The EMBO Journal Vol.16 No.8 pp.2023–2031, 1997

The *Adh-related* gene of *Drosophila melanogaster* is expressed as a functional dicistronic messenger RNA: multigenic transcription in higher organisms

Saverio Brogna[1] and Michael Ashburner

Department of Genetics, Downing Street, University of Cambridge, Cambridge CB2 3EH, UK

[1]Corresponding author
e-mail: sb@mole.bio.cam.ac.uk

Essentially all eukaryotic cellular mRNAs are monocistronic, and are usually transcribed individually. Two tandemly arranged *Drosophila* genes, alcohol dehydrogenase (*Adh*) and *Adh-related (Adhr)*, are transcribed as a dicistronic transcript. From transcripts initiated from the *Adh* promoter, two classes of mRNA are accumulated, one is monocistronic and encodes *Adh* alone, the other is dicistronic and includes the open reading frames of both *Adh* and *Adhr*. The dicistronic transcript is found in polysomes and the *Adhr* protein product is detected by antibody staining. We present evidence that the accumulation of the dicistronic mRNA is controlled at the level of the 3′ end processing.
Keywords: dicistronic mRNA/*Drosophila Adh*/gene duplication/transcription/translation

Introduction

Protein-encoding genes of eukaryotes typically are expressed as monocistronic mRNAs transcribed individually, in contrast to many genes of prokaryotes that are expressed as polycistronic operons. In the protozoan trypanosomes and the nematode *Caenorhabditis elegans*, however, examples have been found recently of genes that are co-transcribed, in a way that resembles a prokaryotic operon (Johnson *et al.*, 1987; Clayton, 1992; Spieth *et al.*, 1993; Zorio *et al.*, 1994). In these examples, the polycistronic RNA is processed into individual mRNAs, each corresponding to a single gene, before translation. The cleavage of the longer polycistronic transcript is accomplished by polyadenylation and *trans*-splicing (Blumenthal, 1995). Such complex processing is thought to be necessary in order to optimize the translation of the internal genes (Spieth *et al.*, 1993; Blumenthal, 1995). Most eukaryotic mRNAs have only one functional open reading frame (ORF), an exception being some viral transcripts (Kozak, 1986; Bonneville *et al.*, 1989) and, possibly, a mammalian transcript (Lee, 1991). In *Drosophila*, polycistronic mRNAs have not yet been described, even though artificial polycistronic mRNAs can be translated efficiently in cultured cells and transgenic flies (Oh *et al.*, 1992; Hart and Bienz, 1996), and there is evidence that *stnA* and *stnB* are co-transcribed (Andrews *et al.*, 1996).

The alcohol dehydrogenase (*Adh*) and *Adh-related* (*Adhr*) genes of *D.melanogaster* are arranged in tandem, and most likely originated from a common ancestor by gene duplication (Schaeffer and Aquadro, 1987; Kreitman and Hudson, 1991; Jeffs *et al.*, 1994). The ADH and putative ADHR proteins are ~40% identical at the amino acid level and both belong to the large family of short-chain dehydrogenase enzymes (Persson *et al.*, 1991). Due to their high sequence divergence, and the observation that point mutations in *Adh* abolish all the ADH activity in the fly, it is considered that *Adh* and *Adhr* have distinct biochemical functions (Jeffs *et al.*, 1994). *Adh* and *Adhr* are arranged as a direct tandem pair of genes, separated by 300 bp or so. This arrangement is found in species as different as *D.melanogaster* and *Scaptodrosophila lebanonensis*, species that would have last shared a common ancestor at least 40 million years ago (Marfany and Gonzalez-Duarte, 1991; Albalat and Gonzalez-Duarte, 1993). In *D.melanogaster*, *Adh* is transcribed from one of two promoters (distal and proximal) depending on the stage of development; transcription in larvae is largely from the proximal promoter, that in adults from the distal (Benyajati *et al.*, 1983b; Savakis and Ashburner, 1985). The discovery of *Adhr* (Schaeffer and Aquadro, 1987), and of its similarity in sequence and intron–exon structure to *Adh*, naturally raised questions about its function. Although there have been hints that it is transcribed (Kreitman and Hudson, 1991; Albalat and Gonzalez-Duarte, 1993), cDNAs have eluded discovery. Moreover, the fact that deletions that include both *Adh* and *Adhr* are homozygous viable (Ashburner *et al.*, 1982; Chia *et al.*, 1985), has somewhat discouraged attempts to discover the function of *Adhr* itself.

Here we describe the structure of the full-length *Adhr* transcript and its expression pattern in adult flies. We also show that the ADHR protein can be detected by antibody staining. The characterization of the full-length transcript shows that *Adhr* is expressed as a functional dicistronic transcript together with *Adh*. Its expression is driven by the distal *Adh* promoter in adult flies and by the proximal *Adh* promoter in embryos. The choice between the monocistronic transcript, encoding only ADH, and the dicistronic transcript, encoding both proteins, probably depends upon the choice of polyadenylation sites, either that immediately downstream of *Adh* or that downstream of *Adhr*. We find that the dicistronic transcript is more abundant in flies carrying a mutant allele of *suppressor of forked* [*su(f)*], a gene that recently has been shown to be involved in the 3′ processing of mRNAs (Mitchelson *et al.*, 1993; Takagaki and Manley, 1994).

Results

The mRNA of Adhr is polycistronic

Preliminary experiments. It has been suggested that the 298 bp between the termination codon of *Adh* and the putative initiator codon of *Adhr* include a TATA box

© Oxford University Press

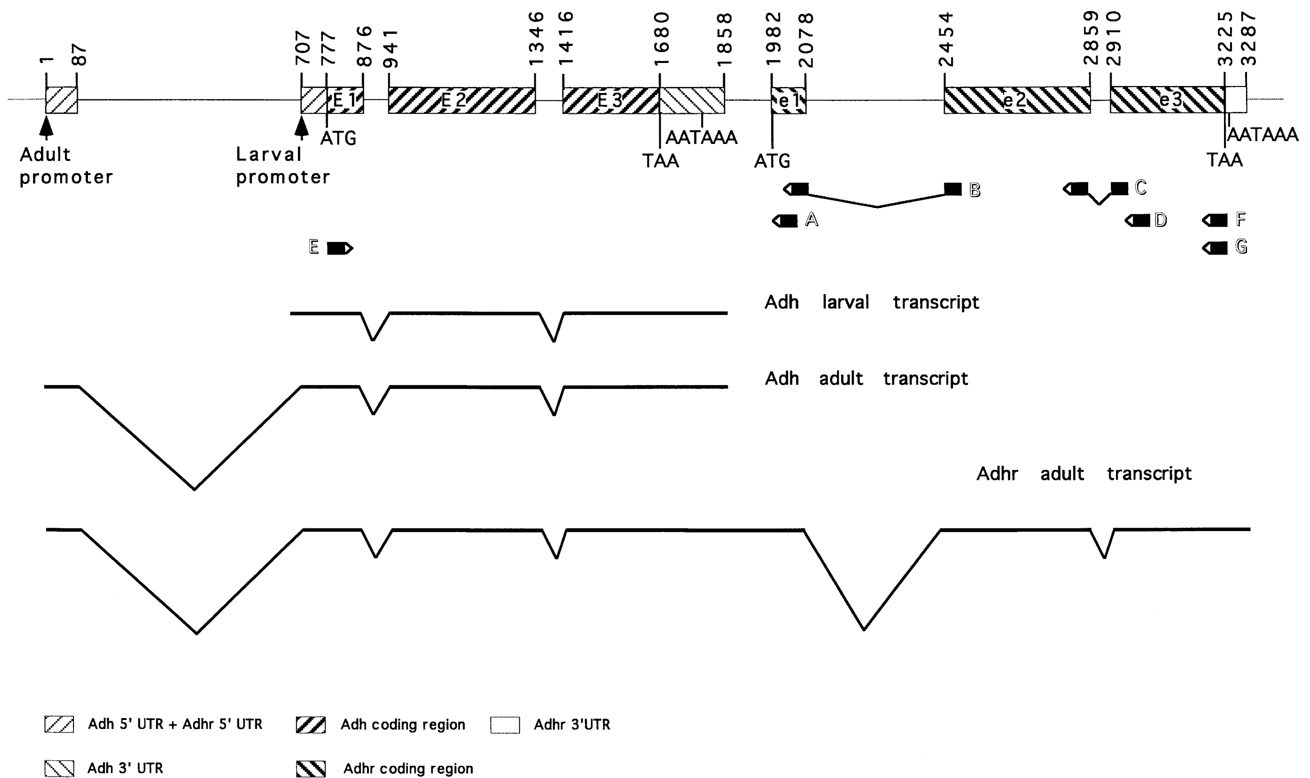

Fig. 1. Schematic diagram of the structure of *Adh* and *Adhr* is shown above. Rectangles represent exons and lines represent introns. A key to the features of the exons is shown. The coding exons named E1, E2 and E3 are in *Adh*. The exons named e1, e2 and e3 are *Adhr* coding exons. The small black rectangles indicate the positions of the primers used. The sequences of the primers are given in Materials and methods. Below is shown a diagram of the two *Adh* cDNAs and the polycistronic *Adh–Adhr* cDNA. The sequence of the adult *Adhr* cDNA has been submitted to the DDBJ/EMBL/GenBank database (accession No. X98338). The sequence of the entire genomic region has been reported by Kreitman and Hudson (1991), accession No. X78384.

for *Adhr* (Schaeffer and Aquadro, 1987; Marfany and Gonzalez-Duarte, 1991; Jeffs *et al.*, 1994). Some 10 bp 3′ to the putative termination codon of *Adhr* there is a sequence that can be interpreted as being a polyadenylation signal sequence (see Figure 1). Were these signals both to be used, then a processed *Adhr* transcript of ~1000 nucleotides would be expected. In preliminary experiments, no transcript of this length could be seen in Northern blots, using either total or poly(A)$^+$-selected RNAs prepared from a variety of developmental stages. Attempts by us and others (M.Kreitman, personal communication) to recover cDNA clones corresponding to *Adhr* were equally unsuccessful. Evidence for an *Adhr* transcript could, however, be obtained by RT–PCR using primers internal to the *Adhr* coding region. In order to determine the structure of this transcript, we attempted to map its 5′ and 3′ ends by RT–PCR, using RNA prepared from adult flies (see Materials and methods). The 3′ end was readily mapped to 43 bp downstream of the putative polyadenylation signal sequence as shown in Figure 1; the 5′ end could not be determined in these experiments. A similar result has been obtained in M.Kreitman's laboratory (personal communication). Subsequently, we were able to detect a rare transcript in adult flies, larvae and embryos. Accurate sizing of this transcript revealed that, contrary to expectation, it was >2000 nucleotides long (Figure 2A). This suggests that the promoter of the *Adhr* gene may be located further upstream than was thought, either within the *Adh* gene or upstream of it. The very low abundance of the transcript, as finally detected in Northern blots (Figure 2B), explains why previous attempts to find it, or its cDNAs, had been so frustrating.

5′ End RACE of the Adhr transcript. Knowing that the 5′ end of the *Adhr* transcript must be well 5′ to the *Adh–Adhr* spacer region allowed us to design primers that would allow a more accurate description of this RNA. In the first experiment, we synthesized single-stranded cDNA from total adult RNA, with a primer located at the very end of the *Adhr* coding region, primer F (Figure 1). The product was tailed and a nested PCR amplification was performed with two primers located in the third and first exons of *Adhr* (primers D and A). The 1200 bp PCR product, obtained in the second cycle of amplification, was cloned and several clones were sequenced. This revealed that the 5′ end of the *Adhr* cDNA also contained the *Adh* ORF and, moreover, that the transcript was initiated from the distal *Adh* promoter. This sequenced PCR product indicated the existence of a processed RNA that included both the *Adh* and *Adhr* exons, as well as the 298 nucleotide intergenic spacer. The predicted size of a processed transcript initiated from the distal *Adh* promoter and terminating 43 nucleotides downstream of the putative *Adhr* polyadenylation signal sequence is 2071 nucleotides, in agreement with the size of the *Adhr* transcript seen on Northern blots.

It is possible that the amplification product could have been synthesized from an *Adh* transcript that happened to

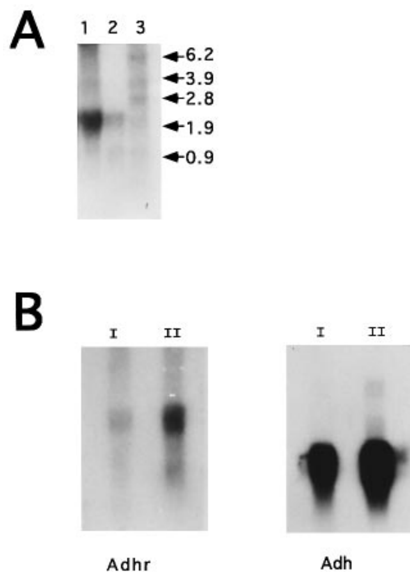

Fig. 2. *Adhr* transcription. (**A**) Northern blot analysis of approximately equal amounts of total RNA isolated from wild-type flies (lane 1) and embryos (lane 2). RNA size standards (Promega) are in lane 3. The probe used is specific for the third exon of *Adhr* (e3), that has been amplified from a clone of the region (pATSac). (**B**) Comparison of the level of accumulation of *Adhr* transcript relative to *Adh*. Adult total RNA (lane I) and adult poly(A)$^+$ RNA (lane II). The *Adhr*-specific probe was the total coding region, isolated as an *Nde*I–*Hin*dIII fragment from the pUCR8.1 plasmid. The *Adh*-specific probe was amplified by PCR from an *Adh* genomic clone. The filter was first hybridized with the *Adhr* probe, then stripped and hybridized with an *Adh* probe. The same amount of radiolabelled probe was used in both hybridizations; the filter probed with the *Adhr* probe was exposed for 6 days at –70°C and the filter probed with the *Adh* probe was exposed for 1 h at room temperature.

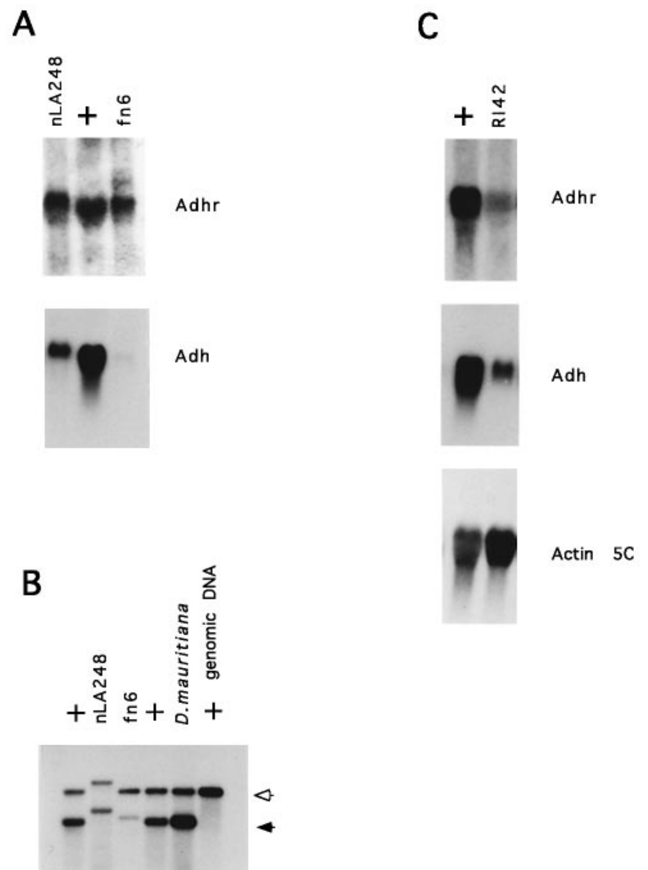

Fig. 3. *Adhr* transcription in *Adh* mutant alleles. (**A**) Northern blot analysis of approximately equal amount of total RNA extract from Adh^{nLA248}/*Df(2L)A63*, wild-type and Adh^{fn6}/*Df(2L)A63*. The *Adhr*- and the *Adh*-specific probes are as in Figure 2B. Note that the transcripts from Adh^{nLA248} and Adh^{fn6} migrate more slowly than that of the wild-type. The *Adh* transcript in Adh^{fn6} RNA is barely visible, as expected from the data of Benyajati *et al.* (1983a). (**B**) *Adhr* RT–PCR with primer E (5′-caagaacgtia/gttttcgtg/tgccgg-3′) and primer G (5′-actgg/atcggccatg/atgccag/atacc-3′) from ~5 µg of adult total RNA. After 3 min at 94°C, PCR conditions were 94°C for 30 s, 65°C for 1 min and 72°C for 2 min for 30 cycles. The PCR reactions were separated on a 1.1% agarose gel and a 123 bp ladder (Gibco-BRL) was used as size markers. The gel was blotted and the filter was hybridized with the *Adhr* probe described in Figure 2B. Note that the bands indicated by the empty arrow correspond to the genomic (or unspliced) product and the black arrow indicates the position of the cDNA fragment.
(**C**) Northern blot analysis of total RNA prepared from wild-type and Adh^{RI42} adult flies. The *Adhr* and *Adh* probes are the same as above, the probe specific for the cytoplasmic *Actin-5C* is the plasmid clone pDmA2. The relative intensity of the signals was estimated with the NIH Image program, from filters exposed for a different length of time.

run through the normal *Adh* polyadenylation site into *Adhr*. It is known that in tissue culture cells *Adh* transcription can terminate well 3′ to its 'normal' site (Benyajati and Dray, 1984). For this reason, a second type of RACE experiment was performed. The single-stranded cDNA in this case was synthesized as before, but this time from poly(A)$^+$ adult RNA rather than from total RNA, and different primers were used for the nested PCR. One of the primers was complementary to a region of the cDNA spanning the 5′ end of the third exon and the 3′ end of the second exon of *Adhr* (primer C), and the other was designed to span the junction of the second and first exon of *Adhr* (primer B). These primers would fail to amplify either a genomic template or an unspliced transcript. The sequenced product of this reaction had the same structure as that seen in the first of the RACE experiments. These experiments indicate that *Adhr* is transcribed as a dicistronic mRNA from the *Adh* distal promoter.

The dicistronic transcript can also be amplified by RT–PCR using one primer at the beginning of the *Adh* coding region (primer E) and a second one complementary to the end of *Adhr* (primer G) directly from adult total RNA from *D.melanogaster* and *D.mauritiana* (Figure 3B).

RACE experiments, like those described above, were also performed with total RNA from embryos. The sequencing of the RACE product showed that the embryonic *Adhr* transcript is also dicistronic, but it is transcribed from the proximal promoter of the *Adh* gene.

Mutations that affect Adh transcripts also affect Adhr transcripts

Some *Adh* null alleles produce an mRNA that is longer than that of the wild-type allele. One of these is Adh^{fn6}, a formaldehyde-induced mutation, where a sequence rearrangement prevents the splicing of the first 65 bp intron of *Adh* (Benyajati *et al.*, 1983a). A second, Adh^{nLA248}, is an X-ray-induced tandem duplication within the *Adh* gene that produces an mRNA ~200 nucleotides longer than that of the wild-type allele (Chia *et al.*, 1995). If *Adhr* was a quite independent transcription unit to *Adh* then neither mutation would be expected to affect the size of the *Adhr* transcript. In fact (Figure 3A), the size of the *Adhr*

transcript is increased in $Adh^{nLA248}$and Adh^{fn6}. The same result was observed by comparing the sizes of the RT–PCR fragments obtained using one primer at the beginning of the *Adh* coding region (primer E; Figure 1) and a second one complementary to the end of *Adhr* (primer G) directly from total cDNA. The PCR fragment corresponding to the *Adhr* cDNA is 1823 bp, ~200 bp longer in Adh^{nLA248} and ~65 bp longer in Adh^{fn6} (Figure 3B).

The Adh promoter also drives the expression of Adhr

The sequence of the *Adhr* cDNA and the increase in size of the *Adhr* transcript seen in the Adh^{nLA248} and Adh^{fn6} strains are consistent with the hypothesis that transcription from the distal promoter of *Adh* results in a dicistronic transcript. If so, then mutations that down-regulate this promoter (as assayed by a lower steady-state level of *Adh* mRNA) would also be expected to reduce the level of *Adhr* transcript. To test this, we used the Adh^{RI42} allele, which carries an insertion of a *copia* element 243 bp 5′ to the distal *Adh* transcription start site. This allele results in an ~3-fold reduction in the steady-state level of *Adh* mRNA, largely (if not entirely) due to reduced transcription from the distal promoter (Dunn and Laurie, 1995). In Adh^{RI42} flies, the levels of both *Adh* and *Adhr* transcripts are reduced, to ~20% of control levels (Figure 3C).

The dicistronic Adh–Adhr mRNA is translated

The experiments discussed so far are consistent with the hypothesis that *Adhr* is co-transcribed, as a dicistronic RNA, with *Adh*. Is this transcript functional with respect to *Adhr*? If so, we would expect to find this transcript associated with polysomes and to be able to detect ADHR protein.

Preliminary data indicated that the dicistronic transcript was associated with polysomes prepared from adult wild-type flies, and ADHR protein could be detected by antibody staining (see below). Taking into consideration the fact that the dicistronic transcript is the only transcript we were able to detect, the data strongly suggested that the dicistroninic transcript is functional.

However, it was possible that the dicistronic transcript is associated with polysomes only if the upstream *Adh* ORF is translated. In order to test if the second ORF coding for *Adhr* is translated, we prepared polysomes from flies carrying a stop codon mutation in the upstream *Adh* ORF. The allele Adh^{nBR114} is an *Adh* null allele where a T→A transversion mutates the 64th codon of *Adh* to a TAA termination codon (Fosset *et al.*, 1990). In Figure 4A and B, we show the profile of the polysomes prepared from wild-type and Adh^{nBR114} flies. The RNA was extracted from the total polysomal fraction and from a heavier component estimated to carry on average more than two ribosomes. The distribution of the monocistronic *Adh* transcript and that of the dicistronic *Adhr* transcript was analysed, by Northern blots, in total RNA, in total polysomal RNA and in RNA from the heavier fraction (Figure 4D). In wild-type (Figure 4D, lanes 1–3), the *Adh* transcript was present in all three fractions, but in Adh^{nBR114} (lanes 4–6) the *Adh* transcript was absent in the RNA from the heavier polysomes (Figure 4D, lane 6). These data indicate that in Adh^{nBR114} the monocistronic *Adh* transcript is associated with not more than two ribosomes,

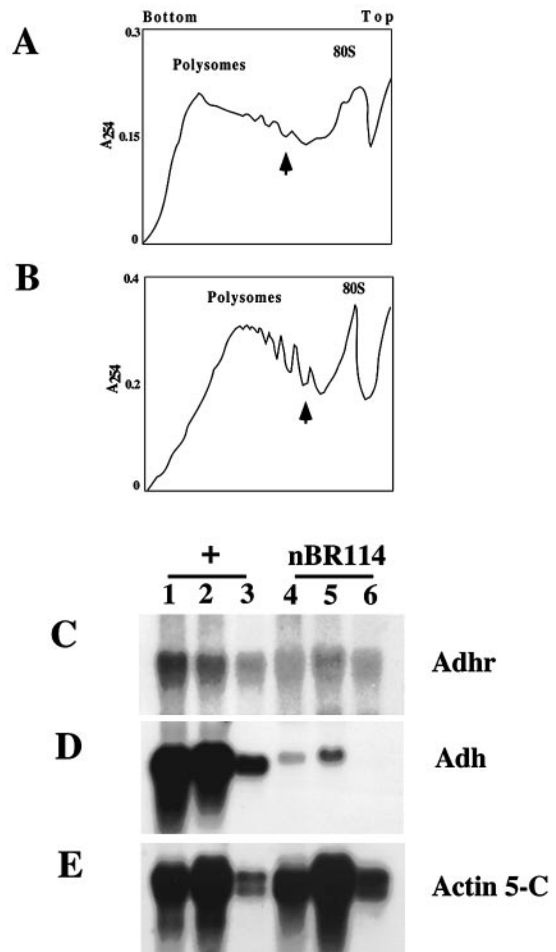

Fig. 4. Sedimentation profiles of polysomes and RNA analysis. (**A**) Profile of polysomes isolated from wild-type adult flies. (**B**) Profile of polysomes isolated from Adh^{nBR114} adult flies. The polysomal fraction were centrifuged through a 15–60% sucrose gradient and the absorbance at 254 nm monitored as described in Materials and methods. The S values are only approximate. The polysomes migrating faster than the point indicated by the black arrow were collected together for RNA extraction. (**C**–**E**) Northern blot analysis of RNA. RNAs from wild-type (lanes 1–3) and Adh^{nBR114} flies (lanes 4–6). Total RNA extract from adult flies (lanes 1 and 4), RNA extracted from the all polysomal fraction (before sucrose gradient sedimentation, lanes 2 and 5) and RNA extracted from polysomes associated with more than two ribosomes (faster than the point indicated by the arrow in A and B, lanes 3 and 6). Note that the amount of RNA in lane 3 is about three times less than in lanes 1 and 2 and the amount of RNA in lane 6 is about two-thirds that of lane 4 (the relative amounts were estimated by comparing the relative intensity of the *Actin* signals with the NIH Image program, from autoradiograms). The *Adhr* and *Adh* probes are described in Figure 2B. The *Actin* probe is as in Figure 3B. Two similar gels equally loaded were transferred and one was probed with *Adhr* (sp. act. ~10^9 d.p.m./μg) and the other was first probed with *Adh* and then stripped and reprobed with *Actin-5C* (both labelled to a sp. act. of ~10^8 d.p.m./μg). The filter in (C) was exposed for 5 days at –70°C and those in (D) and (E) for ~5 h at –70°C.

probably on average one is pausing in the proximity of the initiation codon and one in the proximity of the termination codon (Wolin and Walter, 1988). If the dicistronic transcript is associated with polysomes, only because of the translation of *Adh*, it would not be expected in the RNA extracted from the heavier polysomal fraction of Adh^{nBR114}. In Figure 4A, we show that, in contrast to the monocistronic *Adh* transcript, the *Adhr* transcript is associated with the heavier polysomal fraction (lane 6). It

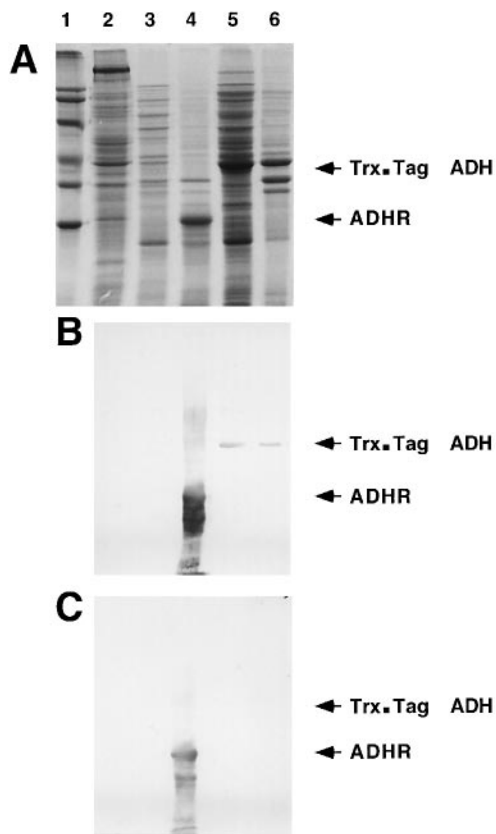

Fig. 5. Anti-ADHR antibody characterization. (**A**) SDS–12% polyacrylamide gel electrophoresis, Coomassie-stained protein markers (29, 36, 45, 66, 97 and 116 kDa) (lane 1), adult flies protein extract (lane 2), soluble fraction of a bacterial extract expressing ADHR (lane 3), ADHR inclusion bodies (lane 4), soluble fraction of a bacterial extract expressing ADH fused to thioredoxin (Trx, lane 5) and ADH-Trx inclusion bodies (lane 6). (**B**) Western blot analysis of a gel loaded as in (A), but without markers. The gel was transferred and probed with a 1:2000 dilution of anti-ADHR antibody affinity purified against the antigen (see Materials and methods). (**C**) Western blot as above but probed with anti-ADHR pre-absorbed against ADH (see Materials and methods).

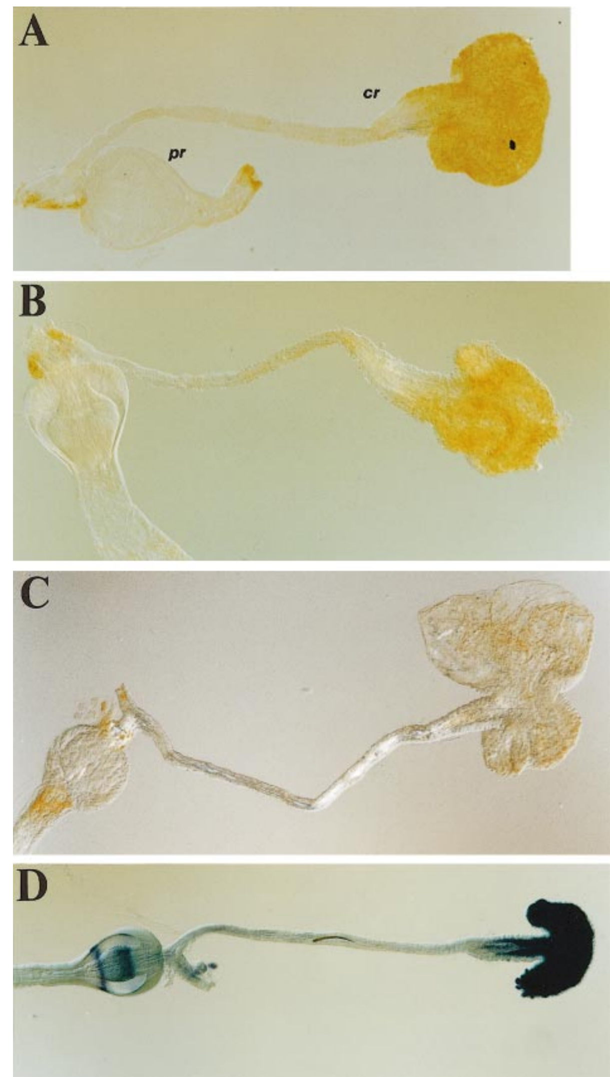

Fig. 6. Immunostaining and histochemistry. (A, B and C) Anterior portion of adult midgut complete with crop (cr) and proventriculus (pr) stained with affinity-purified anti-ADHR antibody (for the staining in B the antibody was pre-absorbed against recombinat ADH, see Results and Materials and methods. (**A**) From wild-type flies, (**B**) from Adh^{nBR114} flies and (**C**) from the deficient strain *Df(2L)TE35B-SR54/ Df(2L)A72*. (**D**) ADH activity in wild-type tissues.

should also be noted that the level of *Adhr* transcript is unchanged in Adh^{nBR114}; the level of *Adhr* transcript is also unchanged in several other nonsense *Adh* alleles which reduce the steady-state amount of *Adh* mRNA (Chia *et al.*, 1995, and S.Brogna, unpublished data).

The data above indicate that the *Adhr* ORF of the dicistronic transcript is being translated, and strongly suggest that *Adhr* translation is initiated by internal initiation in the intergenic region between *Adh* and *Adhr* coding regions.

ADHR protein can be detected

A polyclonal antibody was raised against ADHR protein synthesized in *Escherichia coli* (see Materials and methods) and affinity purified against the antigen. In Figure 5, we show by a Western blot that the affinity-purified antibody is very specific for the antigen (lane 4), but fails to detect any peptide from a total fly protein extract (lane 2), probably due to the low abundance of ADHR in the fly. The antibody has a very weak cross-reaction with recombinant ADH (Figure 5B, lanes 5 and 6). In order to subtract from the antibody the component that cross-reacts with ADH, we pre-absorbed the previously affinity-purified antibody against recombinant ADH expressed in *E.coli* (Figure 5C, lanes 5 and 6), to provide a reagent for immunostaining of adult tissues. Two questions were of interest: can ADHR protein be detected in adult flies? and, if so, does this protein co-localize with ADH?

The answer to both questions is yes. ADHR can be detected by immunostaining in the saccular region of the adult crop (Figure 6A), a pedunculate sac evaginating from the posterior part of the oesophagus. Staining against both ADHR and ADH was absent in the crop (or elsewhere) of flies from a strain deleted for both genes [*Df(2L)TE35B-SR54/Df(2L)A72*] (Figure 6C). ADHR protein can also be detected in the ADH null allele Adh^{nBR114}, and its level does not seems to be different from that in wild-type (Figure 6B). There is also staining for ADHR in the adult fat body, but is hard to say if this is specific, because of a high background in the deletion controls.

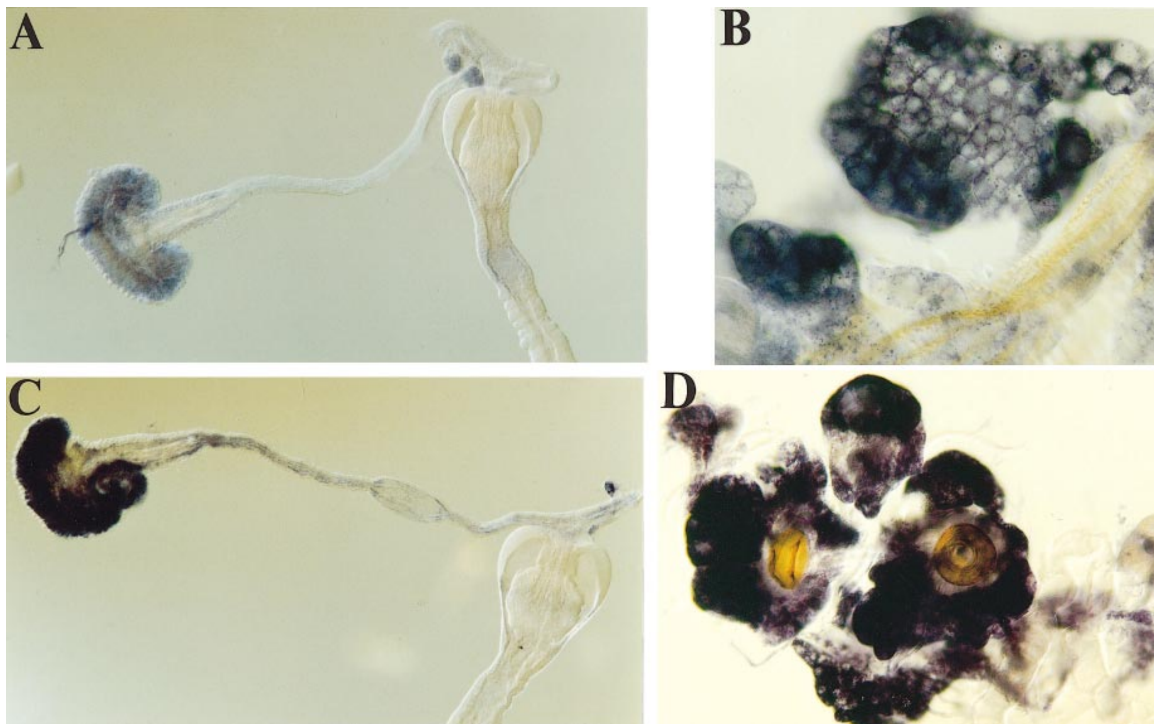

Fig. 7. Tissue distribution. *In situ* hybridization to dissected adult tissues. (A and B) A digoxigenin-labelled *Adhr*-specific probe corresponding to the full coding region (see Figure 2B) detected the transcript in the crop (**A**) and in the fatbody (**B**). (**C** and **D**) Same tissues as in (A) and (B) but hybridized with a digoxigenin-labelled *Adh*-specific probe (see Figure 2B). Note that the ring of cells in the crop expressing *Adh* also express *Adhr*, although this is not evident in this figure.

Tissue distribution of Adh and Adhr transcripts

Confirmation of the co-localization of *Adh* and *Adhr* products was obtained by tissue *in situ* hybridization. In adults, *Adh* is highly transcribed, almost exclusively from the distal promoter. This expression is regulated by the *Adh* adult enhancer (AAE) (Corbin and Maniatis, 1989). *Adh* and *Adhr* transcripts are expressed in the same tissues. This is most obvious in the anterior intestine and fat body, tissues in which ADH is most abundant (Figure 7). No expression of *Adhr* could be seen in any tissue not expressing *Adh.*

In the gut, both genes are most highly expressed in the crop. The expression of both genes is restricted to the saccular portion, where they are expressed in the epithelial cells. There is no expression in the surrounding muscle fibres.

The accumulation of the dicistronic messenger is controlled by a temperature-sensitive post-transcriptional mechanism

Although the same promoter apparently controls the transcription of both the canonical adult *Adh* and the dicistronic *Adh–Adhr* transcripts, the relative amounts of these are very different. The dicistronic transcript is always much less abundant than the monocistronic *Adh* transcript, by as much as 100-fold. A simple hypothesis to account for this difference is that the RNA polymerase would co-transcribe both genes as one single RNA precursor but that cleavage and polyadenylation at the site immediately 3′ to *Adh* is preferred over polyadenylation at the site 3′ to *Adhr*. If so, the relative levels of the monocistronic *Adh* transcript and dicistronic *Adh–Adhr* transcript might be altered in the presence of mutations that affect the 3′ processing of RNA polymerase II transcripts.

Only one such candidate mutation has, so far, been identified in *Drosophila melanogaster*—*su(f)* whose protein is a putative homologue of one of the subunits of the cleavage stimulation factor (CstF), a component required for the cleavage of mRNAs in mammalian cells (Mitchelson *et al.*, 1993; Takagaki and Manley, 1994; Martine *et al.*, 1996). *su(f)* is a vital gene; complete loss-of-function alleles are lethal. Weak mutant alleles are, however, viable and have the effect of suppressing or enhancing some mutations at other loci that result from the insertion of the retrotransposons (Parkhurst and Corces, 1985; Hoover *et al.*, 1993).

Both *Adh* and *Adhr* transcripts were assayed by Northern blots from wild-type and two different strains carrying the $su(f)^1$ allele. Since this allele is known to have a mutant phenotype when raised at high temperature (*Minute*-like bristles) (Schalet, 1972; Lindsley and Zimm, 1992), the flies for this experiment were raised at three different temperatures.

The results (Figure 8) show that $su(f)^1$ has no significant effect on the level of *Adh* transcript, relative to that of the *Actin-5C* transcript. By contrast, the relative level of the *Adhr* transcript is increased in the mutant background, and the extent of this increase is greater under conditions (high temperature) which result in a more severe *su(f)* mutant phenotype. In addition, high temperature alone, on a $su(f)^+$ background, also increases the relative amount of the dicistronic transcript. These data are consistent with the hypothesis that 3′ processing of RNAs is an intrinsically temperature-sensitive process in *Drosophila*.

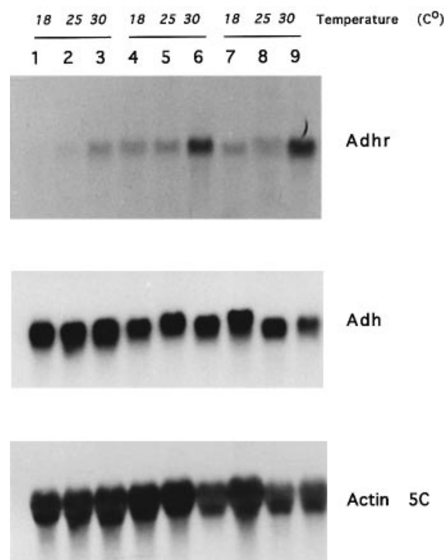

Fig. 8. *Adhr* transcription in *su(f)* mutants. Northern blot analysis of equal amounts of total RNA from wild-type flies (lanes 1, 2 and 3), from *f su(f)*^{1} (lanes 4, 5 and 6) and w^a *su(f)*^{1} (lanes 7, 8 and 9). The RNA is from flies raised in standard condition at 25°C, but then left for 6 days at 18 (lanes 1, 4 and 7), 25 (lanes 2, 5 and 8) and 30°C (lanes 3, 6 and 9). The same filter has been probed sequentially with *Adhr*-specific probe, *Adh*-specific probe and *Actin-5C* probe. The exposure times were 4 days at –70°C for *Adhr*, 2 h at 22°C for *Adh* and 7 h at –70°C for *Actin-5C* (a similar amount of radiolabelled probe was used in all hybridizations).

They also show that the alternative use of the two polyadenylation sites can be regulated by modulating the level of activity of the CstF factor.

Discussion

Polycistronic mRNA in eukaryotes

We conclude from these data that *Adhr* is transcribed as part of a dicistronic mRNA from either the distal or the proximal promoter of the *Adh* gene. The great majority of *Adh* transcripts terminate beyond the polyadenylation signal sequence in the *Adh—Adhr* intergenic sequence. A minority, perhaps 1%, do not and it is these that result in the only detectable *Adhr* transcript. The frequency with which read through occurs appears to depend upon some aspect of the 3′ mRNA processing machinery, since a mutation in one if its components, the *su(f)* gene product, results in an increase of the dicistronic transcript relative to the monocistronic.

Eukaryotic ribosomes are usually considered to be inefficient in their ability to reinitiate translation of 'internal′ reading frames (Kaufman *et al.*, 1987). In trypanosomes and *C.elegans*, where polycistronic transcripts are also known, the need for internal initiation is obviated by post-transcriptional processing and *trans*-splicing of a leader sequence (Blumenthal, 1995). ADHR is most probably translated by internal initiation of ribosomes on the dicistronic mRNA. Indeed, there are cases now known in eukaryotes, including *Drosophila*, where internal ribosome entry, without prior scanning of the 5′ sequence, occurs (Oh *et al.*, 1992; Chen and Sarnow, 1995). Moreover, artificial dicistronic transcripts including an internal ribosome entry sequence (IRES) upstream of an AUG codon have been constructed, and these show translation of the internal ORF (Hart and Bienz, 1996). The IRESes can be considered as functionally equivalent to the Shine and Dalgarno region of prokaryotic mRNAs, but features of their secondary (and presumably tertiary) structure are more important than the sequences themselves (Jackson and Kaminski, 1995). The sequence immediately upstream of the *Adhr* AUG is more conserved than expected, both within the *melanogaster* subgroup and between *D.melanogaster* and *D.pseudoobscura* (Jeffs *et al.*, 1994), but there is not any sequence conservation with more distant species.

The described mechanism of expression of *Adhr* is very unusual in *Drosophila* and in other animals. In a recent paper, it is reported that the *stoned* gene in *Drosophila* also appears to encode a dicistronic mRNA (Andrews *et al.*, 1996). Although the authors could not exclude the possible existence of other forms of RNA and the transcript was not functionally characterized, the finding strongly supports the existence of at least one other polycistronic mRNA in *Drosophila*.

Co-regulation of the Adh–Adhr locus

As our knowledge of the sequence organization of the genome of *D.melanogaster* increases, many pairs of related genes are being discovered. *Adh* and *Adhr* is an interesting example, because of the evidence that the duplication event is rather old. This is evident not only from the relatively low sequence similarity between the genes (49% nucleic acid for the exons; 37% identity of the amino acid sequence) but also from the fact that the duplication is seen in very distant species of *Drosophila*, in all species of the subgenus *Sophora* studied so far, in *D.funebris* (subgenus *Drosophila*) and in *S.lebanonensis* (genus *Scaptodrosophila*). The distance between these genes, measured from the *Adh* terminator codon to the *Adhr* initiator codon, varies between 298 (*D.melanogaster*) and 356 bp (*D.ambigua).* Both genes are transcribed as a common dicistronic mRNA in the sibling species of *D.melanogaster*, and at least for *D.pseudoobscura* preliminary data suggest the same.

The function of Adhr

The function of ADHR is not known. Jeffs *et al.* (1994) have argued that it is not an alcohol dehydrogenase. The reasons for this conclusion were the absence of ADH activity in many missense or nonsense mutant alleles of *Adh* and the fact that ADHR has aspartic acid as residue 14 rather than glycine. Just this substitution is seen in the ethyl methanesulfonate-induced allele Adh^{n11} and it both abolishes ADH activity and binding of the protein to 5′-AMP, a competitive inhibitor of NAD binding. Indeed, purified bacterially synthesized ADHR can bind neither NAD nor NADP (S.Brogna, unpublished data).

ADHR is not an abundant protein, as seen by the level of its transcript on both Northern blots, by *in situ* hybridization to tissues and by immunostaining with a polyclonal antiserum. The complete absence of ADHR in, for example, strains homozygous for a deletion that removes both *Adh* and *Adhr* has no obvious phenotypic consequence, except in embryos and larvae, where we find abnormalities in the development of the gastric caecae (S.Brogna, unpublished data).

The pattern of nucleotide changes of *Adhr* (Kreitman

and Hudson, 1991), the level of sequence conservation of its protein product and the data presented in this study now leave no doubt that ADHR is a functional protein. However, the biochemical function of ADHR remains unknown, and the functional significance of it being expressed via a dicistronic mRNA together with ADH remains to be discovered.

Materials and methods

Strains

The wild-type strain used was Canton-S. The *f su(f)*^{1} and w^a *su(f)*^{1} strains were obtained from the Umea *Drosophila* Stock Center. The mutant strain Adh^{RI42} was from the laboratory of C.Laurie (Duke University, North Carolina), and the Adh^{nBR114} strain was given by Nancy Fossett (Louisiana State University). The other strains used are described in Lindsley and Zimm (1992) or in FlyBase, and are available from Cambridge.

DNA

The plasmid pDmA2, containing the *Actin-5C* gene has been described by Fyrberg *et al.* (1983). The plamid pATSac is a sub-clone of the *D.melanogaster Adh* genomic clone (Jeffs *et al.*, 1994).

RNA extraction and Northern blots

Large and small scale preparations were done according to the protocol described in Ashburner (1989), but the flies were homogenized directly without previous grinding in liquid nitrogen.

The total RNA from polysomes was extracted according to a procedure described by Clemens (1984).

Poly(A)$^+$ selection was performed as described by Sambrook *et al.* (1989), RNA was fractionated in a formaldehyde–agarose gel, transferred by capillary action to a nylon membrane (Hybond-N) and hybridized as described by Yang *et al*. (1993).

RACE amplification

First experiment. For characterization of the 5′ end of the *Adhr* transcripts, single-stranded cDNA was synthesized with Superscript (Gibco-BRL) with primer F (5′-cccaagcttaatcc/atcttcatcattgctc-3′) from 10 µg of total adult RNA. The cDNA was cleaned of the primers and unincorporated nucleotides with a S-400 (Sephacryl, Pharmacia) spin column. The product was tailed with dCTP by using the terminal transferase enzyme (Boeringher). The product was cleaned from unincorporated nucleotides by a G-50 spin column, and one aliquot was used as PCR template. The two nested primers used were primer D (5′-taccccgttttgggaatagt-3′) and A (5′-gccacatagcagacatgctt-3′).

Second experiment. The cDNA was synthesized and tailed as before but from adult poly(A)$^+$ RNA. The first amplification was done with primer C (5′-ggaatagtaaagagggtccgcta-3′) and the second with primer B (5′-tctgtaaaatggccagtttcgcta-3′).

RT–PCR of circular cDNA

The 3′ end was amplified easily from adult circular single-stranded cDNA (J.W.Foster, personal comunication) synthesized from 2 µg of total RNA with phosphorylated oligo(dT) primer. After synthesis, the heteroduplex cDNA–RNA was cleaned as above, denatured in 0.2 M NaOH and neutralized with HCl. Then the cDNA was precipitated and resuspended in 20 µl of buffer containing 50 mM Tris-HCl (pH 8), 10 mM $MgCl_2$, 10 mM $MnCl_2$, 20 µM rATP, 10 mg/ml bovine serum albumin, 1 mM hexamine cobalt chloride and 5 U of T4 RNA ligase (Gibco-BRL). The ligation reaction was performed overnight at room temperature. Usually, 1 µl from the ligation mixture was used for PCR reaction with the two primers, 2079u (5′-cgctatattcttggtcatga-3′) and 2962d (5′-accctctttactattcccaa-3′); specific bands were visible after two rounds of 30 cycles. As with all the other PCR products, these were cloned in pBluescript II KS$^+$ in the *Eco*RV site previously tailed with ddTTP (Holton and Graham, 1990). This method should, in theory, be useful for the amplification of both 5′ and 3′ ends, but under the above conditions molecules longer than ~1000 bp are not amplified efficiently.

Polysome isolation

The method used to isolate adult polysomes is slightly modified from that described by Bradford and Sullivan (1981), and only the steps which we have modified are described. About 300 flies were ground in liquid nitrogen and resuspended in 7 ml of extraction buffer supplemented with Vanidil Ribonucleoside Complex (Gibco-BRL) to give a final concentration of 10 mM. The post-mitochondrial supernatant was centrifuged for 4 h over a 4 ml sucrose cushion in a Beckman SW28 rotor at 27 000 r.p.m. The pellet was resuspended in 0.5 ml of buffer C and supplemented with RNAsin (Promega) to a final concentration of 1 U/ml. About 2 OD units were layered onto linear 15–60% (w/v) sucrose gradients and centrifuged for 3 h at 27 000 r.p.m. in a SW28 rotor. The gradients were pumped through a LKB-uv-MII (Pharmacia) flow cell spectrophotometer and monitored at 254 nm.

Expressions in E.coli, antibody production and protein characterization

The full coding region was PCR amplified from total cDNA using two primers: ADHR.start (gggcatatgttcgatttgacgggca) and ADHR.stop (cccaagcttaatcxtcttcatcattgctc).

The PCR fragments were cloned in pUC18 in the *Nde*I and *Hin*dIII sites. Two of the clones were sequenced to check that no mutations had been introduced during the PCR reaction. The insert of one of the clones (pUCR8.1) was cloned in-frame in the pQE31 vector (Qiagen); in this condition, a six histidine leader is added to the amino-terminus of the protein. The plasmid (pQE8.1) was introduced in the *E.coli* strain M15 (pREP4) and the protein expression was induced by adding 1 mM IPTG. The recombinant protein, which was insoluble, was purified in the presence of 7 M urea by Ni-affinity chromatography according to the Qiagen protocol.

The antibody was produced in a rabbit by injecting 0.5–1 mg of the affinity-purified protein. The immunizations were performed as described by Harlow and Lane (1988). The antibody was column affinity purified against the ADHR protein using a procedure described by Carrol and Laughon (1987).

The entire ADH coding region was fused to thioredoxin by cloning in the pET-32 vector and expressed following the Novagen protocol.

The antibody against the ADHR protein was pre-absorbed to a cross-linked bacterial extract overexpressing the recombinant ADH fused to thioredoxin using a procedure described by Sherman and Goldberg (1992).

The bacterial protein extract was obtained by lysing the cells with lysozyme plus detergent (Sambrook *et al.*, 1989) and the soluble and insoluble fraction were analysed independently. The *Drosophila* extracts were prepared by homogenizing 10 adult flies in 0.2 ml of a buffer containing 0.125 M Tris–HCl pH 6.8, 0.1% Triton X-100, 4 M urea, 5% 2-mercaptoethanol, 1 mM phenylmethylsulfonyl fluoride (PMSF).

The protein extracts were analysed by SDS–PAGE. Western blot analysis was performed by transferring the proteins to nitrocellulose filters with a semi-dry apparatus (Bio-rad). Membranes were blocked using ProtoBlock (National Diagnostics) and probed with a 1:2000 dilution of the primary antibody; the secondary antibody was an anti-rabbit IgG conjugated with alkaline phosphatase (Sigma Immuno Chemicals) diluted 1:10 000.

Histochemistry, immunostaining and in situ hybridization

Tissues were dissected in Ringer's solution and immediately transferred for 30–45 min to an ice-cold solution of 4% formaldehyde in phosphate buffer. Then the material was transferred to a fresh fixation solution and left at room temperature for another 30–60 min. The subsequent treatments were as described by Patel (1994). Affinity-purified antibody, pre-absorbed against fixed tissues, was used in all staining, and a biotinylated anti-rabbit IgG secondary antibody was used according to the instructions provided in the Vectastain ABC Kit (Vector Laboratories). Dissected tissues were stained for ADH activity as described in Ashburner (1989). *In situ* hybridizations to isolated tissues were performed as described in Cubas *et al.* (1991).

Acknowledgements

We thank J.Roote, N.Fossett, C.C.Laurie and the Umea *Drosophila* Stock Centre for providing fly stocks, M.Jacobs for providing oligonucleotide primers and P.Hextall for help in setting up the flow-cell monitor for the polysomes experiments. We also thank R.J.Jackson, and M.Kreitman for discussion and reading of the manuscript and K.O'Hare for helpful discussions. We are grateful to S.Oehler for help, continuous discussions and reading of the manuscript. Additionally, we would like to thank S.C.R.Elgin, S.Russell and S.Collier for critically reading the manuscript. S.B. was supported by the EU Commission. This work was supported by an MRC Programme Grant to M.A. and D.Gubb.

References

Albalat,R. and Gonzalez-Duarte,R. (1993) *Adh* and *Adh-dup* sequences of *Drosophila lebanonensis* and *D.immigrans*: interspecies comparisons. *Gene*, **126**, 171–178.

Andrews,J., Smith,M., Merakovsky,J., Coulson,M., Hannan,F. and Kelly,L.E. (1996) The *stoned* locus of *Drosophila melanogaster* produces a dicistronic transcript and encodes two distinct polypeptides. *Genetics*, **143**, 1699–1711.

Ashburner,M. (1989) *Drosophila: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Ashburner,M., Aaron,C.S. and Tsubota,S. (1982) The genetics of a small autosomal region of *Drosophila melanogaster*, including the structural gene for alcohol dehydrogenase. V. Characterization of X-ray-induced *Adh* null mutations. *Genetics*, **102**, 421–435.

Benyajati,C. and Dray,J.F. (1984) Cloned *Drosophila* alcohol dehydrogenase genes are correctly expressed after transfection into *Drosophila* cells in culture. *Proc. Natl Acad. Sci. USA*, **81**, 1701–1705.

Benyajati,C., Place,A.R. and Sofer,W. (1983a) Formaldehyde mutagenesis in *Drosophila*. Molecular analysis of ADH negative mutants. *Mutat. Res*., **111**, 1–7.

Benyajati,C., Spoerel,N., Haymerle,H. and Ashburner,M. (1983b) The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5′ end in different developmental stages. *Cell*, **33**, 125–133.

Blumenthal,T. (1995) *Trans*-splicing and polycistronic transcription in *Caenorhabditis elegans*. *Trends Genet.*, **11**, 132–136.

Bonneville,J.M., Sanfacon,H., Futter,J. and Hohn,T. (1989) Post-transcriptional trans-activation in cauliflower mosaic virus. *Cell*, **59**, 1135–1143.

Bradford,C.T. and Sullivan,D.T. (1981) Isolation of polysomes from larval and adult *Drosophila melanogaster. Anal. Biochem*., **112**, 176–181.

Carrol,S.B. and Laughon,A. (1987) Production and purification of polyclonal antibodies to the foreign segment of β-galactosidase fusion proteins. In Glover,D.M. (ed.), *DNA Cloning: A Practical Approach*. IRL Press, Oxford, pp. 89–111.

Chen,C. and Sarnow,P. (1995) Initiation of protein synthesis by eukaryotic translational apparatus on circular RNAs. *Science*, **268**, 415–417.

Chia,W., Karp,R., McGill,S. and Ashburner,M. (1985) Molecular analysis of the *Adh* region of the genome of *Drosophila melanogaster. J. Mol. Biol.*, **186**, 689–706.

Clayton,C. (1992) Developmental regulation of nuclear gene-expression in *Trypanosoma brucei*. *Prog. Nucleic Acid Res. Mol. Biol.*, **43**, 37–66.

Clemens,M.J. (1984) Purification of eukaryotic messenger RNA. In Hames,B.D. and Higgins,S.J. (eds), *Transcription and Translation*. IRL Press, Oxford, pp. 211–230.

Corbin,V. and Maniatis,T. (1989) The role of specific enhancer–promoter interactions in the *Drosophila Adh* promoter switch. *Genes Dev.*, **3**, 2191–2120.

Cubas,P., de Celis,J.F., Campuzano,S. and Modolell,J. (1991) Proneural clusters of achaete-scute expression and the generation of sensory organs in the *Drosophila* imaginal wing disc. *Genes Dev.*, **5**, 996–1008.

Dunn,R.C. and Laurie,C.C. (1995) Effects of a transposable element insertion on alcohol dehydrogenase expression in *Drosophila melanogaster. Genetics*, **140**, 667–677.

Fosset,N.G., Arbour-Reily,P., Kilroy,G., McDaniel,M., Mahmoud,J., Tucker,A.B., Chang,S.H. and Lee,W.R. (1990) Analysis of ENU-induced mutations at the *Adh* locus in *Drosophila melanogaster. Mutat. Res*., **231**, 73–85.

Fyrberg,E.A., Mahaffey,J.W., Bond,B.J. and Davidson,N. (1983) Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. *Cell*, **33**, 115–123.

Harlow,E. and Lane,D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Hart,K. and Bienz,M. (1996) A test for cell autonomy, based on dicistronic messenger translation. *Development*, **122**, 747–751.

Holton,T.A. and Graham,M.W. (1990) A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Res.*, **19**, 1156.

Hoover,K.K., Chien,A.J. and Corces,V.G. (1993) Effects of transposable elements on the expression of the *forked* gene of *Drosophila melanogaster. Genetics*, **135**, 507–526.

Jackson,R.J. and Kaminski,A. (1995) Internal initiation of translation in eukaryotes: the picornavirus paradigm and beyond. *RNA*, **1**, 985–1000.

Jeffs,P.S., Holmes,E.C. and Ashburner,M. (1994) The molecular evolution of the alcohol dehydrogenase and alcohol dehydrogenase-related genes in the *Drosophila melanogaster* species subgroup. *Mol Biol Evol.*, **11**, 287–304.

Johnson,P.J., Kooter,J.M. and Borst,P. (1987) Inactivation of transcription by UV irradiation of *T.brucei* provides evidence for a multicistronic transcription unit including a VSG gene. *Cell*, **51**, 273–281.

Kaufman,R.J., Murtha,P. and Davies,M.V. (1987) Translation efficiency of polycistronic mRNAs and their utilization to express heterologous genes in mammalian cells. *EMBO J.*, **6**, 187–193.

Kozak,M. (1986) Bifunctional messenger RNAs in eukaryotes. *Cell*, **47**, 481–483.

Kreitman,M. and Hudson,R.R. (1991) Inferring the evolutionary histories of the *Adh* and *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. *Genetics*, **127**, 565–582.

Lee,S.J. (1991) Expression of growth/differentiation factor 1 in the nervous system: conservation of a bicistronic structure. *Proc. Natl Acad. Sci. USA*, **88**, 4250–4254.

Lindsley,D.L. and Zimm,G. (1992) *The Genome of Drosophila melanogaster.* Academic Press, San Diego, CA.

Marfany,G. and Gonzalez-Duarte,R. (1991) The *Adh* genomic region of *Drosophila ambigua*: evolutionary trends in different species. *J. Mol. Evol.*, **32**, 454–462.

Martine,S., Elliott,K., Mitchelson,A. and O'Hare,K. (1996) Interallelic complementation at the *suppressor of forked* locus of *Drosophila* reveals complementation between suppressor of forked proteins mutated in different regions. *Genetics*, **142**, 1225–1235.

Mitchelson,A., Simonelig,M., Williams,C. and O'Hare,K. (1993) Homology with *Saccharomyces cerevisiae* RNA14 suggests that phenotypic suppression in *Drosophila melanogaster* by *suppressor of forked* occurs at the level of RNA stability. *Genes Dev.*, **7**, 241–249.

Oh,S.K., Scott,M.P. and Sarnow,P. (1992) Homeotic gene Antennapedia mRNA contains 5′-noncoding sequences that confer translational initiation by internal ribosome binding. *Genes Dev.*, **6**, 1643–1653.

Parkhurst,S.M. and Corces,V.G. (1985) *forked*, gypsys and suppressors in *Drosophila*. *Cell*, **41**, 429–437.

Patel,N.H. (1994) Imaging neural subsets and others cell types in whole-mount *Drosophila* embryos and larvae using antibody probes. In Goldstein,L.S.B. and Fyrberg,E.A. (eds), *Drosophila melanogaster: Practical Uses in Cell and Molecular Biology.* Academic Press Inc., San Diego, CA, pp. 446–485.

Persson,B., Krook,M. and Jornvall,H. (1991) Characteristics of short-chain alcohol dehydrogenases and related enzymes. *Eur. J. Biochem.*, **200**, 537–543.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989). *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Savakis,C. and Ashburner,M. (1985) A simple gene with a complex pattern of transcription: the alcohol dehydrogenase gene of *Drosophila melanogaster. Cold Spring Harbor Symp. Quant. Biol.*, **50**, 505–514.

Schaeffer,S.W. and Aquadro,C.F. (1987) Nucleotide sequence of the *Adh* gene region of *Drosophila pseudoobscura*: evolutionary change and evidence for an ancient gene duplication. *Genetics*, **117**, 61–73.

Schalet,A. (1972) Report of A.Schalet. *Drosophila Inf. Service*, **49**, 36–37.

Sherman,M.Y. and Goldberg,A. (1992) Involvement of the chaperonin dnaK in the rapid degradation of a mutant protein in *Escherichia coli*. *EMBO J.*, **11**, 71–77

Spieth,J., Brooke,G., Kuersten,S., Lea,K. and Blumenthal,T. (1993) Operons in *C.elegans*: polycistronic mRNA precursors are processed by *trans*-splicing of SL2 to downstream coding regions. *Cell*, **73**, 521–532.

Takagaki,Y. and Manley,J.L. (1994) A polyadenylation factor subunit is the human homologue of the *Drosophila* suppressor of forked protein. *Nature*, **372**, 471–474.

Yang,H., McLeese,J., Weisbart,M., Dionne,J.L., Lemaire,I. and Aubin,R.A. (1993) Simplified high throughput protocol for Northern hybridization. *Nucleic Acid Res.*, **21**, 3337–3338.

Wolin,S.L. and Walter,P. (1988) Ribosome pausing and stacking during translation of a eukaryotic mRNA. *EMBO J.*, **7**, 3559–3569.

Zorio,D.A., Cheng,N.N., Blumenthal,T. and Spieth,J. (1994) Operons as a common form of chromosomal organization in *C.elegans*. *Nature*, **372**, 270–272.

Received on July 25, 1996; revised on December 13, 1996