Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 2113–2117, March 1991
Biochemistry

A second maternally expressed *Drosophila* gene encodes a putative RNA helicase of the "DEAD box" family

(differential screening/translation initiation factors/chromosome region 31B/female-sterile mutants/vasa gene)

TAMSEN DE VALOIR, MARK A. TUCKER*, ESTHER J. BELIKOFF, LAURA A. CAMP†, CLARE BOLDUC, AND KATHY BECKINGHAM‡

Department of Biochemistry and Cell Biology, Rice University, P.O. Box 1892, Houston, TX 77251

Communicated by John Abelson, December 5, 1990

ABSTRACT **Recently, a family of proteins containing the conserved motif Asp-Glu-Ala-Asp, the "DEAD box" proteins, has been identified. This family is typified by the eukaryotic translation initiation factor eIF4A, and its members are believed to share the functional property of ATP-dependent RNA unwinding. One of the previously identified members of this family (vasa) is the product of a maternally expressed gene from *Drosophila melanogaster* that is known to play a role in the formation of the embryonic body plan. We report here the isolation of a *Drosophila* gene that has an mRNA expression pattern somewhat similar to that of vasa and also encodes a DEAD box protein. We have termed this gene *ME31B* to reflect its maternal (ovarian germ-line) expression and its location within the 31B chromosome region. Comparisons with the other members of this family reveal that although ME31B is most like the protein Tif1/Tif2, which probably represents the *Saccharomyces cerevisiae* version of eIF4A, it is unlikely that ME31B represents the *Drosophila* eIF4A protein *per se*. A search for mutations in the *ME31B* gene has established that the *P* element which causes the female-sterile mutation flipper lies in the 3′ flank of the *ME31B* gene.**

Recently a family of proteins that function as ATP-dependent RNA helicases has been identified (1).These proteins share a domain of about 370 amino acids in which a series of highly conserved motifs are arranged in identical order and with very similar spacing. Two of these conserved sequences represent specialized versions of the A and B motifs previously recognized in other ATP-binding proteins (2, 3). The four-amino-acid sequence Asp-Glu-Ala-Asp (DEAD) is part of the specialized version of the B motif. These "DEAD box" proteins derive from organisms across all evolutionary orders (*Escherichia coli* to mouse), and for some of them, genetic and molecular studies have provided indications as to their functions *in vivo* (4–7). Thus this family of proteins appears to be involved in a wide range of intracellular events in which alteration of RNA secondary structure must play a critical role. These include formation of the initiation complex for mRNA translation, mRNA splicing, and ribosome biosynthesis. Functionally, the eukaryotic translation initiation factor eIF4A, which is the archetypal member of the family, is the one best understood at the molecular level (8–10).

One of the members of this family is the product of the *Drosophila melanogaster* gene vasa (11, 12). The vasa gene is expressed during oogenesis to yield a protein that functions during early embryogenesis in the establishment of anterior–posterior pattern and the proper development of the germ line. Many of the maternally expressed genes such as vasa that play a role in embryonic pattern formation have been identified by mutagenic screening. As a complement to this approach, we used differential gene library screening for direct isolation of *Drosophila* genes that are expressed during oogenesis but not embryogenesis. We have characterized and sequenced one such gene, termed *ME31B*,§ which shows strong expression during oogenesis. *ME31B* has proved, like vasa, to encode a member of the DEAD box protein family.

MATERIALS AND METHODS

Differential Screening with *Calliphora* Probes. ^{32}P-labeled cDNA complementary to poly(A)$^+$ RNA from staged *Calliphora* oocytes or midstage embryos was used for screening of *Drosophila* clone libraries as described previously (13).

cDNA Library Screening. cDNA clones of the *ME31B* mRNA were isolated from the Goldschmidt–Clermont early embryo cDNA library (D. S. Hogness Laboratory, Stanford, CA) prepared in vector λgt10 (14). Genomic subclone A (see Fig. 1) was used to screen this library as described (13).

Analysis of DNA. Embryonic DNA was prepared as described previously (13). Subcloning in vectors Bluescribe M13$^+$ or Bluescript KS$^+$ (Stratagene), Southern hybridizations, and other DNA manipulations were performed by using standard procedures (15).

***In Situ* Hybridizations.** Hybridization of single-strand ^{35}S-labeled RNA probes to whole ovaries was performed as described previously (13). These probes were also used for *in situ* hybridization to the salivary gland polytene chromosomes (16). Hybridizations to whole embryos were performed as detailed by Edgar and O'Farrell (17). An ^{35}S-labeled complementary RNA probe representing the coding strand of subclone B (see Fig. 1) was used.

Analysis of RNA. Total RNA and poly(A)$^+$ RNA were prepared as previously described (13). RNA was denatured in formaldehyde/formamide for electrophoresis in 1.5% agarose/formamide gels (15). Blots were prepared on Genescreen (NEN) membrane and stained with methylene blue (18) to establish that RNA transfer was quantitative. A ^{32}P-labeled complementary RNA probe of the coding strand of subclone B (see above) was used for hybridizations. Conditions for hybridizations were as previously detailed (13) except for the hybridization temperature (50°C) and a stringent wash at 60°C in 0.1% SDS/0.1× SSC (1× SSC = 0.15 M NaCl/0.015 M sodium citrate, pH 7).

DNA Sequencing and Sequence Analysis. The sequence of cDNA clone DAEF5 was obtained from a series of overlap-

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

Abbreviation: eIF, eukaryotic translation initiation factor.

*Present address: Bone Physiology Laboratory, University of Southern California School of Dentistry, Orthopedic Hospital 2400 Flower Street, Los Angeles, CA 90033.
†Present address: Department of Pharmacology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75235.
‡To whom reprint requests should be addressed.

§The cDNA sequence corresponding to the protein sequence reported in this paper has been deposited in the GenBank data base (accession no. M59926).

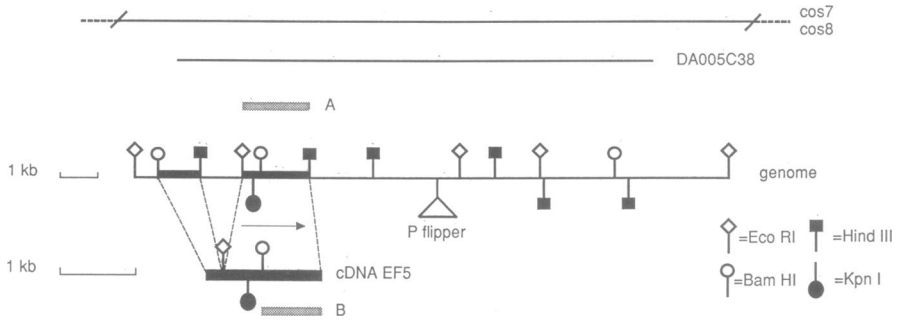

FIG. 1. Map of the *ME31B* region. Structure of 20 kilobases (kb) of genomic DNA containing the *ME31B* gene. The DNA within clone DA005C38 is indicated. Cosmid clones cos 7 and 8 from N. Clegg contain all of the DNA shown and additional flanking DNA. cDNA DAEF5 was obtained as an *Eco*RI insert from a λgt10 clone. The solid bars on the genomic map indicate the restriction fragments within the genomic clones that cross-hybridize to the two *Eco*RI fragments of the cDNA clone. The arrow shows the direction of transcription of the gene. Stippled bars indicate genomic (A) and cDNA (B) subclones that were used for the hybridization studies described here. The position at which the flipper *P* element interrupts the genomic DNA is indicated.

ping subclones. Dideoxy sequencing with Sequenase (United States Biochemical) and a series of custom-synthesized primers was performed. The clone was sequenced throughout on both strands. Sequence analyses were performed at the Molecular Biology Information Resource of Baylor College of Medicine.

RESULTS

Isolation of Genomic and cDNA Versions of the *ME31B* Gene. We set out to screen a series of *Drosophila* clone libraries with probes prepared from (*i*) ovarian poly(A)$^+$ RNA and (*ii*) midstage embryo poly(A)$^+$ RNA to identify clones that hybridized to the former but not the latter. Given the tiny size of the *Drosophila* ovary it would be very laborious to generate the large quantities of poly(A)$^+$ RNA required for this screening from *Drosophila* itself. We therefore used poly(A)$^+$ RNA preparations from the related dipteran *Calliphora erythrocephala*. In this large species, it is possible to prepare workable quantities of poly(A)$^+$ RNA from ovaries with follicles (oocytes) at particular stages of development and thus to screen for clones expressed at different stages of oogenesis. These probes have been used to isolate and characterize several *Calliphora* genes (13).

A series of *Drosophila* clone libraries were screened. The clone (DA005C38) containing *ME31B* originated from a subset of clones from the Maniatis genomic library that were identified by Ambrosio and Schedl (19) as hybridizing to *Drosophila* ovarian poly(A)$^+$ RNA. In our differential screening, it was the only clone of this set that hybridized to ovarian probes but not to a midstage embryo probe. The clone was found to hybridize to probes derived from early, mid-, and late-stage ovarian follicles and to probes from newly laid eggs, indicating that the gene it contains is expressed throughout oogenesis to yield transcripts that persist into the early embryo. The hybridization to probes derived from midstage follicles was strongest, however, suggesting that transcription is most active in the middle stages of oogenesis. Hybridization of a midstage follicle probe to digests of the purified clone DNA revealed that most of the coding sequence of the gene lies within a 1.8-kb *Eco*RI–*Hin*dIII fragment (subclone A; see Fig. 1).

Since our major interest was in identifying maternally expressed genes that might play a role in the early embryo, we first determined whether the gene in this clone was expressed in the germ-line cells of the ovarian follicle (see below). Once we had established that the gene is specifically expressed in the germ line, subclone A was used (*i*) to isolate related early embryonic cDNAs and (*ii*) to examine transcripts in Northern blots of embryonic poly(A)$^+$ RNA. It was thus determined that the longest cDNA clone isolated (DAEF5), which is >1.5 kb in length, probably represents a close to full-length copy of the single polyadenylylated, ≈1.7-kb, transcript from the gene. In collaboration with N. Clegg of the T. Grigliatti Laboratory (University of British Columbia, Vancouver), subclone A was also used to identify two cosmid clones that contain the *ME31B* gene. These cosmids (cos 7 and 8) are part of a walk performed by Clegg in the 31B chromosome region (see below for chromosomal location of *ME31B*). The structural relationship of cDNA clone DAEF5 to our genomic clone (DA005C38) and to cos 7 and 8 is shown in Fig. 1. The *ME31B* gene contains an intron close to its 5′ terminus.

Expression Pattern of *ME31B*. The localization pattern of *ME31B* transcripts in the *Drosophila* ovary was determined by *in situ* hybridization. ^{35}S-labeled complementary RNA probes corresponding to the two strands of subclone A were hybridized individually to whole ovaries. This allowed simultaneous determination of the mRNA localization pattern and the direction of transcription of the gene. The orientation of the gene is shown in Fig. 1. The coding strand probe showed specific hybridization to the germ-line cells of the follicle (the 15 nurse cells and the oocyte) with no hybridization to the somatically derived follicle cells, which encase the germ-line cells (see Fig. 2 *A* and *B*). At all stages that the nurse cells were present, the concentration of transcripts was greater in the nurse cell cytoplasm than in the oocyte cytoplasm (see Fig. 2*A*). However, no specific pattern of localization of the transcripts within either the nurse cell or oocyte cytoplasm was detected (see Fig. 2 *A* and *C*). Interestingly, these studies suggest that a burst of *ME31B* transcription begins at about stage 6/7 of oogenesis, since the nurse cell cytoplasm shows far stronger hybridization during stages 7–10 than at earlier times (see Fig. 2 *A* and *C*). The concentration of transcripts in the mature oocyte cytoplasm is also considerably lower than that seen in stage 7 follicles (see Fig. 2*C*). *In situ* hybridization studies with whole embryos (data not shown) established that *ME31B* transcripts remain uniformly distributed in the early embryo and disappear by germ-band extension.

In developmental Northern blots of total organismal RNA (see Fig. 3) a single abundant 1.6- to 1.7-kb transcript was detected in 0- to 2-hr embryos. This transcript was weakly detectable in gravid females and 2- to 4-hr embryos but was not seen at other stages of the life cycle or in adult males (see Fig. 3). Northern blots of RNA from dissected ovaries, dissected testes, and residual female and male carcasses were also probed (data not shown), and transcripts were detected only in the ovarian RNA. These hybridizations thus indicate that the major tissue expressing the *ME31B* gene is the adult

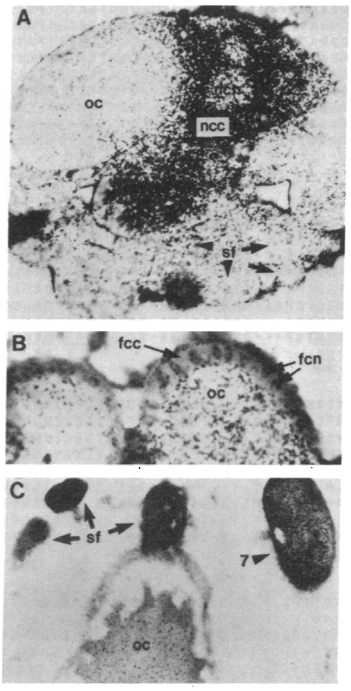

FIG. 2. Ovarian germ-line expression of *ME31B*. Bright-field views of ovarian sections after *in situ* hybridization to *ME31B* mRNA. After autoradiography, sections were stained with methylene blue and basic fuschin. (*A*) *ME31B* transcript accumulation in ovarian follicles of different stages. A section containing small follicles (sf), stages 2–6, and two older follicles (stages ≈7 and 10a) is shown. Note the much stronger hybridization signal seen in the nurse cell cytoplasm (ncc) of the two older follicles. oc, Oocyte cytoplasm. (×220.) (*B*) Germ-line specificity of *ME31B* ovarian transcription: Higher-magnification view of the posterior pole of a stage ≈7 follicle. Silver grains are seen throughout the oocyte cytoplasm (oc) but are absent from the somatically derived follicle cells. fcc, Follicle cell cytoplasm; fcn, follicle cell nuclei. (×600.) (*C*) Presence of *ME31B* transcripts in the mature oocyte. The section shows relative intensity of hybridization signal to the cytoplasm of a mature oocyte (oc), a stage-7 follicle (7), and several small follicles, stages 2–5 (sf). Transcripts are localized fairly uniformly throughout the mature oocyte cytoplasm. (×210.)

female germ line, although they are clearly not sensitive enough to preclude low expression in some other tissue.

Sequence of an *ME31B* cDNA Clone. The longest cDNA clone (DAEF5) was sequenced in its entirety. The cDNA contains an open reading frame of 459 codons, which begins at nucleotide 34 with a methionine codon flanked by a good match to the *Drosophila* translation initiation consensus (20). The clone terminates with nine A residues preceded 20 residues upstream by a polyadenylylation signal. It thus seems likely that this cDNA contains the entire protein coding sequence and 3′ trailer, but it may lack some of the 5′ leader sequences. The complete sequence of this cDNA clone has been submitted to the GenBank data bank.§

Computer searches of protein data banks revealed that the ME31B protein is very similar (37% overall identity) to eukaryotic initiation factor eIF4A, the first member of the DEAD box protein family to be cloned and sequenced (21, 22). Fig. 4 shows the alignment of ME31B with the eight other DEAD box family proteins for which there are complete protein sequences. These nine DEAD box proteins fall into two classes: (*i*) short proteins, with relatively few residues outside the shared domain, and (*ii*) longer proteins, with additional sequence at one or both ends. ME31B belongs to the shorter class and in overall organization is most like the two mouse eIF4A sequences and the Tif1/Tif2 sequence, in that the few additional residues are largely at the amino

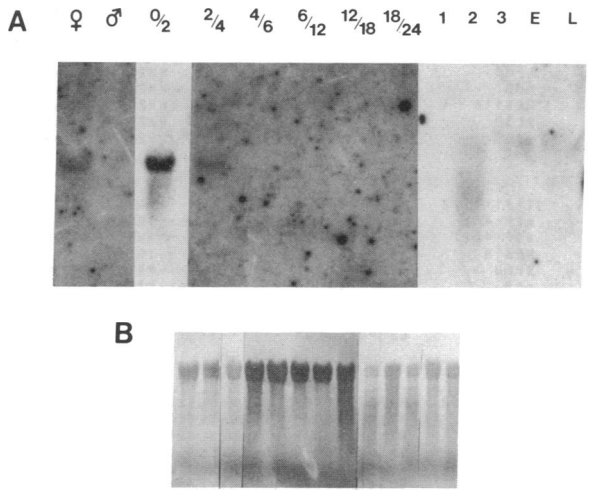

FIG. 3. Developmental expression of *ME31B*. Total RNA from various stages of the life cycle was probed for the presence of *ME31B*-related transcripts, using ^{32}P-labeled complementary RNA prepared from subclone B (see Fig. 1). ♀ and ♂, RNA from adult females and males, respectively; 0/2, 2/4, etc., RNA from embryos at 0–2, 2–4, etc., hr of development; 1, 2, 3, RNA from first-, second-, and third-instar larvae; E and L, RNA from early and late pupae. *A* shows the hybridization patterns; *B* shows the methylene blue staining of the blots prior to hybridization. The intense bands in *B* are the 18S and 28Sα and β bands of the rRNA. Specific hybridization is to a single band of 1.6–1.7 kb as judged by mobility relative to a stained RNA ladder (BRL, Life Technologies, Gaithersburg, MD) that was coelectrophoresed on the gels. Faint hybridization to larger RNA seen in some of the larval and pupal samples is nonspecific hybridization to the rRNA bands. This was also occasionally seen for RNA from other stages of the life cycle.

terminus. *TIF1* and *TIF2* encode the same protein, which is thought to represent the yeast equivalent of eIF4A (6).

The nine DEAD box proteins shown in Fig. 4 show absolute sequence conservation at 46 residue positions and high conservation at 51 additional residues. Recently partial sequences for six additional DEAD box proteins, all of which derive from *Saccharomyces cerevisiae*, have been published. For one of these, the protein SBP4, the sequence of the entire conserved domain is available (7). SBP4 shows identity at 36 of the 46 absolutely conserved residues seen in Fig. 4. The other yeast sequences (24) are partial sequences for five proteins obtained by using polymerase chain reaction (PCR) technology to amplify sequences that contain most of the carboxyl-terminal half of the conserved domain—the region between the DEAD and HRIGR motifs (see Fig. 4). In this region, these five proteins show 84–100% identity at the 25 positions found to be absolutely conserved in the alignments shown in Fig. 4. The comparison of these six additional yeast sequences to the eight original DEAD box family members has permitted Chang *et al.* (24) to identify six further positions in the conserved domain at which there is very high sequence conservation (see Fig. 4). The ME31B protein sequence contains the dominant amino acid at three of these six positions (see Fig. 4).

The percentage of identical residues detected when the conserved domain of ME31B is compared individually to those domains of the other eight DEAD box proteins and SBP4 has been determined. This domain of ME31B is most similar to that of the yeast protein Tif1/Tif2 (39.7% identity). MEB1B shows significantly higher similarity (37–40% identity) in the conserved region to the three proteins that are versions of eIF4A (eIF4AI, eIF4AII, and Tif1/Tif2) than to the other six proteins (24–30% identity).

Chromosomal Localization of *ME31B* and Identification of Candidate Mutations in the Gene. We determined by *in situ*

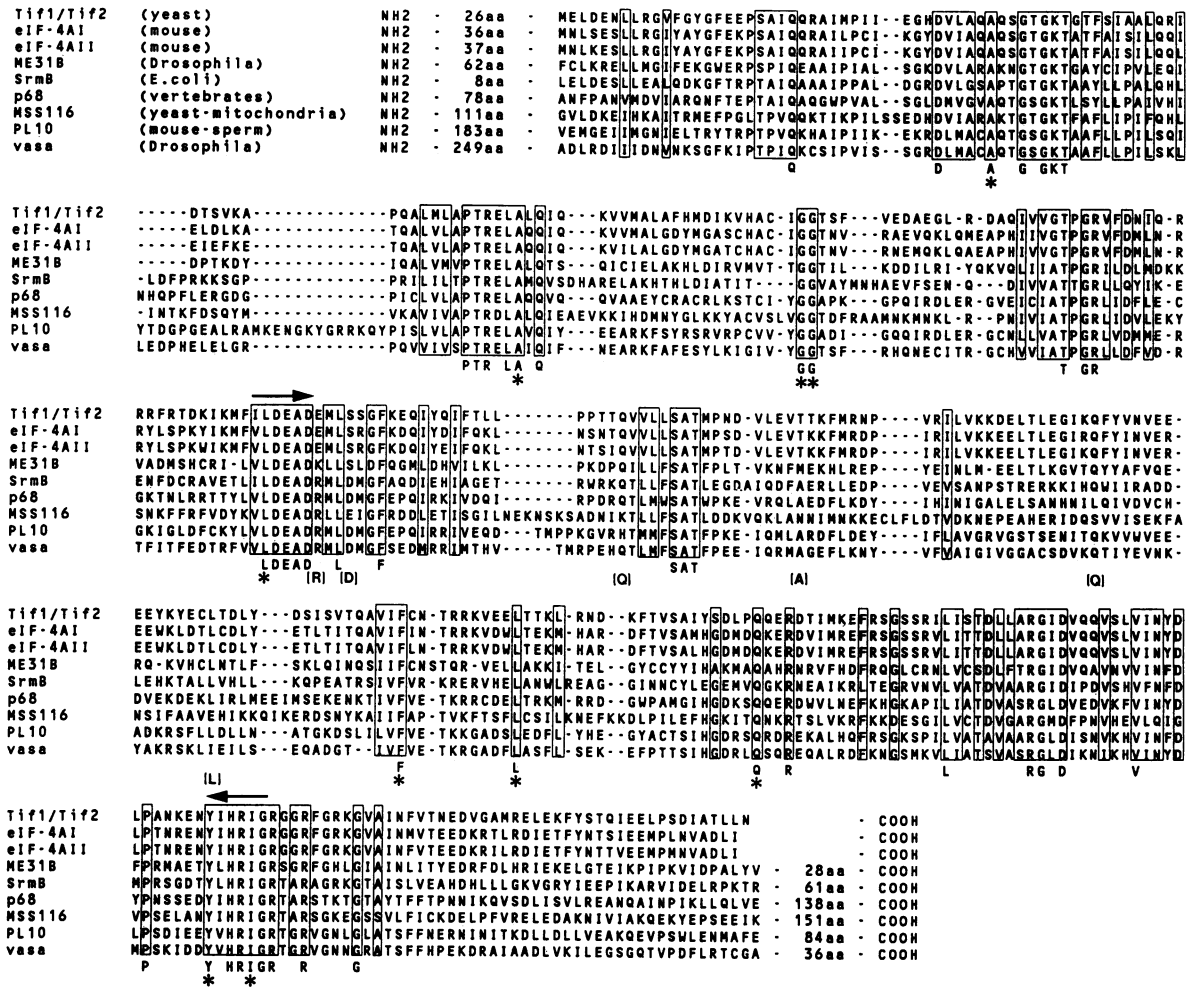

FIG. 4. Alignment of ME31B with the DEAD box protein family. The proteins are arranged according to the number of additional residues they contain outside the region of similarity. The last four proteins have substantial extra amino- and/or carboxyl-terminal domains. Tif1/Tif2, eIF4AI and -II, p68, PL10, and vasa are discussed in the text. SrmB is an *E. coli* product identified by its ability to suppress a defective interaction between a mutant form of ribosomal protein L24 and the 23S rRNA (4). MSS116 is encoded by a nuclear yeast gene and affects the splicing of a subset of the mitochondrial transcripts (5). The sequence for SrmB contains the recently published revisions (4). Boxes enclose residues which in all nine proteins are (*i*) perfectly conserved or (*ii*) highly conserved (23). At 46 positions there is absolute sequence conservation, and these residues are shown below the alignments. Note that for vasa, the penultimate perfectly conserved residue (arginine) is a cysteine in the sequence of Lasko and Ashburner (11). Asterisks below 10 of the conserved residues indicate positions at which absolute conservation is not maintained in the recently sequenced yeast protein SPB4 (7). Arrows above lines 3 and 5 indicate the region of five additional yeast proteins which has been isolated as a result of polymerase chain reaction amplification (24). Comparisons with these additional sequences have demonstrated strong conservation at six additional residues within this region (shown as residues in parentheses below the alignments). Note that in these alignments, the previously recognized modified versions of the A and B motifs of ATP-binding proteins (3) are still perfectly conserved. These are DXXXXAXGXGKT (motif A, line 1) and (V/I)LDEADX(M/L)LXXGF (motif B, line 3).

hybridization to the polytene chromosomes that *ME31B* originates from region 31B-D of chromosome 2. Finer mapping of the gene was achieved by positioning the gene relative to four overlapping deletions for this region (25–27). These are 2L deficiencies *J-der 2, J-der 27, J-der 77*, and *J-der 106*, which delete 31B-31F, 31D/E–31E/F, 31C–31E, and 31D/E–31F2, respectively (J. Tomkiel and D. Sinclair, personal communications). DNA from dead embryos homozygous for each of these deletions was probed in Southern hybridizations for the presence of the 5.8-kb *Eco*RI fragment containing most of the *ME31B* gene (see Fig. 1). Since homozygous embryo preparations are always slightly contaminated with dead wild-type embryos, the presence of a given DNA fragment in such a DNA preparation must be assessed by comparing its hybridization to that of a control DNA fragment known not to be deleted by the deficiency in question. In our case a 3-kb DNA fragment from chromosome 3 was used (M.A.T. and K.B., unpublished observation). This analysis established that *ME31B* is deleted only by *J-der 2*. This positions the gene in region 31B. From the chromosomal walk of N. Clegg (see above), *ME31B* can be further localized as lying ≈30 kb proximal to the *P* element at 31B isolated in the Lis laboratory (28).

As a beginning step in our search for mutations within the *ME31B* gene we have examined a number of female-sterile mutations that lie within the appropriate chromosome region. The mapping of several female steriles generated by T. Schüpbach (ref. 29 and personal communication) relative to the *J-der* deficiencies resulted in the identification of chalice as a potential candidate locus for *ME31B*. Like *ME31B*, this gene is deleted by *J-der 2* but not by the other deletions. In collaboration with N. Clegg, a *P* element-induced mutation (flipper) which was recently generated by C. Berg and D. McKearin in the A. Spradling Laboratory (Carnegie Institution of Washington, Baltimore, MD) has also been identified as a candidate mutation of *ME31B*. Clegg has cloned DNA from the flipper genome that flanks the *P* element associated with the mutation. Using this DNA we have determined that

the *P* element interrupts the 3′ flank of the *ME31B* gene, 3.5 kb from the polyadenylylation site (see Fig. 1). Mapping has revealed that the female sterility and other phenotypic effects of flipper are uncovered by *J-der 2* but not the *J-der 77* deletion. We have established that the flipper and chalice mutations complement one another and therefore represent mutations in separate genes.

DISCUSSION

ME31B is the second maternally expressed *Drosophila* gene that has proved to belong to the DEAD box family. Both *ME31B* and vasa are expressed in the germ-line cells of the ovary to give transcripts that persist into the early embryo. In each case, transcripts show a uniform cytoplasmic distribution during oogenesis and early embryogenesis. Surprisingly, however, given the absence of phenotypic effects in vasa$^-$ males (11), vasa has also proved to be expressed in the adult male germ line (12), with transcription in the germ cells of both sexes beginning in late embryogenesis (12).

Of the DEAD box family members identified to date, six are derived from multicellular eukaryotes: eIF4AI and -II (20, 21), p68(30), PL10 (31), vasa (11, 12), and ME31B. Among these proteins, only the two versions of eIF4A have defined molecular roles. eIF4A acts during the initiation of translation in the binding of mRNA to the 43S preinitiation complex (10). It seems likely that its helicase function (9) is required to melt out secondary structure in mRNA (10). The nuclear protein p68 also has an ATP-dependent RNA helicase activity (32), but nothing is known of the molecular functions of the three remaining proteins (vasa, PL10, and ME31B). Interestingly, all three appear to be abundantly or specifically expressed in germ-line cells. Thus PL10 is encoded by a mouse male germ-line-specific gene, vasa is expressed in both the male and female germ line of *Drosophila* (12), and ME31B is expressed in the *Drosophila* female germ line. This might indicate major roles for this protein family in germ-line-specific functions.

Our comparison of ME31B to the other DEAD box family members provides some hints as to the possible function of this protein. In terms of both overall domain structure and sequence similarity, ME31B is more similar to eIF4AI, -II, and the yeast Tif1/Tif2 protein than it is to the other family members. It is unlikely, however, that ME31B represents the *Drosophila* version of eIF4A itself for two reasons. First, the gene is too limited in its expression pattern to represent a universally required initiation factor. Second, the Tif1/Tif2 protein, which is thought to represent the yeast version of eIF4A, shows high similarity (67% identity) to the two mouse eIF4A sequences, whereas the conservation between these three proteins and ME31B is considerably lower (37–40% identity). ME31B is therefore too divergent to represent the *Drosophila* homologue of this protein.

It is clear that translational control of specific mRNA utilization is operating during late oogenesis and early embryogenesis in *Drosophila*. This phenomenon has been demonstrated for transcripts from both universal housekeeping genes such as the ribosomal protein genes (33) and from key regulatory genes such as bicoid (34). It is possible, therefore, that ME31B represents a translation initiation factor related to eIF4A but with a more specialized role. In addition to the tissue specificity of expression demonstrated here, the protein could show specificity for particular mRNAs or be confined to a discrete subcellular location.

Our screen of all existing uncharacterized female-sterile mutations from the 31B-D chromosome region has identified chalice and flipper as candidate loci for the *ME31B* gene. The *P* element associated with flipper lies within the 3′ flank of *ME31B*, but two aspects of *ME31B* expression are difficult to reconcile with the flipper phenotype. First, whereas *ME31B* expression appears to be largely limited to the ovary, the flipper mutation effects adult size and male fertility in addition to causing female sterility. Second, we have detected *ME31B* transcripts in the tiny ovarian follicles present in homozygous flipper females (C.B. and K.B., unpublished observations). Clearly *P* element transformation experiments are required for a better understanding of the relationship of *ME31B* to the chalice and flipper mutations.

We thank Linda Ambrosio for the gift of her oogenesis-expressed clone collection and Joseph Duffy and Peter Gergen for help with embryonic *in situ* hybridizations. We are very grateful to Nigel Clegg for clones from the 31B region, Donald Sinclair for the *J-der* deficiency stocks, and Dennis McKearin for the flipper mutant stock. The assistance of John Scott with computer analyses at the Baylor College of Medicine Molecular Biology Information Resource is gratefully acknowledged. Thanks go to Trey Boudreau for additional programming assistance. This work was initially supported by Grant HD17688 from the National Institutes of Health and subsequently by Grant NP689 from The American Cancer Society.

1. Linder, P., Lasko, P. F., Ashburner, M., Leroy, P., Nielsen, P. J., Nishi, K., Schnier, J. & Slonimski, P. P. (1989) *Nature (London)* **337**,121–122.
2. Hodgman, T. C. (1988) *Nature (London)* **333**, 22–23, and correction (1988) **333**, 578.
3. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982) *EMBO J.* **1**, 945–951.
4. Nishi, K., Morel-Deville, F., Hershey, J. W. B., Leighton, T. & Schnier, J. (1988) *Nature (London)* **336**, 496–498, and correction (1989) **340**, 246.
5. Séraphin, B., Simon, M., Boulet, A. & Faye, G. (1989) *Nature (London)* **337**, 84–87.
6. Linder, P. & Slonimski, P. P. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2286–2290.
7. Sachs, A. B. & Davis, R. W. (1990) *Science* **247**, 1077–1079.
8. Ray, B. K., Lawson, T. G., Kramer, J. C., Cladaras, M. H., Grifo, J. A., Abramson, R. D., Merrick, W. C. & Thach, R. E. (1985) *J. Biol. Chem.* **260**, 7651–7658.
9. Abramson, R. D., Dever, T. E., Lawson, T. G., Ray, B. K., Thach, R. E. & Merrick, W. C. (1987) *J. Biol. Chem.* **262**, 3826–3832.
10. Pain, V. M. (1986) *Biochem. J.* **235**, 625–637.
11. Lasko, P. F. & Ashburner, M. (1988) *Nature (London)* **335**, 611–617.
12. Hay, B., Jan, L. Y. & Jan, Y. N. (1988) *Cell* **55**, 577–587.
13. Rubacha, A., Tucker, M. A., de Valoir, T., Belikoff, E. J. & Beckingham, K. (1988) *Dev. Biol.* **129**, 449–463.
14. Huynh, T. V., Young, R. A. & Davis, R. W. (1984) in *DNA Cloning Techniques: A Practical Approach*, ed. Glover, D. (IRL, Oxford, U.K.), pp. 49–78.
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
16. Smith, V. L., Doyle, K. E., Maune, J. F., Munjaal, R. P. & Beckingham, K. (1987) *J. Mol. Biol.* **196**, 471–485.
17. Edgar, B. A. & O'Farrell, P. H. (1989) *Cell* **57**, 177–187.
18. MacManus, J. P., Gillen, M. F., Korczak, B. & Nojima, H. (1989) *Biochem. Biophys. Res. Commun.* **159**, 278–282.
19. Ambrosio, L. & Schedl, P. (1984) *Dev. Biol.* **105**, 80–92.
20. Cavener, D. R. (1987) *Nucleic Acids Res.* **15**, 1353–1361.
21. Nielsen, P. J., McMaster, G. K. & Trachsel, H. (1985) *Nucleic Acids Res.* **13**, 6867–6880.
22. Nielsen, P. J. & Trachsel, H. (1988) *EMBO J.* **7**, 2097–2105.
23. Dayhoff, M. O., Schwartz, R. M. & Oncutt, B. C. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington) Vol. 5, Suppl. 3, pp. 345–352.
24. Chang, T.-H., Arenas, J. & Abelson, J.(1990) *Proc. Natl. Acad. Sci. USA* **87**, 1571–1575.
25. Mange, A. P. & Sandler, L. (1973) *Genetics* **73**, 73–86.
26. Sandler, L. (1977) *Genetics* **86**, 567–582.
27. Salas, F. & Lengyel, J. A. (1984) *Dros. Inf. Serv.* **60**, 243–244.
28. Simon, J. A., Sutton, C. A., Lobell, R. B., Glaser, R. L. & List, J. T. (1985) *Cell* **40**, 805–817.
29. Schüpbach, T. & Wieschaus, E. (1989) *Genetics* **121**, 101–117.
30. Ford, M. J., Anton, I. A. & Lane, D. P. (1988) *Nature (London)* **332**, 736–738.
31. Leroy, P., Alzari, P., Sassoon, D., Wolgemuth, D. & Fellous, M. (1989) *Cell* **57**, 549–559.
32. Hirling, H., Scheffner, M., Restle, T. & Stahl, H. (1989) *Nature (London)* **339**, 562–564.
33. Qian, S., Zhang, J., Kay, M. A. & Jacobs-Lorena, M. (1987) *Nucleic Acids Res.* **15**, 987–1003.
34. Driever, W. & Nüsslein-Volhard, C. (1988) *Cell* **54**, 83–93.