

# Mutations in the *mariner* transposase: The D,D(35)E consensus sequence is nonfunctional

(*Drosophila melanogaster*/transposable element/*TcI*)

ALLAN R. LOHE, DANIEL DE AGUIAR, AND DANIEL L. HARTL\*

Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138

Communicated by Melvin M. Green, University of California, Davis, CA, December 6, 1996 (received for review September 6, 1996)

**ABSTRACT** Genetic analysis of eukaryote transposases and comparison with their prokaryote counterparts have been greatly hindered by difficulty in isolating mutations. We describe a simple eye-color screen that facilitates isolation and analysis of mutations in the *mariner* transposase in *Drosophila melanogaster*. Use of ethyl methanesulfonate and site-directed mutagenesis has identified 18 residues that are critical for *in vivo* excision of a target *mariner* element. When the mutations were examined in heterozygous mutant/nonmutant genotypes, more than half of the mutant transposase proteins were found to reduce the activity of the wild-type transposase, as assayed by the frequency of germline excision of a target element. Remarkably, transposase function is obliterated when the D,D(34)D acidic, ion-binding domain is replaced with the consensus sequence D,D(34)E found in the nematode *TcI* transposase and in many other transposases in the superfamily. A number of mutations strongly complement wild-type transposase in a dominant-negative manner, suggestive of subunit interactions in the excision reaction; these mutations are located in a small region that includes part of the D,D(34)D motif. Transposase function also is eliminated by a mutation in the inferred initiation codon and by a mutation in a putative nuclear localization signal.

The transposable element *mariner* is a defining member of the *mariner*/*TcI* superfamily of transposons (1, 2) that are present in an exceedingly wide variety of genomes ranging from fungi to human beings. Members of this superfamily contain a single transposase gene, duplicate a TA dinucleotide pair upon insertion, transpose via a DNA intermediate, and are expressed in both the germline and the soma. Elements in the *mariner* family can be distinguished from those in the *TcI* family by their overall length ( $\approx 1.3$  kb versus 1.6–1.7 kb), length of inverted terminal repeats ( $\approx 30$  bp versus 20–460 bp), and certain differences in conserved amino acids (2). The most striking difference is in the so-called D,D(35)E motif that is shared not only by the *mariner*/*TcI* superfamily but also by eukaryotic retrotransposons, prokaryotic transposases, and retroviral integrases (1, 3, 4). The D,D(35)E motif consists of two aspartic acid residues (D, D) separated from one another by a variable distance (typically  $>90$  residues) followed by a glutamic acid (E) at a distance of 34 or 35 residues. The D,D(35)E motif is an essential component of the catalytic site and is thought to bind a divalent metal ion in the polynucleotidyl transferase reaction (5–7). In eukaryotic transposases of the *mariner*/*TcI* superfamily, the spacing between the last two acid residues is 34, as compared with 35 for prokaryotic transposases. Interestingly, elements in the *mariner* family have

the motif D,D(34)D whereas those in the *TcI* family have the motif D,D(34)E.

The molecular mechanisms of transposition of members of the D,D(35)E superfamily have been determined largely from the study of prokaryote elements, such as Mu, Tn7, and Tn10 (8–10), in which transposase mutations and mutant proteins can be readily isolated. The isolation of mutations in eukaryote transposases is less straightforward, so there has been a greater reliance on sequence comparisons and analogy than on direct experiment (1, 2, 3, 11). These have clearly identified the D,D(34)E domain and a bipartite nuclear localization signal (1, 11). In the N-terminal region of the *TcI* family of transposases, a motif related to the paired domain is combined with a motif that has similarity to a leucine zipper (11, 12), which is consistent with experimental evidence that the N-terminal region of the *TcI* and *Tc3* transposases is involved in DNA binding (12, 13). However, neither of the putative DNA-binding domains is conserved in the *mariner* family.

Identification of functional domains through conserved residues has the obvious limitation that the residues must be conserved. Functional domains that evolve rapidly, perhaps coevolving with other domains with which they interact, will be missed. For example, the inverted repeats of elements in the *mariner* family diverge rapidly in evolution (14, 15), and the DNA-binding domains of the transposase therefore must evolve rapidly as well. Site-directed mutagenesis is an alternative to sequence comparisons, but the approach is of limited utility on a broad scale, because practical considerations require that residues of particular interest be identified in advance, such as the acidic residues of the D,D(35)E motif (8).

We have undertaken a mutational analysis of the *mariner* transposase by exploiting a novel selection scheme based on the two-element system consisting of *Mos1* and *white-peach* ( $w^{p^{ch}}$ ). The *Mos1* element, originally identified in *Drosophila mauritiana*, is one of the few *mariner* elements known to be active (16). Virtually all others are nonfunctional owing to the presence of small deletions or inactivating point mutations. The  $w^{p^{ch}}$  allele has an inactive *mariner* element, designated *peach*, inserted in the 5' noncoding region of the X-linked *white* gene, which produces peach-colored eyes. Excision of the *peach* element by functional transposase supplied in *trans* restores wild-type *white* expression and results in eye color mosaicism. Here we report the isolation and analysis of one site-directed and 20 random ethyl methanesulfonate (EMS)-induced transposase mutations that (i) demonstrate that the D,D(34)D domain cannot be replaced with D,D(34)E, (ii) identify various essential residues in the transposase, and (iii) indicate that the majority of hypomorphic transposase mutations impair the activity of wild-type transposase observed in heterozygous mutant/nonmutant genotypes.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Copyright © 1997 by THE NATIONAL ACADEMY OF SCIENCES OF THE USA  
0027-8424/97/941293-5\$2.00/0  
PNAS is available online at <http://www.pnas.org>.

Abbreviations: EMS, ethyl methanesulfonate;  $w^{p^{ch}}$ , white-peach.  
\*To whom reprint requests should be addressed. e-mail: dhartl@oeb.harvard.edu.

## MATERIALS AND METHODS

**EMS Mutagenesis.** Adult males homozygous for the *P*-element construct,  $P[ry^+, hsp70:Mos1]$ , located at polytene chromosome position 2:33C6 (17), were mutagenized with 25 mM EMS as described (18). Approximately 50,000 F<sub>1</sub> individuals were screened using the  $w^{p^{ch}}$  eye color mosaicism assay, from which a total of 20 mutants in which mosaicism of  $w^{p^{ch}}$  was reduced or absent were isolated and confirmed by additional crosses and DNA sequencing.

**Site-Directed Mutagenesis.** The megaprimer method of site-directed mutagenesis with PCR (19) was applied to the plasmid pAD31-*hsp70:Mos1* (17) to modify the D,D(34)D motif to D,D(34)E. The mutagenic primer D284E (5'-GGCCCCATCCGAGTACCACCTATT C-3') was used with the reverse primer 5'-GGAAACAGCTATGACCCATG-3' to generate an approximately 600-bp PCR product that served as the megaprimer in conjunction with the M13 forward primer (5'-GTAAAACGACGGCCAGT-3') for a second round of PCR to generate the full-length mutant transposase gene. An *SphI-HindIII* fragment from the final PCR product was used to replace the corresponding fragment in pAD31-*hsp70:Mos1*, yielding the plasmid *phsp70:Mos1(D284E)*. Sequence analysis confirmed the D284E modification and also showed that no modifications due to *Taq* misincorporations had taken place. From *phsp70:Mos1(D284E)*, a 2-kb *BamHI-HindIII* fragment was isolated, blunt-ended with Klenow DNA polymerase, and cloned into the *HpaI* site of Carnegie 20 (20), forming pC20*hsp70:Mos1(D284E)*. This plasmid was injected into  $y;ry^{506}$  embryos with wings-clipped helper plasmid (21) as described (20). Three transformed lines were examined. All proved to be defective in both somatic and germline excision. One transformed line with an insertion in chromosome 3 was chosen for further analysis. PCR was used to amplify the *hsp70:Mos1(D284E)* chimeric gene from a single fly in this strain, and the resulting material was sequenced to verify the single site-directed mutation in the gene.

**Genetic Crosses.** Genetic crosses were carried out at 25°C on standard cornmeal/molasses medium. Activity of the mutated transposases was tested in the germline  $w^{p^{ch}}$  reversion assay (17). Sons of the genotype  $w^{p^{ch}}/Y;mutant/+$  were crossed with  $w^-$  females, and the proportion of wild-type  $w^{p^{ch}}$  revertants was scored in the daughters. Possible interactions between wild-type and mutant transposases were examined similarly in *mutant/hsp70:Mos1-182* genotypes. To eliminate maternal and grandmaternal effects of wild-type transposase on the rate of  $w^{p^{ch}}$  reversion (17), mothers carrying the transposase mutations were crossed with fathers carrying *hsp70:Mos1-182*, and the rate of  $w^{p^{ch}}$  reversion was examined in progeny of the heteroallelic sons.

**PCR and Sequencing.** Mutations in the *hsp70:Mos1* gene were identified by DNA sequencing of PCR fragments as described (18). Each mutation was confirmed by sequencing PCR products from at least two independent PCR reactions and by sequencing both strands.

## RESULTS

**Isolation of Mutants.** Excision of the inactive *peach* element from  $w^{p^{ch}}$  by wild-type transposase is very efficient. The number of red ommatidia in each eye of  $w^{p^{ch}};hsp70:Mos1$  flies exceeds 500, even in the absence of heat shock. Hence, a screen using the *peach* excision assay can efficiently recover hypomorphic mutations in the transposase gene. The degree of somatic eye color mosaicism among the mutants could be classified into six categories, ranging from 0 through 100–200 colored ommatidia per eye. Although semiquantitative, the levels of somatic mosaicism are suitable for preliminary screening because they accurately reflect the quantitative levels of  $w^{p^{ch}}$  excision in the germline ( $r = 0.59$ ,  $P < 0.01$  for

all mutants and  $r = 0.52$ ,  $P < 0.05$  for mutants with  $>0$  spots per eye; data not shown). Altogether, the EMS mutagenesis yielded 20 different mutants with reduced or absent  $w^{p^{ch}}$  eye color mosaicism. With the exception of one insertion and one deletion, the mutations were named according to the mutation as identified by DNA sequencing. In addition to the EMS mutations, a D284E replacement converting D,D(34)D into D,D(34)E was created by site-directed mutagenesis. Unexpectedly, transformants carrying D284E yielded no somatic mosaicism of  $w^{p^{ch}}$ .

**Germline Activity of Transposase.** The germline activity of mutant transposases was assayed as the proportion of wild-type revertants among the female progeny of  $w^{p^{ch}}/Y;mutant/+$  males. The results are summarized in Fig. 1 according to the position of the mutation. The reversion rates are based on an average of 450 flies per cross (range 250–1000). The bars, centered at the mean, span  $\pm 1$  binomial standard error. For comparison with the data in Fig. 1, the germline reversion rate in *hsp70:Mos1-182* males is 14.7 percent ( $n = 2269$ ). Corrected for multiple statistical tests, all of the mutants except one (M110I) have a significant effect on the reversion rate. Some yield about 5% wild-type revertants, others about 1–3 percent, many others 0 percent. Among the missense mutations, approximately an equal number completely eliminated germline excision as retained some residual activity. Excision assays carried out on a larger scale might well reveal that some of the apparently inactive elements support a low level of germline excision, because they show a small amount of somatic mosaicism.

The positions and identities of the mutations are indicated across the bottom of Fig. 1. For orientation, above the mutants are listed a number of landmarks in the *mariner* transposase, including the initiator methionine at position 1; the carboxyl residue glutamic acid at 345; the D,D(34)D motif at 156, 249, and 284; and a bipartite nuclear localization signal at 131–133 and 142–146, inferred from a strong similarity with that in a zebrafish *Tc1*-like element (11). The symbol  $\Delta 61$  bp refers to a mutation (*EMS-117*) in which the translational start region, including the methionine initiation codon, is deleted. As might be expected, this mutation behaves as a loss-of-function allele. The symbol +31 bp refers to another mutation (*EMS-96*), a 31-bp insertion of complex structure (see below) that creates a terminator codon; this also behaves as a loss-of-function mutation. Other than W159stop, which is a terminator at codon 159, and M1K, in which two adjacent nucleotides of the methionine initiation codon are altered (ATG  $\rightarrow$  AAA), the rest of the mutations are single-substitution missense mutations. A number of the mutations show a low level of somatic mosaicism (0–1 or 1–5 spots per eye) but no germline reversion (R12Q, S99N, A102V, G292R, and G325E); the limited somatic activity suggests that the mutant transposase retains activity but is too weak to produce germline revertants among a few hundred progeny. Note also that D284E, which shows no somatic mosaicism, also fails to show activity in the germline.

**Effects on Wild-Type Transposase Activity.** Complementation tests were performed among all of the mutants and between the mutants and nonmutant *hsp70:Mos1*. None of the mutations complemented each other in the  $w^{p^{ch}}$  mosaicism assay. However, in heteroallelic combinations with the nonmutant *hsp70:Mos1* transgene, the majority (12/18) of the hypomorphic missense mutations show a net, statistically significant, decrease in the rate of germline excision. The germline reversion rate of  $w^{p^{ch}}$  observed in the progeny of *mutant/hsp70:Mos1* flies is shown for each of the mutants in Fig. 2. There is no significant correlation ( $r = 0.04$ ) between the effect of a mutant *hsp70:Mos1* transgene when present alone (*mutant/+*) and when it is present in a heteroallelic combination with a nonmutant transgene (*mutant/hsp70:Mos1*). In Fig. 2, the bars are centered at the mean and span  $\pm 1$  binomial standard error. The horizontal line indicates

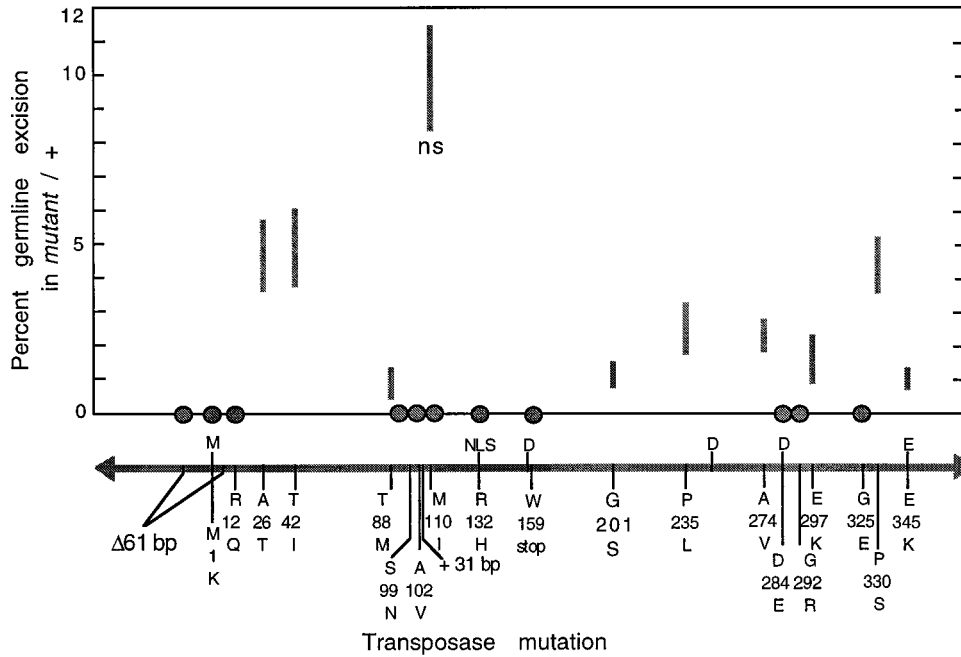


FIG. 1. Germline excision of the *peach* element from  $w^{pCh}$  in males bearing a single copy of the mutant *hsp70: Mos1* transgene. The bars represent the mean  $\pm$  1 binomial standard error; dots indicate no germline excision. The nature and position of the mutations in the *hsp70: Mos1* sequence are depicted along the horizontal axis. The triangles at the ends represent the inverted repeats. M at 1 and E at 345 are the amino and carboxyl ends of the putative transposase. NLS is part of a nuclear localization signal, and the aspartates of the D,D(34)D motif are indicated. The bar labeled ns is not significantly different from the control value of 14.7%. The others are all significant even when adjusted for multiple tests.

a level of reversion equal to the control *hsp70: Mos1-182/+* value of 14.7%. Corrected for multiple statistical tests, nine mutants showed no difference from the reversion rate in the controls; these are the bars without asterisks. The 12 mutants marked with asterisks differ significantly from the control value. The single asterisks indicate significance at the 5% level, and double asterisks indicate significance at the 1% level.

**Dominant-Negative Complementation.** In Fig. 2, the three mutations marked with double asterisks (G201S, D284E, and G292R) also showed a level of germline reversion significantly smaller than even one-half of the value observed in controls. Such a severe reduction is suggestive of dominant-negative complementation between mutant and nonmutant transposase subunits. Compared with the control value of  $14.7 \pm 0.7$  percent, these mutants, when heteroallelic with a nonmutant

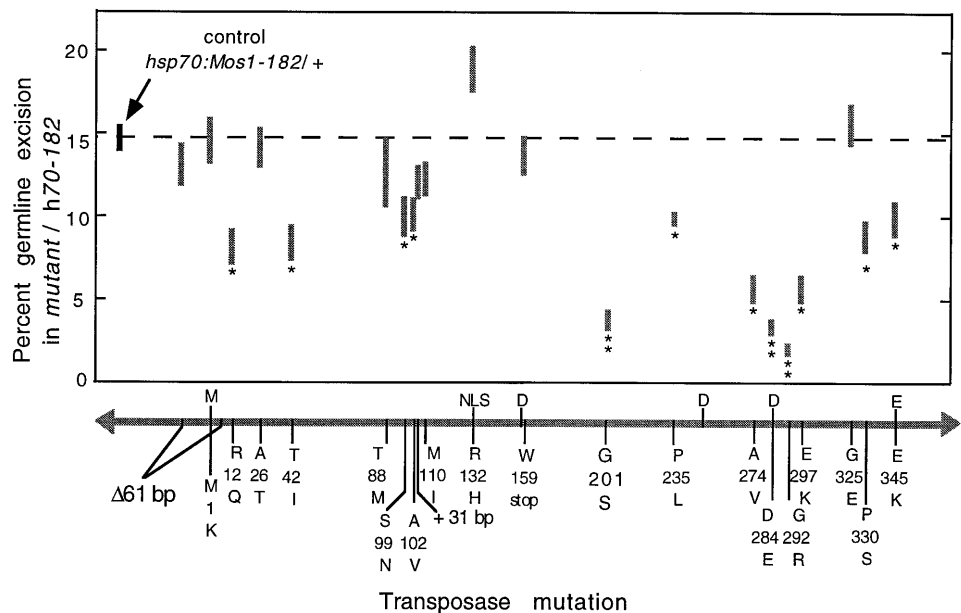


FIG. 2. Germline excision of the *peach* element in heteroallelic *mutant/hsp70: Mos1-182* males laid out as in Fig. 1. The horizontal line indicates a value equal to that observed in *hsp70: Mos1-182/+* controls (14.7%). Unmarked bars are not significantly different from the control value; single asterisks, significantly less than control value; double asterisks, highly significantly less than control value (and also significantly less than one-half the control value). Significance levels corrected for multiple tests.



transposase must enter the nucleus after synthesis in the cytoplasm. A specific bipartite nuclear localization signal, consisting of two clusters of basic amino acids, has been identified in the transposase of the *TcI* family through an element isolated from the zebrafish (11). The *Mos1* element contains a very similar bipartite nuclear localization signal that includes the residues Arg-Arg-Lys at 131–133; hence, the R132H transposase probably is unable to excise the *peach* element because it fails to enter the nucleus.

The D,D(35)E triad of acidic amino acids is present within the active site of numerous recombinases and transposases that mediate phosphoryl transferase reactions, including *mariner* transposase, which has the alternative motif D,D(34)D. In bacterial transposases, the catalytic site functions in two chemical steps, donor cleavage and strand transfer. For the nonreplicative transposition of *Tn10*, the donor cleavage step consists of two single-strand cleavage reactions at each end of the element that proceed sequentially (10). In *Tn7* transposase, failure of the second of the two cleavage reactions by a D114A mutation in the D,D(35)E motif switches the mode of transposition from nonreplicative to replicative (24). In *mariner*, a conservative site-directed change of D,D(34)D to the *TcI* consensus D,D(34)E completely obliterates transposase function. This result is astonishing in view of the apparent common ancestry of the *mariner* and *TcI* families. If the transposable elements share a common ancestry, the change in the ion-binding domain in the *mariner* lineage must have been accompanied by at least one other amino acid replacement that rendered the Asp at position 284 essential. One speculative possibility is the insertion of Tyr at position 285, which is absent in *TcI*. The Asp-Tyr signature is characteristic of the *mariner* family and may have arisen early in its evolution (25). This finding alone might suggest that Glu and Asp-Tyr are optimal in their respective domains, but this hardly would suggest that the single-step mutational intermediates would be nonfunctional.

An additional unexpected effect of D284E is the very strong dominant-negative complementation, comparable to that of G292R, implying a physical interaction between wild-type and mutant transposase subunits (18). All three mutations showing strong dominant-negative complementation (G201S, D284E, and G292R) are located within a 92-residue stretch near the carboxyl end of the transposase, which includes the D(34)D region of the putative catalytic domain, reinforcing the case for

subunit poisoning as the mechanism of the dominant-negative complementation.

We thank Nancy Kleckner for discussions, David Sullivan for advice on isolation of the EMS mutants, Peter Goss for statistical analysis, and Nathan Blow for assistance in DNA sequencing. This work was supported by National Institutes of Health Grant GM33741.

1. Doak, T. G., Doerder, F. P., Jahn, C. L. & Herrick, G. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 942–946.
2. Robertson, H. M. (1995) *J. Insect Physiol.* **41**, 99–105.
3. Henikoff, S. (1992) *New Biol.* **4**, 382–388.
4. Kulkosky, J. K., Jones, S., Katz, R. A., Mack, J. P. G. & Skalka, A. M. (1992) *Mol. Cell. Biol.* **12**, 2331–2338.
5. Engelman, A. & Craigie, R. (1992) *J. Virol.* **66**, 6361–6369.
6. Mizuuchi, K. (1992) *Annu. Rev. Biochem.* **61**, 1011–1051.
7. Craig, N. L. (1995) *Science* **270**, 253–254.
8. Baker, T. A. & Luo, L. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 6654–6658.
9. Kennedy, A. K. & Haniford, D. B. (1995) *J. Mol. Biol.* **256**, 533–547.
10. Bolland, S. & Kleckner, N. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 7814–7818.
11. Ivics, Z., Izsvak, Z., Minter, A. & Hackett, P. B. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 5008–5013.
12. Vos, J. C. & Plasterk, R. H. A. (1994) *EMBO J.* **13**, 6125–6132.
13. van Luenen, H. G. A. M., Colloms, S. D. & Plasterk, R. H. A. (1993) *EMBO J.* **12**, 2513–2520.
14. Bigot, Y., Hamelin, M.-H., Capy, P. & Periquet, G. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 3408–3412.
15. Lohe, A. R., Moriyama, E. N., Lidholm, D.-A. & Hartl, D. L. (1995) *Mol. Biol. Evol.* **12**, 62–72.
16. Medhora, M., Maruyama, K. & Hartl, D. L. (1991) *Genetics* **128**, 311–318.
17. Lohe, A. R., Lidholm, D.-A. & Hartl, D. L. (1995) *Genetics* **140**, 183–192.
18. Lohe, A. R., Sullivan, D. T. & Hartl, D. L. (1996) *Genetics* **144**, 1087–1095.
19. Barik, S. (1993) *Site-Directed Mutagenesis by Double Polymerase Chain Reaction*. (Humana, Totowa, NJ), pp. 277–286.
20. Rubin, G. M. & Spradling, A. C. (1982) *Science* **218**, 348–353.
21. Karess, R. E. & Rubin, G. M. (1984) *Cell* **38**, 135–146.
22. Capy, P., David, J. R. & Hartl, D. L. (1992) *Genetica (The Hague)* **86**, 37–46.
23. Lohe, A. R. & Hartl, D. L. (1996) *Mol. Biol. Evol.* **13**, 549–555.
24. May, E. W. & Craig, N. L. (1996) *Science* **272**, 401–404.
25. Robertson, H. M. & Asplund, M. L. (1997) *Insect Biochem. Mol. Biol.*, in press.