A Family of Genes Clustered at the Triplo-lethal Locus of Drosophila melanogaster Has an Unusual Evolutionary History and Significant Synteny With Anopheles gambiae

Douglas R. Dorer,* Jamie A. Rudnick,†,1 Etsuko N. Moriyama†,‡ and Alan C. Christensen†,2

*Department of Microbiology, Meharry Medical College, Nashville, Tennessee 37208 and †School of Biological Sciences and 
‡Plant Science Initiative, University of Nebraska, Lincoln, Nebraska 68588

Manuscript received March 26, 2003
Accepted for publication May 16, 2003

ABSTRACT
Within the unique Triplo-lethal region (Tpl) of the Drosophila melanogaster genome we have found a cluster of 20 genes encoding a novel family of proteins. This family is also present in the Anopheles gambiae genome and displays remarkable synteny and sequence conservation with the Drosophila cluster. The family is also present in the sequenced genome of D. pseudoobscura, and homologs have been found in Aedes aegypti mosquitoes and in four other insect orders, but it is not present in the sequenced genome of any noninsect species. Phylogenetic analysis suggests that the cluster evolved prior to the divergence of Drosophila and Anopheles (250 MYA) and has been highly conserved since. The ratio of synonymous to nonsynonymous substitutions and the high codon bias suggest that there has been selection on this family both for expression level and function. We hypothesize that this gene family is Tpl, name it the Osiris family, and consider possible functions. We also predict that this family of proteins, due to the unique dosage sensitivity and the lack of homologs in noninsect species, would be a good target for genetic engineering or novel insecticides.

When the Drosophila genome was surveyed for dosage-sensitive regions, only one was found that was both triplo-lethal and haplo-lethal (Lindsley et al. 1972). This locus, located in cytological region 83D4.5-E1.2, was called the Triplo-lethal locus, abbreviated Tpl. Stocks carrying a duplication of Tpl on one homolog and a deficiency on the other homolog are viable and provide a powerful selection for either up or down mutations when crossed to a wild-type fly (Denell 1976). Using this selection, Keppy and Denell (1979) were able to obtain duplications and deficiencies of Tpl, but were unable to isolate point mutations following EMS or formaldehyde mutagenesis. Roehrdanz and Lucchesi (1980) were also unable to isolate point mutations of Tpl following EMS mutagenesis, although they did isolate mutations in the Suppressor of Tpl locus, which has been shown to encode the Ell protein, a general transcription elongation factor (Dorer et al. 1995; Eissenberg et al. 2002).

Denell proposed three hypotheses to explain the lack of point mutations at Tpl: (1) the locus is very small so the mutation rate is very low, (2) the locus does not encode a protein and therefore is less sensitive to single base changes, or (3) the locus consists of a gene cluster with at least partial redundancy, such that mutation of one of the genes does not rescue the lethality of a duplication of the entire cluster. The small size hypothesis predicts that as the number of mutagenized chromosomes increases, the chance that a mutant will be found also increases. However, we have subsequently screened >10^6 chromosomes and still have not isolated point mutations (Dorer and Christensen 1990; our unpublished data). The non-protein-coding hypothesis predicts that transposon insertions or inversion breakpoints would still inactivate the locus, but in spite of considerable effort we have never isolated a Pelement insertional mutation in Tpl (Dorer and Christensen 1990; our unpublished data). Thus the most likely hypothesis is that Tpl consists of a cluster of genes with at least partial redundancy.

Using the complete genomic sequence of Drosophila melanogaster we have tested the prediction that the Triplo-lethal region contains a cluster of genes with high similarity. To do that, first we defined the molecular limits of Tpl by isolating and mapping duplications and deletions and then examining the sequence within those limits for repeated genetic units. We describe here the discovery of a multigene family in the Triplo-lethal region, consistent with the best hypothesis based on the genetic data. Although the proteins encoded by this family are novel, the sequences have features that allow us to make predictions about their function. We predict that a family of genes whose dosage is so critical will be well conserved and show evidence of strong selection on expres-
Osiris 1 (49) - VQADFFSRNQQG.PETRSL- - US-IGRTTCKILATNMSFGPFSPESRRGADG - RGKWLIRVQLGEFAD
Osiris 2 (75) - NMDPFLARTNNSLQCDL- - SGEQ-CDTQNTEDDEP-KQD-YKQEPFARSL- - PETQQRSSLQEP-
Osiris 3 (38) - IRDPFALLKQONLDD - RAKVTSAMNHYLNTIR - ILGDPPDKTEN - GNSMVPDPPDE
Osiris 4 (47) - VENTDLDKLNIVVAAHAS - YCSGRVAMLSSTREYTKLRGQLPDDAS - RKKYQKGTQLRSQSF
Osiris 5 (19) - NQDAQAGY - LRKCMPRNRLNDRSKPDVLRLVEPL - QNNS
Osiris 6 (30) - AF - AQCLSLDS - ICGTOLTLFRKARWPD-NPOIELPGTQMV - NEQRRGKQDLSN
Osiris 7 (37) - DINDSTYPSKDDG - VLDPSLSTPVYLDGQARDQ - FALTEMVDFQ - PDAP
Osiris 8 (44) - WDMWSYRTQGQOSDIMN - SVEAQTLTLLARAFSRASKS - LSEMNPQFSGSSGSEXETRARIPASNDK
Osiris 9 (29) - RPK - EEPDLMRRD - D - VPIFPERERL - - DEIFPVQSL -
Osiris 10 (37) - TEYRGLYHSDGQ - VQDLPSTVLR - VPDQ - - EQ - ETPSQDTRSPSRS
Osiris 11 (46) - TLLRBVDTRAPBF - VQFRAFSDILAPIDA - ENAVFVQTRL - ETARPRPATSENE
Osiris 12 (34) - TLGRRAAYLQKLLSES-D - ATRVAVQATLNRRAKMINN - REGASYTPFCD - PASPVSTRRGKSM
Osiris 13 (42) - TPQRASELGSSASSS - WNPVTEFLKMKL - QEE - NVLPFIND - ENATLXKTEE
Osiris 14 (33) - QLLMRRDLHFL - KHDQGIKL - S - - 
Osiris 15 (30) - LLIEQCNASQGQ- - S - MALEAMRVPVDMKSHDQSFQSNLVRSGNEK - TPITNEARASSADQFL
Osiris 16 (38) - SDELSTITVDFKAHNA - TQVVTLYTVANVEESVGOSRAQDDVKK - VIVDR
Osiris 17 (28) - PEMAVRNYEDIQKND-F - EQ - ICAPVASHGALD - QDS - RQVFLRQAR - QGSETTLQG
Osiris 18 (33) - MRRKQMSLDRNQGQSIQG - SNPSHSMITHSSMSP - GSTQTVFRAIA - PNST - DN

**Figure 1.** Alignment of 18 members of the Osiris multigene family. Identical and strongly conserved amino acids are indicated with black shading, similar residues are shaded gray, and the YXXØ motifs (see text) are boxed. Consensus residues are defined as being present in at least 50% of the proteins. Numbers in parentheses indicate where that number of amino acids is not shown. The signal peptide and the nonconserved region between it and the cysteine motifs are not shown. The revised Osiris names correspond to the provisional CG names as follows: Osiris 1, CG15585; Osiris 2, CG1148; Osiris 3, CG1570; Osiris 4, CG10303; Osiris 5, CG15590; Osiris 6, LD21503; Osiris 7, CG1153; Osiris 8, CG15591; Osiris 9, CG15592; Osiris 10, CG15593; Osiris 11, CG15596; Osiris 12, CG1154; Osiris 13, CG15595; Osiris 14, CG1155, Sp5758; Osiris 15, CG1157; Osiris 16, CG13561; Osiris 17, CG15598; Osiris 18, CG1169; Osiris 19, CG15189; Osiris 20, CG11588; Osiris 21, CG14925; Osiris 22, CG8644; and Osiris 23, CG15538.

**Materials and Methods**

**Drosophila rearrangements and mapping:** Drosophila stocks were previously described (Dorer and Christensen 1990; Dorer et al. 1995) or were obtained from the Indiana University Drosophila Stock Center. Duplications and deficiencies of Tp elements were previously described (Dorer and Christensen 1990; Dorer et al. 1995) or were obtained from the Indiana University Drosophila Stock Center. Duplications and deficiencies of Tp elements were generated by crossing flies carrying both a Δ8-3 source of transposase and a single P element inserted near Tp to Y(3)01086/01086. Survivors were generated by crossing flies carrying both a Δ8-3 source of transposase and a single P element inserted near Tp to Y(3)01086/01086. Inserted in RM62; P{hsneo} l(3)neo33, inserted in castor (Cooley et al. 1988); and RS2/24, inserted in pollux (Vincent et al. 1990). Inverse PCR was used to amplify the DNA flanking Pinduced rearrangements as described (Sparling et al. 1999). DNA sequencing was done by the University of Nebraska DNA Sequencing Facility. Breakpoints were mapped by comparing these sequences to the Drosophila genome using BLAST (Altschul et al. 1997).

**Bioinformatics:** Sequence similarity searches were done using the BLAST server at http://www.ncbi.nlm.nih.gov (Altschul et al. 1990, 1997) or Vector NTI from Informax. Targeting predictions were done using TargetP 1.01 (Nielsen et al. 1997; Emanuelsson et al. 2000). Transmembrane helix predictions were done with TMHMM 2.0 (Krogh et al. 2001). Multiple alignments based on amino acid sequences were generated by MULTICLUSTAL (Yuan et al. 1999) and ClustalW (Thompson et al. 1994). On the basis of this alignment, amino acid distances were estimated by the JTT method (Jones et al. 1992).
Osiris Gene Family Evolution

Figure 1.—Continued.

Phylogenetic relationships were reconstructed with the neighbor-joining (NJ) method (Saitou and Nei 1987). Bootstrap supporting values were estimated from one thousand repetitions of bootstrap sampling, JTT distance estimation, NJ tree reconstruction, and bootstrap analysis were conducted by PHYLIP version 3.6a3 (Felsenstein 2001).

Codon usage bias was measured as the “effective number of codons” (ENC) as developed by Wright (1990). Numbers of synonymous and nonsynonymous substitutions per site were estimated by the Li (1993) method. General statistical analyses were conducted with JMP 5 (SAS Institute).

RESULTS

Molecular mapping of Tpl: Because we and others have been unable to isolate point or transposon insertion mutations in Tpl, in spite of a very powerful selection, we generated duplications and deletions flanking single P-element insertions (Preston et al. 1996). Because the proximal breakpoint of Dp(3;3)Tpl is just distal to the Rm62 gene (Dorer et al. 1990), we chose three single P-element insertions near Rm62 as starting points for new duplications and deficiencies. Of the 24 duplications and 43 deficiencies that resulted, we were able to accurately map 29 of them by inverse PCR and sequencing. The smallest duplication, Dp(3;3)Tpl, was duplicated from Rm62 through Pak, a region of ~334 kb. The smallest deficiency, Df(3R)Tpl, also has its distal breakpoint in Pak, confirming that Tpl is located between Rm62 and Pak.

A novel family of proteins is encoded within the Triplo-lethal region: Examination of this region in the Drosophila genome (Adams et al. 2000) reveals a previously undescribed family of genes that is consistent with the unusual genetic properties of Tpl. This family of genes is located in the 168-kb region from CG15585 through CG15188, which represents roughly half of the Tpl region defined by the breakpoints. Only 7 of the 27 genes in this region are not members of this family (CG31562, NPFR1, CG15589, CG15597, CG15594, CG15593, and CG31556). BLASTP and PSI-BLAST searching has revealed only 3 members of the family outside of this cluster (see below). Of these 23 family members,
18 are shown aligned in Figure 1. The others are described separately below. The complete alignment is given in Supplemental Data at http://www.genetics.org/supplemental/. All of the proteins appear to have endoplasmic reticulum signal peptides. In addition to the signal peptide, these proteins have three conserved domains. The first is near the amino terminus and consists of a pair of cysteines usually separated by seven to nine amino acids. The second consists of four hydrophobic blocks, separated by lysines. A proline is usually present in the first two hydrophobic blocks. This region is predicted to be a transmembrane domain (Krogh et al. 2001) and resembles a stop-transfer anchor, thus these are type I transmembrane proteins with the N terminus outside the cell and the C terminus inside (Goder and Spiess 2001). The third domain is a region rich in conserved histidines and tyrosines, including the highly conserved sequence AQXLAY near the carboxyl terminus. A number of different endocytic signaling motifs include tyrosines (Bonifacino and Dell’Angelica 1999; Royle et al. 2002), and copies of one such signal, YXXØ, are boxed in Figure 1. Because the Isis locus partially rescues the effect of trisomy for this region (Dorer et al. 1993), we call this family of proteins the Osiris family and have named the genes Osiris 1 through Osiris 23 on the basis of their position in the cluster (see Figure 1 legend).

Four protein sequences from the cluster at 83E that appear to be members of the Osiris family cannot be aligned as well and were not shown in Figure 1, but are aligned in Supplemental Data. These four proteins are Osiris 10, Osiris 13, Osiris 14, and Osiris 17. Each includes the transmembrane domain and the conserved tyrosine motif, however Osiris 10 appears to be internally repeated, and Osiris 13 and Osiris 14 are diverged at the ends, perhaps because of errors in predicting the exons. Osiris 17 includes copies of the cysteine motif at both the amino and carboxyl termini, and is predicted to be mitochondrial. Despite these differences, the four are recognizable as members of the family, and annotation errors may account for the differences.
Through BLAST and PSI-BLAST queries of the NCBI database, we have identified three other members of this family, encoded at three different sites elsewhere in the genome. Osiris 21 (CG14925, polytene region 32E) and Osiris 23 (CG15358, 99F) appear to be typical members of the family (Figure 1). Osiris 22 (CG8644, 87E) lacks the N-terminal cysteines, although it is otherwise very similar to the others (supplemental data at http://www.genetics.org/supplemental/). None of these loci are triplo-lethal or located within haplo-insufficient regions (LINDSLEY et al. 1972).

The Osiris gene family is highly conserved between Drosophila and Anopheles: We compared the genes in this region to the A. gambiae genomic sequence (HOLT et al. 2002), and found that Anopheles has an orthologous family of proteins, mostly encoded in two clusters on chromosome 2R. Phylogenetic relationships of the Osiris genes from both species were determined as described in MATERIALS AND METHODS, and the resulting tree is shown in Figure 2. With very few exceptions the closest relative of each gene is its ortholog in the other species, rather than any of the paralogs in the same species. These orthologous pairs were also supported with high bootstrap values (>90%). This suggests that the family diverged by gene duplication before the divergence of Drosophila and Anopheles, but each member of the family has since retained its unique features.

To understand the close relationships between orthologs, we examined the codon usage biases of the genes and the base substitution patterns between orthologous pairs. Codon usage bias was measured as the effective number of codons. It can range from 20 (where only one codon is used for each of the 20 amino acids and thus codon usage is most biased) to 61 (where all possible codons are used and there is no bias). The average ENC for >12,000 genes from D. melanogaster is 49, with a range from 28 to 61 (MORIYAMA 2003). It has been shown that highly expressed genes have high codon usage bias as seen by a low ENC. For example, the average ENC among ribosomal protein genes is 39 (MORIYAMA 2003). The average ENC from the 23 Osiris genes from D. melanogaster is 45.1 with a range from 31 to 61. The ENC from the 22 A. gambiae homologs ranges between 33 and 57 with an average of 40.6. The codon usage bias was significantly correlated between orthologous pairs of genes ($R = 0.52, P = 0.02$). The average G + C contents at the fourfold degenerate positions are 71% (D. melanogaster) and 74% (A. gambiae) and are also significantly correlated between orthologous pairs ($R = 0.54, P = 0.01$). On the other hand, both species have low and uncorrelated G + C content in introns (33% for D. melanogaster and 43% for A. gambiae; $R = 0.12, P > 0.05$). These observations indicate that there is strong selection on these genes and on the maintenance of their expression levels (DURET 2002; HURST et al. 2002).

A significant negative correlation is observed between the numbers of synonymous substitutions per site and the codon usage bias of orthologous pairs ($R = 0.76, P = 0.002$). This again implies that these genes are under translational selection. Remarkably, the orthologous pairs are similar enough to estimate the synonymous substitution rates for 15 of them. The average ratio between the synonymous and nonsynonymous substitutions per site is 0.37 and the ratios range from 0.25 to 0.58. These ratios are $<1.0$ (Figure 3), indicating that these genes are under selection. Not surprisingly, these ratios are higher than those obtained within the Drosophila lineage. BERGMAN et al. (2002), for example, reported that $\sim$90% of such ratios obtained from comparisons among three Drosophila species are under 0.2. Interestingly, nonsynonymous substitution rates are correlated with synonymous substitution rates (Figure 3; Spearman Rho = 0.6, $P = 0.02$). Such correlations have been described in Drosophila (AKASHI 1994) and mammalian genes (MAKALOWSKI and BOGUSKI 1998). In the case of Drosophila genes, translational selection both on codon usage bias and on amino acid substitutions is considered to cause such correlations (AKASHI 1994).

The Osiris gene cluster displays significant synteny: The map locations of most of the A. gambiae orthologs are known (HOLT et al. 2002), and the families maintain
a remarkable degree of synteny in the two species. The genes found in the Tpl cluster in D. melanogaster are found in the same order in A. gambiae in two clusters on chromosome 2. Ten of the family members are in polyteny region 18CD, and another seven are in 15D. A comparison of the Drosophila and Anopheles clusters is shown in Figure 4. The block of genes from Osiris 1 through Osiris 12 retains one of the largest regions of microsynteny found in a comparison of the two genomes (Zdobnov et al. 2002). In this interval, 9 of 11 pairs of orthologous Osiris genes are in the same order. With only two exceptions in each species the genes are all transcribed from the same strand. This synteny also includes the nonfamily members NPFR1 and CG15589/ agCG45916. Similarly, the block from Osiris 14 to Osiris 20 contains seven family members in the same order as their Anopheles orthologs in 15D. Two of the unlinked Drosophila family members (Osiris 21 and Osiris 22) have Anopheles orthologs that are also unlinked to the main clusters. The other unlinked Drosophila family member, Osiris 23, does not have a clear ortholog.

**Osiris family members in other species:** Homologs to the Osiris genes have also been sequenced in the dipteran insects D. pseudoobscura and Aedes aegypti. Not surprisingly, all 23 family members are found in the D. pseudoobscura genome, completely syntenic with the D. melanogaster genes (http://hgsc.bcm.tmc.edu/drosophila). No other insect genomes have been reported to be completely sequenced as of this writing. However, a partial cDNA from A. aegypti (GenBank accession no. BQ789636) encodes a homolog of Osiris 11. The cDNA sequence contains the putative transmembrane domain, and the high degree of conservation of this region is shown in Figure 5A.

We have also observed Osiris genes in four other insect orders. At least 17 different homologs to the Drosophila genes are found by BLAST searches of the first assembly of the honeybee (Apis mellifera) genome (http://titan.biotec.uic.edu/beec/honeybee_project.htm). A honeybee expressed sequence tag (EST; BI512836, BI515793, and BI509951) encodes a probable ortholog to Osiris 14 that contains the signal peptide, the paired cysteines, and the transmembrane domain (Figure 5B). cDNAs of other Osiris genes have been recovered from the
Figure 5.—Alignment of *D. melanogaster* Osiris proteins with homologs from other insect species. (A) Osiris 11. Only the C-terminal 114 amino acids of the *A. aegypti* homolog are available from GenBank. *D. melanogaster*, GenBank accession no. NP_649630; *D. pseudoobscura*, translation of nucleotides 51,261–52,215 of Contig9442 from 1/13/03 sequence release (http://hgsc.bcm.tmc.edu/drosophila); *A. gambiae*, agCP2541, GenBank accession no. EAA08588; *A. aegypti*, translation of partial cDNA clone AeP-230, GenBank accession no. BQ789636. (B) Osiris 14. Only the N-terminal 215 amino acids of the *A. mellifera* homolog are available from GenBank. *D. melanogaster*, GenBank accession no. AAF34808; *D. pseudoobscura*, translation of nucleotides 144,856–145,100, 145,218–145,425, and 145,495–145,851 of Contig5874_Contig1844.5 from 3/26/03 sequence release (http://hgsc.bcm.tmc.edu/drosophila); *A. mellifera*, translation of ESTs BB160010B10-C05, BB160020B10B02, and BB170024B10E-08 (GenBank accession no. BI509951).

Expression: The lethal phase of Tpl aneuploids is late embryonic or early larval, with the tracheae and the gut the first tissues to be affected (SMoyer et al. 2004). Expression data are available from the Berkeley Drosophila Genome Project for some of the *Osiris* genes (http://www.fruitfly.org) and are partly consistent with the phenotype. The embryonic expression of six of these genes (*Osiris* 3, *Osiris* 7, *Osiris* 9, *Osiris* 17, *Osiris* 18, and *Osiris* 20) peaks in stages 13–16, while *Osiris* 19 is detected throughout embryogenesis and with a much broader tissue distribution than the others. *Osiris* 9, *Osiris* 18, and *Osiris* 20 are expressed in the embryonic tracheal system; *Osiris* 9 is expressed in the esophagus and foregut; and *Osiris* 17 is expressed in the esophagus and hindgut. Additional family members known to be expressed in late embryonic stages include *Osiris* 2, *Osiris* 6, *Osiris* 14, and *Osiris* 15, consistent with a late embryonic lethal phase.

Expression of *Osiris* genes in later stages has also been observed. *Osiris* 7 shows two peaks of expression during metamorphosis in addition to one late in embryogenesis (Arbeitman et al. 2002). The *A. aegypti* cDNA homolog to *Osiris* 11 is expressed during metamorphosis (Krebs...
DISCUSSION

The Triplo-lethal locus has been mysterious since its discovery in 1972, primarily because point mutations and transposon insertion mutants have not been isolated (Keppy and Denell 1979; Rohrdanz and Lucchesi 1980; Dorer and Christensen 1990). In 1979 Denell proposed three hypotheses to explain the peculiar genetic properties of Tpl (Keppy and Denell 1979). Subsequent work suggested that a cluster of related genes is the most likely hypothesis. We tested this by molecularly defining the limits of Tpl to ~334 kb and examining the genomic sequence of that region. We found 20 genes that are clustered and closely related and encode a novel family of transmembrane proteins. With 23 total members, this is one of the largest gene families in Drosophila. Of the 1437 sequence similarity groups reported at http://www.fruitfly.org/annot/similarity.html, only 26 groups are larger.

Homologous genes have been found only in insects, and the function of this family is unknown in A. gambiae or any other insect species. We predict that the orthologous families in Anopheles and other insects will be dosage sensitive. Reciprocal crosses between D. melanogaster carrying duplications and deficiencies of Tpl with D. simulans, D. mauritiana, and D. sechellia have shown that Tpl is both triplo- and haplo-lethal to the interspecific hybrids of both sexes (our unpublished data). These genes are also located in one of the longest regions of microsynteny between D. melanogaster and A. gambiae (Zdobnov et al. 2002). Comparison of the sequences in the two species shows that in almost all cases the most closely related gene is the ortholog, rather than any of the paralogs. The low synonymous substitution rate appears to be due to strong codon bias, suggesting selection on expression level. The nonsynonymous substitution rate is lower than the synonymous rate, suggesting selection on function. These two observations and the clustering of 20 of the 23 family members in 83DE are consistent with the Osiris cluster of proteins being the Triplo-lethal locus.

The genetic data for Tpl (Keppy and Denell 1979) led Denell to propose that the functions of the individual genes in the cluster are partially redundant; another possibility is a threshold effect such that duplication or deletion of multiple genes is needed to see lethal dosage effects. Either of these possibilities could explain the inability of point mutants to complement duplications of the entire cluster. Subdivision of the Osiris cluster and transformation with members of the family would allow these hypotheses to be distinguished, as well as allowing a formal demonstration that the Osiris cluster is Tpl. Maintenance of the linkage relationships through-out evolution is reminiscent of the synteny maintained within the homoeotic gene families (Zhang and Nei 1996). It is possible that the tight linkage is maintained in response to selection—perhaps the genes are coordinately regulated or imprinted (Hurst et al. 2002). In any case, the stability of the linkage arrangements and the sequences through long periods of time is unusual.

Without knowing the precise function of these proteins it is hard to know why this genetic region is so dosage sensitive. However, the relative concentrations of membrane proteins can affect rates of association and assembly of complexes (Hoffman and Edelman 1983; Bravo et al. 2000; Keenan Curtis and Kane 2002), leading to dosage sensitivity. The very high sequence conservation in the hydrophobic domain is intriguing. This region is somewhat long for a transmembrane domain, and it is not completely hydrophobic, showing a periodicity of prolines and glycines, as well as lysines. These sequences are conserved in Anopheles, which suggests that this region has been under selection for more than interaction with membrane lipids. This transmembrane domain likely interacts with other proteins—possibly each other in complexes—or with an intramembrane protease such as the rhomboid protein (Urban and Freeman 2002). Given that Tpl-trisomics are partially rescued by hyperoxia (Smoyer et al. 2004), the conserved pair of cysteines is also intriguing. It is possible that the extracellular cysteine pair responds to redox potential and the intracellular tyrosines play a role in signaling. Finally, we suggest that because of the extreme dosage-sensitive lethality, the impossibility of rescuing duplications with point mutants, and the lack of homologs in other phyla, this family of proteins would make an ideal target for genetic modifications or insecticide discovery in Anopheles and other dipteran pests.

We thank Deidre Potter for excellent advice; Larry Harshman, Han Asard, and Bob Weldon for helpful conversations; Laura Smoyer, John Engelman, and Marilyn Cadden for technical assistance; and Joel Eisenberg for critical comments on the manuscript. This work was supported in part by grants to A.C.C. from the National Science Foundation and to D.R.D. from the National Institutes of Health (S06 GM08037-31).

LITERATURE CITED


