11-1-2001

The Functions of the Multiproduct and Rapidly Evolving \textit{dec-1} Eggshell Gene Are Conserved Between Evolutionarily Distant Species of Drosophila

James Charles Badciong  
\textit{Marquette University}

Jeffrey Michael Otto  
\textit{Marquette University}

Gail L. Waring  
\textit{Marquette University}, gail.waring@marquette.edu

The Functions of the Multiproduct and Rapidly Evolving dec-1 Eggshell Gene Are Conserved Between Evolutionarily Distant Species of Drosophila

James C. Badciong
Department of Biochemistry, Medical College of Wisconsin
Department of Biology, Marquette University
Milwaukee, WI

Jeffery M. Otto
Genaissance Pharmaceuticals
New Haven, CT

Gail L. Waring
Department of Biology, Marquette University
Milwaukee, WI

Abstract: The Drosophila dec-1 gene encodes multiple proteins that are required for female fertility and proper eggshell morphogenesis. Genetic and immunolocalization data suggest that the different DEC-1 proteins are functionally distinct. To identify regions within the proteins with potential biological significance, we cloned and sequenced the D. yakuba and D. virilis dec-1 homologs. Interspecies comparisons of the predicted translation products revealed rapidly evolving sequences punctuated by blocks of conserved amino acids. Despite extensive amino acid variability, the proteins produced by the different dec-1 homologs were functionally interchangeable. The introduction of transgenes containing either the D. yakuba or the D. virilis dec-1 open reading frames into a D. melanogaster DEC-1 protein null mutant was sufficient to restore female fertility and wild-type eggshell morphology.
Normal expression and extracellular processing of the DEC-1 proteins was correlated with the phenotypic rescue. The nature of the conserved features highlighted by the evolutionary comparison and the molecular resemblance of some of these features to those found in other extracellular proteins suggests functional correlates for some of the multiple DEC-1 derivatives.

How proteins interact to form complex supramolecular architectures is a central question in both cell and developmental biology. We have been studying formation of the eggshell in Drosophila in an attempt to reveal some of the molecular strategies and motifs that are used to ensure the orderly assembly of this extracellular structure in vivo. During late oogenesis the ovarian follicle cells secrete proteins in a defined temporal order into the space between the oocyte and overlying follicle cells where they assemble into the multilayered eggshell. The oocyte proximal vitelline membrane layer forms during stages 8–10, followed by formation of the inner chorionic layer and tripartite endochorion layer during stages 11–14. The assembly process is surprisingly complex, entailing regulated gene expression and temporally regulated extracellular proteolytic cleavages (Noguerón and Waring 1995; Pascucciet al. 1996; Noguerónet al. 2000), as well as protein trafficking between the eggshell layers (Pascucciet al. 1996; Trougakos and Margaritis 1998; Noguerónet al. 2000).

The identification of eggshell structural genes necessary for proper eggshell assembly has been facilitated by the recovery of female sterile mutants, which display aberrations in eggshell morphology (Diganet al. 1979; Bauer and Waring 1987; Komitopoulouet al. 1988; Savant and Waring 1989). A large number of independent mutations have been linked to the defective chorion -1 (dec-1) locus. The dec-1 gene produces three alternatively spliced mRNAs that encode proproteins of 106, 125, and 177 kD (fc106, fc125, and fc177, respectively, where fc denotes its follicle cell origin). The proproteins are secreted by the follicle cells and become localized in the vitelline membrane layer, where they are cleaved to at least five distinct protein products (Noguerón and Waring 1995). Although generated in the vitelline membrane layer, during late oogenesis the DEC-1 derivatives relocate in a product-specific manner to different eggshell layers as well as to different regions within the same eggshell layers (Noguerónet al. 2000). The distinct spatial distributions of the DEC-1 products suggest functional multiplicity and are consistent with
the multiple \textit{dec-1} functions inferred by genetic data (Waring \textit{et al.} 1990; Mauzy-Melitz 2001).

As a first step toward identifying biologically significant regions within the DEC-1 products, we initiated an interspecific comparison of the \textit{dec-1} gene from \textit{Drosophila melanogaster}, \textit{D. yakuba}, and \textit{D. virilis}. Interspecies comparisons of Drosophila homologs have been used to reveal highly conserved amino acid sequences with potential importance for the structure and function of several proteins (Newfeld \textit{et al.} 1991; Kula \textit{et al.} 1995; Hanna-Rose \textit{et al.} 1997; Burris \textit{et al.} 1998; Takahashiet \textit{al.} 1999). Subsequent mutagenesis experiments targeting the conserved regions validated the utility of such comparative evolutionary analyses (Hanna-Rose \textit{et al.} 1997; Takahashiet \textit{al.} 1999). In this article we cloned and sequenced the \textit{D. yakuba} and \textit{D. virilis} \textit{dec-1} homologs. While the overall organization of the genes is conserved, pairwise comparisons showed extensive sequence divergence within the coding regions. Despite its rapidly evolving nature, highly conserved stretches of amino acids within the DEC-1 products were identified. Although limited, the conserved features appear to be sufficient for function because a highly diverged \textit{D. virilis} \textit{dec-1} transgene was able to rescue the eggshell and female-sterile phenotypes of a \textit{D. melanogaster} DEC-1 protein null mutant.

\textbf{Materials and Methods}

\textit{Drosophila strains}

\textit{D. melanogaster} wild-type (P2) and \textit{fs(1)410} \textit{dec-1} mutant strains have been described previously (Bauer and Waring 1987). The \textit{D. yakuba} and \textit{D. virilis} flies were obtained from the Bowling Green Stock Center (Bowling Green, OH) and the \textit{w;P\{ry\textsuperscript{+}\Delta2-3\}(99B)} flies carrying an endogenous source of transposase activity were obtained from W. Engels (Robertson \textit{et al.} 1988).

\textit{Construction and screening of genomic libraries}

\textit{D. yakuba} genomic DNA was isolated as previously described (McGinnis \textit{et al.} 1983), digested with \textit{BamHI}, and size fractionated by sucrose gradient centrifugation (10–40\%) in a SW41 rotor for 22 hr at 22,000 rpm. Fractions containing DNA fragments in the 12–20 kb size range were pooled, precipitated, ligated to \textit{\lambda} Gem 11 \textit{BamHI} arms.
(Promega, Madison, WI), and packaged according to the manufacturer's instructions. LE 392 host cells were infected with phage from the unamplified library and plated using standard methods (Sambrook et al. 1989). Phage plaques were transferred to Colony/Plaque Screen Hybridization Transfer Membranes (New England Nuclear Research Products, Boston) and hybridized according to the manufacturer's instructions with a subcloned *D. melanogaster* cDNA fragment. The 0.8-kb *D. melanogaster* fragment, derived from the 3' end of the fc106 transcript (Waring et al. 1990), was radiolabeled by nick translation and hybridized to the membranes at 45°.

*D. virilis* genomic DNA was isolated as described by Ashburner (1989), digested with *Hind*III, and size fractionated by electrophoresis in 0.8% agarose gels. DNA fragments in the 9–23 kb size range were excised from the gel and purified using the Prep-A-Gene DNA Purification System (Bio-Rad Laboratories, Hercules, CA). λ arms with *Hind*III ends and central stuffer DNA with *Xho*I ends were purified from a *Hind*III/*Xho*I digest of λ DASHII vector (Strategene, La Jolla, CA), using the Prep-A-Gene kit. Following ligation to the *Hind*III genomic DNA fragments, recombinant DNA was packaged using Gigapack III gold packaging extract (Stratagene). XL1 Blue MRA P2 host cells were then infected with phage from the unamplified library. Plating, phage transfer, and hybridization were performed as described above except that the membranes were hybridized at 65° with a *D. virilis* cDNA probe homologous to the 3' end of the *D. melanogaster* fc106 transcript (see below).

**Isolation of *D. virilis* cDNA probe**

Total RNA was isolated from 20 *D. virilis* stage 9–10 egg chambers as previously described (Spradling and Mahowald 1979). The RNA was ethanol precipitated, resuspended in 3.0 μl of water treated with diethyl pyrocarbonate, and incubated at 68° for 5 min. The RNA template was added to a commercially prepared reverse transcription reaction mix (Ready-To-Go T-primed first strand kit; Pharmacia Biotech, Piscataway, NJ) and incubated for 1 hr at 37°. cDNA synthesis was primed with a *Not*I(dT)18 primer [5’-AACTGGAAGAATTCGCGG(dT)18-3’] contained within the reaction mix. The cDNA reaction mix (10 μl) was amplified in a 100-μl volume containing 50 mm KCl, 10 mm Tris-HCl, pH 9.0, 1% Triton X-100, 2.15 mm MgCl₂, 0.3 mm of each dNTP, 2.5 units of Taq polymerase, and
10–50 pmol each of a forward *D. psuedoobscura* dec-1 primer (5’-GGCGGCAACAAGCAGAAGTCCAAG-3’) and a reverse T-Ad-2 primer (5’-AATCTGGAAGAATTCCGG-3’; Pharmacia Biotech). Using the *D. psuedoobscura* forward primer and an internal *D. psuedoobscura* reverse primer (5’-TGTCTGAAATGCGTCGCCCTCGCCATCCTCTGGTGCTG-3’), a 350-bp *D. virilis* product was amplified from the reverse transcription (RT)-PCR reaction as follows: 1 time (3 min at 95°), 5 times (1 min at 94°; 30 sec at 37°; 30 sec at 45°; 30 sec at 50°; 1 min at 72°), 25 times (1 min at 94°; 1 min at 55°; 2 min at 72°), and 1 time (1 min at 94°; 1 min at 55°; 5 min at 72°). On the basis of DNA sequence obtained from the 350-bp PCR product, a 1.1-kb *D. virilis* cDNA fragment was then amplified from the RT-PCR reaction above using a *D. virilis*-specific forward primer, 5’-TAGACCGAGCTATGGAACCAG-3’, and the T-Ad-2 reverse primer. PCR reactions utilizing these homologous primers were performed as follows: 1 time (5 min at 94°; 1 min at 55°; 2 min at 72°), 29 times (1 min at 94°; 1 min at 55°; 2 min at 72°), and 1 time (1 min at 55°; 5 min at 72°).

**DNA sequencing**

For the *D. yakuba* gene, subcloned restriction fragments and unidirectional deletions of genomic subclones were manually sequenced by the Sanger dideoxy chain termination method. Plasmid subclones containing the unidirectional deletions were made with exonuclease III, using the Erase-a Base system (Promega). Synthetic oligonucleotides were used to prime regions that were not covered by the overlapping deletion subclones. Deletion subclones and an array of oligonucleotide primers were used to sequence the *D. virilis* dec-1 homolog. An automated DNA sequencer (ABI, Foster City, CA) housed at the University of Wisconsin-Milwaukee was used to acquire this sequence data. RT-PCR products spanning the four introns in the *D. virilis* gene were sequenced to establish the exon borders. Computer-assisted sequence analysis and comparisons were done using programs of the University of Wisconsin genetics computer group and Internet resources.

**dec-1 transgenes**

A *D. yakuba* dec-1 transgene consisting of the coding sequence and ~2.8 kb of its 5′ and 1.4 kb of its 3′ flanking DNA was assembled.
from genomic subclones of a 16-kb BamHI dec-1-containing D. yakuba fragment. Briefly, an 8.5-kb dec-1-containing EcoRI/BamHI fragment was subcloned into Bluescript SK (+/−) phagemid. A XhoI/EcoRI-subcloned fragment containing 2.5 kb of additional 5′ flanking DNA was inserted into the Bluescript subclone. The 11-kb XhoI/BamHI fragment was excised and subcloned into a XhoI/BglI cut pCaSperR 4 P-element transformation vector.

A chimeric dec-1 transgene consisting of 2 kb of D. melanogaster 5′ flanking DNA, D. virilis coding sequence, and ~5 kb of D. virilis 3′ flanking DNA was constructed from a mixture of PCR products and genomic clones (Figure 1A). A 14-kb D. virilis XbaI-EcoRV genomic fragment was excised from a subcloned 16-kb HindIII fragment [pGEM-7Z vector (Promega)] and inserted into a pCaSpeR4 P-element transformation vector digested with XbaI and StuI. The 14-kb XbaI-EcoRV fragment contained the dec-1 coding sequence and ~0.6 kb of 5′ and 6 kb of 3′ flanking DNA. A 2.2-kb chimeric dec-1 SacII fragment consisting of ~2 kb of 5′ flanking DNA from D. melanogaster and 180 bp of D. virilis DNA was inserted into a unique SacII site located ~110 bp 3′ of the ATG translation initiation codon in the 14-kb XbaI-EcoRV subclone (Figure 1A). The chimeric SacII fragment was created by abutting D. melanogaster and D. virilis sequences in the 5′-untranslated region (UTR). A SmaI restriction site was engineered into a forward primer complementary to D. virilis 5′-UTR sequences (positioned ~70 bp 5′ of the ATG initiation codon; 5′-CCAGTCCGGGAAAGCG-3′). Using a reverse primer complementary to sequences 170–187 bp 3′ of the initiating ATG (5′-TGCTCACCTTTCTCCTGTG-3′), a 250-bp product was amplified. The 250-bp product was digested with SmaI and SacII, yielding a 180-bp SmaI/SacII fragment that was subcloned into SmaI/SacII-cut pBlueScript. The D. melanogaster 5′ flanking DNA was amplified from a 2.5-kb genomic fragment subcloned in pGEM-5Z (Promega). The forward primer was complementary to plasmid polylinker sequence and contained an engineered XhoI site (5′-GGCGGCGCTCGAGTGATATC-3′), while the reverse primer was complementary to sequences in the 5′-UTR located 53–70 bp 5′ of the initiating ATG codon (5′-CGGAGCTCGTTGGCCGAC-3′). After the ends were blunted with T4 DNA polymerase, the 2-kb amplified product was digested with XhoI and inserted into the D. virilis SmaI/SacII subclone, making use of a XhoI site in the pBlueScript polylinker region and the engineered SmaI site in the 5′-UTR. A 2.2-kb chimeric SacII fragment
was excised from pBluescript using a SacII site in the *D. melanogaster* PCR product derived from the pGEM-5Z poly linker region and the natural SacII site in the *D. virilis* coding region. The insert was subcloned into the unique SacII site of the recombinant XbaI-EcoRV pCaSpeR vector.

![Diagram of construction of dec-1 transgenes](image)

**Figure 1.** Construction of *dec-1* transgenes. (A) A 14-kb *D. virilis* genomic fragment containing the *dec-1* coding region with 0.6 kb of 5′- and ~6 kb of 3′-flanking DNA is shown in the top line. The right arrow indicates the position of the transcription initiation site; the solid circle marks the end of the gene. A 2.2-kb chimeric SacII DNA fragment (shown below) was inserted into the SacII site of the *D. virilis* genomic fragment. The chimeric fragment consisted of ~2 kb of 5′-flanking DNA and 22 bp of 5′-UTR DNA including the transcription initiation site (large arrow) from the *D. melanogaster dec-1* gene (solid line) and 180 bp of *D. virilis dec-1* DNA (dashed line) that included the ATG translation initiation codon (small right arrow). Blunt-ended PCR products from *D. melanogaster* were joined to *D. virilis* PCR products cut at an engineered SmaI(M) site. Xb, XbaI; S, SacII; RV, EcoRV; X, XhoI; M, SmaI. The asterisks indicate sites derived from plasmid poly linker regions. (B) A schematic of the coding region from a *D. melanogaster dec-1* transgene carrying a small deletion (ΔS456–E473). Primer sets 1, 2 and 3, 4 were used to amplify 1.2- and 1.1-kb PCR products, respectively, which were joined via a synthetic SphI site, which was subsequently removed to yield the in-frame 18-amino-acid deletion clone. The triangle above the line indicates the deleted region (not to scale); the triangles below the line indicate the positions of the *dec-1* introns. X, XhoI; A, ApaI; Xba, XbaI.

A mutant *D. melanogaster dec-1* transgene containing a small internal deletion (ΔS456–E473) was created (Figure 1B), using the PCR-based strategy described by Hughes and Andrews (1996). Briefly, a 1.2-kb PCR product abutting the S456 codon and bearing an engineered SphI site at its 3′ end was subcloned into pGEM7z (forward primer 5′-TGGCCGATGATGCGACG-3′; reverse primer 5′-CGTCGAGCATGCAGAGTCTGATCCTGG-3′). Primer sets 1, 2 and 3, 4 were used to amplify 1.2- and 1.1-kb PCR products, respectively, which were joined via a synthetic SphI site, which was subsequently removed to yield the in-frame 18-amino-acid deletion clone. The triangle above the line indicates the deleted region (not to scale); the triangles below the line indicate the positions of the *dec-1* introns. X, XhoI; A, ApaI; Xba, XbaI.

*Genetics*, Vol. 159, No. 3 (November 2001): pg. 1089-1102. [DOI](https://doi.org/10.1534/genetics.159.3.1089). This article is © Genetics Society of America and permission has been granted for this version to appear in e-Publications@Marquette. Genetics Society of America does not grant permission for this article to be further copied/distributed or hosted elsewhere without the express permission from Genetics Society of America.
2.2-kb XhoI-ApaI fragment bearing the engineered SphI site was excised from pGEM7z and exchanged with its wild-type counterpart in a dec-1 genomic clone. A 2.85-kb XhoI-XbaI fragment was isolated from the mutated genomic subclone and exchanged with its wild-type counterpart in a pCasPeR 4 P-element rescuing vector that contained the dec-1 gene and 1.9 kb of 5′ and 1 kb of 3′ flanking DNA (Noguerón et al. 2000). To remove ectopic SphI bases, the deletion construct was digested with SphI, the 4-bp 3′ overhangs were removed with Klenow fragment, and the ends were religated. The precision of the deletion was confirmed by DNA sequencing.

The dec-1 transgenes were injected into w;P{ry+Δ2-3}(99B) preblastoderm embryos using standard methods (Spradling and Rubin 1982). The endogenous source of transposase was removed from transformed flies by genetic crosses, chromosomal linkage of the dec-1 transgene was determined by segregation analyses, and homozygous “transgene” stocks were established. For functional analyses, the dec-1 transgenes were crossed into fs(1)y410 dec-1 null mutant females. Egg chamber proteins were solubilized and analyzed by Western blot analysis as previously described (Noguerón and Waring 1995).

Morphological analysis

Stage 14 egg chambers were isolated from hand-dissected ovaries, fixed in a trialdehyde fixative (Kalt and Tandler 1971), washed in cacodylate buffer containing 0.2 m sucrose, postfixed in Zetterquist's 2% osmium tetroxide, and dehydrated through a graded series of alcohols. Dehydrated samples were embedded in Spurr's low viscosity resin. Thin sections were cut with glass knives, stained with uranyl acetate and lead citrate, and examined with a Hitachi H-600 transmission electron microscope (University of Wisconsin Imaging Facility, Milwaukee, WI).

Results

Cloning dec-1 homologs

Probes from the 5′ and 3′ ends of the D. melanogaster (Dm) dec-1 gene were cross-hybridized to a single 16-kb BamHI genomic fragment in D. yakuba (Dy). A small size-selected BamHI genomic
library was created and screened with a heterologous \textit{Dm} probe from the 3’ end of the \textit{dec-1} gene. Positive recombinant phage were recovered, and an 8.5-kb cross-hybridizing \textit{BamHI-EcoRI} fragment was subcloned into a plasmid vector. After verifying that the cloned \textit{Dy} fragment hybridized to ovarian transcripts of the expected size and temporal specificity, the \textit{dec-1} coding region was sequenced in its entirety.

\textit{D. virilis} (\textit{Dv}) genomic fragments failed to hybridize with \textit{Dm} \textit{dec-1} probes. To generate a homologous \textit{Dv} probe, \textit{dec-1} DNA sequences from more closely related species were compared to identify short highly conserved regions to which hybridizable PCR primers could be designed. By aligning the \textit{D. yakuba} and \textit{D. melanogaster} genes along with segments from a partially sequenced \textit{D. pseudoobscura} \textit{dec-1} homolog (Andersson 1994), two highly conserved stretches of DNA were revealed. Primers complementary to these regions were used to amplify a 350-bp \textit{Dv} fragment. Briefly, oligo(dT)-primed cDNAs were synthesized from total RNA isolated from \textit{Dv} stages 9 and 10 egg chambers, the stages of maximal \textit{DEC-1} protein synthesis in \textit{Dm} (Bauer and Waring 1987), and \textit{Dv} egg chambers (Otto 1995). Using the cDNAs as template, two PCR products in the 1-kb size range were amplified with a \textit{D. pseudoobscura} 5’ \textit{dec-1} primer and the T-Ad-2 3’ primer described in materials and methods. The 350-bp \textit{Dv} fragment was amplified from the RT-PCR products using \textit{D. pseudoobscura} 5’ and 3’ \textit{dec-1} primers. Partial sequencing of the PCR product revealed an open reading frame (ORF) that contained a repeating motif (GQ G/S YGQXY)\textsubscript{6} followed by a stretch of 18 amino acids, which showed 94% identity to a stretch of \textit{D. pseudoobscura} sequences in the fc106 C-terminal region. A homologous 1.1-kb \textit{Dv} probe was then amplified from the RT-PCR reaction described above, using an internal 5’ \textit{Dv} primer complementary to nucleotides within the nonrepetitive 18-amino-acid stretch and the T-Ad-2 3’ primer.

To isolate the \textit{D. virilis} \textit{dec-1} homolog, a size-selected \textit{Dv} \textit{HindIII} fragment subgenomic library was created and screened with the 1.1-kb \textit{Dv} probe. Several positive phage with identical restriction maps were recovered. After Northern blot analysis verified that the phage contained sequences complementary to \textit{dec-1} RNA, a 16-kb \textit{dec-1} fragment was subcloned into a plasmid vector. Smaller
subcloned genomic fragments and a subclone of the 1.1-kb cDNA described above provided the framework for DNA sequencing.

Genomic organization

The major structural features of the *D. yakuba* and *D. virilis* dec-1 genes, including the transcription and translation initiation sites, exons and introns, alternative splice sites, stop codons, and poly(A) addition sites, were inferred from the DNA sequence data based on their alignment with the well-characterized *D. melanogaster* homolog (Waring et al. 1990). Sequence derived from an oligo(dT)-primed cDNA fragment was used to place the poly(A) addition site in the *Dv* homolog. Since a consensus poly(A) addition signal (AAUAAA) was not evident in the 3′-UTR, the authentic poly(A) addition site may be distal to the proposed site. For the *Dv* homolog, all intron-exon borders were verified by sequencing RT-PCR products that spanned the junctions. As shown in Figure 2, the general organization and positions of the introns have been conserved in all three species. In *D. virilis* intron 3 is ~2000 nucleotides or 10 times larger than its *D. melanogaster* and *D. yakuba* counterparts, while intron 4, which encodes the fc177-specific C terminus (Figure 3A), is ~500 bp shorter. Despite size differences, intron 3 shows striking conservation at the nucleotide level around and including the two alternative 3′ splice acceptor sites (Figure 2). The *Dy* and *Dm* genes are identical over a stretch that extends 45 nucleotides 5′ of the 3′-most AG. The large *Dv* intron has not been sequenced in its entirety; nevertheless, 36 of the 38 nucleotides sequenced in this region are conserved.

Evolution of the DEC-1 proteins

In *D. melanogaster*, the dec-1 gene produces three alternatively spliced RNAs that encode three proproteins, fc106, fc125, and fc177. As shown in Figure 3A, the proproteins have identical N termini (944 amino acids) but different C termini. The fc106-specific C terminus has 6 amino acids; the fc125-specific C terminus is 138 amino acids in length, and the fc177-specific C terminus has 604 amino acids. Upon secretion, the proproteins become localized in the assembling vitelline membrane, where they are cleaved in a stage-specific manner to five distinct mature products (Noguerón 1996). The mature products derived from fc106, the most abundant proprotein, are shown at the bottom of Figure 3A.
fc 106 and its cleaved derivatives

The dec-1 ORF in all three species is headed by a signal peptide with a predicted cleavage site that yields proproteins with a N-terminal glutamine residue (Figure 3B). In D. melanogaster the fc106 proprotein begins with amino acid 16 and ends at position 950. Dm fc106 is secreted from the ovarian follicle cells in stage 9 and 10 egg chambers. During late stage 10, fc106 is cleaved within the vitelline membrane to a 25-kD N-terminal derivative, s25, and an 80-kD C-terminal derivative, s80. As oogenesis progresses, s25 becomes localized in the vitelline membrane, inner chorionic, and endochorion layers of the eggshell (Noguerón et al. 2000). We defined the borders of the s25 region by the putative signal sequence cleavage site on the left and by the N terminus of s80 on the right. Alignment of the Dm s25 region [amino acids (aa) 16–280] with its Dy and Dv counterparts revealed amino acid identity at 74 and 43.5% of the positions, respectively. The conserved amino acids within the s25 region were asymmetrically distributed with a disproportionate number near the N terminus. In particular a highly conserved stretch of amino acids (Dm 70–118) showing 81% identity with both the Dy and Dv homologs was highlighted by the three-way comparison. Although the alignment in general indicates extensive sequence divergence between the Dm and Dv homologs, the salient features of the region (e.g., its A/P-rich amino acid content and acidic nature) have been conserved. The high alanine and proline content of the Dm and Dv s25 regions (34 and 33%, respectively) is typical of vitelline membrane components. With isoelectric points of 3.86 and 4.09 (Dm and Dv, respectively), the acidic nature of the s25 segment has also been retained through millions of years of evolution.

The s80 proprotein, generated in late stage 10 egg chambers, remains in the vitelline membrane layer until late stage 13 when it is cleaved to a 20-kD N-terminal product, s20, and a 60-kD C-terminal derivative, s60 (Figure 3A). s20 exits rapidly from the vitelline membrane layer and is taken up by the oocyte where it remains localized in membrane-bound vesicles (Noguerón et al. 2000). Unlike s80, s60 was refractory to Edman degradation; therefore the C-terminal border of s20 could not be delimited by N-terminal sequencing of s60. Since s80 is cleaved extracellularly, a cyclized N-terminal glutamine residue would be consistent with a naturally blocked s60 N terminus. We have tentatively placed the C-terminal
border of s20 next to a conserved Q residue (Dm Q453) on the basis of the inferred size of s20 from SDS-PAGE and the analysis of a deletion mutant discussed later. Alignment of the putative s20 regions (e.g., Dm S281–D452) shows that in all three species the s20 segment is headed by amino acids that fall within a strictly conserved block (Dm A278–V288) that includes the fc106 cleavage site. Overall, the Dy and Dm s20 segments show 86.3% identity, while the Dy and Dm segments are identical at 54.8% of the positions. The s20 segment is extraordinarily rich in charged amino acids. In D. melanogaster, 17% of the amino acid residues are either lysine or arginine and 14% are either glutamic acid or aspartic acid. Approximately equal numbers of acidic (17) and basic (14) residues occur in an alternating pattern in the amino-terminal half; in the C-terminal half basic residues predominate (17 basic, 5 acidic). Despite extensive sequence divergence in the Dy homolog, the basic nature of the segment has been retained (overall pI 10.36 vs. 10.06 in Dm), as has the distribution of the acidic and basic residues within the segment. In the Dv homolog, equal numbers of acidic (14) and basic (14) residues alternate in the N-terminal half; again the C-terminal half is predominately basic (15 K/R vs. 5 E/D). The conservation of the large number and pattern of charged residues within this segment suggests that its ionic nature may be important in mediating protein-protein interactions in the assembling eggshell. While the s20 segment is not retained in the eggshell after its cleavage from s80 in late stage egg chambers, it may help dictate early protein interactions that are crucial for orderly assembly of the eggshell. Although cleaved from s80, the s20 region is retained in the C-terminal fc125 derivative, s95 (Noguerón and Waring 1995).
Figure 2. Genomic organization of dec-1 genes isolated from different Drosophila species. (A) Intron-exon structure and comparative restriction maps of genes isolated from *D. melanogaster* (*Dm*), *D. yakuba* (*Dy*), and *D. virilis* (*Dv*). The positions of the introns are indicated below each restriction map. S, SstI; X, XhoI; B, BamHI; Hc, HincII; H, HindIII; C, ClaI. The scale at the bottom represents 1 kb. (B) Nucleotide sequence conservation surrounding the alternative splice acceptor sites in intron 3. Identical nucleotides in all three species are highlighted in gray; the alternative AG 3′ splice acceptor sites are underlined.

The s60 derivative, defined in *D. melanogaster* as Q453–Q950, has three distinct segments: a short N-terminal region (Q453–Q500), a large central region consisting of perfect and imperfect copies of a 26-amino-acid repeat (R501–Q809), and a 140-amino-acid C-terminal segment. Global alignments of the N-terminal segments show the *Dm* and *Dy* homologs are identical at 79% of the positions while the *Dm* and *Dv* segments are 50% identical. The conserved block of amino acids that heads this region (K<sup>457</sup>–K<sup>481</sup>) is rich in charged residues. Acidic residues outnumber basic residues in the *Dm* homolog (D/E, 21%; K/R, 17%) while basic residues predominate in the *Dv* homolog (K/R, 24%; D/E, 20%). The highly charged nature of this region again suggests a role in protein-protein interactions. While removal of this amino acid block does not appear to be deleterious to *dec-1* function, as we show later, its removal influences DEC-1 protein processing.

The repeat region that encompasses approximately two-thirds of the s60 protein is extraordinarily rich in glutamine and methionine. The *Dm* repeat contains 39 and 13% glutamine and methionine, respectively, while the more divergent *Dv* repeat contains 33%...
glutamine and 12% methionine. On the basis of its glutamine-rich nature it has been proposed (Waring et al. 1990) that this central repeat region may be involved in stabilizing the eggshell via glutamyl-lysine crosslinks much like involucrin, another glutamine-rich protein, stabilizes the keratinocyte envelope (Eckert and Green 1986). Manual alignment of the Dm, Dy, and Dv repeat regions revealed a reoccurring consensus motif (RQW, S/T, E, E/D, QAK, I/A, QQ) that was highly conserved not only between the species but also within the degenerate repeats of each species.
Figure 3. Evolutionary comparison of the DEC-1 proteins. (A) Open reading frames in the Dm dec-1 gene and its three primary translation products, fc106, fc125, and fc177, are denoted by rectangles. The lines within the dec-1 schematic indicate the positions of the 5′- and 3′-UTRs and intron sequences. The fc106-specific C terminus is indicated in black, the fc177-specific C terminus is indicated in gray, and the fc125-specific C terminus is denoted by the stippled rectangle. The hatched box indicates a small region of the ORF that is common to fc125 and fc177. The lines below the fc106 proprotein indicate the regions that correspond to its three cleaved derivatives: s25, s20, and s60. The lightly shaded rectangle within the common N-terminal region demarcates a region consisting of tandem repeats of a 26-aminoacid glutamine-rich sequence. (B) Alignment of the conceptual translation of the fc106 region from D. melanogaster (Dm), D. yakuba (Dy), and D. virilis (Dv). The downward arrow shows the predicted signal sequence cleavage site; the upward arrow between Dm A280 and S281 indicates the presumptive fc106 cleavage site based on N-terminal sequencing of the cleaved C-terminal derivative, s80 (Waring et al. 1990). The repeat region was aligned manually and is offset. Residues that are identical between all three species are highlighted in black while residues that are identical in two of the three species are highlighted in gray. Gaps in the alignment are indicated by dashes. The unique C terminus of fc106 is underlined. The line above S456–E473 shows the amino acids removed in the ΔS456–E473 deletion construct. (C and D) Alignments of the fc177- and fc125-specific C termini, respectively. The line in C denotes sequences that are common to fc125 and fc177 (residues 946–992 in Dm).

Global alignment of the C termini (Dm G810–Q950) revealed 78% identity between Dm and Dy and 36% identity between Dm and Dv. Most of the conservation in the Dv region fell within a block of 29 residues at the C terminus, of which 25 were identical. Although 5 of
the identical residues are encoded by nucleotides contained within alternatively spliced sequences, 19 of the residues are encoded by nucleotides 5′ of the common 5′ donor site of intron 3. This suggests that beyond constraints that may be imposed at the nucleotide level, there may also be functional constraints at the amino acid level. The Dv homolog contains six copies of a repeat (GQ, G/S, Y/S, GQXY) that is missing in the Dm and Dy homologs.

**fc177-specific ORF**

fc177, the largest DEC-1 proprotein, is synthesized in stage 11 and 12 egg chambers and becomes localized in the vitelline membrane layer where it is cleaved to a C-terminal derivative, s85, which consists largely, if not entirely, of fc177-specific C-terminal sequences. After its biogenesis, s85 is released from the vitelline membrane and becomes localized within the endochorionic cavities or spaces (Noguerónet al. 2000). s85 is the only proteinaceous component that has been localized within this chemically ill-defined region. Global alignment of the fc177-specific C termini (Figure 3C) revealed two highly conserved features that may be related to the function and/or localization of s85. The N-terminal region of s85 contains four pairs of strictly conserved cysteine residues (CXC, where X is any amino acid) within a stretch of 36 amino acids (Dm 1007–1042). Aside from a single cysteine residue in the Dm fc106 C terminus, all of the cysteines in all three species are found in the s85 region. The cysteine-rich region in s85 is striking in relation to other endochorion components. Six major endochorion components have been characterized to date: s15, s16, s18, s19, s36, and s38 (Griffen-Sheaet al. 1980; Levine and Spradling 1985; Martinez-Crzuzaet al. 1988; Konsolakiet al. 1990; Swimmeret al. 1990; Vlachouet al. 1997), all of which are devoid of cysteine residues. This suggests that if the DEC-1 cysteine residues are involved in intermolecular disulfide bond formation, homologous interactions or interactions with underlying extracellular matrix components are the most likely. The cysteine-rich region is followed by a 65-amino-acid stretch that shows 94% identity between the Dm and Dv homologs. The region (Dm 1052–1116) is extraordinarily rich in acidic amino acids (17% E; 6% D), which is reflected in its pi of 4.0. Global alignments of the fc177-specific sequences revealed their rapidly evolving nature. Whereas the Dm and Dy regions were approximately the same length (Dm, 603 amino acids; Dy, 627 amino acids) and showed 76% overall identity, the Dv region (455 amino acids) was
punctuated by several gaps ranging in size from 2 to 32 amino acids and displayed only limited homology (35% identity) with its *Dm* counterpart. Small blocks of sequences enriched in basic amino acids (*e.g.*, *Dm* 1451–1466 and *Dm* 1566–1571) and the serine/threonine-rich content of the fc177 C terminus (21 and 24% *Dm* and *Dv*, respectively) were conserved among these otherwise extensively diverged sequences.

**Figure 4.** Restoration of eggshell morphology in a *dec-1* null mutant by a chimeric *Dv dec-1* transgene. (A and B) Transmission electron micrographs (TEM) of late stage 14 egg chambers showing the tripartite nature of the endochorion found in *Dm* (A) and *Dv* (B). (C) TEM of a late stage 14 eggshell from a *Dm* female homozygous for the *fs(1)410 dec-1* null mutant allele. (D) TEM showing restoration of normal endochorion morphology in a late stage 14 egg chamber from *fs(1)410* null mutants carrying the chimeric *Dv dec-1* transgene. Ex, exochorion; en, endochorion; vm, vitelline membrane; icl, inner chorionic layer.

**fc125-specific ORF**

The fc125 proprotein accumulates in the vitelline membrane layer in early stage 10 egg chambers. During late stage 10 it is cleaved to a 95-kD C-terminal derivative, s95, which consists of s80 and fc125-specific C-terminal sequences. The s95 derivative appears to have an essential function since females that lack s95 are sterile (Mauzy-Melitz 2001). Global alignment of the fc125-specific C termini (Figure 3D) shows 77% identity between *Dy* and *Dm* with a *Dy*-specific insertion of 19 amino acids. In the 138 residues of overlap, the *Dm* and *Dy* sequences are identical at 84% of the positions. The C
terminus of the *D. virilis* homolog extends 75 amino acids beyond the ends of its *Dy* and *Dm* counterparts. In the region of overlap, the *Dv* and *Dm* sequences show 41% identity. Charged acidic residues (E/D, 40%) are predominant in the *Dv* C-terminal extension.

**Functional identity of the *Dm*, *Dy*, and *Dv* homologs**

The three-way comparison of the *D. melanogaster*, *D. yakuba*, and *D. virilis* dec-1 genes revealed several blocks of conserved amino acids interspersed among rapidly evolving sequences. Multiple products and multiple genetic functions have been associated with the *Dm* dec-1 locus; therefore it was of interest to determine whether the *Dy* and *Dv* homologs would display the full array of genetic functions in a *Dm* genetic background. Toward this end, *Dy* and *Dv* dec-1 transgenes were introduced into *D. melanogaster* fs(1)410 DEC-1 protein null mutants. DEC-1 protein null mutants are female sterile and produce a defective inner chorionic layer (Margaritis *et al.* 1991) and a disorganized endochorion, which eventually collapses into the underlying vitelline membrane layer during late oogenesis (Bauer and Waring 1987). Preliminary studies showed that 1.9 kb of 5′ and 1 kb of 3′ flanking DNA were sufficient to drive normal expression of a *Dm* dec-1 transgene in the fs(1)410 genetic background and that expression of the transgene was sufficient to rescue the female-sterile phenotype (G. L. Waring, unpublished data). When a *D. yakuba* dec-1 transgene driven by its own regulatory sequences was crossed into fs(1)410 females, fertility was restored. This result was not unexpected considering that the amino acid alignments of the *Dy* and *Dm* dec-1 subregions (Figure 3B) displayed identities that ranged from 74 to 86%. More intriguing were the results with the *D. virilis* transgene where the identities in the subregions ranged from a low of 35% (fc177-specific C terminus) to a high of 55% (s20 region). The *Dv* dec-1 genomic clone contained the dec-1 gene along with ~0.6 kb of 5′ and 5 kb of 3′ flanking DNA. To ensure that the *Dv* transgene contained sufficient 5′ flanking sequence for proper expression, we created a chimeric dec-1 clone consisting of ~2 kb of 5′ flanking DNA from the *Dm* gene linked to the *Dv* dec-1 coding region and 3′ flanking DNA. The *Dm* and *Dv* sequences were merged in the 5′-UTR as described in materials and methods. The chimeric *Dv* transgene rescued the fs(1)410 female-sterile phenotype and restored normal eggshell morphology. Figure 4 (A and B) shows the tripartite nature of the endochorion in wild-type stage 14 egg chambers and Figure 4C...
shows its disruption in homozygous fs(1)410 mutants. Figure 4D shows that restoration of an organized endochorion was correlated with the restoration of genetic function in fs(1)410 mutants containing the chimeric Dv transgene.

**Figure 5.** *D. virilis* DEC-1 proteins are processed in *D. melanogaster* egg chambers. Western blot analysis of SDS-soluble egg chamber proteins from wild-type *D. melanogaster* (lanes 1–3), *Dm fs(1)410* mutant females carrying the chimeric *Dv dec-1* gene (lanes 4–6), and wild-type *D. virilis* (lanes 7–9) incubated with Cfc106 antiserum is shown. The stages are indicated at the top. Fifteen egg chambers/ lane were used in lanes 1–3; the remaining lanes represent 40 egg chambers/lane. The positions of the well-characterized *Dm* DEC-1 proteins (fc106 and its s80 and s60 C-terminal derivatives; fc125 and its s95 C-terminal derivative) are indicated as well as their inferred Dv counterparts. The "extra" bands between fc106 and s95 and s80 in wild-type stage 10 egg chambers are believed to represent fc106-processing intermediates. The intensities of the bands are extremely variable between different stage 10 egg chamber preparations. In addition, these bands comigrate with fc106-processing intermediates that are prominent in *dec-1* mutants that lack the fc106-s80 cleavage site (Badciong 1999).

To verify that normal expression and processing of the Dv DEC-1 proteins were correlated with the phenotypic rescue, staged egg chambers from wild-type *D. melanogaster, D. virilis*, and *fs(1)410* mutants containing the chimeric Dv transgene were collected and analyzed by Western blot analysis. SDS-soluble proteins were separated by SDS-PAGE, transferred to polyvinyl difluoride membranes, and reacted with Cfc106 antiserum. The Cfc106
antiserum is directed toward epitopes in the \textit{Dm} s60 region that stretch from D693 in the central repeat to Q950 at the C terminus (Noguerón and Waring 1995). Since this serum cross-reacts with \textit{Dv} DEC-1 products, we used the Cfc106 antisera to follow the accumulation of \textit{Dv} DEC-1 proteins in the DEC-1 protein null mutant, \textit{fs(1)410}. Lanes 1–3 in Figure 5 show the accumulation of the \textit{Dm} fc125 and fc106 proproteins and their processed derivatives (s95, s80, and s60, respectively) in wild-type egg chambers. Lanes 7–9 show that comparable products accumulate in \textit{D. virilis} egg chambers at the expected times. The slight differences in the mobilities of the \textit{Dm} and \textit{Dv} products observed are consistent with predicted size differences based on the DNA sequence data. Lanes 4–6 show that substantial levels of DEC-1 proteins of \textit{Dv} origin (\textit{i.e.}, fc125, fc106, s95, s80, and s60) accumulate in \textit{fs(1)410} mutant egg chambers carrying the chimeric transgene.

\textbf{Normal processing of s80 requires a block of highly conserved sequences near its site of cleavage}

The interchangeability of the \textit{D. melanogaster} and \textit{D. virilis} dec-1 genes indicates that the functionally significant features of the DEC-1 proteins have been conserved through evolution. To begin functional dissection of the \textit{D. melanogaster} dec-1 gene we focused on the blocks of amino acids that are highly conserved in the \textit{Dv} homolog.

Processing of the \textit{Dv} fc106 and s80 substrates in the \textit{fs(1)410} females suggests that the target sites recognized by the proteolytic machinery responsible for these cleavages have been conserved. Consistent with this interpretation is the block of 11 strictly conserved amino acids around and including the fc106-s80 cleavage site (\textit{A}^{280}/\textit{S}^{281} in \textit{Dm}). The s80-s60 cleavage site has not been determined but the general region has been surmised on the basis of the sizes of the cleavage products. A highly conserved block of amino acids within the predicted region (\textit{K}^{457}–\textit{E}^{473}) suggested a potential target site for the proteolytic machinery. To probe its functional significance and hopefully create a noncleavable substrate for functional analyses, a \textit{D. melanogaster} dec-1 transgene missing amino acids 456–473 was introduced into \textit{fs(1)410} mutant females. Lanes 1 and 2 in Figure 6A show that except for the expected size difference due to the deleted amino acids, processing of fc106 to s80 in the wild-type and transgenic females was indistinguishable. The appearance of a 60-kD band in late...
stage egg chambers from the transgenic females (Figure 6A, lane 6) suggests either that the normal s80-s60 cleavage site lies outside of this block of deleted amino acids or that in its absence a nearby cryptic site can be utilized. The similar sizes of the 60-kD bands in the wild-type and transgenic egg chambers (Figure 6A, lane 5 vs. 6) are compatible either with positioning of the s80 cleavage site C-terminal to the deleted amino acids or with utilization of a cryptic cleavage site in the mutated substrate. Since a subtle size difference between the wild-type and mutant s60 products was resolved on longer gels (Figure 6B, lane 1 vs. 3), it appears that a cryptic cleavage site is used in the transgenic stage 14 egg chambers. In any case there were no discernible functional consequences associated with the removal of these amino acids. Both eggshell morphology (data not shown) and fertility were restored in *fs(1)410* females carrying the mutant transgene. Interestingly, when the mutant transgene was expressed in the presence of wild-type DEC-1 proteins an additional band was clearly resolved in the 60-kD size range (Figure 6B, lane 2). The mobility of the broad upper 60-kD band (lane 2) is consistent with the combined mobilities of the 60-kD derivatives derived from the wild-type *dec-1* gene (lane 1) and the mutant transgene (lane 3). The faster migrating band in lane 2 indicates that the wild type, mutant, or both s80 substrates is (are) cleaved at an additional site when both forms are present. This suggests that cleavage of s80 to s60 involves intermolecular interactions and that a heteromeric association between the mutant and wild-type forms can alter the site of cleavage.
Figure 6. Abnormal processing of Dm s80 in wild-type egg chambers carrying a mutated dec-1 (ΔS\textsuperscript{456–E\textsuperscript{473}}) transgene. (A) Western blot of SDS-soluble proteins from staged egg chambers isolated from wild-type (FMO/w; lanes 1, 3, and 5) and fs(1)410 dec-1 null mutants carrying two copies of the mutated dec-1 (ΔS\textsuperscript{456–E\textsuperscript{473}}) transgene (lanes 2, 4, and 6) incubated with Cfc106 antiserum. The egg chamber stages are indicated at the tops of the lanes; the positions of fc106 and its C-terminal derivatives, s80 and s60, are also indicated. (B) Western blot of SDS-soluble proteins from stage 14 egg chambers isolated from wild type (FMO/w; lane 1), FMO/w females containing two copies of the mutated dec-1 (ΔS\textsuperscript{456–E\textsuperscript{473}}) transgene (lane 2), and fs(1)410 dec-1 null mutants carrying two copies of the mutated dec-1 (ΔS\textsuperscript{456–E\textsuperscript{473}}) transgene (lane 3). The positions of the s80 and s60 wild-type derivatives are shown. The position of the extra band in the mutant stage 10 egg chambers is compatible with a fc106-processing intermediate.

Discussion

Insights into functionally significant domains within proteins are often obtained by searching protein databases for homologous sequences with known function. BLAST searches with different subregions of the dec-1 gene failed to reveal meaningful similarities; thus functional insights were not forthcoming. Since functional domains are often conserved evolutionarily, we undertook an interspecies comparison to identify dec-1 subregions with potential functional significance. D. virilis and D. melanogaster are separated by ~60 million years of evolution (Beverly and Wilson 1984), sufficient time for unconserved sequences to diverge extensively. A number of D. virilis genes have been compared with their D. melanogaster counterparts.
counterparts and, while a high degree of variation in the extent of conservation has been reported, most proteins are identical at 70–80% of the amino acid residues (Yao and White 1991; Clark and Elgin 1992; Zhou and Boulianne 1994; Kulaet al. 1995; Lintermannet al. 1998; Ousleyet al. 1998; Price and Lai 1999; Takahashiet al. 1999). The DEC-1 derivatives with amino acid identities ranging from a low of 35% (s85) to a high of 55% (s20) are among the most highly diverged Drosophila proteins studied to date. Most D. virilis genes have been isolated from genomic libraries on the basis of their complementarity to D. melanogaster DNA probes. However, Schmid and Tautz (1997) have shown that greater than one-third of D. melanogaster clones randomly selected from a cDNA library fail to hybridize with D. virilis genomic DNA. This suggests that the Drosophila genome harbors a substantial proportion of genes with a very high divergence rate. Our failure to detect hybridizing sequences with a variety of D. melanogaster probes from different dec-1 subregions is consistent with its rapidly evolving nature and provided us with an opportunity to analyze a member of this underrepresented set.

Extensive divergence in rapidly evolving genes suggests either that the vast majority of amino acid substitutions have no functional consequences or that the amino acid changes reflect the evolution of species-specific functions. The transformer (tra) gene is one of the most rapidly evolving genes among Drosophila species. Despite extensive divergence at the amino acid level (36% identity between Dm and Dv), transgenic analyses have shown that the Dv tra gene can rescue at least some of the functions of the native Dm gene (O’Neill and Belote 1992). The functional interchangeability of the Dv and Dm tra homologs, although limited, suggests that the rapid evolution is best explained by low functional constraints. Comparisons of interspecific and intraspecific patterns of evolution and polymorphisms at the Drosophila tra gene support this interpretation (McAllister and McVean 2000). The complete rescue of the female-sterile and eggshell phenotype of Dm fs(1)410 mutants with the Dv dec-1 transgene suggests that, like tra, dec-1 has few functional constraints and that although the gene is evolving at a rapid rate, the changes are neutral in function. The diverse locations of the mature DEC-1 derivatives (Noguerónet al. 2000) indicate that they are involved in many different protein interactions. The functional equivalence of the Dm and Dv homologs suggests that this repertoire of protein-protein interactions is also conserved.
Interspecies comparisons have shown that known functional domains are among the regions that are most highly conserved between Dv and Dm homologs (Kula et al. 1995; Lintermann et al. 1998; Ousley et al. 1998; Price and Lai 1999). The conserved features of the dec-1 gene highlighted by our three-way comparison include the modular design of s25 with its alanine- and proline-rich central region, the charged nature of s20, the reoccurring RQWS/TEE/DQAKI/AQQ motif in s60, and the paired cysteines and acidic cluster in s85.

The modular structure of s25 with its proline-rich central region is reminiscent of the small proline-rich proteins that function as cross-bridging proteins in the cornified cell envelope (CE; Steinert et al. 1998a,b). The CE is a specialized structure formed beneath the plasma membrane of terminally differentiating epithelial cells. Like the mature eggshell, the CE provides a physical and chemical barrier that protects the underlying tissue from environmental insults. The CE is an amalgam of several structural proteins that are rendered insoluble by the formation of isopeptide bonds catalyzed by transglutaminases (Steinert 1995). The small proline-rich proteins (SPRs) are a family of related proteins that contain central proline-rich repeats flanked by N and C termini rich in glutamine and lysine residues. By joining themselves via N-ε-(glutamyl)-lysine isopeptide bonds to other CE structural proteins, the small proline-rich proteins serve as cross-bridging proteins. Only glutamine and lysine residues in the N- and C-terminal domains are crosslinked in vivo (Steinert et al. 1998b). The central proline-rich region with its repetitive β-turns (Candi et al. 1999) is thought to function as a spacer, which allows the SPRs to act as flexible cross-bridges between the structural proteins that constitute the backbone of the CE. While the SPRs are involved in stabilization of the CE, several observations suggest that s25 may play a similar role in eggshell assembly:

1. The widespread distribution of s25 in the mature eggshell, including the endochorion, the inner chorionic layer, and the vitelline membrane, is compatible with a role in interlayer protein-protein interactions (Noguerón et al. 2000).
2. dec-1 null mutants fail to form an organized endochorion and the molecular aggregates that do form in the endochorion layer fall into the underlying vitelline membrane layer in late stage 14 egg chambers. dec-1 null mutants carrying mutant transgenes that fail to generate normal s25-like products form an organized
but unstable endochorion that eventually collapses into the underlying vitelline membrane (Badciong 1999; Mauzy-Melitz 2001).

3. The alanine, proline-rich central segment may promote protein-protein interactions by creating a flat hydrophobic surface (Williamson 1994) upon which other eggshell proteins interact. All of the proteins encoded by the known chorion genes as well as the sV23, sV17, and VM32E vitelline membrane proteins are rich in proline and alanine residues (Waring 2000).

4. Using conserved glutamine residues in the regions that flank the central alanine, proline-rich core, s25 may crosslink with itself or with other eggshell proteins that can provide electron-donating lysine substrates. s25 becomes crosslinked into the assembling eggshell during late oogenesis. Essentially devoid of cysteine and tyrosine residues that have been implicated in the crosslinking of the vitelline membrane and chorion layers, respectively (Waring 2000), isopeptide bond formation catalyzed by transglutaminases offers a means by which s25 could become crosslinked. s25, with two conserved lysines, the early endochorion proteins s36 and s38, with several lysine residues embedded in highly conserved central regions (Konsolakiet al. 1990), and other dec-1 derivatives are the most likely lysine donors. As previously noted, the major vitelline membrane proteins as well as the late chorion proteins are lysine poor (Waringet al. 1990).

Although the global alignments indicated that the fc177-specific C terminus s85 is the most highly diverged dec-1 region, it also contains the largest (65 aa) and most highly conserved stretch of amino acids (\(Dm\ E^{1052-1116}\)) in the dec-1 open reading frame. This highly acidic stretch is preceded by a cysteine-rich region embedded in a basic subdomain (\(Dm\ \text{aa} 985-1051, \text{pI} 12.6\)). While the \(Dm\) and \(Dv\) homologs display very little sequence identity in the basic subdomain beyond the highly conserved cysteine residues, its basic nature is conserved (\(Dv\, \text{pI} 13.06\)). The N terminus of s85 with its basic and acidic subdomains resembles the N-terminal domain of type XI and V collagens. The N-terminal propeptides of the type X1 and V collagens are headed by a module rich in basic residues followed by a variable acidic subdomain and a short triple helix. Unlike the major fibril-forming collagens in which the N-terminal propeptide is removed from the main triple helix, collagens V and XI do not undergo complete
cleavage of their N-terminal extensions (Moradi-Ameli et al. 1994; Rousseau et al. 1996). Type XI and V collagens appear to play a role in regulating fibril diameter (Liet al. 1995; Vikkula et al. 1995; Marchant et al. 1996) and the N-terminal noncollagenous sequences seem to be important for this regulatory role. Alternative splicing, involving the exons that encode the acidic subdomain, has been implicated in the regulation. It has been suggested that differential expression of these acidic domains may be important in modulating interactions between collagen XI and other components of the extracellular matrix, thereby influencing heterotypic collagen assembly (Tsumaki and Kimura 1995; Moradi-Ameli et al. 1998).

Like the N terminus of collagens V and XI, the N terminus of s85 may play an important role in regulating a morphogenetic process. dec-1 null mutants carrying a dec-1 transgene in which the fc177-specific C terminus (s85) was removed by the introduction of a premature stop codon failed to organize a tripartite endochorion (Mauzy-Melitz 2001). Rather, aggregates of electrondense chorionic material accumulated in the space between the inner chorionic layer and overlying exochorion. Since s85 localizes exclusively within the endochorion cavities or spaces during normal eggshell morphogenesis (Noguerón et al. 2000), and spaces do not form in its absence, s85 appears to provide a critical structural framework around which the tripartite endochorion forms. Despite the reduced size and extensive divergence of the Dv homolog, the critical features of s85 appear to have been conserved since the Dv dec-1 transgene was able to organize a tripartite endochorion in D. melanogaster. Aside from two small stretches of amino acids near the C terminus, the acidic subdomain is the only other region that displayed identity at six or more contiguous residues in the three-way interspecies comparison. Like the acidic subdomain in the noncollagenous N-terminal sequences of the type V and XI collagens, the acidic domain near the N terminus of s85 may be important in modulating its interactions with underlying components of the extracellular matrix.

Among the highly conserved blocks of amino acids highlighted by the interspecies dec-1 comparison was an 18-amino-acid segment in the region of the inferred s60 N terminus. Although the removal of this conserved segment was without functional consequence, distinct s80 processing patterns were observed depending upon whether the (ΔS\(^{456–473}\)) transgene was introduced into a wild-type or dec-1 null...
genetic background. The distinct processing patterns suggest that the cleavage of s80 is dependent upon intermolecular s80 protein interactions. While it has not been established whether s80 is cleaved by an autocatalytic mechanism or by proteolytic enzymes within the vitelline membrane, intermolecular interactions have been shown to play an important role in both types of processes. Proper cleavage of procollagen by procollagen aminoprotease is dependent upon intermolecular interactions. Deletions in the pro-α segment of procollagen that destabilize the triple helix structure result in cleavage at cryptic rather than the normal processing sites (Arnold et al. 1998). Proteolytic UmuD-like mutagenesis proteins self-process via intermolecular interactions (McDonald et al. 1998). It is believed that UmuD substrate molecules are cleaved within the catalytic cleft of UmuD enzyme molecules and that proper positioning of the cleavage site within the cleft is dependent upon the formation of UmuD “filament-dimers” (McDonald et al. 1999). Our results suggest that although the deleted amino acids are not necessary for intermolecular s80 interactions, different proteolytic cleavage sites are presented to or utilized by the proteolytic machinery, depending upon the nature of the s80 dimer/oligomer. The emergence of an electrophoretically distinguishable form of s60 in wild-type flies with the (ΔS456–E473) transgene correlates with the potential to form heterotypic dimers or oligomers between the wild-type and mutant s80 molecules.

In summary, dec-1 plays a central role in eggshell assembly. A complex series of evolutionarily conserved processing events produces an array of distinct protein products with different spatial distributions. Distinct molecular features of the DEC-1 derivatives have been highlighted by our evolutionary comparison of functionally interchangeable dec-1 genes. By mutagenizing highly conserved blocks of amino acids within the rapidly diverging gene, functional insights should be forthcoming. While the (ΔS456–E473) mutation described in this article was not functionally significant, its analysis revealed potential for previously unsuspected intermolecular interactions. In contrast, deletion of a second block of conserved amino acids (Badciong 1999) produced a dominant negative female-sterile phenotype (to be published elsewhere), thus underscoring the utility of targeting evolutionarily conserved sequences.
Acknowledgments

We thank Dr. Debra Mauzy-Melitz for the electron micrographs, Dr. Heather Owens for use of the imaging facility at the University of Wisconsin-Milwaukee, and members of our laboratory group for helpful discussions. This work was supported by a grant from the National Institutes of Health to G.L.W. (R15GM55952). J.C.B. was supported by predoctoral research fellowships from the Arthur J. Schmitt Foundation and Marquette University; J.M.O. was supported, in part, by a Marquette University Fellowship.

Literature Cited


Badciong J. C., 1999 An evolutionary and genetic analysis of DEC-1 protein function in Drosophila. Ph.D. Thesis, Department of Biology, Marquette University, Milwaukee.


Mauzy-Melitz D., 2001 Genetic dissection of eggshell assembly in *D. melanogaster*. Ph.D. Thesis, Department of Biology, Marquette University, Milwaukee.


Noguerón M. I., 1996 *Processing and distribution of dec-1 eggshell products in Drosophila melanogaster*. Ph.D. Thesis, Department of Biology, Marquette University, Milwaukee.


Takahashi Y., Hirose F., Matsukage A., Yamaguchi M., 1999 Identification of three conserved regions in the DREF transcription factors from


Waring G. L., Hawley R. J., Schoenfeld T., 1990 Multiple proteins are produced from the dec-1 eggshell gene in Drosophila by alternative RNA splicing and proteolytic cleavage events. *Dev. Biol.* **142**: 1–12.

