Genet. Res., Camb. (1995), **66**, *pp.* 9–17 *With 5 text-figures*

Two independent cis-acting elements regulate the sex- and tissue-specific expression of *yp3* in *Drosophila melanogaster*

ELAINE RONALDSON AND MARY BOWNES*
Institute of Cell and Molecular Biology, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JR, Tel: 031 650 5368, fax: 031 668 3870, E-mail: MBOWNES@srv0.bio.ed.ac.uk

(*Received 9 December 1994 and in revised form 16 February 1995*)

Summary

In *Drosophila*, the three *yolk protein* (*yp*) genes are transcribed in a sex-, tissue- and developmentally specific manner, providing an ideal system in which to investigate the factors involved in their regulation. The yolk proteins are synthesized in the fat body of adult females, and in the ovarian follicle cells surrounding the developing oocyte during stages 8–10 of oogenesis. We report here an analysis of the *yolk protein 3* (*yp3*) gene and its flanking sequences by means of P-element mediated germ-line transformation and demonstrate that a 747 bp promoter region is sufficient to direct sex-specific expression in the female fat body and both the temporal- and cell-type-specificity of expression during oogenesis. Two elements that independently govern *yp3* transcription in these tissues have been separated and no other sequences in the upstream, downstream or coding regions have been identified that are autonomously involved in *yp3* expression.

1. Introduction

The yolk proteins of *Drosophila melanogaster* are encoded by three single copy genes on the X chromosome (Barnett *et al.* 1980). The divergently transcribed *yp1* and *yp2* genes are located at 8F-9A and separated by a shared intergenic region of 1225 bp (Hung & Wensink, 1983), whereas *yp3* is isolated with separate regulatory sequences as 12BC (Garabedian *et al.* 1987). Nevertheless, all three genes are coordinately regulated such that transcription occurs only in the fat body of adult females and in the ovarian follicle cells during stages 8–10 of oogenesis (Bownes & Hames, 1978; Brennan *et al.* 1982; Isaac & Bownes, 1982). The mechanisms involved in this regulation must therefore involve sex- and tissue-specific factors that are present at the correct developmental stages for gene activation.

Investigations into *yolk protein* gene control have mainly focused on *yp1* and *yp2* and several cis-acting elements have been identified that are necessary for their correct pattern of expression (Fig. 1). A 125 bp fat body element (FBE) located 200 bp upstream of the *yp1* transcription start site is sufficient to direct the sex-specific fat body expression of both genes (Garabedian *et al.* 1986). However, other sequences within the 1225 bp intergenic region are equally capable of conferring female-specific *yp* transcription in fat body when the FBE is deleted (Abrahamsen *et al.* 1993).

Ovarian expression is regulated by two tissue-specific enhancers, OE1 and OE2. OE1 is a 301 bp sequence located 43 bp upstream of *yp2* directing the correct developmental profile of transcription. Cell type specificity results from the combined positive and negative effects of OE1 regions on *yp* expression in different subpopulations of follicle cells at stages 9–11 of oogenesis. OE2 is located within the first 105 bp of

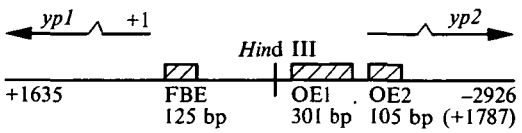


Fig. 1. The *yp1* and *yp2* regions that contain transcription enhancer sequences. The *yp1/yp2* genomic arrangement is shown with horizontal arrows representing the length and direction of transcripts. Introns are indicated by ∧. Boxes show the locations and sizes of the fat body enhancer (FBE), ovarian enhancer 1 (OE1) and ovarian enhancer 2 (OE2).

* Corresponding author.

the *yp2* coding region and acts across the *yp2* promoter to enhance levels of ovarian *yp1* transcripts. In the absence of OE1, OE2 is sufficient to limit *yp1* expression to the ovarian follicle cells but this is not restricted to the usual stages 8–10 of oogenesis. Both ovarian enhancers are therefore required to maintain correct levels of *yp1* and *yp2* transcription in the ovarian follicle cells during limited stages of oogenesis (Logan *et al.* 1989; Logan & Wensink, 1990).

From these data, it is apparent that an understanding of the control of *yp1* and *yp2* transcription is complicated by their shared regulatory elements, and to gain an insight into the mechanisms governing their transcription will also require the isolation and identification of trans-acting regulatory proteins and an analysis of their mode of action. Several transcriptional regulatory proteins have been identified that are implicated in the control of *yp* tissue-specific expression since they bind to the 125 bp FBE. They were initially found to interact with the *Adh* adult enhancer which controls expression of the *Drosophila melanogaster alcohol dehydrogenase* (*Adh*) gene in the adult fat body and were subsequently shown by footprinting assays to have binding sites within the *yp* FBE. They are box B-binding factor-2 (BBF-2) which acts as a transcriptional activator (Abel *et al.* 1992), the CCAAT/enhancer-binding protein (C/EBP) and the adult enhancer factor-1 (AEF-1). C/EBP and AEF-1 have overlapping recognition sites which are thought to be competitively bound since AEF-1 can inhibit binding by C/EBP, thereby repressing its activating potential (Falb & Maniatis, 1992). The only other proteins known to be involved in the regulation of *yp* expression are the male- and female-specific doublesex (DSX) proteins encoded by the differentially spliced *dsx* gene, the last gene in the *Drosophila* somatic sex determination pathway (Slee & Bownes, 1990; Steinmann-Zwicky *et al.* 1990). The bifunctional *dsx* locus must be expressed normally for *yp* expression in the female fat body to occur (Shirras & Bownes, 1987), which requires continuous activity of the *sex-lethal*, *transformer* and *transformer-2* genes in this pathway (Postlethwait *et al.* 1980; Bownes & Nöthiger, 1981; Belote *et al.*. 1985; Bownes *et al.* 1990). The *yps* are thereby the only set of target genes known to be directly under the control of the sex determination hierarchy. More recently, footprinting assays have shown that the FBE contains three DSX protein binding sites (Burtis *et al.* 1991), two of which appear to be required by the male- and female-specific DSX proteins to direct the sex-specific transcriptional regulation of *yp1* and *yp2* (Coschigano & Wensink, 1993). It is interesting to note that all the identified trans-acting factors bind to the FBE despite the fact that tissue-specific expression of *yp1* and *yp2* is maintained in female fat bodies when it is absent (Fig. 1). However, it is possible that these factors can exert their effects via other sequences in the 1225 bp intergenic region since computer searches with the target sequences have shown that this region contains alternative potential binding sites (Bownes *et al.* 1993).

Despite the accumulation of information on the sex- and fat body-specific expression of *yp1* and *yp2*, there have been no reports on the isolation of regulatory proteins that govern tissue-specific ovarian transcription. Several years ago, a heterodimeric DNA-binding protein, yolk protein factor 1 (YPF1), was identified in ovarian and early embryonic extracts that bound with high affinity to a 31 bp fragment located 148 bp downstream of the *yp1* transcription start site (Mitsis & Wensink, 1989*a*, *b*). This has recently been reported to be a *Drosophila* homologue of Ku, a DNA-dependent protein kinase from humans (Jacoby & Wensink, 1994). DNA-dependent protein kinases are involved in the initiation of gene transcription by phosphorylating the carboxyl-terminal domain of RNA polymerase II. This induces a conformational change in pol II such that it switches from being a promoter binding protein and becomes part of a transcription elongation complex. *YPF1* is ubiquitously expressed at low levels throughout *Drosophila* development, except in adult ovaries and early embryos, with the highest level of YPF1 transcripts found in the early stages of oocyte development. The profile of YPF1 transcripts is therefore consistent with the proposed function, that the association of YPF1 with *yp1* may position the kinase subunit adjacent to a transcription initiation complex which would facilitate high levels of *yp* transcription during oogenesis. However, YPF1 has not been implicated in the control of tissue-specific *yp* transcription in the ovary and the factors regulating this are yet to be isolated. This has possibly not been pursued with *yp1* and *yp2* because no single cis-acting element is responsible for their expression in this tissue. The cis- and trans-acting factors involved in the control of *yp* transcription are expected to be similar, if not the same, for *yp3* and we report here the identification of separate and independently acting cis-regulatory regions governing the ovarian and fat body expression of *yp3*. This will enable the factors regulating *yp3* expression to be isolated and identified, and comparisons with *yp1* and *yp2* to commence.

2. Materials and methods

(i) *DNA constructions*

The pERI transformation vector was constructed by cloning a 282 bp *Xho* I–*Pst* I fragment from the *hsp70* promoter region (Ingolia *et al.* 1980) firstly into the polylinker of pBluescript SK (Stratagene, Inc.) and then into pUC19 on a *Kpn* I–*Xba* I fragment. The *hsp70* sequences were subsequently incorporated into the pCaSpeR-AUG-βgal vector (Thummel *et al.* 1988) in the correct orientation, using *Eco*R I and *Bam*H I restriction sites. Constructs *b*, *c* and *d* containing *yp3*

https://doi.org/10.1017/S0016672300034340 Published online by Cambridge University Press

promoter sequences (illustrated in Fig. 2) were cloned into the pCaSpeR transformation vector. Construct *b* was isolated by digestion with *Acc* III and recessed 3′ ends were filled in using Klenow before *Hin*d III digestion. The fragment was then cloned into pBluescript SK cut with *Xho* I, Klenow treated to create blunt ends and *Hin*d III digested. It could then be directionally cloned into pCaSpeR on an *Eco*R I–*Kpn* I fragment. A similar approach was used to clone construct *c*. *Acc* III digestion followed by Klenow end-filling and *Sst* I restriction allowed this fragment to be cloned firstly into pUC19 digested with *Sst* I and *Sma* I. The construct was then transferred to pCaSpeR on an *Eco*R I–*Bam*H I fragment. Construct *d* was isolated by digestion with *Asp* 700 and *Acc* III. The *Acc* III site was blunt-ended and *Eco*R I linkers ligated to this fragment for cloning into the pCaSpeR vector. The orientation of this construct was verified by DNA sequencing. Constructs *e–n* were isolated using appropriate restriction sites and cloned into the *Eco*R I site of the pERI vector by the creation of blunt ends, followed by the addition of *Eco*R I linkers. All construct orientations were verified by dideoxy-nucleotide sequencing.

(ii) *Germ-line transformation*

Transformation procedures were carried out according to standard procedures (Spradling & Rubin, 1982). Embryos from a *white*$^-$ strain of *Drosophila* (Lüning, 1981) were injected with transformation constructs at 400 μg/ml and transposase was provided by the helper plasmid phs70Δ2-3wc, coinjected at a concentration of 100 μg/ml. The resulting transformants were inbred to obtain homozygous stocks.

(iii) *Histochemical assays*

Flies were dissected in Ringer's solution (55 mM NaCl, 40 mM KCl, 7 mM $MgSO_4$, 5 mM $CaCl_2$, 10 mM tricine, 20 mM glucose and 50 mM sucrose) and the unfixed tissues were allowed to stain for β-galactosidase activity, overnight at room temperature, in a solution containing 0·2% X-Gal, 7 mM $K_4[Fe(CN)_6]$, 7 mM $K_3[Fe(CN)_6]$, 1 mM $MgCl_2$, 150 mM NaCl and 10 mM Na_2HPO_4–NaH_2PO_4 buffer (pH 5·8–8·0). Flies transformed with the pERI vector were heat-shocked and stained for β-galactosidase activity to demonstrate that the *hsp70* promoter sequences were functional. Transformants were incubated at 37 °C for 45 min and allowed to recover for 1 h at room temperature prior to histochemical staining.

(iv) *Computer analysis*

Computer searches to identify sequence similarities between the ovarian and fat body regulatory regions of *yp1*, *yp2* and *yp3* were carried out using the Bestfit program search, University of Wisconsin Genetics Computer Group, Sequence Analysis Software Package (Devereux *et al.* 1984), Version 7.

3. Results

(i) *A 747 bp promoter region is sufficient to direct the correct sex-, tissue- and developmental specificity of* yp3 *transcription*

Preliminary investigations to locate regulatory sequences governing *yp3* expression determined that the YP3 coding region with 706 bp of sequence upstream from the transcription start site and 825 bp of 3′ flanking DNA was expressed in the correct tissues of females when transformed into a YP3$^-$ background. The expression was also sex-limited since no YP3 was detected in males of any of the transformed lines (Liddell & Bownes, 1991; Fig. 2*a*).

We decided to separate the upstream, downstream and coding regions in order to localize further any tissue-specific enhancer elements present. It was also of interest to determine whether *yp3* was regulated by a number of elements that could act independently, as appears to occur with the sex-specific fat body expression of *yp1* and *yp2* (Logan *et al.* 1989; Abrahamsen *et al.* 1993) or whether separate elements were interacting to coordinate expression, such as OE1 and OE2 in the regulation of *yp1* and *yp2* ovarian expression (Logan *et al.* 1989).

The *yp3* gene was separated into various 5′, coding and 3′ components using suitable restriction sites (Fig. 2). Fragments containing the *yp3* promoter (Fig. 2*b–d*) were fused upstream of the *E. coli lacZ* (β-galactosidase) reporter gene contained within the P-element-transformation vector, pCaSpeR-AUG-βgal (Thummel *et al.* 1988). A modified version of this vector (pERI) was also constructed. This contained sufficient *Drosophila hsp70* promoter sequences to allow fragments lacking the *yp3* promoter (Fig. 2*e–n*) to be tested for their effects on reporter gene activity (Ingolia *et al.* 1980; Dudler & Travers, 1984). All constructs were introduced into the *Drosophila* germline using standard transformation techniques (Rubin & Spradling, 1982; Spradling & Rubin, 1982) and Southern blot analysis was used to select independently transformed lines for further analysis that contained single copy inserts (data not shown). At least three independently transformed lines were assayed for each construct and reporter gene expression was detected by blue coloration derived from β-galactosidase activity, following histochemical staining of dissected ovaries and fat bodies. As a control, the modified vector was tested to ensure that there were no background levels of *lacZ* expression (Fig. 3*a*, *b*). No enzyme activity was detected unless transformants were subjected to heat-shock, in which case staining occurred in all tissues demonstrating that the hsp*70* promoter sequences used were functional and sufficient to activate gene expression (data not shown).

https://doi.org/10.1017/S0016672300034340 Published online by Cambridge University Press

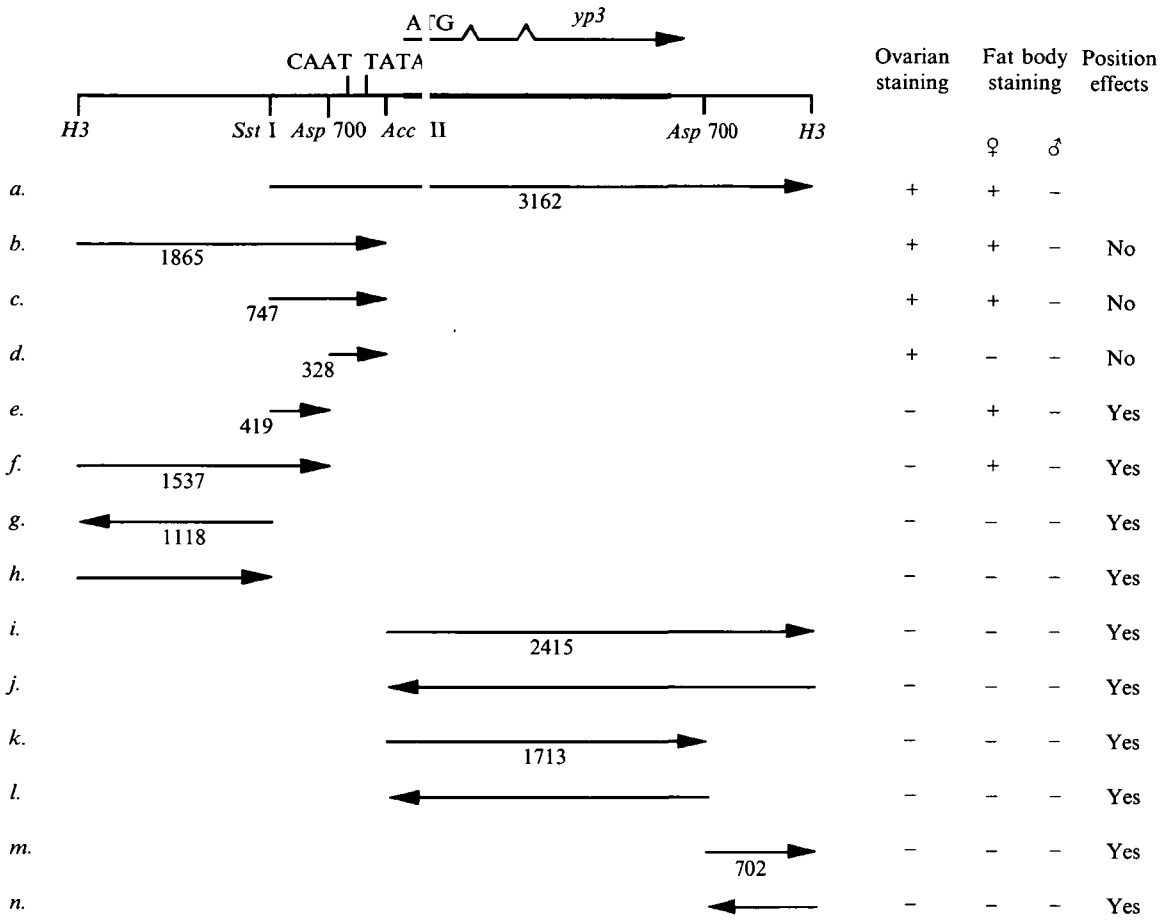


Fig. 2. A summary of the *yp3* constructs tested for transcription enhancer elements and the resulting β-glactosidase expression patterns. At the top is a schematic diagram of the *yp3* gene with the restriction enzyme sites used to create the different reporter gene constructs. The sizes (in bp) of the gene fragments are indicated below the constructs and the orientation, with respect to the reporter gene, represented by the direction of the arrowheads. The ability of all constructs to direct expression was assessed by β-galactosidase activity with the exception of construct *a*, where expression was confirmed by the presence of YP3 protein on SDS-PAGE gels and Western blots (Liddell & Bownes, 1991). Constructs *b–d*, that contained *yp3* promoter sequences, were cloned into the pCaSpeR-AUG-βgal transformation vector and constructs *e–n* were cloned into this vector modified to contain an *hsp70–lacZ* gene fusion. Absence of staining is indicated by – and β-glactosidase activity by +. The influence of chromosomal position effects on gene activity is also indicated.

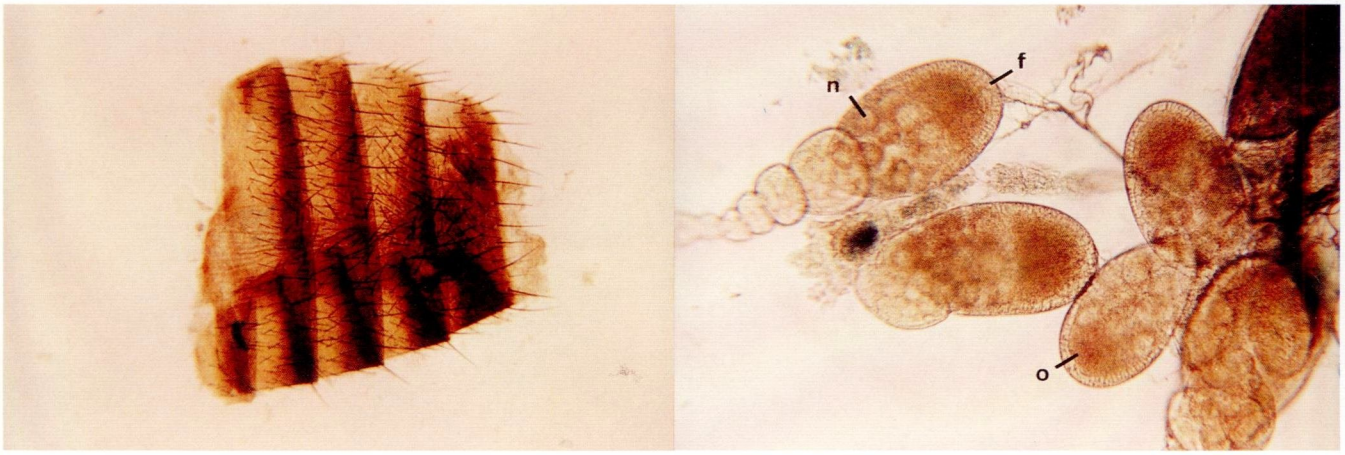
Two of the 12 constructs tested were able to direct reporter gene expression in both the female fat body and the correct stages of ovarian follicle cells (Fig. 3*c*, *d*). The constructs contained 1865 bp (−1822 to +43) and 747 bp (−704 to +43) of *yp3* sequence (Fig. 2*b*, *c*), both of which included the *yp3* promoter (all nucleotide positions are given relative to the *yp3* cap site). These constructs were sufficient to maintain a normal developmental profile of expression. In the fat bodies of staged transformants, the *lacZ* gene was expressed at a low level following eclosion, which increased gradually to reach a maximum after 2–3 d, reflecting the normal pattern of *yp* expression in this tissue (Barnett *et al.* 1980; data not shown). In the ovaries, β-galactosidase activity only occurred in the ovarian follicle cells during stages 8–10 of oogenesis.

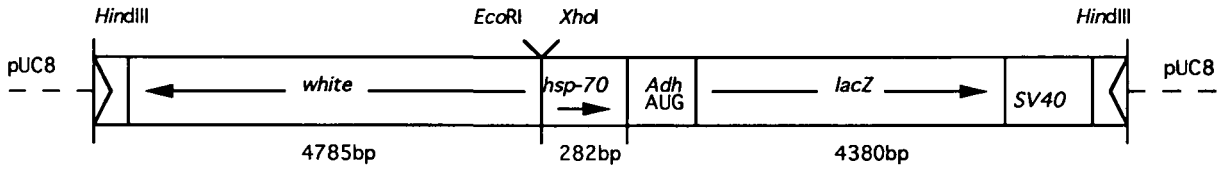
Fig. 3. β-galactosidase activity in female *Drosophila* transformed with *yp3–lacZ* or *yp3–hsp70–lacZ* fusion constructs, shown below the photographs. X-Gal stained abdomen (*a*) and dissected ovary (*b*) of flies transformed with the pERI vector, showing an absence of staining. The oocyte (o), nurse cells (n) and follicle cells (f) are labelled. An example is shown of the fat body (*c*) and ovarian staining (*d*) for which similar patterns are seen in flies transformed with the abdominal cuticle (c) are stained, as are the ovarian follicle cells during stages 8–10 of oogenesis. Numbers indicate the developmental stages of egg chambers (King, 1970). *e* and *f* represent the staining patterns observed in the fat body and ovaries of flies transformed with 328 bp of *yp3* upstream sequence (Fig. 2*d*). Ovarian expression is maintained in the follicle cells of stage 8–10 oocytes but there is no *lacZ* transcription in the fat body. Staining seen in the gut cells (g) is due to the presence of *lacZ* from *E. coli*. Flies transformed with the 419 bp *yp3* upstream fragment (Fig. 2*e*) have reporter gene expression in the fat body (*g*), but transcription in the ovaries is lost (*h*). A similar expression pattern is seen in flies transformed with the construct illustrated in Fig. 2*f*.

https://doi.org/10.1017/S0016672300034340 Published online by Cambridge University Press

A **FEMALE ABDOMEN**

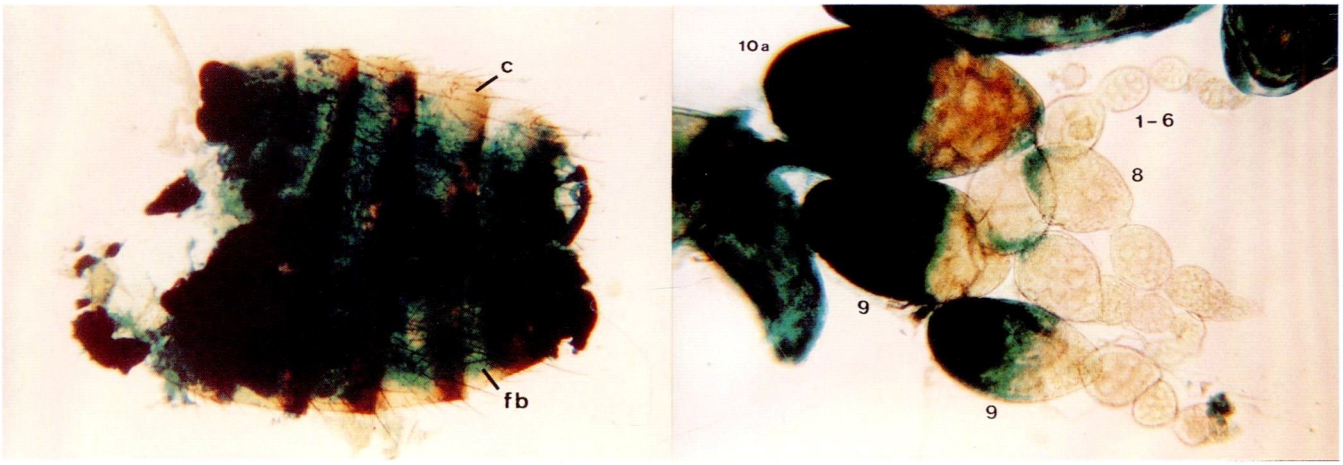
B **DISSECTED OVARY**

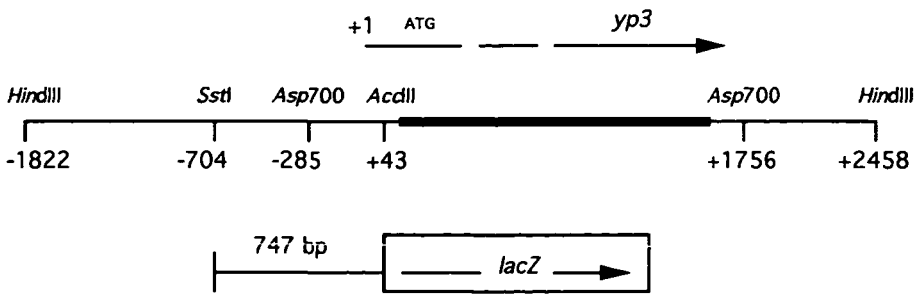




C **FEMALE ABDOMEN**

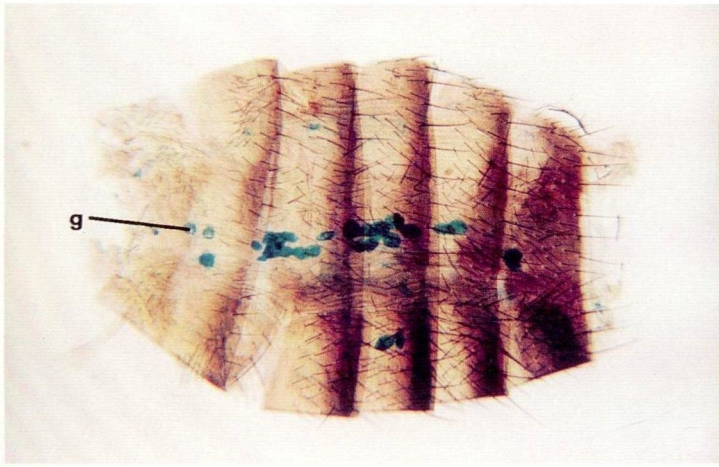
D **DISSECTED OVARY**



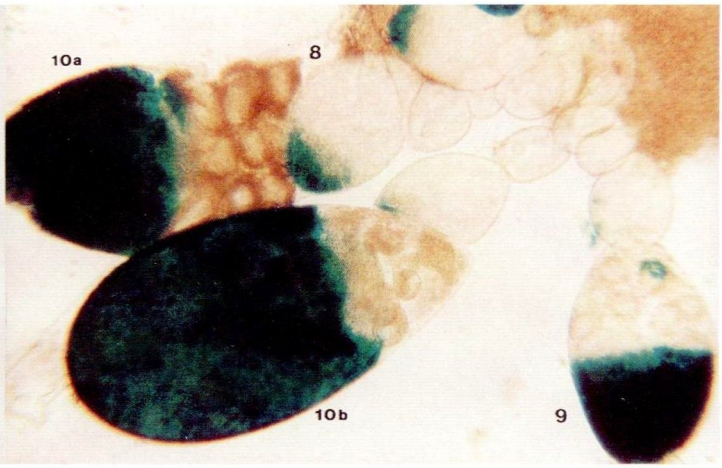


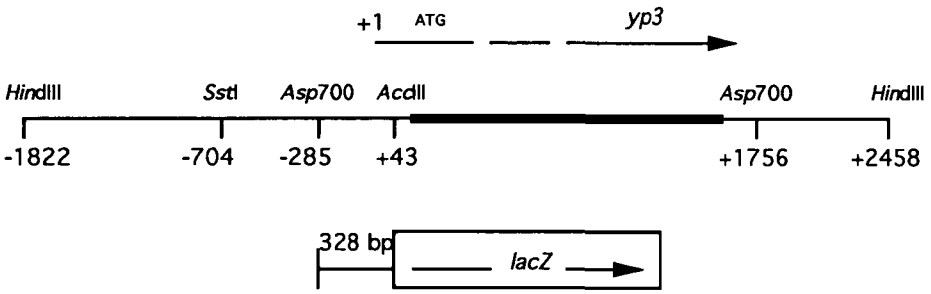
https://doi.org/10.1017/S0016672300034340 Published online by Cambridge University Press

E **FEMALE ABDOMEN**

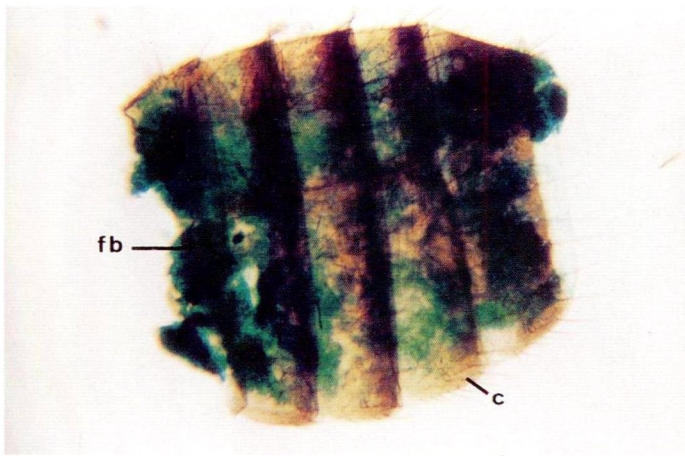


F **DISSECTED OVARY**

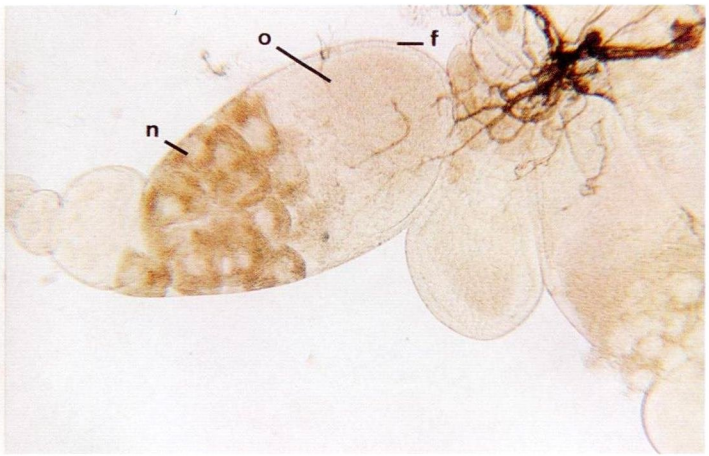


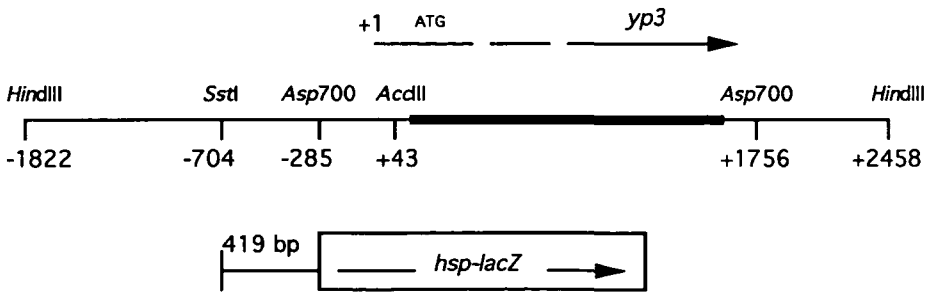


G **FEMALE ABDOMEN**



H **DISSECTED OVARY**





https://doi.org/10.1017/S0016672300034340 Published online by Cambridge University Press

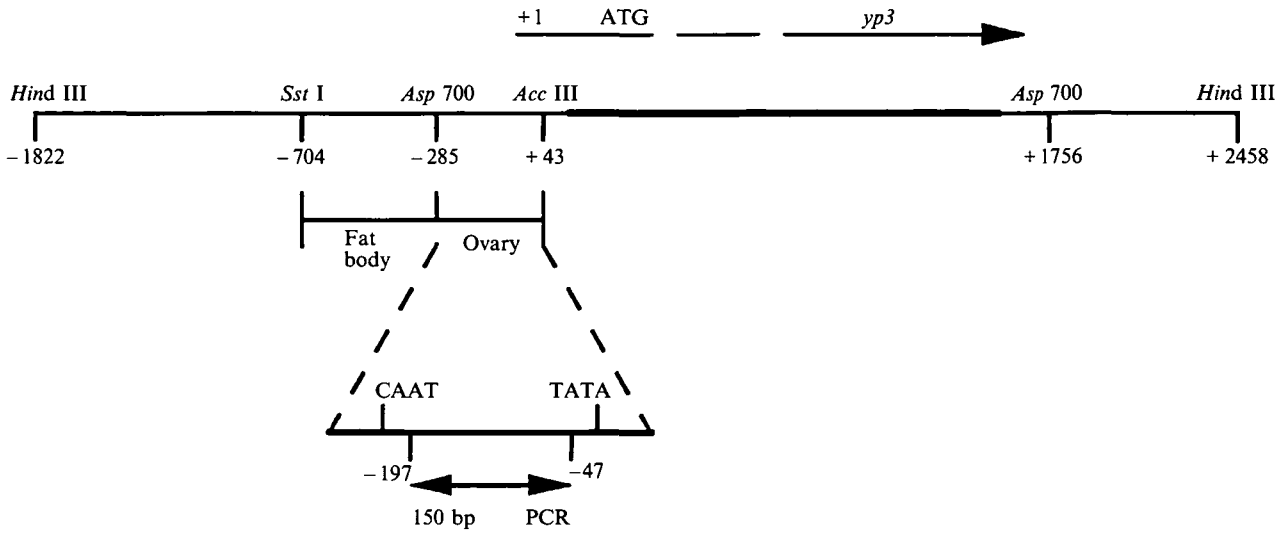


Fig. 4. *Drosophila yp3* gene illustrating the identified ovary and fat body cis-regulatory regions and the 150 bp PCR promoter fragment from OE3 used for the computer sequence comparisons and in gel retardation assays. Nucleotide positions of restrictions enzyme sites and PCR boundaries are given relative to the *yp3* transcription start site.

The constructs were also sufficient to confer the sex-specificity of expression as no β-galactosidase activity was observed in any tissues of the males of transformed lines. The 5′ region absent from the shorter fragment (−1822 to −704) was independently tested for the presence of other enhancer elements, but no β-galactosidase activity was apparent in the fat body or ovaries of lines transformed with this construct (Fig. 2*g*, *h*). Unfortunately, only one transformed line was obtained with this construct in the forward orientation. No *lacZ* expression was detected in any tissues of males or females, but more lines would need to be assessed to eliminate the possibility that this resulted from a chromosomal position effect. Nevertheless, no other constructs lacked expression in one particular line if they contained sequences responsible for directing tissue-specific gene transcription. Furthermore, although this result is mostly based upon this fragment when cloned in a reverse orientation upstream of the reporter gene, any enhancers present would, by definition, be expected to remain functional regardless of orientation (Serfling *et al.* 1985).

Reporter constructs containing the *yp3* coding and downstream non-coding regions were also investigated for their ability to influence gene expression. These regions were tested together, separately and in both orientations (Fig. 2*i–n*) but no staining was observed that resembled the *yp* expression pattern, suggesting that they were not involved in directing tissue-specific *yp3* transcription. It was therefore concluded that the 747 bp region immediately upstream of the *yp3* coding sequence was the only region to contain cis-acting elements autonomously involved in controlling the highly regulated pattern of *yp3* expression. However, it is possible that levels of transcription would be altered if the coding and 3′ sequences were tested in conjunction with identified enhancer elements, or if the 3′ region was tested in a location downstream of the *lacZ* reporter gene. A variety of genes expressed in a cell-type or tissue-specific manner, and developmentally regulated, are known to be influenced by downstream enhancers.

It was noticed that some reporter constructs were subject to chromosomal position effects (Fig. 2*e–n*). Since β-glactosidase activity was observed in different tissues of independently transformed lines, a greater number was analysed to determine whether the staining patterns were related to *yp* expression. Interestingly, these effects were less frequent with fragments that directed transcription from the native *yp3* promoter (Fig. 2*b–d*). This indicates that the *hsp70* promoter may be more susceptible to the influences of nearby enhancers than the *yp3* promoter, which could therefore contain negative regulatory elements or silencers involved in determining the overall pattern of expression. Alternatively, differences may arise simply because of the unusual organization of sites in artificially created fusions.

(ii) *The 747 bp sequence can be separated into two fragments of 328 and 419 bp that govern ovarian and fat body expression respectively and independently*

Since both the fat body and ovarian tissue-specific enhancers were contained within 747 bp of 5′ sequence, this region was analysed further in order to separate the elements controlling *yp3* expression in these tissues.

An *Asp*700 restriction site allowed the isolation of a 328 bp promoter fragment (−285 to +43; Fig. 2*d*) which was sufficient to direct reporter gene expression in the ovarian follicle cells of stage 8–10 oocytes (Fig. 3*f*). No staining was detected in the fat body or any other tissues of lines carrying this construct (Fig. 3*e*). Germ-line transformation of the remaining 419 bp sequence (−704 to −285; Fig. 2*e*) showed that it

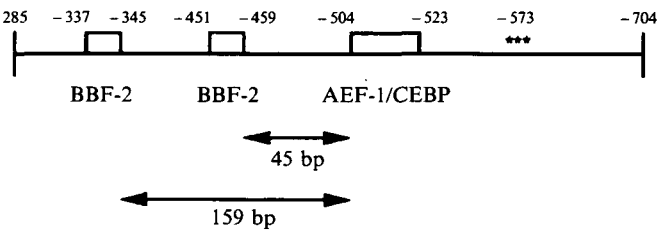
https://doi.org/10.1017/S0016672300034340 Published online by Cambridge University Press

(*a*)

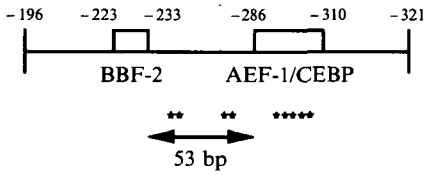
```
                                                  111111111
-336  TTATGCATTTATTTGATCAAATCCGGTGCACAACTACAATGTTGCAATCA          yp1/2 FBE
          |       | | |||||      |||||||  |    |||| ||||||
-546  GCCAGTTCAATTCTAATCAAGGGATCTGCACAAGT....TGTTTCAATCA          yp3 FBE

                                  AEF-1       C/EBP
```

```
      222222222  222222222                          333333333
GCGGAGCCTACAAAGTGATTACAAATTAAAATAATCAGGCG.GCAGCAGGTGCTGCTAAGTCATCA  -221
      |||  ||    | ||  |    | |  |   |     |||  |  |   |     |   ||
ATCCGTACTAGAATACATTTTAAGTGCAGAGAACAAAAATTTGCATTACTTTGGGAATTATATGCA  -434
```

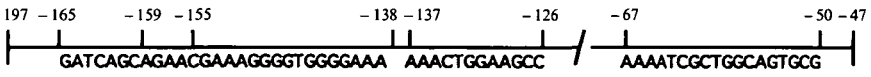
(*b*) (i) *yp3*FBE



(ii) *yp1/2*FBE



(*c*)



aaaatcActgAcGCtgcg	OE1 (–145 to –128)	(14/18 nt (78%)
gatcagcaCaa	OE1 (–106 to –116)	10/11 nt (91%)
caTaacgcaaagggTGCgCAGa	OE2 (+80 to +59)	16/22 nt (73%)
aaCTtggaagcc	OE2 (+36 to +47)	10/12 nt (83%)

Fig. 5. (*a*) Region of *yp3* FBE with sequence similarity to the AEF-1 and C/EBP binding sites in the *yp1/yp2* FBE. The vertical lines indicate nucleotide identities. The regions marked 1111, 2222, 3333, etc. are the sites which bind dsx (Burtis *et al.* 1991). (*b*) (i) The putative AEF-1, C/EBP and BBF-2 binding sites in the *yp3* FBE as identified by computer searches and the corresponding sites, identified by footprinting, in the *yp1/yp2* FBE (ii). The organization of the binding sites is also similar. One of the BBF-2 recognition sequences is separated from the overlapping AEF-1 and C/EBP binding sites by 45 bp in the *yp3* FBE compared with 53 bp in the *yp1/yp2* FBE. Each * in (b) (i) and (b) (ii) represents a dsx binding region. (*c*) Sequence similarities between OE1, OE2 and the OE3 PCR fragment. Sequences from the PCR fragment are given at the top and lower case characters indicate nucleotide identities to sequences from OE1 or OE2. The number of matched nucleotides and the percentage of identity between sequences are shown on the right-hand side of the figure. Nucleotide positions of OE1 and OE2 sequences are given relative to the transcription start site of *yp2* and the PCR nucleotide positions relative to the *yp3* transcription start site.

https://doi.org/10.1017/S0016672300034340 Published online by Cambridge University Press

contained sequences able to direct sex-limited gene expression in female fat bodies (Fig. 3*g*). No staining was detected in the ovaries with this construct (Fig. 3*h*). A similar result was seen in lines transformed with the 1537 bp construct (−1822 to −285; Fig. 2*f*) i.e. the entire upstream region, only lacking the 328 bp promoter fragment.

These observations therefore demonstrated that ovarian expression was lost when the *yp3* promoter region was deleted and it was thus concluded that only the 328 bp fragment contained elements necessary for *yp* transcription in the ovarian follicle cells. We have called this fragment OE3, for ovarian enhancer 3; however it should be noted that formally it has not been proven to be an enhancer since it has not been tested in the reverse orientation.

(iii) *Regions of homology exist between OE1, OE2 and an OE3 PCR fragment*

Computer searches were carried out to identify sequence similarities between OE1, OE2 and an OE3 fragment and revealed that four putative sites of conserved sequence existed (Fig. 5*c*). The sequence from OE3 used in these searches did not include the TATA and CAAT-box promoter elements. This sequence was also isolated as a PCR fragment (Fig. 4) and preliminary data demonstrated specific binding by components of ovarian nuclear extracts in gel-retardation assays (Ronaldson & Bownes unpublished). This fragment was used for the computer comparisons between OE1, 2 and 3. OE1 sequences of 11 and 18 bp had 91 and 78% identity to the PCR fragment respectively and OE2 sequences of 12 and 22 bp were matched to the PCR sequence with 83 and 73% respective sequence similarity (Fig. 5*c*). From the preliminary gel retardation assays, it is likely that protein(s) specifically associating with the PCR fragment are involved in regulating ovarian *yp* transcription, in which case they would also be expected to bind the ovarian enhancers governing *yp1* and *yp2* expression, for which the sequence identities would be candidate binding sites. The chorion proteins are also expressed in the follicle cells around the oocyte. There are no putative binding sites for the chorion factor 1 (CFF1/USP), which has been shown to bind upstream of the chorion gene, in OE3 (Christianson *et al.* 1992).

(iv) *Sequence comparisons between the identified fat body enhancer elements of* yp1, yp2 *and* yp3

Since the *yps* are coordinately transcribed as well as precisely regulated, it is likely that the mechanisms and factors involved will be the same for all three genes. We therefore carried out sequence comparisons between the identified fat body enhancer regions of *yp1*, *yp2* and *yp3* in order to locate conserved regulatory sequences. The *yp1/yp2* 125 bp FBE was compared to the 419 bp *yp3* fat body enhancer and the best sequence similarity occurred with the adjacent and overlapping AEF-1 and C/EBP binding sites shown by DNAseI footprinting assays to be present in the 125 bp FBE (Falb & Maniatis, 1992). There is one gap in the *yp3* sequence but a high level of 90% similarity exists between the matched sequences (Fig. 5*a*). Since BBF-2 was also reported to bind the 125 bp FBE (Abel *et al.* 1992) we searched for the recognition sequence in the *yp3* FBE and found that two putative sites existed, both with 78% sequence similarity and one of which was a similar distance from the AEF-1 and C/EBP binding sites relative to the arrangement seen in the *yp1/yp2* FBE (Fig. 5*b*). Binding studies are underway with the *yp3* FBE and the AEF-1, C/EBP and BBF-2 proteins to verify the significance of these findings.

4. Discussion

We have identified two cis-acting DNA regions necessary for the ovarian and sex-specific fat body expression of the *Drosophila yp3* gene, both of which are located 5′ of the transcription start site. A 328 bp ovarian region (OE3) is sufficient to limit reporter gene expression to the correct follicle cell types at the expected stages 8–10 of oogenesis, reproducing the ovarian *yolk protein* expression pattern (Fig. 3*f*). A neighbouring 419 bp sequence is able to direct *lacZ* expression, from a heterologous promoter, in the fat body cells of adult females (Fig. 3*g*), but not males. Our experiments also demonstrate that these appear to be the only regions essential for governing the observed *yp* expression pattern since the *yp3* coding, 3′ and other 5′ sequences do not contain any elements capable of independently directing the temporal-, tissue- or sex-specificity of expression.

The experimental data demonstrating that the 419 bp fragment is capable of directing sex-specific fat body expression and the presence of potential binding sites for AEF-1, C/EBP and BBF-2 suggest that these factors are involved in regulating female fat body expression of all three *yp* genes and that the interactions which mediate this could thus be of a similar nature. It is thought that AEF-1 and C/EBP compete for their overlapping binding sites since AEF-1 can displace pre-bound C/EBP from its recognition sequence (Falb & Maniatis, 1992). This indicates that small changes in the cellular concentration of either protein could affect a sensitive on/off mechanism to determine whether the associated gene should be active or silent. Such a mechanism could also be used to vary levels of *yp* expression and may be responsible for the observed relationship between nutritional intake and levels of *yp* transcription which subsequently affects the rate of egg production in females (Bownes *et al.* 1988). However, as previously mentioned, it is known that the 125 bp FBE can be deleted from the *yp1/yp2* intergenic

region without the loss of sex-specific fat body expression, and that although computer searches identify other putative AEF-1, C/EBP and BBF-2 binding sites in the intergenic spacer, they are not in the same arrangement as found in the FBE (Bownes *et al.* 1993). This suggests that overlapping recognition sites may not be essential to switch on the *yp* genes in the female fat body, and that the factors involved may also interact over longer DNA sequences. It has also been observed that when a construct confers tissue-specific gene expression in the fat body it is always in the correct sex, i.e. the sex- and tissue-specificity of expression are never separated. A possible explanation for this is seen in the *yp1/yp2* FBE where the binding sites for the regulatory proteins involved in fat body expression overlap with DSX binding sites (Falb & Maniatis, 1992), suggesting that the C/EBP, DSX and possibly BBF-2 proteins could be involved in the formation of a complex necessary to mediate transcription in the female fat body. Again, it has been shown that other putative DSX binding sites exist outside the 125 bp FBE such that the hypothesized interactions could occur with other *yp* fragments capable of directing this pattern of expression. It is also possible that such interactions could regulate *yp3* fat body transcription as several potential DSX binding sites exist in the identified *yp3* FBE (Abrahamsen *et al.* 1993; Bownes *et al.* 1993).

Sequence similarities were also observed between OE1, OE2 and OE3. However, as yet no binding proteins have been identified for any of these sites. Two experimental approaches could be used to determine the significance of these sequence similarities and whether they influence ovarian *yp* expression *in vivo*. Isolation of the DNA-binding protein(s) followed by DNAseI footprinting would identify the sites of DNA-protein interactions, or site-directed mutagenesis followed by germ-line transformation would determine whether these sequences were necessary for enhancer activity. The results obtained from computer searches alone should not be interpreted as conclusive although they can be used to indicate sequences of functional importance when supported by experimental evidence. The fact that OE1 and OE2 are both required for the correct cell type- and developmental specificity of ovarian *yp1/yp2* expression and that the PCR fragment, isolated from the *yp3* ovarian enhancer, contains regions of sequence similarity to both OE1 and OE2 does suggest that these regions could have a role in regulating ovarian *yp3* transcription.

The YPF1, KU homologue, which binds 148 bp downstream of the transcription start site, may be involved in generating high levels of transcription. Interestingly, computer sequence comparisons between *yp3* and the YPF1 binding site identified a 26 bp sequence with 77% sequence identity at a position 140 bp downstream of the *yp3* transcription start site. However, YPF1 has not been implicated in tissue-specific expression, and sequence similarities in the coding region could be important for protein sequence conservation as well as binding transcription factors. Analysis of *yp3* may therefore prove to be valuable in advancing current knowledge of the transcriptional control of all three *yp* genes. An understanding of the regulation of the divergently transcribed *yp1* and *yp2* genes is complicated by their shared regulatory regions but we have shown that the enhancers responsible for directing the fat body and ovarian specificity of *yp3* expression can be separated and act independently. future investigations will be aimed at isolating and characterizing the factors that interact with these sequences, especially the ovarian enhancer, since to date no trans-acting proteins have been isolated that govern expression in the ovarian follicle cells.

We would like to thank Dr R. Slee and Dr C. Simpson for constructive comments on the manuscript, and D. Mauchline for technical assistance. We are also grateful to Dr Simpson for assistance with the figures. This research was supported by an MRC studentship and SERC project grant.

References

Abel, T., Bhatt, R. & Maniatis, T. (1992). A *Drosophila* CREB/ATF transcriptional activator binds to both fat body- and liver-specific regulatory elements. *Genes and Development* **6**, 466–480.

Abrahamsen, N., Martinez, A., Kjaer, T., Søndergaard, L. & Bownes, M. (1993). Cis-regulatory sequences leading to female-specific expression of yolk protein genes 1 and 2 in the fat body of *Drosophila melanogaster*. *Molecular and General Genetics* **237**, 41–48.

Baniahmad, A., Muller, M., Steiner, C. & Renkawitz, R. (1987). Activity of two different silencer elements of the chicken lysozyme gene can be compensated by enhancer elements. *EMBO Journal* **6**, 2297–2303.

Barnett, T., Pachl, C., Gergen, J. P. & Wensink, P. C. (1980). The isolation and characterization of *Drosophila* yolk protein genes. *Cell* **21**, 729–738.

Belote, J. M., Handler, A. M., Wolfner, M. F., Livak, K. J. & Baker, B. S. (1985). Sex-specific regulation of yolk protein gene expression in Drosophila. *Cell* **40**, 339–348.

Bownes, M. & Hames, B. D. (1978). Analysis of the yolk proteins in *Drosophila melanogaster*. *FEBS Letters* **96**, 327–330.

Bownes, M. & Nöthiger, R. (1981). Sex determining genes and vitellogenin synthesis in *Drosophila melanogaster*. *Molecular and General Genetics* **182**, 222–228.

Bownes, M., Scott, A. & Shirras, A. (1988). Dietary components modulate yolk protein gene transcription in *Drosophila melanogaster*. *Development* **103**, 119–128.

Bownes, M., Steinmann-Zwicky, M. & Nöthiger, R. (1990). Differential control of yolk protein gene expression fat bodies and gonads by the sex-determining gene *tra-2* of *Drosophila*. *EMBO Journal* **9**, 3975–3980.

Bownes, M., Ronaldson, E., Mauchline, D. & Martinez, A. (1993). Regulation of vitellogenesis in *Drosophila*. *International Journal of Insect Morphology and Embryology* **22**, 349–367.

Brennan, M. D., Weiner, A. J., Goralski, T. J. & Mahowald, A. P. (1982). The follicle cells are a major site of vitellogenin synthesis in *Drosophila melanogaster*. *Developmental Biology* **89**, 225–236.

Burtis, K. C., Coschigano, K. T., Baker, B. S. & Wensink, P. C. (1991). The doublesex proteins of *Drosophila melanogaster* bind directly to a sex-specific yolk protein gene enhancer. *EMBO Journal* **10**, 2577–2582.

Christianson, A. M. K., King, D. L., Hatzivassiliou, E., Casas, J. E., Hallenbeck, P. L., Nikodem, V. M., Mitsialis, A. & Kafatos, F. C. (1992). DNA binding and heteromerization of the *Drosophila* transcription factor chorion factor1/ultraspiracle. *Proceedings of the National Academy of Sciences, USA* **89**, 11 503–11 507.

Coschigano, K. T. & Wensink, P. C. (1993). Sex-specific transcriptional regulation by the male and female doublesex proteins of *Drosophila*. *Genes and Development* **7**, 42–54.

Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Research* **12**, 387–395.

Dudler, T. & Travers, A. A. (1984). Upstream elements necessary for optimal function of the hsp70 promoter in transformed flies. *Cell* **38**, 391–398.

Falb, D. & Maniatis, T. (1992). A conserved regulatory unit implicated in tissue-specific gene expression in *Drosophila* and man. *Genes and Development* **6**, 454–465.

Garabedian, M. J., Shepherd, B. M. & Wensink, P. C. (1986). A tissue-specific transcription enhancer from the Drosophila yolk protein 1 gene. *Cell* **45**, 859–867.

Garabedian, M. J., Shirras, A. D., Bownes, M. & Wensink, P. C. (1987). The nucleotide sequence of the gene coding for *Drosophila melanogaster* yolk protein 3. *Gene* **55**, 1–8.

Hung, M.-C. & Wensink, P. C. (1983). Sequence and structure conservation in yolk proteins and their genes. *Journal of Molecular Biology* **164**, 481–492.

Ingolia, T. D., Craig, E. A. & McCarthy, B. J. (1980). Sequence of three copies of the gene for the major Drosophila heat shock induced protein and their flanking regions. *Cell* **21**, 669–679.

Isaac, P. G. & Bownes, M. (1982). Ovarian and fat-body vitellogenin synthesis in *Drosophila melanogaster*. *European Journal of Biochemistry* **123**, 527–534.

Jacoby, D. B. & Wensink, P. C. (1994). Yolk protein factor 1 is a *Drosophila* homologue of Ku, the DNA-binding subunit of a DNA-dependent protein kinase from humans. *Journal of Biological Chemistry* **269**, 11 484–11 491.

King, R. C. (1970). *Ovarian Development in Drosophila melanogaster*. New York: Academic Press.

Liddell, S. & Bownes, M. (1991). Investigation of *cis*-acting sequences regulating expression of the gene encoding yolk protein 3 in *Drosophila melanogaster*. *Molecular and General Genetics* **230**, 219–224.

Logan, S. K., Garabedian, M. J. & Wensink, P. C. (1989). DNA regions that regulate the ovarian transcriptional specificity of *Drosophila* yolk protein genes. *Genes and Development* **3**, 1453–1461.

Logan, S. K. & Wensink, P. C. (1990). Ovarian follicle cell enhancers from the *Drosophila* yolk protein genes: different segments of one enhancer have different cell-type specificities that interact to give normal expression. *Genes and Development* **4**, 613–623.

Lüning, K. G. (1981). Genetics of inbred *Drosophila melanogaster*. *Hereditas* **95**, 181–188.

Meyer, K. B. & Neuberger, M. S. (1989). The immunoglobulin kappa-locus contains a second, stronger, B-cell-specific enhancer which is located downstream of the constant region. *EMBO Journal* **8**, 1959–1964.

Mitsis, P. G. & Wensink, P. C. (1989*a*). Identification of yolk protein factor 1, a sequence of specific DNA binding protein from *Drosophila melanogaster*. *Journal of Biological Chemistry* **264**, 5188–5194.

Mitsis, P. G. & Wensink, P. C. (1989*b*). Purification and properties of YPF1, a sequence specific DNA-binding protein from *Drosophila melanogaster*. *Journal of Biological Chemistry* **264**, 5195–5202.

Postlethwait, J. H., Bownes, M. & Jowett, T. (1980). Sexual phenotype and vitellogenin synthesis in *Drosophila melanogaster*. *Developmental Biology* **79**, 379–387.

Rubin, G. M. & Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science* **218**, 348–353.

Serfling, E., Jasin, M. & Schaffner, W. (1985). Enhancers and eukaryotic gene transcription. *Trends in Genetics* **1**, 224–230.

Shirras, A. & Bownes, M. (1987). Separate DNA sequences are required for normal female and ecdysone-induced male expression of *Drosophila melanogaster* yolk protein 1. *Molecular and General Genetics* **210**, 153–155.

Slee, R. & Bownes, M. (1990). Sex determination in *Drosophila melanogaster*. *Quarterly Review of Biology* **65**, 175–204.

Spradling, A. C. & Rubin, G. M. (1982d). Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* **218**, 341–347.

Steinmann-Zwicky, M., Amrein, H. & Nöthiger, R. (1990). Genetic control of sex determination in *Drosophila*. *Advances in Genetics* **27**, 189–237.

Thummel, C. S., Boulet, A. M. & Lipshitz, H. D. (1988). Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. *Gene* **74**, 445–456.