The *fat facets* gene is required for *Drosophila* eye and embryo development

JANICE A. FISCHER-VIZE[1,2,3,*], GERALD M. RUBIN[1] and RUTH LEHMANN[3]

[1]*Howard Hughes Medical Institute, Department of Molecular and Cellular Biology, University of California at Berkeley, Berkeley, California 94720, USA*
[2]*Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA*
[3]*Whitehead Institute for Biomedical Research, Howard Hughes Medical Institute and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA*

*Author for correspondence at present address:Department of Zoology, University of Texas at Austin, Austin, Texas 78712, USA

Summary

In a screen for mutations affecting *Drosophila* eye development, we have identified a gene called *fat facets (faf)* which is required for cell interactions that prevent particular cells in the developing eye from becoming photoreceptors. Analysis of eyes mosaic for *faf*[+] and *faf*[−] cells shows that *faf* is required in cells near to, but outside, normal developing photoreceptors and also outside of the ectopic photoreceptors in mutant facets. *faf* is also essential during oogenesis, and we show that a faf-lacZ hybrid protein is localized via the first 392 amino acids of faf to the posterior pole of oocytes. Posterior localization of faf-lacZ depends on *oskar*. *oskar* encodes a key organizer of the pole plasm, a specialized cytoplasm at the posterior pole of embryos. The pole plasm is required for germ cell formation and contains the determinant of posterior polarity, encoded by *nanos*. Although other pole plasm components are required for localization of *nanos* RNA or for nanos protein function, *faf* is not. We have cloned the *faf* gene, and have shown that it encodes two similar large ($\sim 300 \times 10^3$ M_r) proteins that are unique with respect to other known proteins.

Key words: *fat facets*, *Drosophila,* eye development, oogenesis, neurogenesis, protein localization.

Introduction

Short-range cell interactions guide the assembly of the *Drosophila* compound eye in the eye imaginal disc, as cells sequentially join growing clusters of unit eyes, or facets (recently reviewed by Banerjee and Zipursky, 1990; Rubin, 1991). Many genes have been identified that are involved in particular steps of the assembly pathway, for example, *scabrous, retina aberrant in pattern, rough, Star, Bar, seven-up, sevenless, seven in absentia, bride-of-sevenless,* and *glass* (Baker et al., 1990; Mlodzik et al., 1990a; Karpilow et al., 1989; Tomlinson et al., 1988; Saint et al., 1988; Heberlein and Rubin, 1991; Higashijima et al., 1992; Mlodzik et al., 1990b; Hafen et al., 1987; Carthew and Rubin, 1990; Reinke and Zipursky, 1988; Moses et al., 1989). The products of all of these genes are required in specific cells within the growing facets, where they either participate in cell-cell interactions or control cell determination from within the cell. The large number of such genes supports the model that facet assembly is controlled by cells within the facets (Tomlinson and Ready, 1987a).

In recent studies of the *argos* (Freeman et al., 1992) and *groucho* (*gro*; Fischer-Vize et al., 1992) genes, it was shown that cells outside of the developing facets also influ-

ence facet assembly. Mutations in both genes have a similar eye phenotype; a few cells are inappropriately determined as photoreceptors. *argos* encodes a diffusible factor (Freeman et al., 1992) and *gro* encodes a nuclear protein (Delidakis et al., 1991). Nothing more is certain about the mechanism of either gene product's function in the eye, or their relationship to each other.

Here we describe a gene called *fat facets (faf)*, that also influences facet assembly from outside the facet. *faf* mutant eyes contain one or more ectopic photoreceptor cells in nearly every facet. By analyzing facets mosaic for *faf*[+] and *faf*[−] photoreceptors, we show that *faf* is required in cells near to, but outside the eight photoreceptors in wild-type facets, and outside the extra photoreceptors in mutant facets. Unlike *argos*, *faf* does not appear to be diffusible. Unlike the pleiotropic roles of *gro*, the essential role of *faf* in repression of neural cell determination is limited to the eye disc.

faf is also required during oogenesis. Embryos from *faf* mutant mothers never cellularize, except for the formation of a few primordial germ cells, called pole cells, at the posterior of the embryo. Moreover, we find that a faf-lacZ hybrid protein is localized to the posterior pole of oocytes.

The localization of particular mRNAs and proteins to

opposite poles of the oocyte is the mechanism by which anterior-posterior polarity of the embryo is established (reviewed by St. Johnston and Nüsslein-Volhard, 1992). Posterior polarity depends on ten posterior group genes, nine of which have been organized into a functional hierarchy: *cappuccino (capu)* and *spire (spir), staufen (stau), oskar (osk), vasa (vas), tudor (tud), valois (vls), nanos (nos)* and *pumilio (pum)* (Boswell et al., 1991; Manseau and Schüpbach, 1989; St. Johnston et al., 1991; Ephrussi et al., 1991; Kim-Ha et al., 1991; Ephrussi and Lehmann, 1992; Hay et al., 1990; Lasko and Ashburner, 1990; Lehmann and Nüsslein-Volhard, 1991, 1987; Macdonald, 1992; D. Barker, J. Moore, L.K. Dickinson and R.L., unpublished data). The posterior group genes function through the *nos* gene product, the posterior determinant, by localizing *nos* RNA to the posterior of the embryo, or modulating the function of nos protein. Except for *nos* and *pum*, the posterior group genes are also required for pole cell formation (reviewed by Lehmann, 1992). stau and vas proteins, and *osk* RNA and protein are known to be localized to the posterior of oocytes prior to *nos* RNA, and the localization of each depends on the genes upstream in the hierarchy (St. Johnston et al., 1991; Ephrussi et al., 1991; A. Ephrussi and R.L., unpublished data; Hay et al., 1988b; Lasko and Ashburner, 1990). Nothing is known about the mechanism of protein localization to the posterior of the oocyte.

Here we show that a faf-lacZ hybrid protein containing only the first 392 amino acids of faf is posteriorly localized as a protein, and that its localization depends on *osk*. Unlike the posterior group genes, *faf* is not required for *nos* RNA localization or function. We have cloned the *faf* gene, and have shown that it encodes two large (~300×10^3 M_r) proteins with different carboxy termini, unique with respect to other known proteins.

Materials and methods

Drosophila *genetics*

Fly lines

The enhancer trap line *AE127*, marked with *white*[+], has a P element in the *svp* gene (M. Mlodzik, J. Heilig and G.M.R., unpublished data). The posterior group mutant alleles used were *capu*[RK] and *spir*[RP] (Manseau and Schüpbach, 1989), *stau*[D3], *osk*[150], *osk*[346], *vas*[D1], *vas*[011], *nos*[L7] (Lehmann and Nüsslein-Volhard, 1991), *stau*[HL54], *tud*[WC8], *vls*[RB71] (Schüpbach and Wieschaus, 1986) and *pum*[680] (Lehmann and Nüsslein-Volhard, 1987). All other mutant markers are described in Lindsley and Zimm (1992). Flies were grown on standard food at 25°C.

Generation of faf *alleles*

Six EMS-induced *faf* alleles were isolated by crossing mutagenized *bw; st* males to *bw; st faf*[FO8]*/TM6B* virgins, and screening approximately 27,000 *bw; *st/st faf*[FO8] male progeny for rough eyes. Thirteen additional *faf* alleles were induced with X-rays as follows. *st* males were exposed to X-rays (4000 rads) and then crossed to *bw; marb*[CD4] *st cu sr e faf*[FO8]*/TM6B* virgins. (*marb*[CD4] is another rough eye mutation; J. A. F.-V. and G. M. R., unpublished data). Approximately 30,000 *bw; *st/...faf*[FO8] male progeny were screened for rough eyes. One additional *faf* allele (*faf*[BP]) was isolated in a P-M hybrid dysgenesis screen.

Birm2; st males were crossed to Δ*(2-3)99B* virgins at 16-18°C. (The *Birm2* and Δ*(2-3)99B* chromosomes are described by Robertson et al., 1988). The male progeny (*Birm2/+; Δ(2-3)/st*) were crossed to *bw; marb*[CD4] *st cu sr e faf*[FO8]*/TM6B* virgins at 22ºC, and >100,000 *st/...faf*[FO8] male progeny were screened for rough eyes.

Meiotic mapping of faf

First, *bw; st faf*[FO8]*/th st cu sr e ca* virgins were crossed with *bw; th st cu sr e ca* males. Many progeny of all genotypic classes were individually mated with *bw; st faf*[FO8]*/TM6B* virgins to determine whether or not the recombinant chromosome contained *faf*[FO8]. By this process, *faf*[FO8] was positioned near the tip of chromosome 3R. Next, *faf*[FO8] was found to be distal to a *white*[+] P element (*P[w*[+]*]*) insertion in polytene chromosome region 99A (a gift of the Rubin laboratory enhancer trappers) by crossing *w; st e faf*[FO8]*/P[w+]* virgins to *w; st e faf*[FO8] males and scoring many progeny for the *e, w*[+] and *faf*[−] phenotypes.

Other fly crosses

Standard genetic crosses were used to combine *faf*[FO8] with *AE127*, the posterior group genes with a *faf/lacZ* P element on the X chromosome, and *fafI, fafII* and *fafIII* P elements on chromosomes X and II with *faf*[BX4].

Generation of faf[−] eye clones

Clones of *faf*[FO8]*/faf*[FO8] cells marked by the absence of the *white* gene product (no pigment granules associated with photoreceptor or pigment cells) were generated by crossing *w*[1118]*; st faf*[FO8]*/TM6B* males with *w*[1118]*; P[w*[+]*]90E* virgins and X-irradiating (1000 rads) their progeny as first instar larvae. *P[w*[+]*]90E* is a *white*[+] P element inserted in polytene region 90E (*P[(w, ry)D]3;* Levis et al., 1985). *w*[1118]*; faf*[FO8]*/faf*[FO8] eye clones were observed in ~1/30 flies.

Isolation of faf genomic and cDNA

Unless otherwise noted, all molecular biology was carried out using standard procedures (Sambrook et al., 1989). The starting point for an 83 kb chromosomal walk (see Pirrotta, 1986) through the *faf* locus was a 5.2 kb *Drosophila* genomic DNA fragment adjacent to a P element in polytene region 100D (a gift of the Rubin laboratory enhancer trappers) isolated by 'genomic rescue' (Pirrotta, 1986). We walked within a non-amplified *Sau*3A-partial *Drosophila* genomic DNA library in bacteriophage gem12 (Promega), constructed from genomic DNA of the isogenic *st* strain in which the *faf* alleles were originally induced. The walk was monitored by using the most distal recombinant phage isolated from each step as a probe for in situ hybridization to *faf*[BX4]*/+* polytene chromosomes. After a phage hybridized to both ends of the *faf*[BX4] inversion, we walked another two steps and then attempted to locate, within the walk, DNA lesions associated with each of the homozygous viable *faf* alleles. Blots of genomic DNA prepared from flies homozygous for each *faf* allele and restricted with either *Eco*RI or *Bam*HI were sequentially hybridized with recombinant phage spanning the entire 83 kb walk, and lesions associated with *faf*[BX4] and *faf*[FBB12] were identified.

Seven different partial *faf* cDNAs were isolated from an eye disc cDNA library (constructed by Alan Cowman) using a 4.5 kb *Eco*RI fragment of genomic DNA as a hybridization probe. All of the cDNAs were subcloned into the Bluescript plasmid (Stratagene) as *Eco*RI fragments. The relationships between the seven cDNAs were determined by hybridizing DNA blots of cDNA restriction fragments with probes prepared from each cDNA. The cDNAs were mapped onto the genomic DNA by using cDNAs as

hybridization probes to blots of restricted recombinant phage clones spanning the walk.

DNA sequencing and analysis

The DNA sequence of cDNAs 3-2 (~3.1 kb) and 7-3 (~5.8 kb), subcloned into the Bluescript plasmid (Stratagene), were determined on one strand from a plasmid primer by constructing nested sets of deletions with exonuclease III and S1 nuclease (Erase-A-Base Kit, Promega) and using Sequenase (USB). The sequence of the 3′-most 0.5 kb of cDNA 6-5 was determined starting with a plasmid primer and continually synthesizing 17-mer oligonucleotide primers corresponding to the end of each sequence. The opposite strand of the genomic DNA sequence was determined using 17-mer oligonucleotide primers synthesized based on the cDNA sequence. Additional primers were synthesized to sequence through introns. DNA sequences were compiled and analyzed using the MacVector programs. Database searches for nucleic acid and protein sequence similarities were performed at the NCBI using the BLAST network service (PIR, SwissProt, Genpept, and GPUpdate databases).

Plasmid constructions

DNA manipulations were carried out using standard procedures (Sambrook et al., 1989). The three genomic DNA plasmids (*fafI, fafII, fafIII*) were constructed by piecing together subcloned *Eco*RI and *Bam*HI fragments of various recombinant phage clones, and then cloning them as *Not*I fragments into the P element transformation vector pW8 (Klemenz et al., 1987) in the same orientation as the *white* gene. To construct *faf/lacZ*, *faf* coding sequences ending at the *Bgl*II site after the first 392 amino acids were joined to a *Bam*HI site at the start of the *lacZ* coding region. A 4.2 kb *Xba*I-*Bgl*II fragment of *fafIII* (the 5′-most 4.2 kb with an engineered *Xba*I site at the *Eco*RI site) was ligated with a 4.0 kb *Bam*HI-*Eco*RI fragment of pC4βgal (Thummel et al., 1988; containing *lacZ* coding sequences and poly(A) addition sequences from SV40), into a plasmid vector restricted with *Xba*I and *Eco*RI. The resulting 8.2 kb *faf/lacZ* fragment was then removed as a *Not*I fragment and cloned into pW8. *fafnuc/lacZ* was generated from *faf/lacZ* by cloning sequences encoding a nuclear localization signal (KKKRKV; Kalderon et al., 1984) in frame into the *Xho*I site, 54 amino acids downstream of the ATG. First, two 26-mer oligonucleotides were annealed (5′-TCGAGGGCTAGCGGT-TAACAGATCTC and 5′-TCGAGAGATCTGTTAACCGC-TAGCCC) and cloned into the *Xho*I site as an adaptor for another set of oligonucleotides containing the nuclear localization signal. The plasmid was then restricted within the adaptor with *Nhe*I and *Bgl*II, and annealed 44-mer oligonucleotides with one *Xba*I and one *Bam*HI end (5′-CTAGAAACATGACCCCCCCCAAGAA-GAAGCGCAAGGTGGAGGAC and 5′-GATCGTCCTCCAC-CTTGCCGCTTCTTCTTGGGGGGGGGTCATGTTT) were ligated into the adaptor. In total, 19 amino acids were added.

RNA blot analysis

Total nucleic acid was prepared from 5 ovaries or 20 eye disc complexes by dissecting the tissues in PBS and immediately transferring them to a microfuge tube containing 200 μl of homogenization buffer (50 mM Tris pH 7.5, 50 mM NaCl, 10 mM EDTA, 0.5% SDS, 2 mg/ml proteinase K). The tissues were homogenized with a small teflon homogenizer, incubated at 37°C for 45 minutes and extracted with phenol/chloroform, which was then back-extracted with homogenization buffer. The pooled supernatants were ethanol precipitated, the pellets washed with 100% ethanol and resuspended in formaldehyde gel loading buffer (Sambrook et al., 1989). The total nucleic acid preparations were electrophoresed on formaldehyde gels, blotted and hybridized with ^{32}P-labeled, double-stranded DNA probes according to standard procedures (Sambrook et al., 1989).

P element transformation

w^{1118} embryos were injected according to standard procedures (Spradling, 1986) with each of the *faf* genomic DNA plasmids and *faf-lacZ* hybrid plasmids at 300 μg/ml, and the helper plasmid p 25.7 2-3 at 100 μg/ml. Plasmids for injections were purified on CsCl-EtBr equilibrium density gradients (Sambrook et al., 1989). Transgenic flies were identified by their red eye color, mapped to a chromosome and balanced using standard genetic crosses.

Sections and scanning electron microscopy of adult eyes

Eyes were fixed, embedded in plastic and sectioned essentially as described in Tomlinson and Ready, 1987b. Heads were bisected to ensure penetration of fix. Sections were 1 μm. Scanning electron microscopy was performed as described in Kimmel et al., 1990.

β-galactosidase histochemistry

The procedures followed for β-galactosidase histochemical staining are essentially those of Y. Hiromi and M. Mlodzik (personal communication). Ovaries and eye discs were dissected in PBS, fixed in a depression slide for 15 minutes at room temperature in 1% glutaraldehyde in PBS, and then washed with PBS at least twice for 10 minutes. The wash was replaced by prewarmed (65°C for at least 15 minutes) staining solution (10 mM NaPO$_4$ (pH 7.2), 150 mM NaCl, 1 mM MgCl$_2$, 3 mM K$_4$[Fe(CN)$_6$], 3 mM K$_3$[Fe(CN)$_6$]) and then by prewarmed staining solution to which 1/30 volume of 8% X-Gal in DMSO was added. Ovaries were incubated for 5 hours or less and eye discs overnight at room temperature. The staining reaction was stopped by several rinses in PBS and the tissues were then transferred to 80% glycerol in PBS, and mounted in this solution. Ovaries and eye discs from wild-type (non-transgenic) flies were used as negative controls.

Embryos were dechorionated with 50% bleach, washed with H$_2$O and fixed in a depression slide for 15 minutes in 0.5 ml heptane saturated with 25% glutaraldehyde in PBS. Embryos were transferred to a glass slide and excess heptane blotted. After the heptane evaporated, embryos were transferred to double-sided tape on another glass slide by gently sandwiching the embryos between two slides. Embryos were then covered with PBS, devitellinized with a tungsten needle and transferred to a depression slide. After removing the PBS, embryos were incubated at room temperature in staining solution (above) without X-Gal for 5 minutes, and then in staining solution with X-Gal. *faf/lacZ* embryos began to stain almost immediately. Stained embryos were quickly rinsed in 70% and then 100% ethanol and mounted in 80% glycerol in PBS.

Immunostaining eye discs

Third instar larval eye disc complexes, attached only to the mouth hooks, were dissected in PBS. All antibody incubations and washes between them were performed in 96-well microtiter dishes, and eye discs were transferred from well to well with tungsten hooks (A. Tomlinson, personal communication). For mAb22C10, discs were treated exactly as described in Tomlinson and Ready, 1987a. mAbro staining was performed by a modification of the procedure in Kimmel et al., 1990. Discs were fixed for 45-55 minutes on ice (0.1 M Pipes (pH 7.0), 2.0 mM EGTA, 1.0 mM MgSO$_4$, 4% paraformaldehyde), washed for at least 15 minutes on ice in PBS+0.1% Triton X-100 (PBST), and then incubated in mAbro diluted in PBST for at least 1 hour at 4°C. After three 5

minute washes in PBST, discs were transferred to biotin-conjugated goat anti-mouse secondary antibody (Jackson) diluted in PBST and incubated for at least 1 hour at 4°C. Following three 5 minute washes in PBST, discs were transferred to ABC complex (Vectastain Elite Kit, Vector) diluted in PBST for exactly 1 hour, washed three times for 5 minutes in PBST and put on ice. Discs were incubated in HRP-staining solution (1× PBS, 0.1% Triton X-100, 0.5 mg/ml DAB, 0.02% $CoCl_2$, 0.02% $NiCl_2$, 0.003% H_2O_2) for ~1 minute while monitoring the level of staining with a dissecting microscope, and then transferred to 1× PBS. Discs were transferred to 30% ethanol, the peripodial membranes were removed and then the discs were dehydrated (50, 70, 90, 2× 100% ethanol) and mounted in DPX (Fluka). *faf/lacZ* and *fafnuc/lacZ* eye discs were stained with a monoclonal antibody to β-galactosidase (Promega) by the same procedure used for mAbro staining, except discs were dissected in 0.1 M sodium phosphate (pH 7.2), PLP fix was used (Tomlinson and Ready, 1987a) and 0.1 M sodium phosphate +0.2% saponin was substituted for PBST in all subsequent steps. Convincing staining of faf-lacZ protein behind the morphogenetic furrow required the biotin intermediate amplification step. *AE127* eye discs were similarly stained except that the biotin amplification step was omitted and instead, an HRP-conjugated goat anti-mouse secondary antibody was used (Biorad).

Immunostaining embryos

Embryos were dechorionated in 50% bleach, washed in H_2O and fixed by gentle shaking for 20 minutes in 1 part fix (0.1 M Pipes (pH 7.0), 2 mM $MgSO_4$, 1 mM EGTA, 4% formaldehyde) : 3 parts heptane. The fix was removed and embryos devitellinized by adding an equal volume of methanol and shaking hard. Devitellinized embryos were rinsed twice with methanol, then washed three times for 30 minutes with BBT (1× PBS, 0.1% BSA, 0.1% Tween20) and incubated at 4°C overnight with preadsorbed primary antibody (anti-vas polyclonal rabbit sera or anti-hb polyclonal rat sera) diluted in BBT. Embryos were then washed four times for 30 minutes with BNT (PBS + 0.2% goat serum) and incubated for 2 hours in preadsorbed biotin-conjugated goat anti-rabbit (for anti-vas) or biotin-conjugated goat anti-rat (for anti-hb) secondary antibody (Jackson) diluted in BBT. After at least three 30 minute washes in PBT (PBS + 0.1% Tween 20) embryos were incubated for exactly 1 hour in ABC complex (Vectastain Elite Kit, Vector), washed three times in PBT, then put on ice. Embryos were incubated for a few minutes in HRP staining solution (0.5 mg/ml DAB +0.006% H_2O_2 in PBT) and the reaction stopped by washing with PBT. Embryos were dehydrated with ethanol (30, 50, 70, 90, 95, 2× 100%), rinsed for 30 seconds in acetone and mounted in Araldite (Ladd).

Hoechst and phalloidin staining embryos

Embryos were double-stained with Hoechst 33258 (Sigma) and rhodamine-conjugated phalloidin (Molecular Probes) exactly as described in Wieschaus and Nüsslein-Volhard (1986).

RNA in situ hybridization

RNA was detected in whole embryos and ovaries using digoxigenin-labeled DNA probes exactly as described in Ephrussi et al. (1991), except that the methanol/DMSO step was omitted for the embryos, the levamisole was omitted from the developing buffer, and the hybridization buffer was 50% formamide, 5× SSC, 100 µg/ml sonicated salmon sperm DNA, 50 µg heparin, 0.1% Tween 20. Digoxigenin-labeled double-stranded DNA probes were prepared from 200 ng of DNA as described in Ephrussi et al. (1991). The probes were a 2.1 kb *Eco*RI fragment of *faf* cDNA 3-2, a 1.9 kb fragment of *nos* cDNA from pN5 (Wang and Lehmann, 1991),

a 0.8 kb fragment of a *cyclinB* cDNA and linear Bluescript plasmid (Stratagene) as a negative control. Embryos and ovaries were dehydrated and mounted using exactly the same procedure as for the immunostained embryos.

Results

Identification of faf alleles

The two original *faf* alleles (*faf^FO8* and *faf^FBB12*) were identified as mutations causing rough eyes in an extensive screen of the *Drosophila* third chromosome for recessive viable mutations with abnormal eye morphology (J.A.F.-V., R.W. Carthew and G.M.R., unpublished data; see Fischer-Vize et al., 1992). Subsequently, 20 additional alleles, generated chemically, with X-rays, or by hybrid dys-

Table 1. fat facets *alleles*

Alleles	Eye phenotype	Female sterility	Cytology
EMS-induced			
faf^FO8	strong	sterile	
faf^FBB12	strong	sterile	
faf^B3	strong	sterile	
faf^B4	weak	fertile	
faf^B5	weak	sterile	
faf^B6	weak	sterile	
faf^B7	strong	sterile	
faf^B8	strong	sterile	
X-ray-induced			
faf^BX1	strong	sterile	no defect
faf^BX3	weak	sterile	no defect
faf^BX4	strong	sterile	In(3R)100E/98D
faf^BX5	lethal	n.a.	no defect
faf^BX6	lethal	n.a.	no defect
faf^BX7	weak	fertile	N.D.
faf^BX8	strong	sterile	no defect
faf^BX9	lethal	n.a.	*
faf^BX10	lethal	n.a.	no defect
faf^BX11	weak	fertile	no defect
faf^BX13	lethal	n.a.	*
faf^BX14	lethal	n.a.	N.D.
faf^BX15	strong	sterile	N.D.
hybrid dysgenesis-induced			
faf^BP	lethal	n.a.	Df(3R)100EF

Twenty-two independent alleles of *faf* were generated using EMS, X-rays, or P-M hybrid dysgenesis (Materials and methods). Except for the two original alleles, *faf^FO8* and *faf^FBB12*, all of the *faf* alleles were identified as flies with rough eyes carrying a mutagenized chromosome in *trans* to *faf^FO8* (Materials and methods). All of the strong *faf* alleles also result in female sterility. The eyes of flies expressing *faf^BX3* and *faf^B5* appear weakly rough. The eyes of flies carrying each of the other weak alleles vary from weakly rough to wild-type. Only three of the six weak *faf* alleles that result in a weakly rough eye also cause female sterility. This implies that there are different requirements for *faf* function, either for levels of faf activity or for specific functional domains of faf protein, in the eye and ovary. No intermediate maternal effect phenotypes were observed; the phenotypes of embryos from *faf* mutant mothers carrying any of the sterility-causing alleles are similar to the null phenotype, and females homozygous for three of the alleles that cause weak eye phenotypes, *faf^B4*, *faf^BX7* and *faf^BX11*, display wild-type fertility. Six of the X-ray-induced *faf* alleles cause lethality when homozygous, presumably because their lesions affect essential loci outside *faf*. Salivary gland polytene chromosomes of most of the X-ray alleles and the hybrid dysgenesis allele were examined for visible lesions (Materials and methods). Asterisks (*) indicate complicated rearrangements involving the tip of chromosome 3R. (n.a. means 'not applicable' and N.D. means 'not determined'.)

genesis (Table 1), were identified as rough-eyed flies carrying the mutagenized chromosome in *trans* to *faf^FO8* (Materials and methods). Based on the severity of eye roughness of homozygotes, the *faf* alleles were categorized as strong or weak (Table 1 and see below).

Meiotic mapping positioned *faf* very close to the tip of chromosome 3R (Materials and methods). *faf* was specifically localized to polytene region 100E by analyzing the polytene chromosomes of several *faf* alleles (Table 1) and finding that *faf^BX4* is an inversion with breakpoints in polytene regions 98D and 100E, and *faf^BP* is deficient for the entire 100EF region (data not shown).

The faf *mutant phenotype*
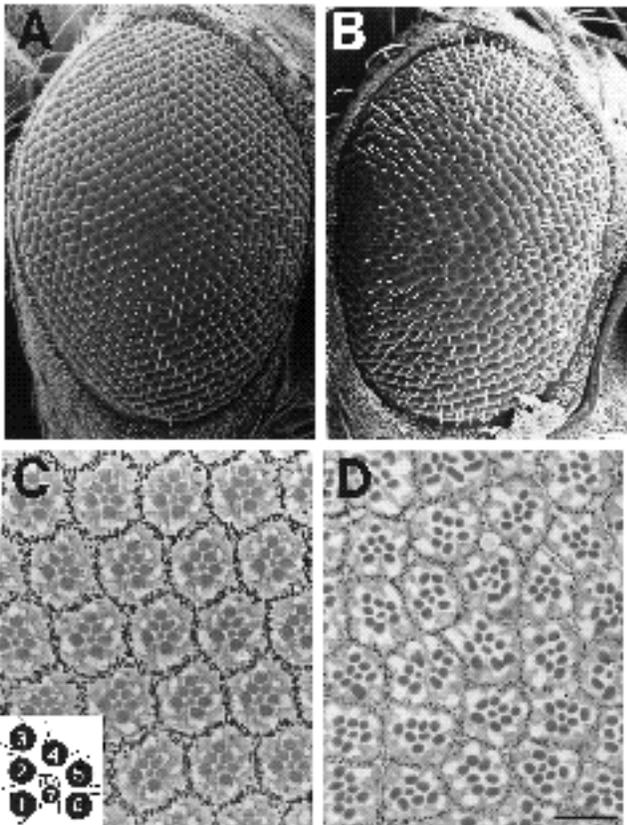
Fig. 1. *faf* mutant eye phenotype. (A,B) Scanning electron micrographs (Materials and methods) of wild-type and *faf^FO8* eyes, respectively, show that the hexagonal facet array of *faf* mutant eyes is irregular and that bristles are sometimes absent or multiplied. (C,D) Tangential sections (Materials and methods) through wild-type and *faf^FO8* eyes, respectively. Photoreceptor cells may be identified by their rhabdomeres, which appear as darkly staining circles. Photoreceptor cell R7 is visible in this plane of section. The most obvious defect in the *faf* mutant retina (D) is the appearance of one or two extra outer photoreceptor cells in most facets (~88%). There are also facets with too few photoreceptors (~2%), facet fusions, rhabdomere malformations and wild-type facets (~10%). In addition, the orientations of the facets are non-uniform. The bar (in D) is 20 μm in C,D, and 100 μm in A,B.

Eye phenotype

The eyes of *faf* mutant flies appear rough due to irregularities in the normally precise hexagonal array of facets (Fig. 1A,B). In order to investigate the *faf* mutant phenotype beneath the eye surface, the retinas of *faf* flies were examined in tangential sections. The wild-type retina is a hexagonal lattice of identical facets, or ommatidia (Fig. 1C). An ommatidium has eight photoreceptor cells (R-cells), each uniquely positioned within a trapezoidal array. There are six outer photoreceptor cells, R1-R6, with large light-gathering organelles, called rhabdomeres, and two inner R-cells, R7 and R8, with small rhabdomeres. The trapezoids are oriented in one direction and are symmetrical with respect to an equator. Many defects are present in strong *faf* mutant retinas, the most striking of which is the appearance of one, two or occasionally three extra outer photoreceptor cells in most facets (Fig. 1D and legend).

The retinas of all strong *faf* mutants appear similar (data not shown). The retinas of all weak *faf* alleles are also similar to each other and generally have as few as 1% mutant facets (data not shown; Table 1 legend). The eye defects of the strong alleles are likely to represent the complete loss-of-function (null) eye phenotype of *faf*, as the strong mutants (including the viable inversion *faf^BX4* that breaks within the *faf* transcript (see below)) show the same phenotype when homozygous or in *trans* to the *faf^BP* deficiency (data not shown).

Maternal effect phenotype

All of the strong *faf* alleles and three of the weak alleles cause female sterility (Table 1 and legend). Females homozygous for any of these alleles have apparently normal ovaries, but lay eggs that never form cuticle and never hatch. In order to examine the mutant embryos in more detail, embryos of wild-type and homozygous *faf^FO8* mothers were stained consecutively with the DNA stain Hoechst to reveal nuclei, and with rhodamine-conjugated phalloidin, which stains f-actin and thus allows visualization of cell membranes (Materials and methods). No normal embryos from *faf^FO8* mothers were ever observed. After the fertilized egg is laid, the nucleus normally undergoes 14 synchronous cycles of division prior to cellularization (see Foe and Alberts, 1983). By division cycle 10, the nuclei have migrated to the egg periphery to form the 'syncytial blastoderm' (Fig. 2A) and the primordial germ cells, called pole cells, located at the posterior pole of the embryo, form (Fig. 2B). The rest of the nuclei continue synchronous divisions until cycle 14, when cellularization produces the 'cellular blastoderm'. In wild-type embryos, f-actin is associated with each nucleus at syncytial blastoderm (Fig. 2B), and then with cell membranes at cellular blastoderm (Fig. 2F; Karr and Alberts, 1986). In embryos from *faf* mutant mothers (hereafter called *faf* mutant embryos), no normal syncytial blastoderm embryos are observed. The embryo shown in Fig. 2C,D appears to be at least in cycle 10 because pole cells are present (Fig. 2D), but most of the nuclei have not migrated to the periphery (Fig. 2C). Also commonly observed were embryos in which patches of asynchronously dividing nuclei had migrated to the periphery (Fig. 2G). Except for the pole cells (which are fewer in number and

spread out instead of grouped together as in wild-type - see Fig. 11D, below), no cellularization of *faf* mutant embryos was ever observed (Fig. 2H).

In summary, the most obvious defects in *faf* null mutants are in the eye and in the embryo. In the eye, the most striking abnormality is the appearance of ectopic outer photoreceptors. *faf* mutant embryos never reach normal syncytial blastoderm and except for some pole cells, never cellularize.

Developmental defects in faf *mutant larval eye discs*

To determine if defects in early stages of ommatidial assembly contribute to the *faf* eye phenotype, wild-type and strong *faf* mutant larval eye discs were stained with the neural specific antibody mAb22C10 (Fujita et al., 1982). mAb22C10 reveals the sequence of photoreceptor cell assembly (R8, R2/5, R3/4, R1/6, R7) as each R-cell expresses the mAb22C10 antigen when it acquires neural identity (Fig.3A; Tomlinson and Ready, 1987a). As ommatidial development proceeds in a posterior-to-anterior wave in the eye disc, facets at all stages of photoreceptor cell assembly are observed in one larval disc. The most striking defect observed in stained *faf* mutant larval eye discs is the appearance of one or two ectopic photoreceptor cells in most developing facets at the time when R3/4 normally begin to stain with mAb22C10 (Fig.3B,C). These cells may be the mystery cells, which are normally positioned between R3 and R4 in a precluster that emerges from the morphogenetic furrow, but then disappear into the surrounding pool of undifferentiated cells without expressing neural antigens (Tomlinson and Ready, 1987a; Wolff and Ready, 1991a).

In order to determine if the ectopic photoreceptors are indeed positioned between R3 and R4, we examined the expression of the *rough* (*ro;* Tomlinson et al., 1988; Saint

et al., 1988) and *seven-up* (*svp;* Mlodzik et al., 1990b) genes in *faf* larval eye discs. Both *ro* and *svp* are normally expressed in R3 and R4 (Kimmel et al., 1990; Mlodzik et al., 1990b; Fig.3 legend). In *faf* mutant eye discs, both genes are also expressed in one or two cells adjacent to R3 and R4. Thus, the ectopic R-cells in *faf* mutant larval eyes are likely to be the mystery cells.

faf[+] *function is required outside normal and ectopic photoreceptors*

In order to determine which cells require *faf[+]* function to exclude the ectopic R-cells from the precluster, marked *(w[−])* clones of *faf[FO8]* mutant cells in wild-type eyes were generated using X-ray induced somatic recombination

Fig. 3. The mystery cells become ectopic photoreceptors in *faf* mutant eye discs. Wild-type (A) and *faf[FO8]* (B,C) larval eye discs were stained with mAb22C10 (Materials and methods) to reveal the sequence of photoreceptor cell assembly. mAb22C10 stains photoreceptor cell cytoplasm and the discs are photographed to show the apical tips of the photoreceptor cells near the morphogenetic furrow. The arrows point along the direction of the morphogenetic furrow. Five-cell clusters containing R8,2,5,3,4 are apparent in wild-type (A) discs. At a similar distance from the furrow in *faf* mutant discs (B), clusters containing 6 or 7 cells are seen. The cell(s) which appear to be the extra photoreceptors (see below), situated between R3 and R4, are indicated by asterisks (*). Improperly spaced clusters (C) also appear occasionally in *faf* mutant discs. (D) *faf* mutant larval eye discs stained with an antibody to the ro protein (Materials and methods), which is expressed in the nuclei of R2,5,3,4. One or two ectopic nuclei (*), apparently located between R3 and R4, express ro. (E) *faf* mutant larval eye discs carrying one copy of the enhancer trap P element *AE127,* which is inserted in the *svp* gene and expresses β-galactosidase in the nuclei of R3,4,1,6, were stained with an antibody to β-galactosidase (Materials and methods). A cluster is shown containing an ectopic nucleus (*), apparently between R3 and R4. As the mystery cells are normally located between the R3/4 pair, the ectopic photoreceptor cells are likely to be improperly determined mystery cells. *ro* and *svp* appear to be otherwise appropriately expressed, suggesting that many of the photoreceptor cells in *faf* mutant discs acquire their normal identities. The bar (in C) is 10 μm in all panels.

Fig. 4. *faf*[+] function is required outside of the photoreceptor cells. (A) A tangential section through a typical w^- faf^- clone in a w^+faf^+ eye, generated by X-ray-induced somatic recombination (Materials and methods). The w^- clone is recognizable by the absence of pigment granules in the pigment cells forming the hexagonal lattice, and also near each photoreceptor (black dots). The large arrow points to the w^+ ectopic cell in a phenotypically mutant mosaic facet at the clone border. The small arrow indicates a phenotypically wild-type facet at the clone border. (B) A tabulation of data obtained by examining both phenotypically wild-type and phenotypically mutant mosaic facets in many different clones. The R-cell genotypes in 209 wild-type mosaic facets in 20 separate clones were scored. (For technical reasons, a smaller number of R8 cells was scored.) As there are normally ~10% wild-type facets in an *faf* mutant eye, if there were no requirement for *faf* in particular R-cells, those R-cells would be expected to be w^+faf^+ at a frequency of ~55% [50%+(1/2)(10%)]. In contrast, if *faf* were required in particular R-cells in order to form a wild-type facet, those R-cells would be expected to be *faf*[+] at a frequency of ~95% [100%-(1/2)(10%)=95%]. The data indicate that all R-cells are *faf*[+] at a similar frequency, averaging at 65%. This slightly higher than random (55%) frequency suggests that *faf* is required in cells closely related by lineage to (near to) the R-cells (see Tomlinson et al., 1988; Reinke and Zipursky, 1988; Carthew and Rubin, 1990; Mlodzik et al., 1990b.) Phenotypically mutant facets were also examined, and the tabulation of the frequency of w^+faf^+ outer R-cells in 78 mosaic facets at the borders of 18 different clones is shown. In contrast to the results obtained with the wild-type mosaic facets, the frequency of *faf*[+] outer R-cells in the mutant mosaic facets (35%) is lower than random [50%-(1/2)(10%)=45%], suggesting that either R-cells within these facets or cells related by lineage to them must be *faf*[-] in order to produce a phenotypically mutant facet. (C-I) Phenotypically mutant mosaic facets in which one or two extra outer photoreceptors are neatly added. In these facets, knowing that the extra cells are the mystery cells, located between R3 and R4, it is possible to identify all of the outer R-cells. We assume that no gross rearrangement of the R-cells occurs in the eye disc after the assembly stages that we have observed. The ectopic cell (labeled 'm' for mystery cell) and also sometimes either R3, R4 or both, are w^+faf^+. Only the w^+ R3,4, and m-cells are labeled. The facets in E and F also contain w^+ R8's (not visible in this plane of section). The facet in H has two ectopic R-cells, only one of which is w^+. These facets show that *faf* expression in the ectopic cells or in the cells they contact (R8,3,4) is not sufficient to exclude the mystery cells from the facet. We observed a total of five facets like those shown in C and D, six additional facets in which the ectopic cell was w^+, and four facets where R8 could be scored definitively as w^+. No phenotypically mutant facets were observed in which all R-cells were w^+. We interpret this as a reflection of the proximity of the *faf*-requiring cells to the mystery cells, and thus the low probability of the cell(s) requiring *faf*[+] function being w^- while the mystery cells and all of the other R-cells in a mutant facet are w^+.

(Materials and methods). Within the patch of *w⁻faf⁻* cells, the retina shows the *faf* mutant phenotype, and outside the *w⁻* clone, the retina appears wild type (Fig.4A). As there are no lineage restrictions on cells within particular facets (Ready et al., 1976; Lawrence and Green, 1979; Wolff and Ready, 1991b), there are mosaic ommatidia at the clone borders, containing both *w⁺faf⁺* and *w⁻faf⁻* photoreceptor cells (Fig. 4A). Some of the mosaic facets are phenotypically wild type (that is, there are two inner and six outer R-cells) and some are typical *faf* mutant facets, with more than six outer R-cells (Fig.4A).

Examination of both the phenotypically wild-type and mutant mosaic facets shows that *faf⁺* function is not required in photoreceptor cells. First, the genotypes of the different photoreceptors in phenotypically wild-type mosaic ommatidia were scored to determine if there is a tendency for particular photoreceptors to be *faf⁺*. None of the eight photoreceptors was *faf⁺* nearly as frequently as expected if *faf⁺* function is required in a particular photoreceptor cell (Fig. 4B and legend). Second, we find that in the phenotypically mutant facets, the ectopic photoreceptor cells are not always *faf⁻*, and the R-cells that they contact (R3, R4 and R8) are also not always *faf⁻*, consistent with the analysis of the wild-type mosaics (above). Thus, it is not the absence of *faf⁺* function in the ectopic cells, or the cells they contact, that results in their misdetermination as photoreceptors.

The analysis of phenotypically wild-type and mutant facets suggests that *faf* is required in cells closely related by lineage to (that is, near to) the photoreceptors. Each of the photoreceptors in the phenotypically wild-type mosaic ommatidia are *faf⁺* at a higher than random frequency (Fig. 4B and legend). In contrast, the outer photoreceptors within the mutant mosaic ommatidia are *faf⁺* at a lower than random frequency (Fig. 4B and legend).

In summary, we interpret these observations as indicating that cells near to, but outside the normal or ectopic photoreceptors in a particular facet must be *faf⁺* in order to prevent the neuralization of the mystery cells.

Cloning the faf *gene*

Having localized *faf* to polytene region 100E (above), we isolated *Drosophila* genomic DNA adjacent to a P element inserted in 100D (Materials and methods) and used this DNA as a starting point for an 84 kb chromosomal walk within a bacteriophage lambda *Drosophila* genomic DNA library, towards the telomere of 3R (Materials and methods, Fig. 5A). Progress in the walk was monitored by hybridizing, in situ, phage clone probes from each step to polytene chromosomes of the homozygous viable inversion *faf^BX4* (Table 1, Materials and methods). A phage clone hybridized to both ends of the inversion about 55 kb into the walk (Fig. 5A; data not shown). An additional 29 kb of genomic DNA was isolated distal to the inversion breakpoint. The *faf^BX4* breakpoint was localized to a 1.8 kb *Eco*RI-*Bam*HI fragment of genomic DNA and a small deletion in a neighboring 1.5 kb *Bam*HI fragment was found to be associated with *faf^FBB12* (Fig. 5A; data not shown - see Materials and methods). A 4.5 kb *Eco*RI fragment of genomic DNA containing the region of both mutant lesions was used to isolate many cDNAs from an eye imaginal disc

cDNA library (Materials and methods). All of the cDNAs are either identical to or contained within the three shown in Fig. 5A (Materials and methods). DNA sequencing (below) of the three cDNAs as indicated in Fig. 5A revealed that the cDNAs represent two transcripts of at least ~8500 and ~8900 nt, which differ in their 3′ ends. (We do not know whether the 5′ end of the cDNA 3-2 represents the actual start of the *faf* transcript.) This result correlates with an analysis of RNA blots (Fig. 5B) hybridized with the 4.5 kb *Eco*RI fragment, which shows two transcripts of slightly different size in the appropriate range. Conceptual translation of the two transcripts reveals long open reading frames differing at their 3′ termini (see below).

In order to show more conclusively that the cDNAs correspond to the *faf* gene, transgenic *Drosophila* lines were generated by P element transformation with the three different fragments of genomic DNA shown in Fig. 5A (Materials and methods). One copy of any of the three genomic DNA fragments complements completely both the eye and female sterile mutant phenotypes of *faf^BX4* (data not shown; see Materials and methods). In addition, blots of ovary and eye disc RNA probed with all regions of the smallest complementing genomic DNA fragment *(fafIII)* reveal only the two transcripts that hybridize to the 4.5 kb *Eco*RI fragment of genomic DNA (Fig. 5A). Thus, these are the only two RNAs transcribed within *fafIII* that are present in eye discs or ovaries.

Predicted faf *protein sequences*

The DNA sequences of three overlapping partial cDNAs were determined (Fig. 5A and Materials and methods). In order to confirm the cDNA sequences, the corresponding genomic DNA sequence was determined on the opposite strand A composite sequence is shown in Fig. 6. The genomic DNA sequence shows that the *faf* gene is composed of at least 17 exons. Conceptual translation of the two classes of cDNAs reveals proteins of 2711 and 2747 amino acids, differing slightly in their carboxy termini due to an alternate splice near the 3′ end of the primary transcript. Apart from their large size, the faf proteins have few remarkable features. At amino acids 262-290, there is a potential leucine zipper structural domain, shown to be a dimerization site within transcription factors (Landschulz et al., 1988). As the leucine repeat is not adjacent to a particularly basic (DNA-binding) domain, faf is unlikely to be a transcription factor. In addition, there is a likely PEST sequence (Rogers et al., 1986), often found in rapidly degraded proteins, at amino acids 18-34. No other significant similarity with another protein or DNA sequence in multiple databases has been found (Materials and methods).

Expression of faf-lacZ *hybrid proteins in the eye disc*

In order to determine which cells in the ovary and eye disc express *faf,* we 'tagged' the faf protein with the *E. coli* gene encoding β-galactosidase *(lacZ)* by constructing the two hybrid genes shown in Fig. 7A (Materials and methods). One hybrid, *faf/lacZ*, contains the *faf* promoter through the first 392 amino acids of the faf protein (Fig. 7A legend), fused in frame to sequences encoding lacZ. The other hybrid gene, *fafnuc/lacZ*, is identical to *faf/lacZ*, except that an oligonucleotide containing a nuclear localization signal
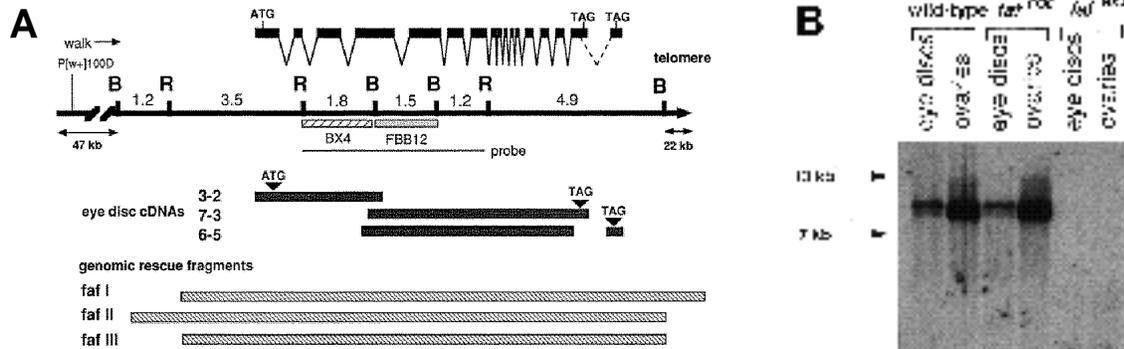
Fig. 5. Cloning the *faf* gene. (A) The thick black line is a restriction map (B, *Bam*HI; R, *Eco*RI) of the *fat facets* locus within approximately 83 kb of cloned *Drosophila* genomic DNA, isolated by walking within a bacteriophage genomic DNA library from a P element inserted in polytene region 100D (Materials and methods). The numbers above the line represent the approximate sizes of the restriction fragments indicated. The hatched and stippled bars just beneath the line indicate the restriction fragments containing one *faf^BX4* inversion breakpoint and a small deletion associated with *faf^FBB12*, respectively, determined by hybridizing recombinant phage probes to DNA blots of restricted genomic DNA from flies carrying these mutant alleles (Materials and methods). The line labeled 'probe' indicates the 4.5 kb *Eco*RI fragment of genomic DNA used to screen an eye disc cDNA library by hybridization (Materials and methods). Three of the seven different cDNAs isolated are shown as shaded bars extending approximately beneath the genomic regions to which they correspond. The lengths of the cDNAs are ~3.1 kb (3-2), ~5.8 kb (7-3) and ~6.2 kb (6-5). The other four cDNAs isolated are partial products of those shown, as determined by DNA blot hybridization experiments (Materials and methods). The complete DNA sequences of 3-2 and 7-3 were determined, as well as ~0.5 kb of the 3′ end and ~0.2 kb at the 5′ end of 6-5, revealing a long open reading frame and two differentially spliced 3′ ends (Fig. 6; Materials and methods). The cDNA sequences were confirmed by sequencing the opposite strand of genomic DNA, which revealed seventeen exons, as shown at the top of the figure (Fig 6; Materials and methods). Due to alternate splicing, the sixteenth exon is shorter in 6-5 than in 7-3, and the seventeenth exon is contained only in 6-5. ATG and TAG indicate start and stop codons, respectively. The long hatched bars indicate the approximate extents of three different genomic DNA fragments (*fafI, fafII* and *fafIII* are ~17, 14 and 13 kb, respectively) introduced into the *Drosophila* genome by P element transformation (Materials and methods). *fafI* extends from the indicated *Eco*RI site through 4.6 kb downstream of the indicated *Bam*HI site, *fafII* extends between the two *Bam*HI sites indicated, and *fafIII* extends between the indicated *Eco*RI and *Bam*HI sites. All three genomic DNA fragments (2 lines of *fafI*, 3 lines of *fafII* and 1 line of *fafIII*) complement the rough eye and female infertility phenotypes of *faf^BX4* (Materials and methods). Although only one *fafIII* line obtained could be tested, we are confident that the ~13 kb *fafIII* fragment contains the entire *faf* gene because its 5′ end is identical to that of *faf I* and its 3′ end is identical to that of *fafII*. (B) An autoradiograph of an RNA blot revealing the two *faf* transcripts. Each lane contains total RNA prepared from 20 third instar larval eye imaginal disc complexes or 5 ovaries of each of the three *Drosophila* lines indicated (Materials and methods). The autoradiograph shown was following hybridization with a ^{32}P-labelled double-stranded DNA probe corresponding to the 4.5 kb *Eco*RI fragment shown in (A). The bottom panel is a lighter exposure of the first four lanes. Identical results were obtained when the blot was stripped of old probe and hybridized with the upstream 3.5 kb *Eco*RI fragment or the downstream 6.1 kb *Bam*HI fragment of DNA (see A). In order to approximate the sizes of the transcripts shown, the blot was also probed with labelled DNA corresponding to a spectrin gene (~13 kb) and a non-muscle myosin gene (~7 kb; Materials and methods). Both *faf* transcripts are expressed in wild-type and *faf^FO8* eye discs and ovaries. As neither the size nor the level of *faf* mRNA appears to be affected in *faf^FO8*, this allele is likely to be a point mutation within the coding sequences. As expected (see A), only one truncated transcript is detected in *faf^BX4* RNA.

was cloned in frame just after the first 54 amino acids encoded by *faf* (Materials and methods). Nuclear localization of an antigen greatly facilitates the identification of individual cells expressing it in the eye disc. Both constructs were introduced into the *Drosophila* genome by P element transformation (Materials and methods). Ten independent lines (five containing each construct) were analyzed as described below and showed similar results, except that *fafnuc/lacZ* lines showed nuclear expression.

Eye discs of third instar transgenic larvae were histochemically stained for β-galactosidase activity (Materials and methods). Staining was seen within the entire region ahead of the morphogenetic furrow including the antennal disc (Fig. 7B), but only after prolonged incubation. Staining eye discs with antibodies to β-galactosidase, which is more sensitive than histochemical staining, revealed an

even lower level of faf-lacZ protein behind the furrow as well (Fig 8A-C; Materials and methods). Because the antibody stain in *fafnuc/lacZ* lines is nuclear, it is possible to identify the cells that express faf-lacZ protein. *fafnuc/lacZ* eye discs stain with anti-β-galactosidase antibodies in all cells ahead of and behind the morphogenetic furrow, but not within the furrow (Fig. 8A-F). Therefore, faf-lacZ hybrid protein must be rapidly degraded in cells as they enter the furrow.

faf-lacZ protein was also consistently detected in other discs and larval and adult tissues (i.e. fat body, gut, larval ovary and testes, adult male sex organs; data not shown).

faf-lacZ protein, but not mRNA, is localized to the posterior pole of oocytes

Ovaries of transgenic females were also histochemically

```
   1 AAACAAAAACAATTTTAAATTAAACTCAAACGCACACACAAAAAGAGAGGGTTCTGTGGGTGAGCGTGCGTGCGTGTGAGTGGGTGAGAGCGTAATTTGTGTGCTTTTGGAGCATTATTT
 121 GGCGCTTGCTGCCCCGTCTATTAAATAATATTGATTTTTCATATTCCAACCCAGCGCTCTCTTACACACACGCAGGACCACTCACACACCCAAAACTGAGCTTCGTACACATGTGGGCCT
 241 TTTCGGTCTCTCTCTTAGCGCGCTGTATTTATCCGCCTCTTATTGCCTCTTTCGCTTCGTTGCACTTTATAGATTTTTGCAAACTGGCCGCTGACCAGGCATTGGGAAAAAGGACAATAAGA
 361 GGGAGCCCTGAGCAAGAAGTCCAACTAACCGACCAACTGTGCCTCAAGAAACAGCAAATCAGCGGGGAACTCCCAAACTGGACAAGCAGAAGCCCGGATACCTCACATTTGCCATAAGAC
 481 ACCAGTCAAAAAACGACGCGCACTGCAGAACGCATTAGGATTAGCCCGGCACTAGAGCTCGTGCATTAACATCCGCTTCCGGTGCTCCGCAGAACGAGTTGGCTGGAGAGGATACCCCGC
 601 CACCATGACGTTCGACACTCGTAGGCACACCACCGGACAGCCGGGAAGCACAGCGCCCAGCTCAAGCAGCAGCACCACCAGCACAACAACCACCACAACGAGTCCGGCGCAGAGTGCCGG
   1         M   T   F   D   T   R   R   H   T   T   G   Q   P   G   S   T   A   P   S   S   S   S   S   T   T   S   T   T   T   T   T   T   T   S   P   A   Q   S   A   G
 721 ATCAGGATCGGGCATAGGAACAGGCACAGGAACAGTCGCCAACTCGAGTCTGCCAGGAGGAGGATCAGGTAGCTTGGACGGCAACCAGGATCAACAGCCGGCCACAGACAGTCAGAGGCTC
  40         S   G   S   G   I   G   T   G   T   G   T   V   A   N   S   S   L   P   G   G   G   S   G   S   L   D   G   N   Q   D   Q   Q   P   A   T   D   S   Q   S   S
 841 TGATGACGTGGCCGCTTCCTTGTCAGCGAACAGCGTGGACAGCACCATCACAATAGTgtaagtggtcgtgctgccgtatttgcatgactccgcccaattagtcgggtttttgtgcgctcag
  80         D   D   V   A   A   S   L   S   A   N   S   V   D   S   T   I   T   I   V
 961 agtatgaccgccaccccctaaagatccccaacattcgtgtgtttacttttgcggtagtaattcttccacatgcattctttgtttgtctatggttgggccagcacaatgttcgtaattcatt
1081 taactatgtttcagatgattttggccaggccactggttagtgaccataaatgtggcaataaatgttatctttcttttttcttcgataggtgcacttggctgctgtccagttttagcaatga
1201 acttaccaagctttactaatattaccttacccggctattactatctcattttgaatgcccgtccacggccatcttatcgcttgtccatcaaaataagctaacactgaaaattacaaaaa
1321 aaaaaaaaaaaaatgtatgcgttttggcatttatttttaatagaaataaccatttaataaacagaaatgtttatggaagggtgtgctttgatttgattgacccctaagaacttttttttat
1441 gtgtattatttaattgttattttaatattgaataattgtccatttaattaaatagGCCTCCAGAGAAGCTTATATCCTCGTTTCCCACGACGAAACTGAGATCACTAACGCAAAAGAT
  99                                                              P   P   E   K   L   I   S   S   F   P   T   T   K   L   R   S   L   T   Q   K   I
1561 ATCCAATCCGCGTTGGGTGGTGCCGGTGCTGCCTGAACAGGAATTGGAGGTTCTCTTAAATGCAGCCATCGAACTGACCCAGGCTGgtaagtgtgtccaaacgatttatttgataccctc
 120         S   N   P   R   W   V   V   P   V   L   P   E   Q   E   L   E   V   L   L   N   A   A   I   E   L   T   Q   A
1681 taaagtcaatttatttgcagGTGTGGATCACGACTGCGAGCCATGTGTGGAGTTTTATCGCAATGGACTGAGCACTTCCTTTGCCAAGATCCTAACCGACGAGGCGGTGAATTCGTGGAA
 148                             G   V   D   H   D   C   E   P   C   V   E   F   Y   R   N   G   L   S   T   S   F   A   K   I   L   T   D   E   A   V   N   S   W   K
1801 GAATAACATCCACCACTGTATCCTTGTGTCTTGCGGCAAGCTTCTACATCTAATCGCCATCCACATGCAGCGAGACAATCCCTATCTGCTTGACTTGCTGGCTATTGTTTTTGATCCGGA
 182         N   N   I   H   H   C   I   L   V   S   C   G   K   L   L   H   L   I   A   I   H   M   Q   R   D   N   P   Y   L   L   D   L   L   A   I   V   F   D   P   E
1921 GAACAAGTTCAACACATTTAACGCAGGCCGTCAACCGGACTGCTTCGCAGGCCGGACTACATCTGGGGTCAGTTGGACAGCAACAAAATGTACGCCCGTCCGCCGCCGGAGCCCAAGAA
 222         N   K   F   N   T   F   N   A   G   R   Q   P   D   C   F   A   A   P   D   Y   I   W   G   Q   L   D   S   N   K   M   Y   A   R   P   P   P   E   P   K   N
2041 TGCCGCGTGGCTGGCTTGTGGATCTTATTAACCGCTTCGGACAGTTGGGAGGCTTTGACAATCTACTGGAACGGTTTAACATCGGACTGGAACTGCTGAAGCGCAACCAAAACAAGTGCAC
 262         A   R   G   W   L   V   D   L   I   N   R   F   G   Q   L   G   G   F   D   N   L   L   E   R   F   N   I   G   L   E   L   L   K   R   N   Q   N   K   C   T
2161 TGGCAAGAATATTAGTGTCGAGGGTAGGGTTGAAAATGGGGCTCAGGATAATCGCCTCACCCTTCATACACAGCCTCTTGCGTCCATTCGGTCAATGTTATGAGCTTCTAATGCC
 302         G   K   N   I   S   V   E   G   R   V   E   N   G   A   Q   D   N   R   L   T   L   A   L   I   H   S   L   L   R   P   F   G   Q   C   Y   E   L   L   M   P
2281 TGCCACTATAGCCAAGTACTTTATGCCCACCTGGAACGTTGTGTTAGACCTTCTGGACAGCTTCACAGACGAGGAGTTAAAGCGAGAAGTGAAACCTGAGGGACGCAATGACTACATCAA
 342         A   T   I   A   K   Y   F   M   P   T   W   N   V   V   L   D   L   L   D   S   F   T   D   E   E   L   K   R   E   V   K   P   E   G   R   N   D   Y   I   N
2401 TGGGATTGTAAAGTCAGCCCGCTTACTGGCCAGCCGCTTGACTGGCCAGGAGGAGCTAATCCGAGATCTTGAGATGTTTCGCCTAAAGATGATCCTGCGTCTGCTGCAGGTGTCCAGCTT
 382         G   I   V   K   S   A   R   L   L   A   S   R   L   T   G   Q   E   E   L   I   R   D   L   E   M   F   R   L   K   M   I   L   R   L   L   Q   V   S   S   F
2521 CAACGGCAAGATGAATGCCCTTAATGAGATCAACAAGGTGCTCTCCTCGGTCGCATATTTCTCCCATCGATCCCAACCTTTGCCCCACTGTATGCCCGAAGACGAAATGGACTGGCTGAC
 422         N   G   K   M   N   A   L   N   E   I   N   K   V   L   S   S   V   A   Y   F   S   H   R   S   Q   P   L   P   H   C   M   P   E   D   E   M   D   W   L   T
2641 GGCGGATCGCATGGCGCAATGGATAAAGTCGTCAGATGTATTGGGTGTCGTGCTCAAGGACTCGCTACACCAACCGCAGTATGTGGAAAAATTGGAAAAGATCATCCGCTTTCTTATTAA
 462         A   D   R   M   A   Q   W   I   K   S   S   D   V   L   G   V   V   L   K   D   S   L   H   Q   P   Q   Y   V   E   K   L   E   K   I   I   R   F   L   I   K
2761 AGAGCAGGCACTAACGTTGGATGATTTGGATGCAGTTTGGAGAGCGCAGGCCGGCAAACACGAGGCAATCGTGAAGAACGTTCACGATCTACTGGCCAAACTTGCATGGGACTTTACTCC
 502         E   Q   A   L   T   L   D   D   L   D   A   V   W   R   A   Q   A   G   K   H   E   A   I   V   K   N   V   H   D   L   L   A   K   L   A   W   D   F   T   P
2881 CGAGCAACTGGATCACCTCTTCGAGGCGGTTCCAGgtgagttttctgctttaaaagtgtattaagtaagtttaagtttaaattatgattattacattttccgcgatggtttgtatataaa
 542         E   Q   L   D   H   L   F   E   A   F   Q
3001 tctcatgaacaatgaaaacaatattaatttttatttcagGCCAGCATGACCACTGCCAATAAGCGGCAACGGGAAAGACTTCTTGAGCTAATACGCCGCTTGGCTGAGGATGACAAAAAT
 553                                                 A   S   M   T   T   A   N   K   R   Q   R   E   R   L   L   E   L   I   R   R   L   A   E   D   D   K   N
3121 GGTGTGATGGCCCAAAAGGTGCTTAAGCTGTTTTGGACTCTGGCCCACAGTCAGGAGGTGCCGCCAGAGGTGCTTGATCAGGCGCTGGGCGCACACGTTAAAATCTTGGACTACAGCTGC
 590         G   V   M   A   Q   K   V   L   K   L   F   W   T   L   A   H   S   Q   E   V   P   P   E   V   L   D   Q   A   L   G   A   H   V   K   I   L   D   Y   S   C
3241 TCGCAGGAGCGGGATGCCCAAAAGACCATTTGGTTGGATAAGTGCGTTGACGAACTTAAGTCAGGCGATGGATGGGTTCTGCCCGCCTTGCGTCTAATCCGGGATATTTGTTGCCTATAT
 620         S   Q   E   R   D   A   Q   K   T   I   W   L   D   K   C   V   D   E   L   K   S   G   D   G   W   V   L   P   A   L   R   L   I   R   D   I   C   C   L   Y
3361 GATACCACGACAAATCATGCCCAGCGCACTCAAACGAGTACGAATCGTCAGCAGGTGATTGAACGACTACAGAATGATTATTCCTTGGTCATTCTGGTCAACTAAGTTTGACTGCATAC
 660         D   T   T   T   N   H   A   Q   R   T   Q   T   S   T   N   R   Q   Q   V   I   E   R   L   Q   N   D   Y   S   L   V   I   L   V   T   N   S   L   T   A   Y
3481 ATGGAAAAGGTACGCCAAATGGTAACGGACTCACCAGGCCTGGATGCAACCAGAATCCTAATTGACGGAAGATTTCCGCATCACGTGCAAATAGCGGAGCGACTGGAGTTCCTTAAGTTC
 700         M   E   K   V   R   Q   M   V   T   D   S   P   G   L   D   A   T   R   I   L   I   D   G   R   F   P   H   H   V   Q   I   A   E   R   L   E   F   L   K   F
3601 CTGCTGAAAGATGGACAGCTATGGCTGTGCGCAGATCAGGCAAAGCAGATTTGGCATTGCTTGGCGGTCAATGCAGTTTTTCCGGCAGATCGTGAGGAGTGTTTTAGATGGTTTGGCAAG
 740         L   L   K   D   G   Q   L   W   L   C   A   D   Q   A   K   Q   I   W   H   C   L   A   V   N   A   V   F   P   A   D   R   E   E   C   F   R   W   F   G   K
3721 TTGATGGGTGAGGAGCCTGACTTGGATCCTGGCATAAACAAAGACTTCTTCGAGAACAACATCCTGCAACTTGATCCGCATCTAACCGAAAGCGGAATCAAGTGCTTTGAGCGCTTC
 780         L   M   G   E   E   P   D   L   D   P   G   I   N   K   D   F   F   E   N   N   I   L   Q   L   D   P   H   L   L   T   E   S   G   I   K   C   F   E   R   F
3841 TTTAAGGCTGTCAATTCCAAAGAGGACAAGCTGAAGGCGATACATAGAGGTTACATGCTGGACAACGAGGATCTAATAGGAAAGGACTACTTGTGGCGGGTCATTACAACTGGAGGCGAG
 820         F   K   A   V   N   S   K   E   D   K   L   K   A   I   H   R   G   Y   M   L   D   N   E   D   L   I   G   K   D   Y   L   W   R   V   I   T   T   G   G   E
3961 GAAATCGCCAGCAAGGCCATTGATCTACTTAAGGAAGTGTCGACGGCATTGGGCCCACGTCTGCAGGAGAATATCGCAGAGTTTCATGAAATGTTCATTGGTGAGTGCTGCTCCCGCCTG
 860         E   I   A   S   K   A   I   D   L   L   K   E   V   S   T   A   L   G   P   R   L   Q   E   N   I   A   E   F   H   E   M   F   I   G   E   C   C   S   R   L
4081 CGCACACACTATGGAAATATGTTATATTGGGCAAGACGCAGCTGCAAGAGGAATTAGATGCCCCCGATCAGTCGGATAATACCAACGACGAGTCTAAAGATTCAAAAATGCGTTTCATC
 900         R   T   H   Y   G   N   I   V   I   L   G   K   T   Q   L   Q   E   E   L   D   A   P   D   Q   S   D   N   T   N   D   E   S   K   D   S   K   M   R   F   I
4201 GAGGCGGAGAAGATGTGCCGAATTTTAAAGGTGCTGCAAGAATACGTAAAGGAGTGTGATCGTTCTTTCAGCGGCGATCGTGTTCACCTGCCACTTAGCCGGGTTACACGGGGCAAAAAC
 940         E   A   E   K   M   C   R   I   L   K   V   L   Q   E   Y   V   K   E   C   D   R   S   F   S   G   D   R   V   H   L   P   L   S   R   V   T   R   G   K   N
4321 ACCATCTTATATATCCGCTTCCAGAACCCTGGCAGATCATCTGACGACATGGAGATCGTTACACACGACAACGAGACGATGGCCGCTTTTAAGAGAAATCTACTTAAGCGAATCAAGGGA
 980         T   I   L   Y   I   R   F   Q   N   P   G   R   S   I   D   D   M   E   I   V   T   H   S   N   E   T   M   A   A   F   K   R   N   L   L   K   R   I   K   G
4441 ACTTCGACGGCCAACATAAAGGTTGATCTCTTCTATGCTAATGACGAGATGATCGGGGGTTTCCGATGAAATAAACCCACTTTATCAGTACACTATCCGGGATAAGATGAATCTTACAGCA
1020         T   S   T   A   N   I   K   V   D   L   F   Y   A   N   D   E   M   I   G   V   S   D   E   I   N   P   L   Y   Q   Y   T   I   R   D   K   M   N   L   T   A
4561 AAACTCACGCCCGTGGGTACGGGTCTAGCCAGCAGTCCCGACTCCTCGAGCGACTCGAGCACAGGATCTCCTCCGCGTCCCTGCCCTGACATGCAGCGTGTGGAGTCAGAAAGCACGCTA
1060         K   L   T   P   V   G   T   G   L   A   S   S   P   D   S   S   S   D   S   S   T   G   S   P   P   R   P   C   P   D   M   Q   R   V   E   S   E   S   T   L
4681 CCCGGTGTGATCATCTCGCAAAACTACCAGTACACAGAGTTCTTTCTAAAACTGGGCAGTGATCTCGAGCACCGGTCGTCTTCGGACAGTGCAAAAGTTTTACTACACCTG
1100         P   G   V   I   I   S   Q   N   Y   Q   Y   T   E   F   F   L   K   L   Y   Q   L   G   S   D   L   E   H   G   R   L   R   D   S   A   K   V   L   L   H   L
4801 CTTCCCTGCGACCGCCAGACGATCCGTCAGCTAAAGATTATGTGCAAGGTGCCGAAGGCTGCCGTAACGGTGGCTGTGACGGGTGACAAGATCGCTAAGGATGAGGAGGAGAAACTGTAC
1140         L   P   C   D   R   Q   T   I   R   Q   L   K   I   M   C   K   V   P   K   A   A   V   T   V   A   V   T   G   D   K   I   A   K   D   E   E   E   K   L   Y
4921 CCCACTGAGCAGGCGGGGATTGAAGATGAAGAGgtaagactatagttagtcgagttagtcattcaggaaagctataatggttatttaatttaaattgtttttctgtatcaatctcttaat
1180         P   T   E   Q   A   G   I   E   D   E   E
5041 atctcttcttttagGAACACTGCACGCCAGAGCAAATGTTTTTGCACCCCACACCTGCTCAAGTGCTATACAACCTTAGTGTGTTGCACGGACTGCTGATTCCCGCGCTAGATCCACTA
1191                         E   H   C   T   P   E   Q   M   F   L   H   P   T   P   A   Q   V   L   Y   N   L   S   V   L   H   G   L   L   I   P   A   L   D   P   L
5161 GGAGAAAGTGCCCTACTAGTGCAGTCTGCATGGATGCATTCTGGCTGCGCTCATTCGTCTGGAACTGTTGACCAAAAACAATTTTCTGCCTAGCGCCGATATGCACACAAAGCGTGCG
1226         G   E   S   A   L   L   V   Q   S   A   W   M   H   S   G   C   A   H   F   V   L   E   L   L   T   K   N   N   F   L   P   S   A   D   M   H   T   K   R   A
5281 TCCTTCCAGTGTGTGTTGAGGTTGGCAAAGCTGTTCCTGTACATAGTTGGCAGTGTATTGTCTCGTGTAGGCGATGAACCCATGATCTGCGACCTTGACAACGGATCCCGTTCCCAGGTT
1266         S   F   Q   C   V   L   R   L   A   K   L   F   L   Y   I   V   G   S   V   L   S   R   V   G   D   E   P   M   I   C   D   L   D   N   G   S   R   Q   V
5401 GACATCCTGAAGCAGAACTTCTCCACGATGCCCAGTAGCTCACAGGGAACATTGAGGGCGATCTCCGCGAAACTTGCTGTGATCCTTGCCCGCGAGATGCTTTCAGCCAGCCCAGAAGGT
1306         D   I   L   K   Q   N   F   S   T   M   P   S   S   S   Q   G   T   L   R   A   I   S   A   K   L   A   V   I   L   A   R   E   M   L   S   A   S   P   E   G
5521 GATCGGTGCCGAACGCTATTTAGCTCCACACTGCAGTGGTCATGTCCCGACATTTCAACCATTAAGGCGGTAGTGCAACTGGCCTGGGCCTCGTCTTGCGGAAATCTTCAGGCTCTAGGC
1346         D   R   C   R   T   L   F   S   S   T   L   Q   W   S   C   P   D   I   S   T   I   K   A   V   V   Q   L   A   W   A   S   S   C   G   N   L   Q   A   L   G
5641 AATAGTAGGCGGCGACTTTGAAGACGGAGGTGATCGTGCCGGACGGGCAAGACTTTAGCATGTGCAAGGAGGCGGCTAGAGGTTCTCACCATTTCGTTTATTCTGAATCCGAGTGCCAACGAG
1386         N   S   S   G   R   R   L   *   R   R   *   S   C   R   T   G   K   T   L   A   V   C   K   E   A   A   *   V   L   T   I   S   F   I   L   N   P   S   A   N   E
5761 GCGTTGACTAGCGATCCGAACTGGCCAAAGTTCATCACCTCCATTGTTCTGAAGAACCCGCTTCGCCACGTGCGTCAGGTGGCCTCTGAGCAGTTGTTTTTGGCTTCTACCTATTGCGCT
1426         A   L   T   S   D   P   N   W   P   K   F   I   T   S   I   V   L   K   N   P   L   R   H   V   R   Q   V   A   S   E   Q   L   F   L   A   S   T   Y   C   A
5881 GGGGATCGGCGACCGTTCGTATACATGGTTAACTTGCTGGTTGGCGCGTTGAAGACGCTGGTTCCCCAGTACGAGTCAACTTGCGCTGAGTTCTTCTCGGTGCTGTGTCGGACCCTTTCT
1466         G   D   R   R   P   F   V   Y   M   V   N   L   L   L   V   G   A   L   K   T   L   V   P   Q   Y   E   S   T   C   A   E   F   F   S   V   L   C   R   T   L   S
```

6001 TATGGATGCATCTACAACTGGCCGCTGCAGATTAGCGAGGGATTGTTAGGCGATGAGATTAAGTGGCTGCAACGCATTCGCGAAAACGTCCACGCCACCGGCGACACACAAGTGCATGAG
1506 Y G C I Y N W P L Q I S E G L L G D E I K W L Q R I R E N V H A T G D T Q V H E

6121 GAGCTCCTTGAAGGCCACCTCTGCCTTGCCAAAGAGCTAATGTTTTTCCTTGGCGCCGATTCAAAGGCGCAGCTGAACGAGCTCATCCACGAACTGATCGACGACTTCCTGTTCACGGCC
1546 E L L E G H L C L A K E L M F F L G A D S K A Q L N E L I H E L I D D F L F T A

6241 TCTCGCGAATTTCTGCATTTACGGCGTCATGGCAGTCTACGGCAGGACACCGTCCCGCCACCAGTTTGTGGCAGTCCGCACACAATCGCCGCTGCTTGCGATCTCCTCATTGCCTTGTGC
1586 S R E F L H L R R H G S L R Q D T V P P P V C R S P H T I A A A C D L L I A L C

6361 CAGCTCTGTGTTCCTAATATGAAGCTACTCACCAATACACTTATAGACTTCGTCTGCACGGgtgagtgttggaatcgtaagatcatttcatgattaaatgtttcctcaatttttgtagA
1626 Q L C V P N M K L L T N T L I D F V C T   D

6481 TACTGATCCGTTGCGCGAATGGGATTACCTGCCGCCGGTGGGCGCCAGGCCAACTAAAGGTTTCTGCGGACTAAAGAACGCAGGTGCCACTTGCTACATGAATTCGGTGTTGCAGCAGCT
1647 T D P L R E W D Y L P P V G A R P T K G F C G L K N A G A T C Y M N S V L Q Q L

6601 TTACATGGTACCAGCGGTCCGGGTGGGTATTCTGCGAGCCCACGGCGGTGCCACCACCGACGGCGAGGATTTCAGTGGCGATTCCGATTTGACGGGAGGCGGTCTTGGATCGGCCCTATT
1687 Y M V P A V R V G I L R A H G A A T T D G E D F S G D S D L T G G G L G S A L F

6721 CTCGGGTCCCGCTTCTGCCCTAGTCAGCCTCCCCTCGTCGTCGTCAACTATCGAAGACGGTTTGCACGATGTGCGAAAGAACTATCACGTTGTGATCCTGAAACATGTGCAGGCCATATT
1727 S G P A S A L V S L P S S S T I E D G L H D V R K N Y H V V I L K H V Q A I F

6841 CGCCCACCTGGGCCACAGTGCGTTGCAGTATTACGTTCCACGTGGTCTCTGGACACATTTAAGtacaaaattgtttcattgattggcatcacaatgatattaatatattataattttttg
1767 A H L G H S A L Q Y Y V P R G L W T H F K

6961 tttgctagGCTACTGGGAGAGCCCGTGAATCTTCGCGAGCAGCAGGATGCCGTGGAGTTCTTCATGTCCTTGCTTGAAAGTCTCGATGAAGGACTTAAGGCGCTAGGCCAGCCACAGCTT
1788   L L G E P V N L R E Q Q D A V E F F M S L L E S L D E G L K A L G Q P Q L

7081 ATGAATGCCACTCTGGGCGGTTCGTTCAGCGACCAGAAAATCTGCCAAGAGTGTCCCCATCGCTACTCCAAAGAGGAACCATTTAGTGTATTTAGTGTCGATATTAGGAATCATAGCTCA
1825 M N A T L G G S F S D Q K I C Q E C P H R Y S K E E P F S V F S V D I R N H S S

7201 TTAACCGAATCTCTGGAGCAGTATGTCAAGGGAGAGCTGCTCGAGGGAGCCGATGCATACCATTGTGATAAATGTGATAAAAAAGtaatgttctcatctctgatattcatatacaattaa
1865 L T E S L E Q Y V K G E L L E G A D A Y H C D K C D K K

7321 tttttattaacatcattttttgtgtccaacagGTAGTTACCGTTAAGCGAGTGTGCGTTAAGAAGCTGCCACCCGTGTTAGCCATTCAACTAAAGCGTTTCGAATACGACTACGAACGCGT
1893   V V T V K R V C V K K L P P V L A I Q L K R F E Y D Y E R V

7441 TGTGCGATTAAATTTAATGATTATTTTGAATTTCCACGTATCTTGGATATGGAACCATACACA.GTGTCTGGCCTAGCTAAGCTAGAGGGCGAGGTGGTGGAAGTGGGTGATAATTGTCAA
1923 C A I K F N D Y F E F P R I L D M E P Y T V S G L A K L E G E V V E V G D N C Q

7561 ACAAACGTTGAGACGGACGAAGTACGAACTGACCGGCATTCTGGTGCACAGCGGACAGGCTTCCGGCGGTCACTACTTCAGTTACATACTCTCCAAgtgcgtaccataacaatgttacagt
1944 T N V E T T K Y E L T G I V V H S G Q A S G H Y F S Y I L S K

7681 tttttttaagactaattcaatacaacacctgctccagAAACCCAGCAAACGGCAAGTGTCAGTGGTATAAGTTTGACGACGGCGAAGTCACCGAGTGCAAGATGCACGAGGACGAGGAAA
1982   N P A N G K C Q W Y K F D D G E V T E C K M H E D E E

7801 TGAAGGCGGAGTGCTTTGGCGGCGAGTACATGGGCGAAACCTACGACAACAATCTTAAGCGCATGCAGTACCGTCGACAGAAGCGCTGGTGGAATGCCTACATGTTGTTCTACACCCGCT
2002 M K A E C F G G E Y M G E T Y D N N L K R M Q Y R R Q K R W W N A Y M L F Y T R

7921 GTGATCAAACTCCCGTGCAGTATGAACCCAGCGTGGAGCAGTTGTCACTCGCAGAAAGTCGAAACATGGTTTTACCATTGCCAAAGCCCATTGAGCGCAGCGTGCGgtgagtggaaatag
2042 C D Q T P V Q Y E P S V E Q L S L A E S R N M V L P L P K P I E R S V R

8041 tgcttgaccactatacagatacttttttaattatttttttttcgattgcagACACCCAGAACATTCGGTTTCTCCACTCCCGCAGCATTTTTTCCGTGGAGTTCTTTAACTTTATCAAAAAGT
2082   H Q N I R F L H S R S I F S V E F F N F I K K

8161 TGGTCAGCTGCAATCTACTTTCTGCTCGGAGCAATAAGATCGtcagtttgtattcaagtggtttttatttcccgtttaaaaaaactattatttttttaacagACTCCTGCTGCAGAAGAGC
2101 L V S C N L L S A R S N K I   T P A A E E

8281 TTTCACTGCTGGGCGTCCAATTAGCATCGCAGTTTCTTTTCCACACCGGCTTCCGCACGAAAAAGTCTCTGCGGAGGCCCCGTTATGGAATGgtgagtagcttgggaggtattgactcggg
2135 L S L L G V Q L A S Q F L F H T G F R T K K S L R G P V M E W

8401 tataattaggatacatcttctgctttacagGTATGACGCGCTCTCACACCACATACGTTCCTCGGCACTAGTTCGCAAGTGGTTCGCTAATCATGCGCTCCTCTCGCCGCCATCGCGTTT
2161   Y D A L S H H I R S S A L V R K W F A N H A L L S P P S R L

8521 AGGCGAGTACATTCTGATGGCTCCGTCTCCGGATGTCCGTACTGTCTTTGTCAAGTTGGTGGTTTTTTTCTGCCATTTTGCCATCAACGATGAACCTCTGACTGGCTATGACGGCGCTAA
2181 G E Y I L M A P S P D V R T V F V K L V V F F C H F A I N D E P L T G Y D G A N

8641 TCTCTGCGAGCAGGTACTCATCAGCGTGTTGCGCCTCTTGAAATCCGAGGCAGCCGACTATGGCAAACACCTGCCCCACTATTTTAGCCTTTTCAGTATGTACGTGGGACTAGGAACACG
2221 L C E Q V L I S V L R L L K S E A A D Y G K H L P H Y F S L F S M Y V G L G T R

8761 GGAAAAACAGCAACTGCTTAGGgtgagaactatttacaataagattgcttgccattctaaacgattatgttcgtttgcagCTCAATGTGCCGCTGCAGTTTATTCAGGTGGCGTTGGATG
2261 E K Q Q L L R   L N V P L Q F I Q V A L D

8881 ATGGCCCCGGTCCGGCAATAAAATACCAGTATCCAGAGTTCAGCAAGCTACACCAAGTGGTTTCGCACCTAATACGCTGCAGCGATGTAAGCGAAAAATGCCAGAGCTCCAACCAGAACG
2289 D G P G P A I K Y Q Y P E F S K L H Q V V S H L I R C S D V S E K C Q S S N Q N

9001 CCCGTCCACTGTCCAATCCGTTCAAAGATCCCAATGTTGCGCACGAGGAGCTGACGCCCCTGTCCACTGAGTGCATGGACCTGCTCTTTAACCGCACGGGGtaaggtgcatggcatttta
2322 A R P L S N P F K D P N V A H E E L T P L S T E C M D L L F N R T G

9121 ctgataactgtgctcagttcagagagcatttcaagccatttccttatgaaagaaacgtacagtgagtgcttacgatccgttttgtcatttttttttcagCTACATCAAGAAAGTGATCGA
2362   Y I K K V I E

9241 AGACACTAACGTGGGCGACGAGGGTCTGAAGCTGTTGCAATACTGCAGCTGGGAGAATCCACATTTCTCTCGCGCAGTGCTAACCGAATTGTTGTGGCAGTGTGGGTTCGCCTATTGTCA
2376 D T N V G D E G L K L L Q Y C S W E N P H F S R A V L T E L L W Q C G F A Y C H

9361 CGACATGCGTCACCACACTGATCTGCTTCTGAACATCCTGATCGACGACTCGTGGCAGCACCATCGCATACACAATGCTCTCAACGGAGTGGCGGAGGAACGTGAGGGTCTGTTAGA
2402 D M R H H T D L L L N I L L I D D S W Q H H R I H N A L N G V A E E R E G L L E

9481 GACGATACAGCGGGCGAAGACGCACTACCAGAAGAGGGCGTACCAAATAATCAAGTGCCTGACGCAACTTTTCCATAAGTCGCCGATTGCGCTGCAGATGCTCCACACAAATTCGAACAT
2442 T I Q R A K T H Y Q K R A Y Q I I K C L T Q L F H K S P I A L Q M L H T N S N I

9601 TACTCGGCACTGGAGCATTGCAGTGGAGTGGCTGCAAGGCGAACTAGATCGGCAGCGGGGCATCGGTTGCCAGTACAATTCATACTCGTGGTCGCCGCCGCGCAGAGTAACGACAACAC
2482 T R H W S I A V E W L Q G E L D R Q R G I G C Q Y N S Y S W S P P A Q S N D N T

9721 TAATGGCTACATGCTAGAGCGATCGCAGTCTGCTAAAAACAGTGGTCCATGGCCTTCGAGACTCTGTCCAGATGAGGTCAGCGAAAAAACGgtatctaaaacataacaattatattaaag
2522 N G Y M L E R S Q S A K N T W S M A F E L C P D E V S E K T

9841 aatactttattctattgatcgttcattataatagGATGAAAATAATGAGCCGAACTTGGAGACGAACATGGATGGAACAAATCGGAACCGGTGGCTCAGCCAGGAGGAGTGCTTGAGGG
2562   D E N N E P N L E T N M D G T N K S E P V A Q P G V L E G

9961 ATCTACCGGCGGCACAGAGCAGTTGCCCGAAAATAAGACGCCAACCACAAGCAGCCCGTCGACTGCCGCCTGGCAGCGCGTGGGGATAGCAACGCTATTCCACGACTTTCGCGTCAGCT
2581 S T G G T E Q L P E N K T P T T S S P S T A A W P A R G D S N A I P R L S R Q L

10081 ATTTGGGGCTTACACTTCGACTGGCAGTGGCTCCACTAGTGGAGGTTCCGCGCGCAACTTCCGCCTTGACGACAACGGCGGGTAGTGGGGCTAACAGCGAAACAGAAAGCAGCGCCCAGGA
2621 F G A Y T S T G S G T S G G S A P T S A L T T A G S G A N S E T E S S A Q E

10201 AACAACCGGCGAGACGACCATTAACGGACTGACAAACAGCCTGGACCAAATGGAGATAACGGCCAAAAAGgtaactcgggcgaacaatgtctaggacaagtgtccatctctataaggtgtgt
2661 T T G G E T T I N G L T N S L D Q M E I T A K K V T R A N N V *

10321 gtccgttgccgcaattgaagagttcgagacgtctaaaaacagtcccagaaaactagtaaacaCaacaaacacattttacactggaactccaatctgtatctttctatctgtcttcttaa

10441 tacgtggatagaagatatgtagtttagtcaaactaaacatgtcctcccgacgttcttttccatttatatttatgtttactaaaagacactatgtactttaacaagaacgttggttaatat

10561 aaaaatattggttcgcttcgactgtttaccccttcggttttcgaagatccacgaccttcttatttaactcttctttaatatttctaccatacacaacatcaccctcggtaacccttctgcg

10681 tttcggtgtttaatcctttattgtatatgaatctttttaatccttcagtaacaacacagcaggtgtcgtgaagtcgcccaaaggaggcaagctcccccccggtgctgtttcaccatcaccagttgccgcccagag

10801 caacaacttaaactgaagcagcaacaacagcagcagcagcaacaacagcatcagcaacttcacgatcccggcacaaatgagatgcacttgccatcccccggatctgtcttcccctttggctt

10921 tttatttcctttactgcaagttttctctacttatccatgcaaaaccacgcacgccacaatcagtccacacacacaaaatcacagggggtccaacaaattgaatgtcaacgaagtcaacgaa

11041 aactagttgaaatacgatggatatgtttcttaattcgtactttttaaataaaaaaactaagctgagaatcttaaaaatcattgtttttttgtatgttatattattgaaa

11161 agagtgtttttttggttcttcagcgaattatttcaacgaaaaatatacaaatacaaaatatcagtatatacatttcatataaaaaaagatatcctagattaaagtacataactaatcgtatta

11281 gactttgacctattaagatactactttatgatccactttttgaggcgtgaatgaacccgaaaacactcgaagcaagattgcattacgcttagctggcacccgtagcactgggcgatagatt

11401 gtcaacctccactaaccgaactcttttttgttggctattcacagAAGTGCCGCAGGGTGATAATAAAAAAGCTAGTGGAAAGTAAGGACGAAGAGGACGCAACTaCTGCTACCACAGCGGC
2705   K C R R V I I K K L V E S K D E E D A T S A T T A A

11521 CACCACAGAGGTCACAACATCGCCAGCGACAGCATCGCAACGGCAGCAACTTTAGAGaTGCCGGAATGAGTGAGCTCACGACGATGGTAGAGAAAAATCTTATAATAAGCCAAGAAAATC
2710 T T E V T T S P A T A S Q R Q Q L *

11641 CACAAGCGAAAAGCTCATTGCAATAAAATGTATTCGCAACCTAGCTGGAGGAGAAGATGATGGTATCCATAGTGTTAAATGTAACCAAAGCAGCCTGACAATGAGAACGGTGTGGACAGG

11761 TGTTTACTAGTGGGGATCCCCACCGTCGAAAACGAGACTTAATGGGTAAATACAGAAAAAGAGCAGAATCCGTAGATATTAACAAAGGATACTTGGTTGGAGCGATGCGAGTCAAGGAAC

11881 GAAAGTCTGGTAGATAAGGAGATCAAAAACAAGCAAAGGTGTTTATTAGCTATCGAACTTTATAATTACCTTTAATAATAAACAAATACACCAAT

Fig. 6. For legend see p. 996

Fig. 6. DNA and predicted protein sequences of the *faf* gene. A composite genomic and cDNA sequence of 11,975 nt of the *faf* locus is shown, including the region containing all three cDNAs shown in Fig. 5A, and 360 nt upstream. The sequenced nucleotides are numbered in plain type at the left, starting with 360 nt upstream of the beginning of cDNA 3-2, indicated by the **G** at position 361. cDNA 3-2 has 4 additional bases (CACG) upstream of that G that are not contained in the genomic sequence, which could either be an artifact of cDNA cloning, or could indicate an intron/exon boundary. We have not determined whether the G at nt 361 is the actual transcription start site. The predicted protein sequence is shown using the single letter amino acid code, the letters placed beneath the middle nucleotide of each codon. Stop codons are indicated with asterisks (*). The numbers in bold at the left refer to the protein sequence. The ATG at position 605 is probably the start of translation as there are no other ATGs upstream in the 3-2 cDNA sequence, there are stop codons in all three reading frames within the 3-2 cDNA sequence upstream of nt 605, and the sequences upstream of the ATG are a good match with the *Drosophila* translation initiation consensus sequence (Cavener, 1987). The PEST sequence similarity at amino acids 18-34 (Rogers et al., 1986) and the leucine-zipper

motif at amino acids 269-290 (Landschulz et al., 1988) are underlined. Introns are in lowercase. All of the intron/exon boundaries conform to the splice junction consensus sequence. Efforts to determine the sequence of the intron located between nt 7503 and 7504 were abandoned because plasmid subclones containing this region and of a suitable size for sequencing were unstable and produced various deletions with at least one end within the intron. The sequence of this region was determined on both strands of two independent cDNAs (7-3 and 6-5). The **c** at nt 10383 is the 3′ end of cDNA 7-3. As 7-3 is not polyadenylated, we do not know the exact location of the actual 3′ end of this spliced form of the *faf* mRNA. cDNA 6-5, however, is polyadenylated just after the final nt shown. The **C** at nt 11578 indicates the end of the genomic DNA sequence determined. There are eight single base differences between the genomic and cDNAs, which are indicated by underlined, bold letters. The sequence shown is the genomic sequence, and the corresponding changes in the cDNA are **G**-T (2638), **C**-T (4347), **G**-C (7605), **C**-T (8500), **T**-C (8563), **C**-T (8980), **T**-C (9754), **C**-G (10111) and **A**-T (11504). All the differences are third base codon substitutions except for position 11565, which is a first base substitution changing the amino acid from S to T.

---

stained for β-galactosidase activity. Each *Drosophila* ovary is composed of about twenty ovarioles, which are strings of egg chambers in successive stages of maturity. Each egg chamber contains one oocyte and fifteen nurse cells, which provide the oocyte with RNAs and proteins, including pole plasm components which are transported to the oocyte posterior pole. β-galactosidase activity stain is seen throughout oogenesis within the nurse cell-oocyte cluster (Fig. 7C-E). When the oocyte becomes morphologically distinct from the nurse cells, cytoplasmic stain appears to be confined to the nurse cells (Fig. 7C). Later, high levels of stain are seen at the posterior pole of oocytes (Fig. 7D). The posterior staining persists throughout oogenesis (Fig. 7E), and after the egg is laid, faf-lacZ protein at the posterior becomes incorporated into the pole cells (Fig. 7F). The β-

galactosidase activity in the pole cells is not due to zygotic transcription, as embryos from wild-type females mated with *faf/lacZ* males do not show pole cell staining (data not shown).

To determine whether *faf* mRNA, as well as faf-lacZ protein, is posteriorly localized in oocytes, we used a *faf* cDNA probe for whole mount in situ hybridization to ovaries of wild-type females (Materials and methods). The staining pattern observed is identical to that seen in ovaries of *faf/lacZ* transformants histochemically stained for β-galactosidase activity, except that no posteriorly localized mRNA is seen in oocytes (Fig. 9). Thus, faf-lacZ protein, not *faf* mRNA, is posteriorly localized, and the first 392 amino acids of faf protein is sufficient for localization.
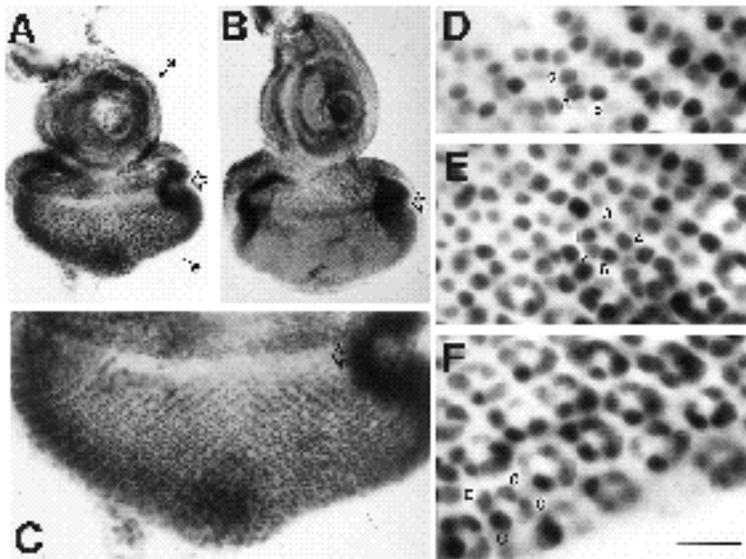
Fig. 8. faf-lacZ expression in third instar larval eye-antennal discs. (A) An eye-antennal disc from a *fafnuc/lacZ* transgenic third instar larvae stained with an anti-β-galactosidase antibody (Materials and methods). Nuclear staining is seen throughout both the antennal portion of the disc (a), and the eye portion (e), except within the morphogenetic furrow, indicated by the arrow. (B) A wild-type eye-antennal disc stained with an antibody raised against a ubiquitous nuclear antigen (gro; see Materials and methods) is shown to demonstrate that nuclear staining present in the furrow (arrow) is readily visible. (C) A close-up of the eye portion of the disc in A, showing the nuclear staining within assembling facets and the absence of stained nuclei within the furrow (arrow). (D-F) Enlargements of progressively more posterior portions of the disc shown in A and C (the direction of the furrow is straight across the top of all panels), showing that *fafnuc/lacZ* is expressed in the nuclei of all photoreceptors (1-8) and cone cells (c). We could not specifically find stained mystery cell nuclei,

probably because they are not usually apically located. However, stained basal nuclei of undetermined cells were observed (not shown). Also, peripodial membrane nuclei were stained. The bar (in F) is 10 μm in D-F, 40 μm in C and 120 μm in A and B.

Fig. 7. Expression of faf-lacZ in eye discs and ovariesassayed by β-galactosidase activity staining. (A) The two *faf-lacZ* hybrid genes introduced into the *Drosophila* genome by P element transformation are shown (Materials and methods). *faf/lacZ* contains the first 4.4 kb of the *fafIII* genomic DNA fragment (Fig. 5A), including ~2.1 kb upstream of the putative transcription start site (indicated by the arrow), the 5′-untranslated sequences and DNA encoding the first 392 amino acids of the faf protein, including the first two exons (black boxes), two introns and part of the third exon. The *faf* coding sequences are fused in frame after amino acid 392 to *E. coli* sequences encoding lacZ (stippled box), followed by termination sequences from SV40 (Materials and methods). *fafnuc/lacZ* is identical to *faf/lacZ* except that an oligonucleotide containing an 11 amino acid nuclear localization signal (hatched inverted triangle) was cloned in frame just after the first 54 amino acids (Materials and methods). (B) A histochemically stained (Materials and methods) third instar larval eye-antennal disc from a *fafnuc/lacZ* transgenic line shows β-galactosidase activity in the antennal portion (a) of the disc and in the eye portion (e), anterior to the morphogenetic furrow (mf). In C-J, posterior is to the right. (C-E) Portions of stained ovarioles from a *faf/lacZ* transgenic female are shown. β-galactosidase activity is detected in the germarium (g in panel C) and in the egg chambers at all subsequent stages of oogenesis. When the oocyte (o) becomes morphologically distinct from the nurse cells (n), the β-galactosidase activity is apparently restricted to the nurse cells, as the oocyte appears unstained at stage 8 (S8). (Stages are according to Mahowald and Kambysellis, 1980). (Some faf-lacZ protein must be present in stage-8 oocytes, however, because in *fafnuc/lacZ* lines, stain is seen concentrated in stage-8 oocyte nuclei). At stage10 (D), β-galactosidase activity is detected not only in the nurse cells, but also at the oocyte posterior pole. The posterior staining persists through the final stage of oogenesis (E-stage 14). The same results were obtained in vls^{RB71}/vls^{RB71}, tud^{WC8}/tud^{WC8}, pum^{680}/pum^{680} and nos^{L7}/nos^{L7} mutant females (Materials and methods). (F) A stained embryo from a *faf/lacZ* transgenic mother. Maternal β-galactosidase activity persists at the posterior (see text) and becomes incorporated into pole cells (pc). (G,H) Stage-10 (G) and stage-11 (H) oocytes from an osk^{150}/osk^{346} mutant femalecontaining a *faf/lacZ* P element. β-galactosidase activity is present in the nurse cells, but there is none at the oocyte posterior pole. The same result was obtained in $capu^{RK}/capu^{RK}$, $spir^{RP}/spir^{RP}$ and $stau^{D3}/stau^{HL54}$ females (Materials and methods). (I) A portion of an ovariole from a weak *vas* mutant female (vas^{011}/vas^{011}) carrying a *faf/lacZ* P element. Particulate β-galactosidase staining is seen throughout the ovariole, and posteriorly localized stain is present in stage 10 oocytes. Similar results were obtained with vas^{D1}/vas^{D1}, a strong allele. (J) A portion of an ovariole from a vas^{D1}/vas^{D1} female carrying a *faf/lacZ* P element, showing that the particulate staining is outside of the nucleus. We do not yet understand the significance of the particulate staining.

Fig. 9. *faf* transcripts in ovaries. *faf* transcripts were detected in wild-type ovaries by whole-mount in situ hybridization of a digoxigenin-labeled *faf* cDNA probe (Materials and methods). *faf* transcripts were detected in the germarium (G) and in the egg chambers in subsequent stages of oogenesis. At stage 8 (S8), the oocyte (o) is easily distinguishable from the nurse cells (n) and *faf* transcripts appear to be restricted to the nurse cells, through stage 10. At stage 10, the time that faf/lacZ protein first appears at the oocyte posterior, no posteriorly localized *faf* transcripts are detected.

Localization of faf-lacZ protein in oocytes depends on osk

We next asked whether posterior localization of faf-lacZ

protein depends on the posterior group genes. Ovaries of females homozygous mutant for each of nine posterior group genes, that also carry a copy of the *faf/lacZ* P element, were stained for β-galactosidase activity (Materials and methods). β-galactosidase activity was detected at the oocyte posterior in *vas* and mutants downstream, but not in *osk* or mutants upstream in the hierarchy (Fig. 7G-I and legend). As stau protein and *osk* RNA and protein are posteriorly localized earlier than faf-lacZ protein, and vas protein is localized at about the same time, *faf* appears to fit into the posterior group hierarchy.

Insights into the specific roles of some of the posterior group genes were recently obtained using transgenic flies containing an *osk* gene whose RNA is mislocalized to the anterior pole of oocytes (Ephrussi and Lehmann, 1992). Ephrussi and Lehmann have shown that *osk* RNA at the anterior is sufficient to localize *nos* RNA and also to recruit all of the pole plasm components required for the formation of functional pole cells, which include vas and tud, but not capu, spir, stau or vls. Thus, the authors conclude that capu, spir and stau are required only to tether *osk* gene products to the posterior pole, while osk, vas and tud are essential for pole plasm formation and *nos* RNA localization. (The role of vls is uncertain.)

We find that faf-lacZ protein colocalizes with *osk* gene products at the anterior pole in embryos of transgenic mothers carrying one copy of *faf/lacZ* and one copy of anteriorly localized *osk* (Fig. 10). As faf-lacZ localization in oocytes is independent of genes downstream of *osk* in the hierarchy, and genes upstream of *osk* appear to be required only to tie *osk* RNA to the posterior pole, we conclude that faf-lacZ protein localization is likely to depend only on *osk* gene products.

Pole plasm components are posteriorly localized in faf *mutant embryos*

The observation that faf-lacZ protein localization depends on *osk* suggested that posterior localization of downstream gene products (vas protein and *nos* RNA) may depend on *faf. cyclinB* mRNA may also be considered downstream of *faf* as it is posteriorly localized, dependent on *vas* and genes upstream (Raff et al., 1990). Using *nos* and *cyclinB* cDNA probes for in situ hybridization, and an antibody to vas protein, we find that all of these gene products are posteriorly
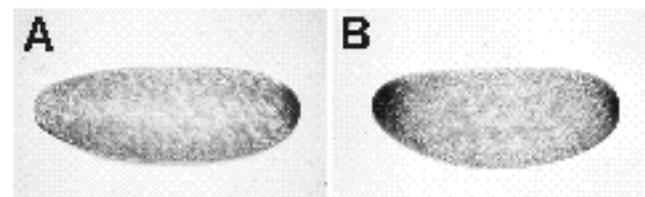
Fig. 10. Posterior localization of faf-lacZ protein depends on *osk*. Pre-syncytial blastoderm embryos from mothers with one copy of the *faf/lacZ* P element (A), or from mothers with one copy of the *faf/lacZ* P element and one copy of a P element containing an anteriorly localizing *osk* gene (B) were stained with an anti-β-galactosidase antibody (Materials and methods). Posterior is to the right in both panels. faf-lacZ protein colocalizes with both wild-type and anterior *osk* RNA in B.
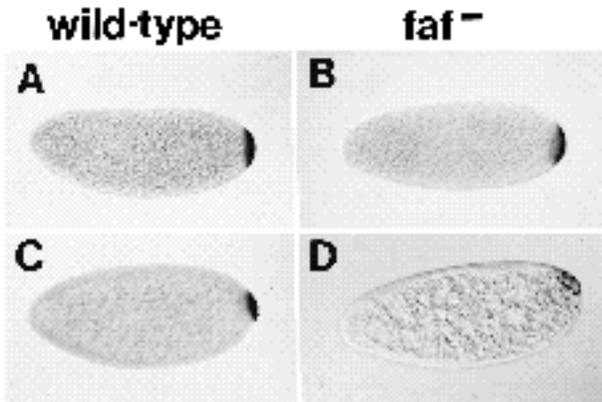
Fig. 11. *nos* RNA is posteriorly localized in *faf* mutant embryos. *nos* transcripts were detected in 0-4 hour embryos from wild-type (A,C) and *faf^FO8* (B,D) mothers by in situ hybridization of whole embryos with a digoxigenin-labeled *nos* cDNA probe (Materials and methods). (A,B) Posterior localization of *nos* transcripts appears identical in wild-type (A) and *faf* mutant (B) presyncytial blastoderm embryos. *nos* RNA is incorporated into pole cells in wild-type embryos (C) and also in *faf* mutant embryos (D). The pole cells in *faf* mutant embryos are fewer in number and more spread out than in wild-type. The same results were obtained with all *faf* alleles that cause female sterility (see Table 1) and with *faf^BX4/faf^BP*, the allele combination most likely to be null. The same results were obtained when embryos from *faf^FO8* mutant mothers were hybridized with a *cyclinB* probe, or stained with anti-vas antibodies (Materials and methods).

localized in *faf* mutant embryos, and are incorporated into pole cells (Fig.11).

nos *function is unimpaired in* faf *mutant embryos*

nos protein facilitates the expression of genes that direct the abdominal segmentation pattern by preventing translation of *hunchback (hb)* maternal mRNA in the posterior of the embryo (Tautz, 1988; Struhl, 1989; Irish et al., 1989; Hülskamp et al., 1989). Thus, hb protein normally accumulates only in the anterior half of the embryo. By staining *faf* mutant embryos with an antibody to hb protein, we find that maternal hb protein is distributed normally in *faf* mutant embryos (data not shown; Materials and methods). Thus, *faf* is unlikely to be required for *nos* function.

Discussion

The *faf* gene encodes two similar large proteins specifically required during two different developmental processes in *Drosophila:* compound eye assembly and oogenesis. In the larval eye disc, cells outside the assembling facets require *faf* for short-range cell interactions that prevent the mystery cells from becoming photoreceptors. Maternal faf protein appears to be required for nuclear migration and cellularization in early embryogenesis. In addition, a faf-lacZ protein containing only the first 392 amino acids of faf is posteriorly localized in oocytes and its localization depends on *osk*.

faf *and cell interactions in the eye disc*

The *faf* gene appears to function in one of several negative regulatory mechanisms in place in the larval eye disc which ensure that each facet contains only eight photoreceptor cells. We have shown that *faf* is involved in a process whereby the mystery cells are prevented from becoming photoreceptors by cells near to but outside the photoreceptors and mystery cells. argos, a diffusible protein and gro, a nuclear protein, appear to play roles in facet assembly similar to faf (Freeman et al., 1992; Delidakis et al., 1991; Fischer-Vize et al., 1992). Further experiments are required to determine if these three genes function in the same or different pathways. In contrast, analysis of *svp* mutants shows that photoreceptors R3 and R4 also repress neuralization of the mystery cells (Mlodzik et al., 1990b).

Although our results suggest that faf is required in cells near the developing photoreceptors and mystery cells, we have not identified these cells. Presumably, when the mystery cells are influenced by the negative regulatory mechanism involving faf, they are in contact with the cells in which faf is required. The precise time during eye development when faf function influences mystery cell fate is unknown. The first structure to emerge from the morphogenetic furrow is a rosette in which 10-15 cells, including the R8,2,5,3,4 precursors and the mystery cells, form a ring around four or five core cells. The ring then opens and preclusters containing R8,2,5,3,4 precursors and the mystery cells are formed (Wolff and Ready, 1991a). Thus, the cells that require faf to prevent neurogenesis of the mystery cells could be the core cells or cells adjacent to the mystery cells within the ring. Alternatively, if the faf-dependent process occurs later, uncommitted epithelial cells surrounding the precluster could be involved. We also cannot rule out the possibility that faf acts anterior to the morphogenetic furrow.

The expression of *ro* and *svp* by the mystery cells in *faf* mutant eye discs is consistent with the hypothesis that they are R3/4 subtype photoreceptors. In *gro* mutant eye discs, it was shown more conclusively that the mystery cells become R3/4 cells (Fischer-Vize et al., 1992). This repeated observation is interesting because experiments with *ro* show that the normal R3 and R4 cells require inductive cues from R2 and R5 in order to join the facet (Tomlinson et al., 1988). The mystery cells are located between R3 and R4 and are not in contact with R2 and R5. Their determination as R3/4 subtype photoreceptors suggests either that the inductive signal from R2/5 is permissive for photoreceptor cell determination and R3/4 subtype determination is a type of default state, or that R3/4 can send the same instructive signals as R2/5 (see Fischer-Vize et al., 1992).

The pattern of faf-lacZ expression in the larval eye disc

The expression pattern of the faf-lacZ hybrid protein in the larval eye disc is unusual; faf-lacZ is expressed in all cells anterior and posterior to the morphogenetic furrow. This pattern is consistent with the conclusion that faf is required in cells outside the early developing ommatidial clusters. It is possible that faf expression in cells within the assembling facets could be required for faf-dependent aspects of

eye morphogenesis other than excluding the mystery cells, such as orientation and spacing of facets.

The absence of faf-lacZ protein within the morphogenetic furrow indicates that it must be rapidly degraded in cells as they enter the furrow and suggests two possibilities. First, normal faf protein in the eye disc may be unstable, and factors that stabilize it or activate its transcription may be absent within the morphogenetic furrow. Alternatively, there may be a factor present in the furrow that actively degrades faf protein. faf-lacZ protein turnover may be mediated by the PEST sequence contained within the first 392 amino acids of faf. Some caution is warranted in interpreting lacZ hybrid protein expression patterns. However, as lacZ is generally more stable than endogenous *Drosophila* proteins in similar experiments (for example, see Mlodzik et al., 1990a), the normal faf protein is likely to be absent within the furrow.

Functions of faf *in oogenesis*

The localization of faf-lacZ protein to the posterior of oocytes and its dependence on *osk* suggests that faf has an important function at the posterior pole. (Although there may be stability differences between faf-lacZ and normal faf protein, we assume that the the normal faf protein is similarly localized.) As *nos* RNA is localized to and maternal hb protein is eliminated from the posterior of *faf* mutant embryos, a role for faf in abdominal pattern formation, which could have been obscured by the earlier role of *faf* in nuclear migration or cellularization, appears unlikely. Perhaps *faf* plays a role in pole plasm formation. Two pole plasm components essential for pole cell formation, osk and vas proteins, are also localized to the posterior pole of oocytes (Hay et al., 1988b; Ephrussi et al., 1991; Ephrussi and Lehmann, 1992). In addition, vas protein has been shown to be part of the polar granule, a particle composed of RNA and protein, associated with pole cells (Hay et al., 1988b). Although pole cells do form in *faf* mutant embryos, there are fewer than the wild-type number and they are irregularly positioned. Thus, *faf* could play a role in pole cell determination, development or function. Although the pole cell defects observed could be a secondary consequence of the failure of nuclear migration or cellularization in *faf* mutant embryos, this would not preclude a specific role for faf in pole cells as well. Further experiments are required to determine if *faf* mutant pole cells are functional, and to determine if faf protein is part of the polar granule.

The defects in *faf* mutant embryos indicates that faf plays a role in nuclear migration and possibly also in cellularization. It therefore seems puzzling that faf-lacZ is detected only at the posterior pole of oocytes and embryos. The absence of localized faf protein in *faf* mutant embryos does not cause the global embryogenesis defects as localization depends on *osk* and *osk* mutant embryos cellularize normally. The defects in embryogenesis caused by *faf* mutations could be accounted for if some maternal faf protein is distributed throughout the oocyte or embryo. There is some evidence for non-localized faf-lacZ protein in early embryos stained with antibodies to -galactosidase (data not shown). However, as the faf-lacZ protein may be more stable than endogenous faf, an anti-faf antibody would be useful for investigating this possibility. Interestingly,

embryos from *vas* and *vls* mothers not only lack pole cells, but sometimes also display defects in cellularization (Schüpbach and Wieschaus, 1989; Hay et al., 1990; Lehmann and Nüsslein-Volhard, 1991).

References

Baker, N. E., Mlodzik, M. and Rubin, G. M. (1990). Spacing differentiation in the developing *Drosophila* eye: A fibrinogen-related lateral inhibitor encoded by *scabrous*. *Science* **250**, 1370-1377.

Banerjee, U. and Zipursky, S. L. (1990). The role of cell-cell interactions in the development of the *Drosophila* visual system. *Neuron* **4**, 177-187.

Boswell, R. E., Prout, M. E. and Steichen, J. C. (1991). Mutations in a newly identified *Drosophila melanogaster* gene, *mago nashi*, disrupt germ cell formation of mirror-image symmetrical double abdomen embryos. *Development* **113**, 373-384.

Carthew, R. W. and Rubin, G. M. (1990). *seven in absentia,* a gene required for specification of R7 cell fate in the *Drosophila* eye. *Cell* **63**, 561-577.

Cavener, D. (1987). Comparison of the consensus sequences flanking translational start-sites in *Drosophila* and vertebrates. *Nucl. Acids Res.* **15**, 1353-1361.

Delidakis, C., Preiss, A., Hartley, D. A. and Artavanis-Tsakonas, S. (1991). Two genetically and molecularly distinct functions involved in early neurogenesis reside within the *Enhancer of split* locus of *Drosophila melanogaster. Genetics* **129**, 803-823.

Ephrussi, A. E., Dickinson, L. K. and Lehmann, R. (1991). *oskar* organizes the germ plasm and directs localization of the posterior determinant *nanos. Cell* **66**, 37-50.

Ephrussi, A. E. and Lehmann, R. (1992). Induction of germ cell formation by *oskar. Nature* **358**, 387-392.

Fischer-Vize, J. A., Vize, P. D. and Rubin, G. M. (1992). A unique mutation in the *Enhancer of split* gene complex affects the fates of the mystery cells in the developing *Drosophila* eye. *Development* **115**, 89-101.

Foe, V. E. and Alberts, B. M. (1983). Studies of nuclear and cytoplasmic behaviour during the five mitotic cycles that precede gastrulation in *Drosophila* embryogenesis. *J. Cell. Sci.* **61**, 31-70.

Freeman, M., Klämbt, C., Goodman, C. S. and Rubin, G. M. (1992). The *argos* gene encodes a diffusible factor that regulates cell fate decisions in the *Drosophila* eye. *Cell* **69**, 963-975.

Fujita, S. C., Zipursky, S. L., Benzer, S., Ferrus, A. and Shotwell, S. L.

(1982). Monoclonal antibodies against the *Drosophila* nervous system. *Proc. Natn. Acad. Sci. USA* **79**, 7929-7933.

Hay, B., Jan, L. Y. and Jan, Y. N. (1988a). A protein component of *Drosophila* polar granules is encoded by *vasa* and has extensive sequence similarity to ATP-dependent helicases. *Cell* **55**, 577-587.

Hay, B., Ackerman, L., Barbel, S., Jan, L. Y. and Jan, Y. N. (1988b). Identification of a component of *Drosophila* polar granules. *Development* **103**, 625-640.

Hay, B., Jan, L. Y. and Jan, Y. N. (1990). Localization of *vasa*, a component of *Drosophila* polar granules, in maternal-effect mutants that alter embryonic anteroposterior polarity. *Development* **109**, 425-433.

Hafen, E., Basler, K., Edstroem, J. E. and Rubin, G. M. (1987). *sevenless*, a cell-specific homeotic gene of *Drosophila*, encodes a putative transmembrane receptor with a tyrosine kinase domain. *Science* **236**, 55-63.

Heberlein, U. and Rubin, G. M. (1991). *Star* is required in a subset of photoreceptor cells in the developing *Drosophila* retina and displays dosage-sensitive interaction with *rough. Dev. Biol.* **144**, 353-361.

Higashijima, S., Kojima, T., Michiue, T., Ishimaru, S., Emori, Y. and Saigo, K. (1992). Dual *Bar* homeo box genes of *Drosophila* required in two photoreceptor cells, R1 and R6, and primary pigment cells for normal eye development. *Genes Dev.* **6**, 50-60.

Hülskamp, M., Schröder, C., Pfeifle, C., Jäckle, H. and Tautz, D. (1989). Posterior segmentation of the *Drosophila* embryo in the absence of a maternal posterior organizer gene. *Nature* **338**, 629-632.

Irish, V., Lehmann, R. and Akam, M. (1989). The *Drosophila* posterior-group gene *nanos* functions by repressing *hunchback* activity. *Nature* **338**, 646-648.

Kalderon, D., Roberts, B. L., Richardson, W. D. and Smith, A. E. (1984). A short amino acid sequence able to specify nuclear location. *Cell* **39**, 499-509.

Karpilow, J., Kolodkin, A., Bork, T. and Venkatesh, T. (1989). Neuronal development in the *Drosophila* compound eye: *rap* gene function is required in photoreceptor cell R8 for ommatidial pattern formation. *Genes Dev.* **3**, 1834-1844.

Karr, T. L. and Alberts, B. M. (1986). Organization of the cytoskeleton in early *Drosophila* embryos. *J. Cell Biol.* **102**, 1494-1509.

Kim-Ha, J., Smith, J. L. and Macdonald, P. M. (1991). *oskar* mRNA is localized to the posterior pole of the *Drosophila* oocyte. *Cell* **66**, 23-35.

Kimmel, B. E., Heberlein, U. and Rubin, G. M. (1990). The homeodomain protein *rough* is expressed in a subset of cells in the developing *Drosophila* eye where it can specify photoreceptor cell subtype. *Genes Dev.* **4**, 712-727.

Klemenz, R., Weber, U. and Gehring, W. J. (1987). The *white* gene as a marker in a new P-element vector for gene transfer in *Drosophila*. *Nucl. Acids Res.* **15**, 3947-3959.

Landschulz, W. H., Johnson, P. F. and McKnight, S. L. (1988). The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759-1764.

Lasko, P. F. and Ashburner, M. (1988). The product of the *Drosophila* gene *vasa* is very similar to eukaryotic initiation factor-4A. *Nature* **335**, 611-617.

Lasko, P. F. and Ashburner, M. (1990). Posterior localization of *vasa* protein correlates with, but is not sufficient for, pole cell development. *Genes Dev.* **4**, 905-921.

Lawrence. P. A. and Green, S. M. (1979). Cell lineage in the developing retina of *Drosophila. Dev. Biol.* **71**, 142-152.

Lehmann, R. and Nüsslein-Volhard, C. (1987). Involvement of the *pumilio* gene in the transport of an abdominal signal in the *Drosophila* embryo. *Nature* **329**, 167-170.

Lehmann, R. and Nüsslein-Volhard, C. (1991). The maternal gene *nanos* has a central role in posterior pattern formation of the *Drosophila* embryo. *Development* **112**, 679-691.

Lehmann, R. (1992). Germ plasm formation and germ cell determination in *Drosophila. Curr. Opin. Genet. Devel.*, in press.

Levis, R., Hazelrigg, T. and Rubin, G. M. (1985). Separable cis-acting control elements for expression of the *white* gene of *Drosophila. EMBO J.* **4**, 3489-3499.

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

Macdonald, P. M. (1992). The *Drosophila pumilio* gene: an unusually long transcription unit and an unusual protein. *Development* **114**, 221-232.

Manseau, L. J. and Schüpbach, T. (1989). *cappuccino* and *spire:* two unique maternal-effect loci required for both the anteroposterior and dorsoventral patterns of the *Drosophila* embryo. *Genes Dev.* **3**, 1437-1452.

Mlodzik, M., Baker, N. E. and Rubin, G. M. (1990a). Isolation and expression of *scabrous,* a gene regulating neurogenesis in *Drosophila. Genes Dev.* **4**, 1848-1861.

Mlodzik, M., Hiromi, Y., Weber, U., Goodman, C. and Rubin, G. M. (1990b). The *Drosophila seven-up* gene, a member of the steroid receptor gene superfamily controls photoreceptor cell fates. *Cell* **60**, 211-224.

Moses, K., Ellis, M. C. and Rubin, G. M. (1989). The *glass* gene encodes a zinc-finger protein required by *Drosophila* photoreceptor cells. *Nature* **340**, 531-536.

Pirrotta, V. (1986). Cloning *Drosophila* genes. In *Drosophila: A Practical Approach.* (ed. D. B. Roberts), pp. 83-110. Oxford: IRL Press.

Raff, J. W., Whitfield, W. G. F. and Glover, D. M. (1990). Two distinct mechanisms localise cyclin B transcripts in syncytial *Drosophila* embryos. *Development* **110**, 1249-1261.

Ready, D. F., Hanson, T. E. and Benzer, S. (1976). Development of the *Drosophila* retina, a neurocrystalline lattice. *Dev. Biol.* **53**, 217-240.

Reinke, R. and Zipursky, S. L. (1988). Cell-cell interaction in the *Drosophila* retina: the *bride-of-sevenless* gene is required in photoreceptor cell R8 for R7 cell development. *Cell* **55**, 321-330.

Robertson, H. M., Preston, C. R., Phillis, R. W., Johnson-Schlitz, D. M., Benz, W. K. and Engels, W. R. (1988). A stable genomic source of P element transposase in *Drosophila melanogaster. Genetics* **118**, 461-470.

Rogers, S., Wells, R. and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**, 364-368.

Rubin, G. M. (1991). Signal transduction and the fate of the R7 photoreceptor in *Drosophila. Trends Genet.* **7**, 372-377.

Saint, R., Kalionis, B., Lockett, T. J. and Elizur, A. (1988). Pattern formation in the developing eye of *Drosophila melanogaster* is regulated by the homeobox gene, *rough. Nature* **334**, 151-154.

Sambrook, G., Fritsch, E. F. and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Schüpbach, T. and Wieschaus, E. (1986). Maternal-effect mutations altering the anterior-posterior pattern of the *Drosophila* embryo. *Roux's Arch. Dev. Biol.* **195**, 302-317.

Schüpbach, T. and Wieschaus, E. (1989). Female sterile mutations on the second chromosome of *Drosophila melanogaster.* I. Maternal effect mutations. *Genetics* **121**, 101-117.

Spradling, A. C. (1986). P-element mediated transformation. In *Drosophila: A practical approach.* (ed., D. B. Roberts), pp. 175-197. Oxford: IRL Press.

St. Johnston, D., Beuchle, D. and Nüsslein-Volhard, C. (1991). *staufen,* a gene required to localize maternal RNAs in the *Drosophila* egg. *Cell* **66**, 51-63.

St. Johnston, D. and Nüsslein-Volhard, C. (1992). The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**, 201-219.

Struhl, G. (1989). Differing strategies for organizing anterior and posterior body pattern in *Drosophila* embryos. *Nature* **338**, 741-744.

Tautz, D. (1988). Regulation of the *Drosophila* segmentation gene *hunchback* by two maternal morphogenetic centers. *Nature* **332**, 281-284.

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

Tomlinson, A. and Ready, D. F. (1987a). Neuronal differentiation in the *Drosophila* ommatidium. *Dev. Biol.* **120**, 366-376.

Tomlinson, A. and Ready, D. F. (1987b). Cell fate in the *Drosophila* ommatidium. *Dev. Biol.* **123**, 264-275.

Tomlinson, A., Kimmel, B. E. and Rubin, G. M. (1988). *rough,* a *Drosophila* homeobox gene required in photoreceptors R2 and R5 for inductive interactions in the developing eye. *Cell* **55**, 771-784.

Wang, C. and Lehmann, R. (1991). *Nanos* is the localized posterior determinant in *Drosophila. Cell* **66**, 637-648.

Wieschaus, E. and Nüsslein-Volhard, C. (1986). Looking at embryos. In

Drosophila*: A Practical Approach*. (ed. D. B. Roberts), pp. 199-226. Oxford: IRL Press.

Wolff, T. and Ready, D. F. (1991a). The beginning of pattern formation in the *Drosophila* compound eye: the morphogenetic furrow and the second mitotic wave. *Development* **113**, 841-850.

Wolff, T. and Ready, D. F. (1991b). Cell death in normal and rough eye mutants of *Drosophila. Development* **113**, 825-839.

Note added in proof

The GenBank Accession numbers for the sequences reported are: L04958 (short cDNA), L04959 (long cDNA), L04960 (genomic DNA).

dev 2008

Fig. 2. The phenotype of embryos from *faf* mutant mothers. Embryos, collected after 4 hours of egg-laying, from wild-type (A,B,E,F) and *faf^FO8* mutant (C,D,G,H) mothers were labeled consecutively with the DNA stain Hoechst and rhodamine-conjugated phalloidin, which stains f-actin, and photographed using fluorescent light with either fluorescein (A,C,E,G) or rhodamine (B,D,F,H) filters to show nuclei and f-actin filaments, respectively (Materials and methods). The same embryo is shown in each of the horizontal pairs of panels. (A,B) A wild-type embryo fixed during the mitosis that concludes cycle 11. Nuclei are at the periphery (A), and as cellularization does not occur until cycle 14 (below), f-actin is seen associated with each nucleus (B; see Karr and Alberts, 1986). The posterior of the embryo at a deeper focal plane is shown to the right of B, where the pole cells (pc), which form at cycle 10, are visible. (C,D) A *faf* mutant embryo, in which few nuclei have migrated to the periphery (C) although pole cells have formed (D), indicating that cycle 10 has already been reached. (E,F) A wild-type embryo at cycle 14 (cellular blastoderm). Cellularization is apparent as f-actin outlines the cell membranes (F). (G,H) The most mature-appearing *faf* mutant embryos seen. Patches of asynchronously dividing nuclei have migrated to the periphery (G), but no cellularization is apparent (H).