Molecular and Mutational Analysis of a Gelsolin-Family Member Encoded by the *flightless I* Gene of *Drosophila melanogaster*

Heinz Gert de Couet,*,† Keith S. K. Fong,* A. G. Weeds,‡ P. J. McLaughlin§ and George L. Gabor Miklos*†

*Department of Zoology, and Department of Genetics and Molecular Biology, The University of Hawaii, Honolulu, Hawaii 96822, †Integrative Biology Laboratory, Research School of Biological Sciences, Australian National University, Canberra City, ACT 2601, Australia, ‡Medical Research Council, Laboratory of Molecular Biology, Cambridge CB2 2QH, England, §Department of Biochemistry, University of Edinburgh, Edinburgh EH8 9XD, Scotland and *The Neuroscience Institute, San Diego, California 92121

Manuscript received April 10, 1995
Accepted for publication July 28, 1995

ABSTRACT

The *flightless* locus of *Drosophila melanogaster* has been analyzed at the genetic, molecular, ultrastructural and comparative crystallographic levels. The gene encodes a single transcript encoding a protein consisting of a leucine-rich amino terminal half and a carboxyterminal half with high sequence similarity to gelsolin. We determined the genomic sequence of the *flightless* landscape, the breakpoints of four chromosomal rearrangements, and the molecular lesions in two lethal and two viable alleles of the gene. The two alleles that lead to flight muscle abnormalities encode mutant proteins exhibiting amino acid replacements within the S1-like domain of their gelsolin-like region. Furthermore, the deduced intron-exon structure of the *D. melanogaster* gene has been compared with that of the *Caenorhabditis elegans* homologue. Furthermore, the sequence similarities of the *flightless* protein with gelsolin allow it to be evaluated in the context of the published crystallographic structure of the S1 domain of gelsolin. Amino acids considered essential for the structural integrity of the core are found to be highly conserved in the predicted *flightless* protein. Some of the residues considered essential for actin and calcium binding in gelsolin S1 and villin VI are also well conserved. These data are discussed in light of the phenotypic characteristics of the mutants and the putative functions of the protein.

The analysis of mutations that impair the flight abilities of *Drosophila melanogaster* and their rescue by germline transformation has not only provided a significant foundation for the understanding of early myogenesis, myofibrillar assembly, and the mechanics of muscle contraction, but has also led to significant insights into the neuronal and behavioral basis of motor control. Furthermore, the transgenic methodologies that have been brought to bear on the *Drosophila* genome have allowed rapid dissection of function at the genetic, molecular, and behavioral levels. In this context our laboratories have subjected one such locus, termed *flightless-I*, located on the X chromosome in sub-division 19F, to a comprehensive analysis at a number of different levels (Miklos and de Couet 1990). We have previously demonstrated by genetic complementation analysis that several independently isolated viable *flightless* mutations are all alleles of *flt-I* (Miklos and de Couet 1990). Furthermore, the *flightless* phenotype is revealed when a number of lethal mutations are made heterozygous with any of the viable alleles of the locus. In addition, individuals heterozygous for overlapping chromosomal deficiencies uncovering the locus during the pupal stage, indicating that the *flightless* gene product is also essential for the survival of the organism to adulthood. This lethality can be rescued by germline transformation of the appropriate constructs of the genomic *flightless* landscape.

At the ultrastructural level, several *flightless* mutants, while undergoing apparently normal embryonic and larval development, nevertheless exhibit abnormal flight muscle morphology. The flight muscle fibers of these viable *flightless* mutants show severely disrupted Z-discs and are often associated with “striated bundles” that have the structural appearance of Z-band material (Koana and Hotta 1978; Miklos and de Couet 1990). The myofibrils of affected muscles fray at their peripheries, indicating problems during myoblast differentiation. Although neither cell degeneration nor cell death is apparent, muscle architecture is abnormal and flies are incapable of flight.

At the molecular and cell biological levels, severe disruptions of the *flightless* transcription unit lead to death during embryogenesis (Miklos and de Couet 1990). Furthermore, germline clone analysis reveals that the embryo does not undergo the normal processes of cellularization. Although nuclei migrate to their positions in the egg cortex following 13 rounds of mitotic divisions, they lose their position at the periphery before cellularization is complete. Abnormal mesodermal invagination ensues; defective gastrulation and finally
death follows (Perrimon et al. 1989), indicating a maternal requirement for the flightless gene product.

The protein product encoded by the flightless gene consists of two distinct regions. The aminoterminus carries 17 repeats of a leucine-rich motif implicated in macromolecular interactions, while the carboxyterminal half of the molecule shares significant sequence similarity with the gelsolin-like family of actin severing and nucleating factors. We have isolated homologues of this gene from both the nematode Caenorhabditis elegans and from Homo sapiens (Campbell et al. 1993). The cytogenetic location of the human gene is known (Chen et al. 1995), and the locus is included in rearrangements that lead to clinical phenotypes.

Both the existence of lethal alleles of the flightless gene in D. melanogaster and the high degree of sequence conservation between the worm, fly and human homologues are indicators of a potentially central function of this molecule throughout the metazoa. We have therefore begun to subject this locus to a thorough molecular investigation, and we are studying the cell biological aspects of gene dosage effects in transgenic organisms, as well as analyzing various mutant alleles and making comparisons to similar molecules in other phyla. Detailed crystallographic information is now available for both gelsolin-like domains (McLaughlin et al. 1993; Markus et al. 1994) and for leucine-rich repeat domains (Kob and Deisenhofer 1993). In addition, the in vitro mutagenesis studies of the gelsolin gene may allow us to perform the analysis of D. melanogaster mutants into a molecular perspective.

In this communication we report the genomic sequence, the location of several chromosomal rearrangements affecting the locus, and the molecular lesions of two lethal alleles of the gene. We also examine the intron/exon boundaries of the gene in an evolutionary context. Finally, we present the DNA sequence analysis of two viable alleles of the gene. These data indicate that a region of the S1-like domain of the molecule implicated in the binding of a calcium ion in gelsolin may be responsible for the observed flightless phenotype.

MATERIALS AND METHODS

Stocks and chromosomes: Details of D. melanogaster stocks and the cytogenetic characteristics of the chromosomal rearrangements are as previously described (Schaefer and Leefvre 1976; Homyk and Sheppard 1977; Leefvre and Watkins 1976; Green et al. 1987; Miklos et al. 1987; Perrimon et al. 1989; Miklos and de Couet 1990). To avoid strain-specific polymorphisms, DNA was extracted from specially constructed stocks in which the rearrangement and lethal-bearing chromosomes were heterozygous with the wild-type chromosome from which they were derived. Thus, Df(1)Ga104, Df(1)Kc4263, and l(1)HC183 were placed over the M56a (Amherst) chromosome, whereas l(1)D14 (Eken et al. 1985) and Df(1)2/19B were heterozygous with their parental y neet-9^mei-41^9m chromosome. In cases where the parental chromosomes were unavailable, heterozygotes consisted of a lethal-bearing or a rearranged chromosome together with a Canton-S chromosome.

Molecular protocols: Genomic DNA from 0.5 to 3 g of adult female flies or from frozen heads of flies of appropriate genotypes, as well as DNA from bacteriophage A and from plasmids, was purified by standard methods (Sambrook et al. 1989). Two D. melanogaster genomic DNA libraries in bacteriophage A EMBL4 (a gift from Vincenzo Pirrotta) were screened with subcloned DNA fragments. The libraries consisted of size-selected embryonic DNA from either Canton-S or Oregon-R strains. Southern blotting, colony hybridization, plaque-lift techniques, and DNA hybridization were performed according to Sambrook et al. (1989). DNA sequencing utilized the chain-termination method (USB Sequenase kit) and d-32P-dATP. Compressions were resolved using dITP and the inclusion of 40% formamide in sequencing gels. Genomic DNA was sequenced on both strands by a combination of subcloning fragments of genomic DNA from phage A into M13mp10 phagemids followed by primer-directed second strand synthesis at specific sites using custom synthesized oligonucleotides. The analysis of translational products, database screening, and sequence alignments were performed using PCGENE software (IntelliGenetics Inc., Mountain View, CA) and the University of Wisconsin GCG suite of sequence analysis programs.

Northern blotting: Total cellular RNA was extracted from developmentally staged whole animals by routine methods (Churgin et al. 1979). poly(A)^+ RNA was selected from this fraction by oligo-dT-cellulose chromatography (Pharmacia poly A purification kit). Before electrophoresis, RNA samples were glyoxylated at 60°C for 1 hr and separated through 1% agarose gels in 10 mM phosphate buffer, pH 7.0 at 3 V/cm in a circulating buffer tank. The relative amounts of poly (A)^+ RNA in each sample were normalized by hybridization with a D. melanogaster adh probe.

Sequence analysis of flightless mutations: Specific regions of the flightless landscape from flies carrying the mutant alleles fil^- and fil^- (Miklos and de Couet 1990) were amplified as overlapping sections of the translated part of the transcription unit by the PCR. Oligonucleotide primer pairs of 25-30 nucleotides were used to cover the transcribed region. This strategy also incorporated all introns, as well as 120 nucleotides upstream of the translational start site. The appropriate DNA fragments were amplified by Taq polymerase (Perkin-Elmer, Norwalk, CT) using standard protocols for template denaturation, annealing, and primer extension. Amplified DNA products were isolated by agarose gel electrophoresis, purified with Gene Clean (Bio101, La Jolla, CA), and subsequently cloned into the pCRII plasmid vector (Invitrogen, San Diego, CA). Inserts of genomic DNA were completely sequenced using the flanking and internal primers on either strand. Any alterations in the DNA from the mutants were verified by reamplification of the respective target DNA using the PCR followed by direct sequencing of the amplified DNA with appropriate primers or sequencing of the cloned product.

RESULTS

Molecular and cytogenetic characterization of the flightless landscape: The various subregions of the base of the X chromosome were accessed by microdissection of subdivisions 19 and 20, by construction of a chromosomal mini-library in bacteriophage A (Miklos et al. 1988), and by mapping of the single-copy DNA inserts against three large deficiencies denoted Df(1)B57, Df(1)US54, and Df(1)C4, which together span the entire
region (Figure 1). Finer mapping was achieved for the flightless region by utilizing a panel of smaller chromosomal deficiencies: Df(1)16-129, Df(1)GA104, Df(1)2/19B, Df(1)17-257, and Df(1)GE263. This localization was achieved using quantitative Southern hybridizations involving DNA from individuals heterozygous for a given deficiency and DNA from appropriate nondeficiency controls. Chromosomal walks in subdivision 19F yielded a genomic landscape of the order of 250 kb (G. L. G. Miklos unpublished results). The four chromosomal deficiencies Df(1)17-257, Df(1)GA104, Df(1)GE263, and Df(1)16-129 delimited the flightless locus to an ~16-kb region (Figure 1).

We now present the breakpoint data for all the rearrangements within this landscape, determined by Southern hybridizations to genomic DNA from individuals carrying rearranged and normal chromosomes.

DNA from each of the different genotypes was cleaved with a variety of restriction endonucleases to determine the approximate location of the lesions (Figure 2). As well as the breakpoint determinations, we found that the lethal allele, denoted l(1)D44, represents a deficiency of ~400 bp. Furthermore, Df(1)2/19B is not only deficient for a region outside and more proximal to the landscape shown here but has an insert of ~5.4 kb, which is very probably due to a mobile element (data not shown). These results are summarized in the lower panel of Figure 1.

We also examined genomic DNA from a number of other flightless mutants, namely the lethals l(1)W2, l(1)HC183, and l(1)En3, as well as from two viable alleles, flil and flil0, but were unable to detect any changes within the boundaries of the locus using Southern analysis (data not shown).

**Transcription of the flightless locus:** Although the genomic region between the breakpoints of the deficiency chromosomes was found to encompass four independent transcription units, as defined by Northern blot analysis and by cDNA library screening, only one represents the flightless transcription unit as shown by germline transformation and rescue of the mutant phenotype (Campbell et al. 1993; G. L. G. Miklos unpublished results). A single 5.1-kb mRNA species emanates from the region as estimated from Northern analysis, which is in good agreement with the minimum message length of 4.7 kb (minus the poly(A) + tail) as estimated from alignment of the cDNAs with the genomic land-
scape. Northern blots of developmental stages of *D. melanogaster* probed with fragments X and Y revealed that this transcript is present throughout all life stages, with the highest abundance being in adult flies (data not shown).

**DNA sequence of the *flightless* genomic landscape:** To determine the relationship between the various cDNAs and the genomic landscape presented in this communication and to provide the foundation for analyzing the relevant regulatory regions, as well as an understanding of the of the transgenic rescue experiments, we sequenced Canton-S genomic DNA derived from the area corresponding to the two EcoRI fragments X, Y, comprising 5906 bp (Figure 3). The cDNA and genomic sequences were aligned and revealed the existence of three introns within the protein coding region of the gene. These introns separate codons 18/19, 1120/1121, and divide codon 1070 between the second and third base, respectively. (See also Figure 4). The nucleotide sequences at the intron/exon junctions conform to the *D. melanogaster* consensus (Mount *et al.* 1992). The DNA sequence from Oregon R strains reveals several differences between it and Canton-S (Figure 4) affecting the sequence and length of the third intron, and a double transition from GG to AA, affecting the third and first positions of codons 1067 and 1968. This leads to a conservative replacement of Ala encoded by codon 1068 with Thr in the Oregon-R strain. This position is occupied either by alanine in the human and fly sequences, and by threonine in the *C. elegans* homologue, and thus probably represents a selectively neutral protein polymorphism. A further length polymorphism was discovered in the untranslated trailer sequence, and five additional synonymous transitions were found clustered in the sequence between nucleotide positions 1396 and 1540. None of these affect the conceptual translation product of the transcription unit. Interestingly, the frequency and distribution of these polymorphisms is highly nonrandom. All the observed differences occur within regions of short nucleotide repeats (2-, 3-, 4-mers).

Analysis of the genomic sequence upstream from the 5' end of several cDNA clones revealed no regions with obvious sequence similarity to eukaryotic promoter consensus sites. However, a T/A-rich region upstream from the presumed transcription initiation site may substitute for the TATA box.

We have previously cloned cDNAs of the *C. elegans* homologue of the *flightless* gene (Campbell *et al.* 1993), and fortuitously, this homologue lies within the recently released 2.2 Megabases of the *C. elegans* genomic sequence project (Sulston *et al.* 1992; Waterston *et al.* 1992). This has allowed us to compare the genomic organizations of the *D. melanogaster* and *C. elegans* genes as well as to place them within the context of the predicted protein structures.

While the *D. melanogaster* gene is only split by three introns, its *C. elegans* homologue has 13, the first of which is in the 5' untranslated region (Figure 5). Of the 16 introns that occur in both genes, only one is held in common. Five of the remainder occur within the leucine-rich repeat domain, and, with the exception of intron IV, all of these split the individual leucine-repeats into subdomains. One intron is found near the carboxyterminal end of the leucine-rich-repeat-region, and an additional one is located within the bridging domain separating the gelsolin-like and leucine-rich-repeat regions of the protein coding sequence. Five additional introns occur within the gelsolin-like domains, which for the *flightless* gene we have denoted F2, F4, F5, and F6. We have thereby followed the conventional domain nomenclature adopted for gelsolin (G1-G6) and villin (V1-V6). Of the remaining gelsolin-like domains G4, G5, G6 are commonly thought to be a duplication product of G1, G2, G3. However, if the "introns early" hypothesis has validity (see discussions in Logsdon and Palmer 1994; Stolzefus *et al.* 1994), then in its simplest form the distribution of introns in the *flightless* gene makes this hypothesis unlikely. Other alternatives are that introns are a more recent acquisition ("introns late"), that F4, F5, and F6 represent captured domains added to an F1, F2, F3 core, or in fact that some introns were early and others later additions to a genome (Karn *et al.* 1983).

For these reasons we extended our search for introns held in common between different members of the gelsolin family to human villin and gelsolin. Complete genomic structural data are available only for villin (Pringault *et al.* 1991) and its presumed paralogous quail from *D. melanogaster* (S. Mahajan-Miklos and L. Cooley, personal communication), and none of the introns in the human genes coincide with those found in either the *D. melanogaster* or the *C. elegans* *flightless* homologues.

**Analysis of mutant alleles:** DNA sequence analysis of the two viable mutant alleles *flyO* and *flyI* revealed single base changes affecting the first positions of codons 523 and 601, respectively (Figure 4). In both cases, the mutations represent transversions from G → A. Both lead to the same conservative amino acid replacement of Gly → Ser. Interestingly, both mutations affect the F1 domain of the gelsolin-like portion of the protein, a domain for which the three-dimensional structure has now been elucidated in the gelsolin protein (McLaughlin *et al.* 1993), allowing us to evaluate the possible effects of these lesions at the functional level in *vivo*.

DNA from the D44-bearing chromosome displayed a restriction fragment with an altered length when probed with fragment X that contains the 5' end of the *flightless* transcription unit, including the leucine-rich repeats of the predicted protein (Figures 1, 2 and 4). Additional Southern hybridization indicated that the lesion resides within the proximal 1 kb of this fragment. To determine the precise nature of the rearrangement,
Molecular Analysis of flightless-I

FIGURE 3.—Nucleotide sequence of the genomic DNA from the flightless region of the Canton-S strain of D. melanogaster and its conceptual translation corresponding to EcoRI fragments X and Y in Figure 1. Corresponding cDNA sequence is shown in bold letters. Translational STOP signals and the polyadenylation consensus sequence are underlined. The sequence has been deposited in GENBANK under the accession number U28044.
we amplified DNA from this deficiency-bearing chromosome by primer-directed PCR. DNA sequencing revealed that nucleotide positions 233–624 are deleted in deficiency D44, which comprise sequences from 207 bases upstream of the presumed transcription initiation site, including the translational start site at position 505, and 66 bases of the first intron (Figure 4). The deficiency is therefore likely to represent a null mutation. We have found that the 5' flanking sequence AAACC at the deletion site is duplicated twice in the deficiency, possibly as a consequence of a transposition event involving target site duplication before imprecise excision, which may have led to the deficiency.

**DISCUSSION**

The flightless gene encodes a novel member of the growing superfamily of gelsolin-related proteins, linked to a series of leucine-rich domains. Gelsolins are actin filament-severing and -capping proteins that are regulated by both calcium and phosphoinositol (reviewed by VanDeKerkhove 1990; Hartwig and Kwiatkowski 1991; Weeds and Maciver 1993). The gelsolin family of proteins is composed of a series of repeating segments of ~130 amino acids, each characterized by its pattern of conserved residues. Gelsolins, villin, and flightless contain six domains. Equivalent severing proteins from lower eukaryotes, severin from Dictyostelium discoideum and fragmin from Physarum polycephalum, contain only three segments; in this they are similar to certain capping proteins from mammalian tissues, known as Cap G (Yu et al. 1990; Prendergast and Ziff 1991; Dabiri et al. 1992; Mishra et al. 1994), and an
actin-modulating protein from an earthworm (GIEBING et al. 1994). Sequence data are now also available for quail, a D. melanogaster gene that encodes a villin homologue (MAHajan-MIKLoS and COOLEY 1994) and a Drosophila homologue of gelsolin (HEINTZELMAN et al. 1993; STELLA et al. 1994).

The three-dimensional structures of G1 of gelsolin (MCLaughlin et al. 1993) and the equivalent V1 of villin (MARKUS et al. 1994) have been solved. These show that the characteristic segmental repeat reflects the underlying apolar core of the structural fold (MCLaughlin et al. 1993), as originally predicted by MATSudAIRA and JANmey (1988). This has been further supported by structure analysis of segment 2 of severin (Schnuchel et al. 1994). The structure of the complex of G1 with actin has also provided details of the contact residues at the interface and insights into the mechanism of filament severing. Interestingly, two members of the gelsolin family possess additional domains also found in other proteins, which convey specific aspects of their physiological roles. Villin carries a somewhat inappropriately named “headpiece” at its carboxyterminal end that is responsible for its actin-bundling activity.

The flightless protein exhibits 17 repeats of a common leucine-rich sequence motif at its amino terminal end. These repeats appear to be related to those of porcine ribonuclease inhibitor where they represent a new class of α/β fold (KOBE and DEsENHOFER 1993, 1994). Similar repeats are present in many types of proteins where they mediate homotypic and heterotypic macromolecular interactions (reviewed by KOBE and DEsENHOFER 1994), e.g., in cell-adhesion molecules (KRantz and ZIPURsKI 1990; LOPEZ, 1994), developmentally important receptors (KEITH and GAY 1990; CHANG et al. 1993) and transmembrane receptors (SCHNEIDER and SCHWEIGER 1991).

Intron/exon structure and protein domain organization of flightless: The location of intervening sequences in coding genes is thought by some to be related to the domain structure of the proteins they encode. In addition, the location of some introns is clearly conserved over long evolutionary periods of time and is retained even after duplication events. The current debate over the question of whether introns appeared early or late during evolution is therefore fueled by comparative structural data. Studies on human villin (PRINGault et al. 1991) concluded that the size and distribution of 14 introns in the regions encoding V1-V3 and V4-V6 did not support the postulated hypothesis of duplication from a precursor gene consisting of only three segments, which itself arose by triplication of an original single V1-like domain. Reevaluation of these data in the context of the known three-dimensional structure of V1 indicates that most of the introns split the proposed structural domains of the protein. Such cases are compatible with the “introns late” hypothesis. Only one intron (out of a total of 14) is located in accordance with the “introns separating domains” hypothesis. Similarly, none of the introns found in the gelsolin-like part of the C. elegans and D. melanogaster flightless genes corresponds to each other or to other known members of the gelsolin family. As was pointed out (STOLZPUS et al. 1994), chimeric proteins with a relatively short evolutionary history are less instructive in terms of intron/exon evolution, because exon shuffling may have favored chance correspondences between protein secondary structure and the distribution of exons. Despite this qualification, none of the introns found in either the C. elegans or D. melanogaster flightless genes coincides with domain boundaries of the protein or with recognizable secondary structure motifs. Although predictions relating to the structural location of introns suffer from the difficulty of defining the boundaries of a domain, in the case of leucine-rich repeat motifs, these have been clearly established.

The number and distribution of introns within the flightless homologues in C. elegans and D. melanogaster is open to a number of interpretations. First, the precursor that gave rise to the modern flightless gene may have arisen very early during metazoan evolution, and its original intron/exon structure was not preserved. The second interpretation is that introns disrupted the sequence after the more complex gene evolved. Gelsolin-like domains exist in protists (severin, fragmin), but prokaryotic homologues have yet to be discovered. If the emergence of gelsolin-like proteins, including flightless, was coupled to the appearance of actin-based contractile structures in eukaryotic evolution, then their presence in prokaryotes seems unlikely. Proteins containing leucine-rich repeats occur ubiquitously among eukaryotes and some prokaryotes (VENkATESAN et al. 1991), but their restricted distribution among pathogenic prokaryotes suggests acquisition from eukaryotic hosts.

Developmental implications of flightless expression: Most of the gene products involved in egg-polarity and the components of the cytoskeleton that provide a framework for their correct distribution in the egg are produced by nurse cells, which remain connected to the oocyte until the end of oogenesis. Since these nurse cells produce maternal products, mutations in genes that are expressed in the ovaries can lead to female sterility. Not surprisingly, a number of maternally transcribed genes are found to encode cytoskeletal proteins (reviewed by COOLEY and THEURKAUF 1994; KNOWLES and COOLEY 1994). Cytoskeletal proteins are also involved in the complex mechanics of cell division after fertilization of the egg. In Drosophila, cellularization of the syncytial blastoderm and migration of the nuclei to the cellular periphery takes place after 14 rounds of cell division. Several genes are known whose products are required at this stage and mutations at these loci are associated with disruptions of the actin cytoskeleton (SHEIJTER and WIESCHAUS 1993). The flightless gene
clearly belongs to this category. The maternal contribution of gene products by these loci can only be assessed by germline clone analysis. Maternal wild-type transcript can be supplied to rescue homozygous mutant embryos into extended periods well beyond embryogenesis.

Germline clones homozygous for the lethal alleles of flightless fail to develop into cellularized embryos (Perrimon et al. 1989), indicating that the gene product is already necessary at early embryonic stages and that sufficient transcript or translation product is maternally supplied to rescue homozygous mutant embryos into the later stages of larval development. The flightless gene may therefore represent a pivotal locus required for basic cellular processes.

**Structural homology of the flightless protein with gelsolin:** Despite the overall sequence similarity of the conceptual flightless gene product with members of the gelsolin family, analysis of the predicted sequence in the context of the known three-dimensional structure of the gelsolin G1-actin complex and that of villin V1 (McLaughlin et al. 1993; Markus et al. 1994) reveals several distinct differences, the functional significance (or lack thereof) of which is not yet known. The hydropobic core residues believed to be essential for structural integrity of G1 and V1 are conserved in all equivalent positions in the predicted flightless protein (Figure 6). There are, however, variations in residues of flightless corresponding to those in G1 that form the actin contact. All of the apolar residues in the hydrogen-bonding region of the actin involved in the gelsolin interface are completely conserved in both flight muscle actin 88F and in direct muscle isoforms (data not shown). The extent to which these sidechains contribute to high affinity actin binding is unclear; with a single exception they are conserved in the calcium-binding sites are different. These include sidechains equivalent to Asp<sup>50</sup> (Leu<sup>506</sup> in F1), Gln<sup>92</sup> (Leu<sup>555</sup>), Gln<sup>97</sup> (Lys<sup>560</sup>) and Asp<sup>109</sup>→Asp<sup>110</sup> (Arg<sup>567</sup>→Asn<sup>568</sup>). [It should be noted that the residues of actin involved in the gelsolin interface are completely conserved in both flight muscle actin 88F and in direct muscle isoforms (data not shown).] The extent to which these sidechains contribute to high affinity actin binding is unclear; with a single exception they are conserved in V1, but there are more substantial changes in segment 1 of both severin and the D. melanogaster villin-like quail gene product (Mahajan-Miklos and Cooley 1994).

The substitution of the residue equivalent to Gln<sup>97</sup> by a basic charge would be expected to disrupt the intramolecular calcium site. This is the calcium trapped in the G1:actin complex and probably also in the EGTA stable gelsolin:actin complex (Weeds et al. 1995). The presence of this calcium leads to at least 1000-fold enhancement of affinity of G1 for actin (Bryan 1988) but...
is not required for the severing activity of G1–3 or G1–2. The other carboxylate that is coordinated to this calcium is that of Asp$^{56}$, which, in G1, also forms a salt bridge with Arg$^{55}$ in the apolar core of the structure (Figure 7). The importance of this salt bridge has been highlighted in the related position of G2, where mutation of the conserved Asp to the corresponding amide causes the rare Finnish Familial Amyloidosis (Haltia et al. 1990; Maury et al. 1990). The structure analysis of G1 offers an explanation for the origin of this disease (McLaughlin et al. 1993). In flightless, no salt bridge could form in this position because the residue equivalent to Arg$^{55}$ is replaced by Glu.

The sequence similarity of the flightless protein with the gelsolin family extends to the G2 domain, which is essential for both severing activity and filament side-binding of gelsolin (Yu et al. 1990; Way et al. 1992a,b). Based on analogy with severing segment 2, this domain should contain a similar structural fold, but there is no information about its actin contacts, which are expected to be significantly different from those in G1. Preliminary data indicate that the human flightless protein is a phosphoinositide-regulated actin-binding protein that lacks severing activity (D. Kwiatkowski, personal communication).

Functions of the flightless protein: While the functional properties of the flightless protein are just beginning to be evaluated, the function of gelsolin has been probed at different levels. Domain-interchange and domain-deletion experiments have been carried out between gelsolin and villin (Finidori et al. 1992) and between Cap G and gelsolin (Yu et al. 1991), and their effects examined in vivo and in vitro. These experiments have revealed which parts of the gross architecture of these capping proteins contribute to the in vivo activities. At a finer level the in vitro mutagenesis experiments on gelsolin have revealed sites required for actin binding and severing (Way et al. 1989, 1990, 1992a,b). The generation of a gelsolin knock-out mouse (Witke et al. 1995) corroborates the view that the protein is non-essential for the development and maintenance of mammalian tissues since this homozygous mutant mouse is viable and fertile. The same applies to severin from D. discoideum, which is also dispensable under laboratory conditions (André et al. 1989). The analysis of the flightless gene is potentially significant to this field, since the rapid manipulability of the Drosophila system allows functional evaluation of both flightless and the recently cloned D. melanogaster gelsolin (Heintzelman et al. 1993). The discovery of an additional member of the gelsolin family in invertebrates may provide new insights into gelsolin function.

Mutational analysis: The flightless gene produces a single 5.1-kb transcript during development. Several chromosomal rearrangements truncating the transcript or eliminating the transcriptional activity of the gene, such as those analyzed in this study, all have lethal phenotypes. The complex rearrangement 2/19B exhibits an insertion within the 3' region of the gene, which may either truncate the gene product or compromise transcriptional efficiency and mRNA stability. The small deficiency D44 deletes the 5' end of the gene, including the translation initiation site and upstream sequences. The D44 rearrangement was generated by hybrid dysgenesis (Eeken et al. 1985), likely to have mobilized P elements. Our molecular analyses revealed a short duplication in the deletion site that is the hallmark of a transposition event, and the deletion is likely to have been caused by imprecise excision of a mobile element.

The existence of tissue- and stage-specific mutant phenotypes on the other hand, namely embryonic lethality or adult flightlessness, are explicable a number of ways. First, the product of this gene may be over-utilized during morphogenesis of flight muscles and mutations affecting the transcriptional efficiency may lead to flightlessness without affecting other functions of the adult fly. This hypothesis is supported by our finding that a single flightless transgene inserted into the heterochromatin of chromosome IV is able to rescue the lethal phenotype of two overlapping deficiencies uncovering the fliI locus but does not fully restore flight ability (Campbell et al. 1993). The transcriptional activity of the transgene in these lines is probably below normal, as assessed by the reduced expression of the eye-color marker gene in the same context. In contrast, transgene insertions at other chromosomal locations fully restore flight in the lethal allele fliI$^{w2}$ (R. Maleszka, S. D. Hanes, R. L. Hackett, H. G. De Couet, G. L. G. Miklos, unpublished results).

Point mutations may affect a function exclusively in a particular tissue. In the context of the presumed actin-binding and -severing properties of the flightless protein, this may include flight-muscle-specific protein-iso-
forms, calcium concentration, or other protein factors interacting with either actin or flightless. What may be an analogous phenomenon is observed in mutants at the \( \text{fl}a \) locus encoding \( \alpha \)-actinin (FYRBERG et al. 1990). Null alleles of \( \alpha \)-actinin lead to larval lethality, whereas several viable alleles specifically cause flightlessness. It is possible that in both cases functionally overlapping proteins present in tissues other than indirect flight muscle compensate for a loss of function caused by mutations in \( \alpha \)-actinin or flightless.

Now that the three-dimensional structure of the G1 domain of gelsolin in complex with actin has been solved (MC LAUGHLIN et al. 1993), this information can be more easily placed into a biological context in the Drosophila system. Two viable mutations are located in the F1 domain \((\text{fl}20)\) in which Gly\(^{325}\) (equivalent to Gly\(^{385}\) of G1) is replaced by Ser, and \( \text{fl}41 \) in which Gly\(^{462}\) (equivalent to Gly\(^{442}\) of G1) is replaced by Ser, and we therefore looked for possible structural implications of these substitutions. Both glycine residues are highly conserved in the gelsolin family and provide flexibility for turns in the structure. By analogy with G1, they are localized close to each other on two separate peptide loops at the surface of the molecule (Figure 7). Because the substitutions are conservative and at the surface of the domain, it is not possible to draw specific conclusions. The molecular changes observed in these G1 mutations. This possibility is being further explored by genetic screens for second site suppressor mutations. The evidence that function is perturbed by single conservative substitutions in both flightless and Finnish gelsolin demonstrates the need for more detailed functional analysis of the mutant proteins. The key to understanding the cellular function of the gelsolin family may be provided by detailed analysis of this family of Drosophila proteins.

**LITERATURE CITED**


