GAGGGCGCCT TTGCTATTGC CGTAAAGTCA AAATACTTTC CCGGAGAGTG
      721 TGTGGCGTCG CGGCGTAGTT CGCCCTTGCT AGTGGGAATC AAGACAAAAA CACGCCTAGC
      781 CACAGACCAC ATTCCAATTC TGTACGGAAA AGATGACAAG AAGCTCTGCA CCGATCAAGA
      841 TGCCGACTCT GGAAAACCTC AAGATATTCG CCCACATGGA CAATCTCGTG AACTGCCAGT
      901 GCTTCCCCGT TCGGAAAGCA CTTCTGAGTT TATGCCCTTG GAAGAGAAGG AAGTTGAGTA
      961 CTTTTTCGGA TCGGACGCCT CAGCCGTCAT AGAGCACACT AACCGGGTCA TCTATTTGGA
     1021 GGACGACGAT GTTGCTGCTG TTCGGGATGG AACTTTGAGT ATACATCGCC TAAAGAAGAG
     1081 CCTGGATGAT CCGCACGCTC GCGAAATCAC TACCCTAAAA ATGGAAATTC AACAGATCAT
     1141 GAAGGGAAAC TATGACTATT TTATGCAAAA GGAGATTTTC GAGCAGCCCG ACTCCGTGGT
     1201 GAACACAATG CGCGGTCGCG TCCGCTTCGA TGGTAACGCC ATAGTGCTCG GCGGGATCAA
     1261 AGACTACATT CCTGAAATCA AACGCTGTCG ACGCCTGATG TTGATTGGAT GTGGCACATC
     1321 TTACCACAGC GCTGTAGCCA CTAGGCAGCT GCTCGAAGAA CTCACAGAGC TTCCCGTGAT
     1381 GGTTGAGCTG GCTTCCGACT TTTTAGACCG AAACACTCCT ATTTTTCGAG ACGACGTCTG
     1441 CTTTTTTATA TCGCAGTCCG GAGAGACTGC CGACACCCTG ATGGCCTTAC GTTACTGTAA
     1501 GCAGCGAGGA GCCCTGATTG TGGGCATTAC GAATACCGTA GGCAGCAGCA TATGTCGGGA
     1561 ATCGCATTGT GGAGTGCACA TTAATGCCGG ACCAGAGATA GGCGTGGCCT CGACCAAGGC
     1621 ATACACCTCC CAATTCATTT CCCTGGTGAT GTTCGCTCTA GTTATGTCCG AAGATCGACT
     1681 GTCACTGCAA CAGCGACGGC TTGAGATTCT GCAGGCGTTG TCCAAGCTCG CGGATCAAAT
     1741 CCGAGACGCT CTGCAGCTGG ACTCCAAAGT TAAAGAACTG GCCAAAGACC TATACCAACA
     1801 CAAGTCGCTT CTGATAATGG GTAGGGGCTA CAACTTTGCC ACTTGCCTAG AAGGTGCATT
     1861 GAAAGTCAAA GAGTTGACTT ACATGCACAG CGAGGGCATC ATGGCCGGTG AATTGAAGCA
     1921 CGGCCCACTG GCTCTCGTAG ACGACTCCAT GCCCGTGCTG ATGATTGTTT TGCGGGACCC
     1981 CGTTTACGTA AAGTGCATGA ACGCTCTACA GCAGGTCACA TCCCGCAAAG GATGCCCGAT
     2041 TATTATCTGC GAGGAGGGAG ACGAGGAGAC CAAGGCTTTC TCCTCCCGCC ATCTAGAGAT
     2101 TCCTCGCACC GTCGACTGCC TGCAAGGAAT TCTCACCGTT ATCCCAATGC AACTACTGTC
     2161 TTATCATATT GCCGTGCTTC GCGGATGCGA CGTTGACTGT CCTAGAAACT TAGCAAAGTC
     2221 TGTGACAGTT GAGTAAACAC CTATTTAATT GGAACTCCGA CTTTGTACCT CTACATGATA
     2281 ATCTCCACAT TTTGTTGTGA AAACTTCCAA ACAAGACACA TCTACAAAAT ACCATTATTA
     2341 TAAAATATCT CCATATTTTC GGATAAACCG TACTCAATTT TATTGAATAA AAATATTATT
     2401 TCGTTCAGAA AAGTAGTAAA GGTTAAAAAA AAAAAAAAAA
```

Appendix B: Primers used and Tm Information.

Primer Name	Gene	Sequence	Tm ° C
GFP-F GFP-R	GFP	caagagtgccatgcccgaag gacagggccatcgccaattg	55
Grp84-01 Grp84-02	Grp84	acgcttctcgctgatggac gtcgcagtaactggattgagt	50-65
Gfat1F Gfat1B	Gfat1	tcttggcttgtctgcgtgctacc aaccgagttttggctgctgat	60
Gfat2F Gfat13	Gfat1	tcgcgttcgcattttcacaag ccaacgtgggtcaggacgggttc	60
Gfat4F Gfat4B	Gfat1	actacaaggttcaacattctg gggaaatgaattgggaggtgt	62
Gfat5F Gfat5B	Gfat1	ctgatggccttacgttactgt atgtgtccaagcaatgtgatg	60
Gfat15 Gfat6B	Gfat1	ctcgcgcgcaatcaacatacgatt caagtgtaagatggtaggtgc	60
Gfat10F Gfat6B	Gfat1	gtaagcacagcaatggtt caagtgtaagatggtaggtgc	60
Gfat14F Gfat24R	Gfat1	gccacttgcctagaaggtgga cggcaatatgataagacagtagtagc	62

Appendix C: Transgene Gal4 drivers

Bloomington Stock Number	Genotype	Description
4414	P{Act5C-GAL4}25F01	Actin ubiquitous
5143	P{tubP-GAL4}LL7	Tubulin ubiquitous
1878	W[*];P{w[+mW.hs]=GawB}T80/CyO	Ubiquitous in third instar imaginal discs
1967	W[*];P{w[+mW.hs]=GawB}34B/CyO	GAL4 in embryonic salivary glandsposterior midgut, eye-antennal, haltere, leg and wing imaginal discs.

REFERENCE LIST

References

Adams, M.D., et al, 2000. The Genome Sequence of Drosophila Melanogaster. Science 287, 2185-2195.

Alberts, B., Johnson, A., Lewis, J., et al. 2002. Molecular Biology of the Cell. Garland Science, New York.

Badet, B., et al, 1987. Glucosamine Synthetase from Escherichia Coli: Purification, Properties, and Glutamine-Utilizing Site Location. Biochemistry 26, 1940-1948.

Corradini, N., et al, 2007. Drosophila Melanogaster as a Model for Studying Protein-Encoding Genes that are Resident in Constitutive Heterochromatin. Heredity 98, 3-12.

Dimitri, P., et al, 2003. Vital Genes in the Heterochromatin of Chromosomes 2 and 3 of Drosophila Melanogaster. Genetica 117, 209-215.

Eberl, D.F., Duyf, B.J., Hilliker, A.J., 1993. The Role of Heterochromatin in the Expression of a Heterochromatic Gene, the Rolled Locus of Drosophila Melanogaster. Genetics 134, 277-292.

Fitzpatrick, K.A. 2005. Genetic and molecular characterization of chromosome three heterochromatin in Drosophila melanogaster. Simon Fraser University, Burnaby B.C.

Fitzpatrick, K.A., et al, 2005. A Genetic and Molecular Profile of Third Chromosome Centric Heterochromatin in Drosophila Melanogaster. Genome 48, 571-584.

Fletcher, J.C., Thummel, C.S., 1995. The Ecdysone-Inducible Broad-Complex and E74 Early Genes Interact to Regulate Target Gene Transcription and Drosophila Metamorphosis. Genetics 141, 1025-1035.

Gelbart, W.M., et al, 1997. FlyBase: A Drosophila Database. the FlyBase Consortium. Nucleic Acids Res 25, 63-66.

Graack, H.R., Cinque, U., Kress, H., 2001. Functional Regulation of Glutamine:Fructose-6-Phosphate Aminotransferase 1 (GFAT1) of Drosophila Melanogaster in a UDP-N-Acetylglucosamine and cAMP-Dependent Manner. Biochem J 360, 401-412.

Grewal, S.I., Jia, S., 2007. Heterochromatin Revisited. Nat Rev Genet 8, 35-46.

Hart, G.W., Housley, M.P., Slawson, C., 2007. Cycling of O-Linked Beta-N-Acetylglucosamine on Nucleocytoplasmic Proteins. Nature 446, 1017-1022.

Hilliker, A.J., 1976. Genetic Analysis of the Centromeric Heterochromatin of Chromosome 2 of Drosophila Melanogaster: Deficiency Mapping of EMS-Induced Lethal Complementation Groups. Genetics 83, 765-782.

Hoskins, R.A., et al, 2007. Sequence Finishing and Mapping of Drosophila Melanogaster Heterochromatin. Science 316, 1625-1628.

Irion, U. and Leptin, M., 1999. Developmental and Cell Biological Functions of the *Drosophila* DEAD-Box Protein Abstrakt. Curr Biol 9, 1373-1381.

Kato, N., et al, 2002. Glucosamine:Fructose-6-Phosphate Aminotransferase: Gene Characterization, Chitin Biosynthesis and Peritrophic Matrix Formation in Aedes Aegypti. Insect Mol Biol 11, 207-216.

Koryakov, D.E., Zhimulev, I.F., Dimitri, P., 2002. Cytogenetic Analysis of the Third Chromosome Heterochromatin of Drosophila Melanogaster. Genetics 160, 509-517.

Koundakjian, E.J., et al, 2004. The Zuker Collection: A Resource for the Analysis of Autosomal Gene Function in Drosophila Melanogaster. Genetics 167, 203-206.

Lindsley, D.L., et al, 1972. Segmental Aneuploidy and the Genetic Gross Structure of the Drosophila Genome. Genetics 71, 157-184.

Lippman, Z., Martienssen, R., 2004. The Role of RNA Interference in Heterochromatic Silencing. Nature 431, 364-370.

Lu, B.Y., et al, 2000. Heterochromatin Protein 1 is Required for the Normal Expression of Two Heterochromatin Genes in Drosophila. Genetics 155, 699-708.

Marchant, G.E., Holm, D.G., 1988. Genetic Analysis of the Heterochromatin of Chromosome 3 in Drosophila Melanogaster. II. Vital Loci Identified through Ems Mutagenesis. Genetics 120, 519-532.

Marshall, S., 2006. Role of Insulin, Adipocyte Hormones, and Nutrient-Sensing Pathways in Regulating Fuel Metabolism and Energy Homeostasis: A Nutritional Perspective of Diabetes, Obesity, and Cancer. Sci STKE 2006, re7.

Merzendorfer, H., 2006. Insect Chitin Synthases: A Review. J Comp Physiol [B] 176, 1-15.

Moussian, B., et al, 2007. Assembly of the Drosophila Larval Exoskeleton Requires Controlled Secretion and Shaping of the Apical Plasma Membrane. Matrix Biol 26, 337-347.

Moussian, B., et al, 2005. Involvement of Chitin in Exoskeleton Morphogenesis in Drosophila Melanogaster. J Morphol 264, 117-130.

Ostrowski, S., Dierick, H.A., Bejsovec, A., 2002. Genetic Control of Cuticle Formation during Embryonic Development of Drosophila Melanogaster. Genetics 161, 171-182.

Renault, N., King-Jones, K., Lehmann, M., 2001. Downregulation of the Tissue-Specific Transcription Factor Fork Head by Broad-Complex Mediates a Stage-Specific Hormone Response. Development 128, 3729-3737.

Sambrook, J., Fritsch, E.F., Maniatis, T. 1989. Molecular Cloning: A laboratory manual. Cold Spring Harbour Laboratory Press, .

Schulze, S.R., et al, 2006. Heterochromatic Genes in Drosophila: A Comparative Analysis of Two Genes. Genetics 173, 1433-1445.

Schulze, S.R., et al, 2005. A Genetic and Molecular Characterization of Two Proximal Heterochromatic Genes on Chromosome 3 of Drosophila Melanogaster. Genetics 169, 2165-2177.

Sinclair, D.A., et al, 2000. Essential Genes in Autosomal Heterochromatin of Drosophila Melanogaster. Genetica 109, 9-18.

Sinclair, D.A., Lloyd, V.K., Grigliatti, T.A., 1989. Characterization of Mutations that Enhance Position-Effect Variegation in Drosophila Melanogaster. Mol Gen Genet 216, 328-333.

Smith, C.D., et al, 2007. The Release 5.1 Annotation of Drosophila Melanogaster Heterochromatin. Science 316, 1586-1591.

Tasaka, S.E., Suzuki, D.T., 1973. Temperature-Sensitive Mutations in Drosophila Melanogaster. XVII. Heat- and Cold-Sensitive Lethals on Chromosome 3. Genetics 74, 509-520.

Yamazaki, K., et al, 2000. Cloning and Characterization of Mouse Glutamine:Fructose-6-Phosphate Amidotransferase 2 Gene Promoter. Gene 261, 329-336.

Yasuhara, J.C., Wakimoto, B.T., 2006. Oxymoron no More: The Expanding World of Heterochromatic Genes. Trends Genet 22, 330-338.