Nucleic Acids Research, Vol. 20, No. 12 3063–3067

Dbp73D, a *Drosophila* gene expressed in ovary, encodes a novel D-E-A-D box protein

Lisa F.Patterson+, Michel Harvey[1] and Paul F.Lasko[1,*]
Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK and [1]Department of Biology, McGill University, 1205 Avenue Docteur Penfield, Montréal, Québec H3A 1B1, Canada

Received March 11, 1992; Revised and Accepted April 28, 1992 GenBank accession no. M74824

ABSTRACT

Proteins of the D-E-A-D family of putative ATP-dependent RNA helicases have been implicated in translation initiation and RNA splicing in a variety of organisms from *E. coli* to man. The *Drosophila vasa* protein, a member of this family, is required in the female germ line for fertility and for specification of germ line and posterior positional information in progeny embryos. We report the isolation of another D-E-A-D gene from *Drosophila*, which, like *vasa*, is expressed in germ line tissue. The predicted amino acid sequence of this new gene, *Dbp73D*, contains all of the highly conserved helicase motifs, but is otherwise the farthest-diverged member of the family so far identified.

INTRODUCTION

Since the initial identification of the D-E-A-D family of putative RNA unwinding proteins, the number of known genes encoding such products has greatly increased (1, 2). Thus far, only in *S. cerevisiae* has there been an effort directed toward finding these genes, and at least fourteen different D-E-A-D box genes have now been reported in that organism (2, 3). The protein of this type which has been best studied biochemically is eukaryotic initiation factor-4A, which is part of the cap-binding protein complex, has an ATPase activity, and as part of the complex acts as an RNA helicase, serving to unwind secondary structure at the 5′ end of the messenger RNA molecule (4–7). This complex has recently been implicated in oncogenesis, as overexpression of the cap-binding subunit in NIH 3T3 cells results in the transformed phenotype (8).

Many of the other D-E-A-D proteins are larger than eIF-4A, with molecular weights on the order of 70,000 daltons or greater, rather than the 46,000 daltons characteristic of the translation initiation factor. These larger proteins include a conserved domain including most of the residues of eIF-4A, flanked by relatively unique amino acid sequences on the amino- and carboxy-terminal ends. Eight of these larger proteins, including the three members of the related D-E-A-H subfamily, have been implicated as involved in RNA splicing in *S. cerevisiae* (2, 9–16). The *Drosophila* protein vasa, also a member of this class, is a component of ribonucleoprotein complexes called polar granules, which are determinants for germ cell formation in the embryo (17–20). Mutations in *vasa* in the maternal germ line lead to somatic pattern deletions and absence of germ cells in progeny embryos (21).

Here we report the isolation of another D-E-A-D gene from *Drosophila*. The gene, *Dbp73D*, is expressed in both sexes; in ovaries, its transcription is limited to germ-line cells. The putative RNA helicase protein encoded by *Dbp73D* includes all of the previously described highly-conserved sequence motifs of the D-E-A-D family of proteins, but also contains a number of unique features which will be discussed in detail below.

MATERIALS AND METHODS

Analysis of *Dbp73D* transcripts

RNA was extracted by either of two methods: homogenizing tissues in 4 M guanidinium isothiocyanate, 0.1 M Tris, pH 7.5, 1% 2-mercaptoethanol, 0.5% N-lauroylsarcosine, 0.5% diethylpyrocarbonate, extracting twice with phenol:chloroform, and precipitating by adding 0.04 volumes 1 M acetic acid and 0.5 volumes absolute ethanol. Alternatively, RNA was extracted by the method of Sambrook, *et al.* (22). Tissue was homogenized in a buffer consisting of 50 mM NaCl, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.5% SDS, and 200 μg/ml proteinase K, and incubated at 37° for 1 hr. The homogenate was extracted twice with phenol:chloroform, and the RNA was precipitated once with ethanol, and a second time with 4 M LiCl. Polyadenylated RNA was selected by oligo-dT cellulose column chromatography (Pharmacia). The RNA was separated on a 1% agarose/formaldehyde gel.

Whole-mount in situ hybridization to Drosophila tissues

The procedure followed was that of Tautz and Pfeifle (23) with the following modifications. Ovaries were dissected in Ringer's

* To whom correspondence should be addressed

+ Present address: Duke Medical School, Durham, NC, USA

solution, fixed in 4% paraformaldehyde in PBS, and permeabilized in 90% methanol:10% DMSO as described by Ephrussi et al. (24). Proteinase K treatment was for 1 hr at room temperature at a concentration of 50 μg/ml. Embryos (0−24 hr) were collected on apple juice-agar plates, collected, their chorions removed with 50% commercial bleach, and fixed according to the protocol of Tautz and Pfeifle (23). They were then dehydrated through 30%, 50% and 70% ethanol, and stored in 70% ethanol at −20°. For hybridizations, the embryos were rehydrated at room temperature and treated according to the procedure of Tautz and Pfeifle (23) as modified by Ephrussi, *et al.* (24). All hybridizations were carried out overnight, with digoxigenin-labelled DNA probes, at 45° in microcentrifuge tubes in a hybridization solution of 5×SSC, 50% formamide, 100 μg/ml denatured salmon sperm DNA, and 50 μg/ml sodium heparin. Visualization of the sites of hybridization was accomplished by reaction with an alkaline phosphatase conjugated anti-digoxigenin antibody (1:2000 dilution for 1 hr), washing, and staining with bromochloroindolyl phosphate/nitro blue tetrazolium for 45 minutes.

Nucleic acid hybridizations

Southern and Northern transfers were performed using standard techniques (22). Hybridizations were carried out in 2×SSCP, 5×Denhardt's, 0.1% SDS, 9% dextran sulfate overnight at 65°, washes were twice in 1×SSC, 0.5% SDS, then once in 0.1×SSC, 0.1% SDS, all at 65°. Radiolabelled DNA probes were made by the oligonucleotide-primed labelling method of Feinberg and Vogelstein (25). *In situ* hybridizations to polytene chromosomes were performed with biotinylated probes according to the method described by Ashburner (26).

DNA sequencing

The entire cDNA and genomic sequence was determined on both strands. All sequencing was performed using the dideoxynucleotide-based chain-termination method (27), using either Sequenase (United States Biochemical) or T7 DNA polymerase (Pharmacia). The sequence was mostly determined using subclones into the M13mp18 and M13mp19 single-stranded phage vectors (28), and two nested series of *ExoIII* deletions (29) in the plasmid vector BlueScript (Stratagene), and was completed by using newly-synthesized oligonucleotide primers on existing subclones.

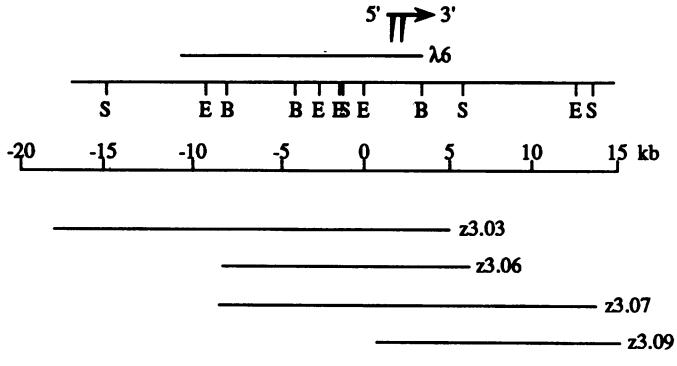

Figure 1. Restriction map of the *Dbp73D* chromosomal region. The arrowed line represents the *Dbp73D* transcript, as defined by the complementary DNA clone p6. The straight lines indicate the extents of the *Drosophila* inserts of five lambdaphage clones selected from a genomic library. E, restriction site for *Eco*RI; B, restriction site for *Bam*HI; S, restriction site for *Sal*I.

RESULTS

The original *Dbp73D* clone (λ6) was isolated fortuitously in an oligonucleotide-based screen of a genomic phage library (30) for novel rhodopsin genes. Restriction fragments of λ6 were then used to isolate a cDNA clone, p6, from a library made from larval transcripts, and to complete a short chromosomal walk of 33 kb. Nucleotide sequence comparisons indicated no significant homologies between the rhodopsin oligonucleotide and *Dbp73D*.

The chromosomal region surrounding the *Dbp73D* gene is illustrated in Figure 1. The *Dbp73D* gene is present in a single copy in the *Drosophila* genome, at polytene chromosome location 73CD on the left arm of chromosome *3* (Figure 2). Its single 2.4-kb transcript is expressed in embryos, larvae and adults, but is undetectable in RNA prepared from adult heads (Figure 3). The lack of expression in heads, and the reported ovarian expression of other *Drosophila* D-E-A-D genes, such as *vasa*

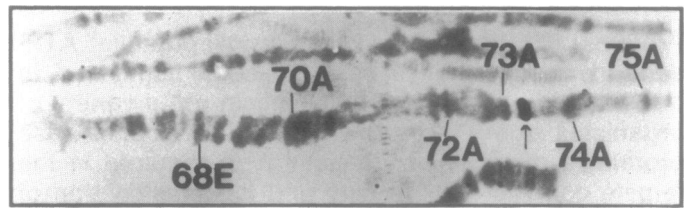

Figure 2. *In situ* hybridization of p6 sequences to larval salivary gland polytene chromosomes. The site of hybridization is marked with an arrow, and neighboring major bands defining chromosome subdivisions are identified.

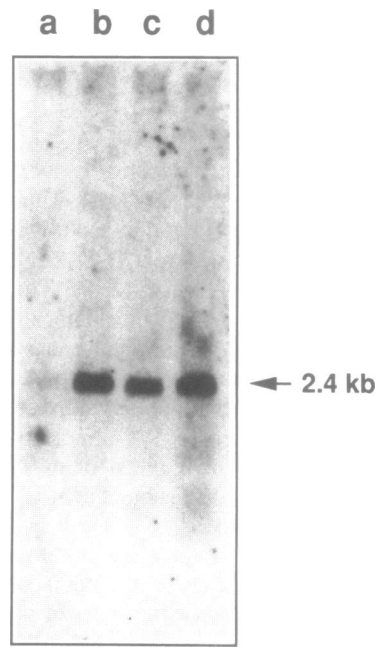

Figure 3. Expression pattern of the *Dbp73D* gene. Polyadenylated RNA was prepared from the following tissues: a) heads, b) thoraxes and abdomens, c) whole *glass*⁻ flies, d) third-instar larvae and separated on an agarose-formaldehyde gel as described in Materials and Methods. 3μg was loaded in lanes a−c, 8μg in lane d. The transfer was probed with the radiolabelled 2.3-kb insert of the p6 complementary DNA clone. Autoradiographic exposure was for 6 days. Relative molecular weight was estimated from comparison of migration distances with commercially radiolabelled RNA size standards (Bethesda Research Laboratories; not shown).

and *ME31B* (18, 19, 31), suggested that a more detailed investigation of the expression of *Dbp73D* in the female germ line was warranted. This was carried out by the method of whole-mount *in situ* hybridization.

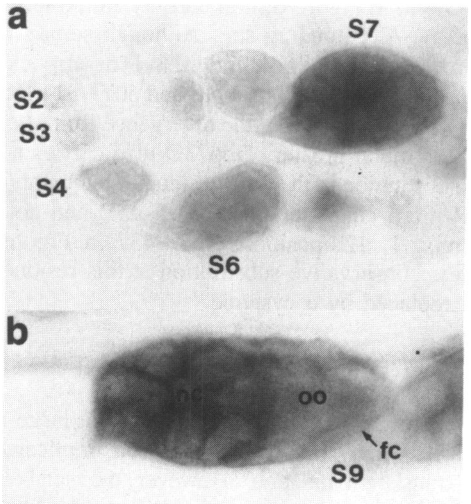

Figure 4. a) Egg chambers hybridized with a digoxigenin-labelled probe for the *Dbp73D* transcript, and visualized with alkaline phosphatase as described in Materials and Methods. S2–S7 refers to the developmental stages of oogenesis as described by King (32). **b)** A stage-9 egg chamber hybridized as in **a**. Abbreviations: nc, nurse cells; oo, oocyte; fc, follicle cells.

In the two panels of Figure 4 is illustrated the transcription of *Dbp73D* in egg chambers. Expression of the gene is very low in the initial stages of oogenesis, but begins to increase at about stage 6 in the germ-line cells (panel a; oogenesis stages are those of ref. 32). The peak of ovarian expression is reached in the nurse cells of the stage 9 and 10 egg chamber (panel b). Transcript appears to be exported from the nurse cells into the oocyte at about this time, but is not localized within the oocyte. The somatic follicle cells transcribe little, if any, *Dbp73D* message. The transcript is present at a low level throughout the cleavage-stage embryo (data not shown); as this signal becomes undetectable by the syncytial blastoderm stage, it probably corresponds to the exported maternal message. No obvious spatial regulation of *Dbp73D* expression is observed during the remainder of embryonic development.

Dbp73D encodes a D-E-A-D-box protein

The nucleotide sequence of *Dbp73D* was determined, and is listed in Figure 5. The gene is of a rather simple organization, with three exons separated by two small introns of 54 and 60 bp. It encodes an open reading frame of 572 amino acids, giving a weight of 64,799 daltons for the predicted polypeptide. Both the amino-terminal and carboxy-terminal ends of the protein are rich in hydrophilic residues. The amino-terminal end of the protein is very highly charged, with 25 acidic and 14 basic amino acids among the initial 91. The short unique carboxy-terminal domain is quite basic, with 9 basic (and 2 acidic) amino acids in the 30

```
   1 ATATCACATTTTTAACGTTCTCAATAAATAAATAAATAACAAAAAATGAAGAGGGTTATATGCTAATTTGCATTTAATTATTTGAAGCCGAAAAACTTTTACATACAATATAAACGCGGC  120
 121 AAAGTACAAAGGTCACACCGCGCGATAGGCATACACTTCTTAATATCGAAGTGCCACAGAGCCCGATATCGATAATATGTTTTTGTAAAACGTAAACAAAGCAGCCGCAACACGTGCATT  240
 241 TCATTCTCTCTCGTTTAGTTTAATAAAAACCGTGTAGATAATGGAATTATTTACTGTTAACAGGTAACTAAACTTGTGTTTATTACCTACAGTCATTTTAAATCAATTACACCCCAGATA  360
                                      M  E  L  F  T  V  N  R                    (intron 1)                                   Y
 361 TACCGAGGATTTGAAAGAACAAAAAGATGGGGCCCAAGGCACCAACAACGAGGATGAAATACTGCAAAAATTGCTTAAGAAGGCTGCAAAACGCAAAAGGAAACATGAAGCCATAGAAGT  480
       T  E  D  L  K  E  Q  K  D  G  A  Q  G  T  N  N  E  D  E  I  L  Q  K  L  L  K  K  A  A  K  R  K  R  K  H  E  A  I  E  V
 481 TGTAGAAACACCAATCCCGGAAAAAGAAACCTCAGATGTGAAGGAGTCGGAATCCAAGGAGGAGCAGGTGGAAGAACCAGAAAAGCCCCTAGAAGTTGTCCAAGAGGAAGTAGTTCCGTC  600
       V  E  T  P  I  P  E  K  E  T  S  D  V  K  E  S  E  S  K  E  E  Q  V  E  E  P  E  K ·P  L  E  V  V  Q  E  E  V  V  P  S
 601 TAATGAGTTTCAAGTTCTTGGAGGCGATGATTCGGCGGCCAAGAAGAAGAAAGTGCAGATGCAGCTGCCCAATTGGCTGGCACATCCAACCATCATTGAAGGTGGAAGCCTGCAGCCGGA  720
       N  E  F  Q  V  L  G  G  D  D  S  A  A  K  K  K  K  V  Q  M  Q  L  P  N  W  L  A  H  P  T  I  I  E  G  G  S  L  Q  P  E
 721 GGAGGAAGTGCCCGCCTCCGAGGCAATTGACCAGCTCGATTACCTGGAGAAGTACACATGTCAAGCTCTCAAGCAGATGAAGATTAAGCGGCTTTTTCCAGTCCAAAAGCAGGTCATCCC  840
       E  E  V  P  A  S  E  A  I  D  Q  L  D  Y  L  E  K  Y  T  C  Q  A  L  K  Q  M  K  I  K  R  L  F  P  V  Q  K  Q  V  I  P
 841 ATGGATCCTGGAGGCACACGCTAAGCCTCCTCCCTTTCGTCCCCGAGATATTTGTGTATCTGCTCCTACAGGCAGTGGAAAAACCTTGGCCTTTGCCATACCAATTGTCCAGTTGCTATC  960
       W  I  L  E  A  H  A  K  P  P  P  F  R  P  R  D  I  C  V  S  A  P  T  G  S  G  K  T  L  A  F  A  I  P  I  V  Q  L  L  S
 961 GCAGCGGGTGGATTGTAAAGTAAGGGCTTTAGTCGTCCTTCCTGTGGCAGAGCTGGCCCTACAGGTTTATCGGGTCATCAGTGAGTTGTGCAGCAAAACGGAGCTGGAGGTTTGCCTTCT 1080
       Q  R  V  D  C  K  V  R  A  L  V  V  L  P  V  A  E  L  A  L  Q  V  Y  R  V  I  S  E  L  C  S  K  T  E  L  E  V  C  L  L
1081 ATCCAAGCAGCACAAACTAGAGGATGAGCAGGAGAAGCTGGTGGAGCAATACAAAGGCAAATACTACTCCAAGGCAGACATTGTGGTGACCACACCAGGTGAGTTAAACAAACCGAGAGA 1200
       S  K  Q  H  K  L  E  D  E  Q  E  K  L  V  E  Q  Y  K  G  K  Y  Y  S  K  A  D  I  V  V  T  T  P  G          (intron 2)
1201 AGTCTTGTAGTTTCATTATGTATGTAATTTTCTTACAGGCCGCTTGGTTGACCATTTGCACGCCACGAAAGGGTTCTGCCTGAAGAGCCTAAAATTTCTGGTCATTGACGAGGCCGATCG 1320
                                               R  L  V  D  H  L  H  A  T  K  G  F  C  L  K  S  L  K  F  L  V  I  D  E  A  D  R
1321 AATAATGGATGCAGTTTTCCAAAACTGGCTTTATCATCTGGACAGTCATGTAAAGGAAACCACCGACCAGCTGCTGGCCGGCACTCAAGCACCTCTATGCTATGCAGAACTTCAAGCCAG 1440
       I  M  D  A  V  F  Q  N  W  L  Y  H  L  D  S  H  V  K  E  T  T  D  Q  L  L  A  G  T  Q  A  P  L  C  Y  A  E  L  Q  A  S
1441 CTTTGGCAAGCAACCACATAAGCTCCTGTTCTCGGCGACATTGTCTCAAGATCCCGAGAAGCTGCAGGATCTACGACTCTTTCAGCCTCGTCTGTTCGCCACCGTACTTACCATGCCCGT 1560
       F  G  K  Q  P  H  K  L  L  F  S  A  T  L  S  Q  D  P  E  K  L  Q  D  L  R  L  F  Q  P  R  L  F  A  T  V  L  T  M  P  V
1561 GTTGAAGGACGCAACGGAAGAAGGTGCTGATACCGAAGCGCTTACAGATCCAGGACAATTTGTGGGCAGATACACAACGCCGGCGGAGTTAACCGAGCAGTACTGTGTGACGGAGCTGCG 1680
       L  K  D  A  T  E  E  G  A  D  T  E  A  L  T  D  P  G  Q  F  V  G  R  Y  T  T  P  A  E  L  T  E  Q  Y  C  V  T  E  L  R
1681 ACTTAAACCCCTCACTGTTTTTGCCTTGGTGGAAAAGTACAAATGGAAGCGATTCTTGTGCTTCACCAACAGTTCGGATCAGGCAACTCGACTAACATTCGTGCTGAAAGTGCTATTTCA 1800
       L  K  P  L  T  V  F  A  L  V  E  K  Y  K  W  K  R  F  L  C  F  T  N  S  S  D  Q  A  T  R  L  T  F  V  L  K  V  L  F  Q
1801 AAAGTATAGCACCAAAGTGTCTGAATTGTCGGGAAATCTCTCGGCTAAGGTTCGAAATGAACGACTAAGGGACTTTGCTGCTGGGAAAATCAATGGACTGATCTGCTCAGATGCACTGGC 1920
       K  Y  S  T  K  V  S  E  L  S  G  N  L  S  A  K  V  R  N  E  R  L  R  D  F  A  A  G  K  I  N  G  L  I  C  S  D  A  L  A
1921 GCGTGGTATCGATGTGGCAGACGTGGATGTTGTGCTCTCATATGAGACACCTCGCCATATCACGACATACATTCATCGAGTGGGCCGAACCGCCTCGAGCGGGAAGAAAGGGCACCGCCG 2040
       R  G  I  D  V  A  D  V  D  V  V  L  S  Y  E  T  P  R  H  I  T  T  Y  I  H  R  V  G  R  T  A  S  S  G  K  K  G  H  R  R
2041 TCACCGTGCTCACGGAACAAGATATGACTTTATTCAAAAAAATACTAAGCGATGCGAACAAGGGATTGGGTGAGGAAATCCACGTTTCTCCAGATATTGAGATTCAACATGCAGTTGAAT 2160
       H  R  A  H  G  T  R  Y  D  F  I  Q  K  N  T  K  R  C  E  Q  G  I  G  *
2161 ACAAAGAGGCTTTACCGCGCCTGCGCTCGGAAAAGGTGAAGAACAAGAACCAAAAGATGGCCGAGAAGAATCGTGTGGCCACGCAAGGCATTGATCCACAAGAAACAGGAGGAAACGGCC 2280

2281 ACAGTTCGTCCACTGACGTTGATGGAAAAGTTGCAAATCAAAGCGAATGAAATCGTGCAATCATCAAAGAAATCCTCCGAAACGAAAAATTCCAAGACCAAGGCAGACAAAACAAAATAC 2400

2401 CAGCCTAAAAGGAAACCAAGAAGCAAATCATTGCCAAACAACTTAAGGCCATCCGAGAAC 2460
```

Figure 5. Genomic and complementary DNA sequences of the *Dbp73D* gene. The amino acid sequence of the predicted protein (one-letter code) is shown below the nucleotide sequence. Potential CAT and TATA boxes are marked with a solid underline. The two introns are labelled and underlined (wavy underline). The 5′ end of the p6 complementary DNA clone is at nucleotide 267, and the 3′ end of the cDNA clone is at nucleotide 2448.

residues from the end of the helicase domain to the termination codon.

Overall, using the BESTFIT program (gap weight 3.00, length weight 0.10; ref. 33) the predicted *Dbp73D* product was compared with other D-E-A-D proteins, and similarity scores between 44.2% (with Mss116) and 52.9% (with Tif) were obtained. Furthermore, the helicase domain of Dbp73D shares 36 of the 47 amino acid identities common to the initial eight members of the D-E-A-D helicase family (1,34). Five additional residues differ only by conservative substitutions (isoleucine for leucine in two cases; valine for leucine in one; tyrosine for phenylalanine in another, serine for threonine in the last). The predicted amino acid sequences for *Dbp73D* and four other RNA helicase genes are compared in Figure 6, showing clearly the identity of Dbp73D as a member of this protein family. Dbp73D appears to represent a novel member of the D-E-A-D protein class, however, as it has no particularly strong homology with any previously characterized individual member of the family. Of the 11 amino acids that differ in Dbp73D from the consensus sequence, seven are also different in Spb4, a protein involved in the assembly of the 60S ribosomal subunit in yeast (16). This may suggest a closer relationship between these two proteins than among the family in general, although their similarity is a moderate 50.3%.

Although Dbp73D is a member of the D-E-A-D family, in its sequence the number of residues between conserved motifs is often greater than in the other proteins; for example, positions 34–40, 215–233, and 323–353 in Figure 6. Some of these additional amino acids would be expected to confer relative structural differences to the *Dbp73D* protein, such as the eight amino acid sequence containing three proline residues from positions 33–40 in Figure 6, immediately amino-terminal to the A-motif of the ATP-binding site. Although some of the other proteins, most notably Spb4 and Mss-116, also exhibit such aberrant spacing (positions 114–118 and 307–314 for Mss-116, positions 313–324 for Spb4), the divergence from the consensus in Dbp73D is much greater. Dbp73D also carries a valine for isoleucine substitution in the third residue of the highly conserved H-R-I-G-R motif characteristic of D-E-A-D and not D-E-A-H proteins (refs. 1, 12; positions 433–437 in Figure 6). Spb4 carries a less conservative substitution at this residue, with the isoleucine replaced by a cysteine.

DISCUSSION

The sequence of *Dbp73D* is very similar to a large number of gene products, many of which have been implicated in RNA splicing or translation initiation . However, the metabolic function of *Dbp73D* has not been investigated in this study. Some screening for mutations has been carried out in the surrounding chromosomal region (35), leading to the identification of four genes in 73CD: two which are required for viability, one (*dark*

```
                  1                                                                                                     100
PL10 (183aa)      VEMGEIIMGN IELTRYTRPT PVQKHAIPII KE........ .KRDLMACAQ TGSGKTAAFL LPILSQIYTD GPGEALRAMK ENGKYGRRKQ YPISLVLAPT
Vasa (249aa)      ADLRDIIIDN VNKSGYKIPT PIQKCSIPVI SS........ .GRDLMACAQ TGSGKTAAFL LPILSKLLED PHELELGR.. .......... .PQVVIVSPT
MSS116 (110aa)    GVLDKEIHKA ITRMEFPGLT PVQQKTIKPI LSS....... EDHDVIARAK TGTGKTFAFL IPIFQHLINT KFDSQYM... .......... .VKAVIVAPT
Dbp73d (140aa)    LDYLEKYTQA LKQMKIKRLF PVQKQVIPWI LEAHAKPPPF RPRDICVSAP TGSGKTLAFA IPIVQLLSQR VDCK...... .......... .VRALVVLPV
SPB4  (11aa)      FSLLPWIRTG LDVMGFETMT PVQASTIPML AG........ .NKDVVVDSV TGSGKTAAFV IPVLEKVVKE EANTSKFKKA H......... .FHSLIIAPT
Consensus(4/5)    ------I--- ---------T PVQ---IP-I ---------- ---D----A- TGSGKT-AF- -PI------- ---------- ---------- --------PT

                  101                                                                                                   200
PL10              RELAVQIYEE ARK.....FS YRSRVRPCV. .VYGGADIGQ QIRDLERGC. .......HLL VATPGRLVDM MERGK.IGLD FCKYLVLDEA DRMLDMGFEP
Vasa              RELAIQIFNE ARK.....FA FESYLKIGI. .VYGGTSFRH QNECITRGC. .......HVV IATPGRLLDF VDRTF.ITFE DTRFVVLDEA DRMLDMGFSE
MSS116            RDLALQIEAE VKKIHDMNYG LKKYACVSL. .V.GGTDFRA AMNKMNKLRP .......NIV IATPGRLIDV LEKYSNKFFR FVDYKVLDEA DRLLEIGFRD
Dbp73d            AELALQVYRV IS.......E LCSKTELEVC LLSKQHKLED EQEKLVEQYK GKYYSKADIV VTTPGRLVDH LHATKGFCLK SLKFLVIDEA DRIMDAVFQN
SPB4              RELSRQIESV VL.......S FLEHYPSDLF PIKCQLLVGT NEATVRDDVS NFLRNRPQIL IGTPGRVLDF LQMPAVKT.S ACSMVVMDEA DRLLDMSFIK
Consensus(4/5)    RELA-QI--- ---------- ---------- ---------- ---------- ---------- --TPGRL-D- ---------- -----V-DEA DR-LD--F--

                  201                                                                                                   300
PL10              QIRRIVEQDT MPPK...... .......... ....GVRHTM MFSATFPKEI QMLARDFLDE ...YIFLAVG RVG.STSENI TQKVVWVEEA DKRSFLLDLL
Vasa              DMRRIMTHVT MRPEH..... .......... .......QTL MFSATFPEEI QRMAGEFNYV ...LRFVAIG IVG.GACSDV KQTIYEVNKY AKRSKLIEIL
MSS116            DLETISGILN EKNSK..... .......... ..SADNIKTL LFSATLDDKV QKLANNIMNK KECLFLDTVD KNEPEAHERI DQSVVISEKF AN.SIFAAVE
Dbp73d            WLYHLDSHVK ETTDQLLAGT QAPLCYAELQ ASFGKQPHKL LFSATLSQDP EKLQDLRLFQ PRLFATVLTM PVLKDATEEG ADT....EAL TDPGQFVGRY
SPB4              DTEKILRLLP KQR....... .......... .......RTG LFSATMRSA. ...GSDIFKT GLRNPVRITV NSKNQAPSSL KLNYCVVNPA EKLQLLVSIL
Consensus(4/5)    ----I----- ---------- ---------- --------T- -FSAT----- ---------- ---------- ---------- ---------- ----------

                  301                                                                                                   400
PL10              NATGKD.... ..SLILVFVE TK........ .......... .......... ...KGADSLE DFL...YHEG YACTSIHGDR SQRDREEALH QFRSGKS.PI
Vasa              SEQADG.... ....TIVFVE TK........ .......... .......... ...RGADFLA SFL...SEKE FPTTSIHGDR LQSQREQALR DFKNG.SMKV
MSS116            HIKKQIKERD SNYKAIIFAP TV........ .......... .......... ...KFTSFLC SILKNEFKKD LPILEFHGKI TQNKRTSLVK RFKKDES.GI
Dbp73d            TTPAEL.... ....TEQYCV TELRLKPLTV FALVEKYKWK RFLCFTNSSD QATRLTFVLK VLFQKYSTKV SELSGNLSAK VRNER...LR DFAAG.KING
SPB4              NNYKFK.... ...KCIVYFP TCVSVSYFYS FIQY...... .......... .......LGK RNI...LVNE VEIFSLHGKL QTSARTKTLT AFTDSLSNSV
Consensus(4/5)    ---------- ---------- T--------- ---------- ---------- --------L- ---------- ------HG-- ----R---L- -F----S---

                  401                                              450
PL10              LVATAVAARG LDISNVKHVI NFDLPSDIEE YVHRIGRTGR VGNLGLATSF (116aa)
Vasa              LIATSVASRG LDIKNIKHVI NYDMPSKIDD YVHRIGRTGR VGNNGRATSF ( 69aa)
MSS116            LVCTDVGARG MDFPNVHEVL QIGVPSELAN YIHRIGRTAR SGKEGSSVLF (185aa)
Dbp73d            LICSDALARG IDVADVDVVL SYETPRHITT YIHRVGRTAS SGKKGHRRHR ( 21aa)
SPB4              LFTTDVAARG IDIPDVDLYI QLDPPTNTDM FMHRCGRTGR ANRVGKAITF (233aa)
Consensus(4/5)    L--T-V-ARG -D---V--V- ----P----- Y-HR-GRT-R -G--G----F
```

Figure 6. Alignment of the amino acid sequences in the conserved region of D-E-A-D box proteins for Dbp73D and four other family members: PL10 from mouse (39), vasa from *D. melanogaster* (18, 19), Mss-116 from *S. cerevisiae* (15), and Spb4 from *S. cerevisiae* (16). The consensus line lists amino acids which are identical in four of the five proteins. The numbers of amino acids in parenthesis gives the lengths of the amino- and carboxy-terminal regions of that protein which are not shown. Note the additional amino acids in the Spb4 and Dbp73D sequences from positions 150–157 in the alignment, and from positions 323–353.

body) which is involved in cuticular pigmentation, and one (*plucked*) in which a mutation results in fewer abdominal tergite bristles. It would be instructive to determine whether the *Dbp73D* clone can rescue the phenotype of either of the lethal mutations. The immediate chromosomal region surrounding *Dbp73D* has not yet been screened for recessive genes affecting fertility, nor has this region been saturated for recessive lethal mutations. Two available deficiencies, *Df(3L)st b11* and *Df(3L)st j11*, should make such screens possible (36). Another potential approach to isolating mutations in this gene would employ one of the PCR-based transposon mutagenesis procedures recently described (37, 38).

The D-E-A-D family of putative RNA helicase genes is likely to be very large in Drosophila, as it is in *S. cerevisiae* (2, 3). Such a large number of genes implies a large degree of specificity in the RNA substrates with which they may interact. Like *vasa* and *ME31B* (31), *Dbp73D* is expressed in the germ line tissue of the ovary; its expression pattern in this tissue is strikingly similar to that reported for *ME31B* (31), with expression beginning at stage 6. As the *Dbp73D* and *ME31B* genes are expressed in the female germ line at the time that patterning of transcripts within the oocyte is taking place (25, 40), it will be of interest to determine the roles these genes may play in oocyte development. Dbp73D differs from other D-E-A-D proteins in its novel spacing of the conserved sequence motifs, so it is possible that it is involved in genetic regulatory pathways unrelated to those which may be mediated by previously-described genes of this type. Analysis of the *Dbp73D* protein, isolation of mutations in the *Dbp73D* gene, and identification of the RNA molecules it may bind *in vivo* should assist in the understanding of the roles played by this gene family in developmental processes.

ACKNOWLEDGMENTS

The original *Dm73D* clone was isolated in the laboratory of Charles Zuker, and characterization of it began in Cambridge where it was supported by an MRC Programme Grant to Michael Ashburner. We thank J. Trenear for polytene chromosome hybridizations, and the Sheldon Biotechnology Centre at McGill for synthesis of oligonucleotides. This work was also supported by operating grants from the Natural Sciences and Engineering Council of Canada and the National Cancer Institute of Canada, with funds provided by the Canadian Cancer Society, to P.L. L.F.P. was a fellow of the Churchill Foundation.

REFERENCES

1. Linder,P., Lasko, P.F., Ashburner,M., Leroy,P., Nielsen,P.J., Nishi,K., Schnier,J., and Slonimski,P.P. (1989) *Nature* **337**, 121–122.
2. Wasserman,D.A. and Steitz,J.A. (1991) *Nature* **349**, 463–464.
3. Chang,T.-H., Arenas,J. and Abelson,J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1571–1575.
4. Ray,B.K., Lawson,T.G., Kramer,J.C., Cladaras,M.H., Grifo,J.A., Abramson,R.D., Merrick,W.C. and Thach,R.E. (1985) *J. Biol. Chem.* **260**, 7651–7658.
5. Sonenberg, N. (1988) *Prog. Nuc. Acids Res. Mol. Biol.* **35**, 173–207.
6. Lawson,T.G., Lee,K.A., Maimone,M.M., Abramson,R.D., Dever,T.E., Merrick,W.C. and Thach,R.E. (1989) *Biochemistry* **28**, 4729–4734.
7. Rozen,F., Edery,I., Meerovitch,K., Dever,T.E., Merrick,W.C. and Sonenberg,N. (1990) *Mol. Cell. Biol.* **10**, 1134–1144.
8. Lazaris-Karatzas,A., Montine,K.S., and Sonenberg,N. (1990) *Nature* **345**, 544–547.
9. Chen,J.-H. and Lin,R.-J. (1990) *Nucleic Acids Res.* **18**, 6447.
10. Burgess,S., Couto,J.R. and Guthrie,C. (1990) *Cell* **60**, 705–717.
11. Company,M., Arenas,J. and Abelson,J. (1991) *Nature* **349**, 487–493.
12. Ruby,S.W. and Abelson,J. (1991) *Trends Genet.* **7**, 79–85.
13. Jamieson,D.J., Rahe,B., Pringle,J. and Beggs,J.D. (1991) *Nature* **349**, 715–717.
14. Strauss,E.J. and Guthrie,C. (1991) *Genes Dev.* **5**, 629–641.
15. Séraphin,B., Simon,M., Boulet,A., and Faye,G. (1989) *Nature* **337**, 84–87.
16. Sachs,A.B. and Davis,R.W. (1990) *Science* **247**, 1077–1079.
17. Hay,B., Ackerman,L., Barbel,S., Jan,L.Y., and Jan,Y.N. (1988) *Development* **103**, 625–640.
18. Lasko,P.F. and Ashburner,M. (1988) *Nature* **335**, 611–617.
19. Hay,B., Jan,L.Y., and Jan,Y.N. (1988) *Cell* **55**, 577–587.
20. Lasko,P.F. and Ashburner,M. (1990) *Genes Dev.* **4**, 905–922.
21. Schüpbach,T. and Wieschaus,E. (1986) *Roux Arch. Dev. Biol.* **195**, 302–317.
22. Sambrook,J., Fritsch,E.F., and Maniatis,T. (1989) Molecular Cloning: A Laboratory Manual, second edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
23. Tautz,D. and Pfeifle,C. (1989) *Chromosoma* **98**, 81–85.
24. Ephrussi,A., Dickinson,L.K., and Lehmann,R. (1991) *Cell* **66**, 1–20.
25. Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.* **132**, 6–13.
26. Ashburner,M. (1989) *Drosophila*: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
27. Sanger,F., Nicklen,S., and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
28. Yanisch-Perron,C., Vieira,J., and Messing,J. (1985) *Gene* **33**, 103–119.
29. Henikoff,S. (1987) *Meth. Enzymol.* **155**, 156–165.
30. Maniatis,T., Hardison,H.C., Lacy,E., Lauer,J., O'Connell,C., Quon,D., Sim,G.K., and Efstratiadis,A. (1978) *Cell* **15**, 687–701.
31. deValoir,T, Tucker,M.A., Belikoff,E.J., Camp,L.A., Bolduc,C. and Beckingham,K. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2113–2117.
32. King,R.C. (1970) Ovarian Development in *Drosophila melanogaster*. Academic Press, New York.
33. Devereux,J., Haeberli,P., and Smithies,O. (1984) *Nucleic Acids Res.* **12**, 387–395.
34. Schnier,J. (1989) *Nature* **340**, 246 (correction).
35. Belote,J.M., Hoffmann,F.M., McKeown,M., Chorsky,R.L. and Baker,B.S. (1990) *Genetics* **125**: 783–793.
36. Belote,J.M. and McKeown,M. (1985) *Drosophila Inform. Serv.* **61**, 33–34.
37. Ballinger,D.G. and Benzer,S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9402–9406.
38. Kaiser,K. and Goodman,S.F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1686–1690.
39. Leroy,P., Alzari,P., Sassoon,D., Wolgemuth,D. and Fellous,M. (1989) *Cell* **57**, 549–559.
40. Kim-Ha,J., Smith,J.L., and Macdonald,P.M. (1991) *Cell* **66**, 23–35.